



**HAL**  
open science

# Provisionnement en assurance non-vie pour des contrats à maturité longue et à prime unique : application à la réforme Solvabilité 2

Geoffrey Nichil

► **To cite this version:**

Geoffrey Nichil. Provisionnement en assurance non-vie pour des contrats à maturité longue et à prime unique : application à la réforme Solvabilité 2. Mathématiques générales [math.GM]. Université de Lorraine, 2014. Français. NNT : 2014LORR0200 . tel-01751162

**HAL Id: tel-01751162**

**<https://hal.univ-lorraine.fr/tel-01751162>**

Submitted on 29 Mar 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



## AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : [ddoc-theses-contact@univ-lorraine.fr](mailto:ddoc-theses-contact@univ-lorraine.fr)

## LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

[http://www.cfcopies.com/V2/leg/leg\\_droi.php](http://www.cfcopies.com/V2/leg/leg_droi.php)

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>



UNIVERSITÉ  
DE LORRAINE



UFR MIAE  
École Doctorale IAEM Lorraine  
Université de Lorraine  
D.F.D. Mathématiques

---

Thèse

présentée pour l'obtention du titre de

Docteur de l'Université de Lorraine

en Mathématiques

par

Geoffrey Nichil

---

**Provisionnement en assurance non-vie pour des contrats à  
maturité longue et à prime unique**  
-  
**Application à la réforme Solvabilité 2**

---



Soutenue publiquement le 19 décembre 2014.

Membres du Jury :

**Président du jury**

**Rapporteur**

**Rapporteur**

**Examineur**

**Examineur**

**Directeur de thèse**

**Co-Directeur de thèse**

**Co-Directeur de thèse**

H. Albrecher, Professeur Université de Lausanne

S. Loisel, Professeur ISFA Lyon

V. Scheepers, Commissariat aux Assurances

J. Printems, Maitres de conférences Créteil

P. Vallois, Professeur à l'Université de Lorraine

S. Herrmann, Professeur à l'Université de Bourgogne

M. de Calbiac, Directeur CAMCA Assurance

---

Institut Élie Cartan Nancy, Laboratoire de Mathématiques, B.P. 239, 54506  
Vandoeuvre-lès-Nancy Cedex

---

# Remerciements

Avant de débiter, je tiens à adresser quelques remerciements à toutes celles et ceux qui, directement ou indirectement, ont contribué à la réalisation de ces travaux.

Tout d'abord, je ne saurais remercier suffisamment Pierre Vallois, Professeur à l'Université de Lorraine, pour m'avoir fait l'honneur de diriger cette thèse. Durant ces trois années de collaboration, Monsieur Vallois s'est montré très investi dans l'avancement de nos travaux. J'espère d'ailleurs que notre collaboration ne s'arrêtera pas là. Je tiens également à remercier Samuel Herrmann pour avoir accepté de co-diriger cette thèse.

Je tiens ensuite à exprimer toute ma gratitude à Martial de Calbiac, Directeur de CAMCA Assurance à Luxembourg, qui m'a non seulement permis de réaliser cette thèse professionnelle dans le cadre du dispositif CIFRE (Conventions Industrielles de Formation par la REcherche), mais aussi de mettre en application les problématiques statistiques objet de la présente thèse. Son expérience et sa réputation dans le domaine de l'assurance m'ont été bénéfiques.

Merci à Messieurs Hansjoerg Albrecher et Stéphane Loisel d'avoir accepté d'être mes rapporteurs de thèse; merci également à eux pour leurs remarques utiles. Je ne peux qu'être honoré de la présence de Valérie Scheepers et de Jacques Printems dans mon jury. Merci à eux de me faire cet honneur.

Enfin, je remercie celles et ceux qui se sont intéressés à mon travail et qui ont contribué à leur manière à l'aboutissement de cette thèse : mes collègues de CAMCA Assurance, le groupe de statistiques de l'IECL (Institut Elie Cartan de Lorraine), les autres doctorants (Roberto, Romain, Paul, Yann, etc.) sans oublier mes amis et ma famille, avec une pensée toute particulière pour mes "relectrices" et "traductrice" attentives.

Merci enfin à Stuart pour la traduction ;)



# Table des matières

## Version française

---

---

### I PARTIE 1 : Introduction générale

---

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Généralités et plan de la thèse</b> | <b>3</b>  |
| <b>2</b> | <b>Etat de l'art</b>                   | <b>13</b> |
| <b>3</b> | <b>Conclusion et perspectives</b>      | <b>27</b> |

---

---

---

### II PARTIE 2 : Provisionnement face au risque de défaut des emprunteurs

---

|          |   |           |
|----------|---|-----------|
| <b>4</b> | <b>Provisionnement face au risque de défaut des emprunteurs</b> | <b>33</b> |
| 4.1      | Défaut de paiement des emprunteurs                              | 34        |
| 4.2      | Modélisation de la provision                                    | 36        |
| 4.3      | Distribution de la provision                                    | 43        |
| 4.4      | Estimations des paramètres                                      | 46        |
| 4.5      | Preuves   | 50        |
| <b>5</b> | <b>Compléments au modèle de provisionnement</b>                 | <b>65</b> |
| 5.1      | Garantie caution via les “credit default swap”                  | 65        |
| 5.2      | Provision via les options d'échange                             | 67        |
| 5.3      | Justifications des hypothèses                                   | 70        |
| 5.4      | Application numérique   | 76        |

---

---

---

### III PARTIE 3 : Besoin de solvabilité lié au risque de provisionnement

---

|          |   |            |
|----------|---|------------|
| <b>6</b> | <b>Besoin de solvabilité lié au risque de provisionnement dans un contexte d'ORSA</b> | <b>111</b> |
| 6.1      | Solvabilité 2   | 111        |
| 6.2      | Besoin de solvabilité lié au risque de provisionnement                                | 114        |
| 6.3      | Preuves techniques  | 130        |
| <b>7</b> | <b>Compléments numériques au calcul du besoin de solvabilité</b>                      | <b>139</b> |
| 7.1      | Espérances asymptotique et théorique  | 139        |
| 7.2      | Distribution empirique de la provision  | 142        |
| 7.3      | Besoin de solvabilité empirique   | 147        |

## Version anglaise

---

### IV PART 1 : General introduction

---

|           |                                       |            |
|-----------|---------------------------------------|------------|
| <b>8</b>  | <b>Thesis overview</b>                | <b>159</b> |
| <b>9</b>  | <b>Existing methods</b>               | <b>167</b> |
| <b>10</b> | <b>Conclusion and future projects</b> | <b>181</b> |

---

### V PART 2: Provisioning against borrowers default risk

---

|           |  |            |
|-----------|--|------------|
| <b>11</b> | <b>Provisioning against borrowers default risk</b> | <b>187</b> |
| 11.1      | Borrowers' default                                 | 187        |
| 11.2      | Loss provision modeling                            | 189        |
| 11.3      | Results  | 196        |
| 11.4      | Parameters estimation                              | 199        |
| 11.5      | Proof  | 203        |

---



|  |            |
|--|------------|
| <b>12 Additional information concerning the provisioning model</b> | <b>217</b> |
| 12.1 Guarantee caution with the “credit default swap”              | 217        |
| 12.2 Provision with exchange option                                | 219        |
| 12.3 Justification of the assumptions                              | 221        |
| 12.4 Numerical application   | 228        |

---

**VI PART 3 : Solvency need resulting from provisioning risk in a ORSA context**

---

|  |            |
|--|------------|
| <b>13 Solvency need resulting from provisioning risk in a ORSA context</b> | <b>261</b> |
| 13.1 Solvency 2  | 261        |
| 13.2 Solvency Need resulting from provisioning risk                        | 264        |
| 13.3 Technicals proofs   | 279        |
| <b>14 Additionnal information concerning the solvency need calculation</b> | <b>289</b> |
| 14.1 Asymptotic and theoretical means                                      | 289        |
| 14.2 Empirical distribution of provision                                   | 292        |
| 14.3 Empirical solvency need   | 297        |

*TABLE DES MATIÈRES*

---

# Table des figures

## Version française

---

### Figures de la PARTIE 1

---

|     |   |    |
|-----|---|----|
| 1.1 | Illustration de la garantie proposée par CAMCA Assurance. . . . .   | 6  |
| 1.2 | Processus de recouvrement. . . . .                                  | 9  |
| 1.3 | Provision (négative) concernant les recours. . . . .                | 9  |
| 1.4 | Provision (positive). . . . .                                       | 9  |
| 2.1 | Provision individuelle via le processus de Poisson composé. . . . . | 18 |
| 2.2 | Méthodes collectives Vs Méthodes individuelles. . . . .             | 20 |
| 2.3 | Bilan choqué. . . . .   | 22 |
| 2.4 | Risque de provisionnement. . . . .                                  | 24 |

---

### Figures de la PARTIE 2

---

|      |  |    |
|------|--|----|
| 4.1  | Majoration de la durée des prêts. . . . .                          | 37 |
| 4.2  | Ensembles $A_{t_0-}^i$ et $A_{t_0+}^i$ . . . . .                   | 40 |
| 4.3  | Ensembles $C_{t_0-}^i$ et $C_{t_0+}^i$ . . . . .                   | 48 |
| 4.4  | Synthèse des estimations. . . . .                                  | 50 |
| 5.1  | Les “credits defaults swap”. . . . .                               | 66 |
| 5.2  | Illustration des versements dans un “credit default swap”. . . . . | 67 |
| 5.3  | Déclaration d’un sinistre. . . . .                                 | 71 |
| 5.4  | Capital restant dû Vs Intérêts et intérêts de retard. . . . .      | 72 |
| 5.5  | Remboursement d’une mensualité non payée. . . . .                  | 72 |
| 5.6  | Nombre de prêts par génération. . . . .                            | 76 |
| 5.7  | Statistiques descriptives sur le portefeuille considéré. . . . .   | 76 |
| 5.8  | Histogramme du montant des prêts. . . . .                          | 77 |
| 5.9  | Histogramme de la durée des prêts. . . . .                         | 78 |
| 5.10 | Explication boxplot. . . . .                                       | 79 |
| 5.11 | Boxplot du montant des prêts. . . . .                              | 79 |

---

TABLE DES FIGURES

---

|      |   |     |
|------|---|-----|
| 5.12 | Boxplot de la durée des prêts. . . . .  | 80  |
| 5.13 | Poids (en montant de prêts) par génération. . . . .                               | 80  |
| 5.14 | Boxplot du montant des prêts par génération. . . . .                              | 81  |
| 5.15 | Boxplot de la durée des prêts par génération. . . . .                             | 81  |
| 5.16 | Boxplot du capital restant dû par génération. . . . .                             | 82  |
| 5.17 | Ratio CRD sur montant des prêts. . . . .  | 83  |
| 5.18 | Poids (en CRD) par génération. . . . .  | 83  |
| 5.19 | Poids (en CRD) par génération - Graphique. . . . .                                | 83  |
| 5.20 | Nombre de sinistres. . . . .  | 84  |
| 5.21 | Prêts échus. . . . .  | 85  |
| 5.22 | Synthèse estimation paramètres. . . . .   | 86  |
| 5.23 | Espérance théorique Vs espérance simulée. . . . .                                 | 88  |
| 5.24 | Espérance et intervalle de confiance en fonction de la date d'analyse. . . . .    | 89  |
| 5.25 | Espérance de la provision en fonction des générations. . . . .                    | 90  |
| 5.26 | Ratio CRD Vs Ratio Espérance (%). . . . .   | 91  |
| 5.27 | Ratio CRD Vs Ratio espérance (graphiques). . . . .                                | 91  |
| 5.28 | Nombre de sinistres prévus. . . . .   | 92  |
| 5.29 | Espérance de la provision sous hypothèse de run-off. . . . .                      | 93  |
| 5.30 | Comparaison des méthodes de provisionnement. . . . .                              | 95  |
| 5.31 | Ratio CRD Vs Ratio espérance en fonction des méthodes de provisionnement. . . . . | 96  |
| 5.32 | Ratio Sinistres / Primes en fonction des méthodes de provisionnement. . . . .     | 97  |
| 5.33 | Sensibilité de l'espérance face au taux d'intérêt. . . . .                        | 100 |
| 5.34 | Sensibilité de l'espérance face aux paramètres $\mu$ et $\sigma$ . . . . .        | 102 |
| 5.35 | Comparaison des différents sensibilités en fonction de $t_0$ . . . . .            | 103 |
| 5.36 | Probabilité de nullité de la charge d'un sinistre. . . . .                        | 105 |

---

**Figures de la PARTIE 3**

---

|     |   |     |
|-----|---|-----|
| 6.1 | Ensemble $A_h$ . . . . .  | 115 |
| 7.1 | Espérance de la provision en fonction de $h$ . . . . .                      | 140 |
| 7.2 | Espérance de la provision en fonction de $\mu$ . . . . .                    | 141 |
| 7.3 | Distribution empirique ; $\mu = -0,01$ . . . . .                            | 143 |
| 7.4 | QQ plot ; $\mu = -0,01$ . . . . .   | 144 |
| 7.5 | Distribution empirique ; $\mu = -0,04$ . . . . .                            | 145 |
| 7.6 | QQ-plot ; $\mu = -0,04$ . . . . .   | 145 |
| 7.7 | Distribution empirique et QQ-plot ; $\mu = -0,07$ et $\mu = -0,1$ . . . . . | 146 |
| 7.8 | Besoin de solvabilité en fonction de $\mu$ . . . . .                        | 148 |

---

## Version anglaise

---



---

### Illustration PART 1

---

|     |   |     |
|-----|---|-----|
| 8.1 | Guarantee offered by CAMCA Assurance. . . . .                   | 161 |
| 8.2 | Recovery procedure. . . . .                                     | 164 |
| 8.3 | Provision (negative) concerning recoveries. . . . .             | 164 |
| 8.4 | Provision (positive). . . . .                                   | 164 |
|     |   |     |
| 9.1 | Individual provision with the compound Poisson process. . . . . | 171 |
| 9.2 | Collective methods Vs individual methods. . . . .               | 173 |
| 9.3 | Balance sheet stressed. . . . .                                 | 175 |
| 9.4 | Provisioning risk. . . . .                                      | 177 |

---



---

### Illustration PART 2

---

|       |   |     |
|-------|---|-----|
| 11.1  | Duration capped. . . . .                                    | 190 |
| 11.2  | Sets $A_{t_0-}^i$ and $A_{t_0+}^i$ . . . . .                | 193 |
| 11.3  | Sets $C_{t_0-}^i$ and $C_{t_0+}^i$ . . . . .                | 200 |
| 11.4  | Estimations. . . . .  | 203 |
|       |   |     |
| 12.1  | The “credit defaults swap”. . . . .                         | 218 |
| 12.2  | Illustration of payment in a “credit default swap”. . . . . | 219 |
| 12.3  | Submission of a claim. . . . .                              | 223 |
| 12.4  | Unpaid capital Vs Unpaid interest. . . . .                  | 223 |
| 12.5  | Reimbursement of a missing monthly installment. . . . .     | 224 |
| 12.6  | Number of loans per grouping. . . . .                       | 228 |
| 12.7  | Descriptive statistics on contracts involved. . . . .       | 228 |
| 12.8  | Histogram of the loan amount. . . . .                       | 229 |
| 12.9  | Histogram of the duration of the loans. . . . .             | 230 |
| 12.10 | Boxplot of loan amount. . . . .                             | 231 |
| 12.11 | Boxplot of loan duration. . . . .                           | 232 |
| 12.12 | Weight (in terms of loan amount) by grouping. . . . .       | 232 |
| 12.13 | Box-plot of loan amount by grouping. . . . .                | 233 |
| 12.14 | Box-plot of loan duration by grouping. . . . .              | 233 |
| 12.15 | Box-plot of outstanding amount owed. . . . .                | 234 |

---

---

TABLE DES FIGURES

---

|       |   |     |
|-------|---|-----|
| 12.16 | Ratio OAO over loan amount. . . . .   | 234 |
| 12.17 | Weight (in OAO) by grouping. . . . .  | 235 |
| 12.18 | Weight (in OAO) by grouping - Graphic. . . . .                                | 235 |
| 12.19 | Number of claims. . . . .   | 236 |
| 12.20 | Loans matured. . . . .  | 236 |
| 12.21 | Synthesis estimation parameters. . . . .                                      | 237 |
| 12.22 | Theoretical mean Vs simulated mean. . . . .                                   | 239 |
| 12.23 | Mean and confident interval based on the analysis date. . . . .               | 240 |
| 12.24 | Provision mean in relation to groupings. . . . .                              | 241 |
| 12.25 | Ratio OAO Vs Ratio mean (%). . . . .  | 242 |
| 12.26 | Ratio OAO Vs Ratio mean (graphic). . . . .                                    | 242 |
| 12.27 | Number of predicted claims. . . . .   | 243 |
| 12.28 | Provision mean under run-off assumption. . . . .                              | 244 |
| 12.29 | Comparison of provisioning method. . . . .                                    | 246 |
| 12.30 | Ratio OAO Vs Ratio mean in function of provisioning method. . . . .           | 247 |
| 12.31 | Ratio Claim / Premium in function of provisioning method. . . . .             | 248 |
| 12.32 | Sensitivity of the provision mean with regard to $\tau$ . . . . .             | 251 |
| 12.33 | Sensitivity of the provision mean with regard to $\mu$ and $\sigma$ . . . . . | 253 |
| 12.34 | Comparison of sensitivities in function of $t_0$ . . . . .                    | 254 |
| 12.35 | Probability of nullity of a claim amount. . . . .                             | 256 |

---

**Illustration PART 3**

---

|      |   |     |
|------|---|-----|
| 13.1 | Set $A_h$ . . . . .   | 265 |
| 14.1 | Provision mean according $h$ . . . . .  | 290 |
| 14.2 | Provision mean according $\mu$ . . . . .                                      | 291 |
| 14.3 | Empirical distribution ; $\mu = -0.01$ . . . . .                              | 293 |
| 14.4 | QQ plot ; $\mu = -0.01$ . . . . .   | 294 |
| 14.5 | Empirical distribution ; $\mu = -0.04$ . . . . .                              | 295 |
| 14.6 | QQ-plot ; $\mu = -0.04$ . . . . .   | 295 |
| 14.7 | Empirical distribution and QQ-plot ; $\mu = -0.07$ and $\mu = -0.1$ . . . . . | 296 |
| 14.8 | Solvency need in function of $\mu$ . . . . .                                  | 298 |

---



# Table des notations et des abréviations

|  |   |    |
|--|---|----|
| CAMCA Assurance                            | Compagnie d'assurances non-vie . . . . .  | 5  |
| $t_0$                                      | Date d'analyse . . . . .  | 9  |
| $C_{i,j}$                                  | Montant de sinistres payés jusqu'à l'année $i+j-1$ concernant la génération $i$ . . . . . | 13 |
| $X_{i,j}$                                  | Montant de sinistres payés l'année $i+j-1$ concernant la génération $i$ . . . . .         | 13 |
| Chain-Ladder                               | Méthode collective de provisionnement . . . . .   | 13 |
| Chain-Ladder pondéré                       | Méthode collective de provisionnement . . . . .   | 14 |
| Cape code                                  | Méthode collective de provisionnement . . . . .   | 15 |
| GLM  | Generalized linear model . . . . .  | 16 |
| London Chain                               | Méthode collective de provisionnement . . . . .   | 15 |
| Modèle de Benktander                       | Méthode collective de provisionnement . . . . .   | 15 |
| Modèle de Bornhuetter-Ferguson             | Méthode collective de provisionnement . . . . .   | 15 |
| Modèle de HFB                              | Méthode collective de provisionnement . . . . .   | 16 |
| Modèle de Hertig                           | Méthode collective de provisionnement . . . . .   | 16 |
| Modèle de Jab Chain                        | Méthode collective de provisionnement . . . . .   | 16 |
| Modèle de Mack                             | Méthode collective de provisionnement . . . . .   | 16 |
| Modèle de Munich Chain                     | Méthode collective de provisionnement . . . . .   | 16 |
| Modèle de Schnauss                         | Méthode collective de provisionnement . . . . .   | 16 |
| Modèle des moindres carrés de F. de Vylder | Méthode collective de provisionnement . . . . .   | 15 |
| Processus de poisson composé               | Méthode individuelle de provisionnement . . . . .   | 17 |
| v.a.                                       | Variable aléatoire . . . . .  | 18 |
| i.i.d.                                     | Indépendante et indentiquement distribuées . . . . .                                      | 18 |
| Shot noises processes                      | Méthode individuelle de provisionnement . . . . .   | 18 |
| Solvabilité 2                              | Réforme européenne du secteur des assurances . . . . .                                    | 20 |
| CSR  | Capital de solvabilité requis . . . . .   | 20 |



|                    |  |    |
|--------------------|--|----|
| EIOPA              | European Insurance and Occupational Pensions Authority . . . . .                             | 20 |
| ORSA               | Own Risk Solvency Assesment . . . . .  | 22 |
| $B$                | Ensemble des emprunteurs . . . . .   | 36 |
| $D_j$              | Durée du prêt de l'emprunteur $j$ . . . . .  | 36 |
| $I_j$              | Date de souscription du prêt de l'emprunteur $j$ . . . . .                                   | 36 |
| $M_j$              | Montant du prêt de l'emprunteur $j$ . . . . .  | 36 |
| $h$                | Durée maximale des prêts . . . . .   | 36 |
| $i$                | Mois de souscription . . . . .   | 36 |
| $t^*$              | Unité de temps : un mois . . . . .   | 36 |
| $(B_t)_{t \geq 0}$ | Mouvement Brownien . . . . .   | 38 |
| $(R_t)_{t \geq 0}$ | Mouvement Brownien géométrique . . . . .   | 38 |
| $\gamma$           | Coefficient représentant l'apport personnel et les coûts liés à la vente du bien . . . . .   | 38 |
| $\mu, \sigma$      | Coefficients du mouvement Brownien géométrique . . . . .                                     | 38 |
| $\tau$             | Taux exponentiel des emprunts . . . . .  | 37 |
| $S_t^1$            | Montant dû par l'emprunteur à la date $t$ . . . . .  | 38 |
| $S_t^2$            | Montant de la revente du bien à la date $t$ . . . . .  | 38 |
| $T_j^i$            | Date du fin du prêt de l'emprunteur $j$ ayant souscrit un prêt le mois $i$ . . . . .         | 39 |
| $\alpha$           | Coefficient intervenant dans la charge d'un sinistre . . . . .                               | 39 |
| $\phi_t(s)$        | Fonction intervenant dans la charge d'un sinistre . . . . .                                  | 39 |
| $f$                | Fonction intervenant dans la charge d'un sinistre . . . . .                                  | 39 |
| $p$                | Coefficient lié aux intérêts de retard . . . . .   | 39 |
| $A_{t_0+}^i$       | Ensemble de déclaration d'un sinistre dans l'intervalle de temps $]t_0, t^*i + h[$ . . . . . | 40 |
| $A_{t_0-}^i$       | Ensemble de déclaration d'un sinistre dans l'intervalle de temps $[t^*i, t_0]$ . . . . .     | 40 |
| $\Lambda_i$        | Intensité du PPP . . . . .   | 41 |
| PPP                | Processus ponctuel de Poisson . . . . .  | 41 |
| $P_{t_0}^i$        | Provision calculée à la date $t_0$ et concernant les prêts souscrits le mois $i$ . . . . .   | 41 |

TABLE DES FIGURES

---

|                        |   |    |
|------------------------|---|----|
| $P_{t_0}$              | Provision calculée à la date $t_0$ .....  | 42 |
| $F_1$                  | Fonction intervenant dans le calcul de la moyenne de la provision .....   | 43 |
| $\Phi$                 | Fonction de répartition d'une v.a. gaussienne centrée et réduite .....  | 43 |
| $F_2$                  | Fonction intervenant dans le calcul de la variance de la provision .....  | 44 |
| $F_3$                  | Fonction intervenant dans le calcul de la variance de la provision .....  | 44 |
| $\tilde{d}$            | Fonction intervenant dans le calcul de la variance de la provision .....  | 44 |
| $C^{i,i'}$             | Coefficient intervenant dans le calcul de la variance de la provision .....   | 45 |
| $N_i(A)$               | Nombre de prêts souscrits la génération $i$ et inclus dans l'ensemble $A$ .....                                     | 45 |
| $g$                    | Fonction intervenant dans le calcul de la variance de la provision .....  | 45 |
| $F_M(m)$               | Fonction de répartition empirique des montants prêts  | 46 |
| $C_{t_0+}^i$           | Ensemble des prêts ayant atteint leur maturité dans l'intervalle $]t_0, t^*i + h]$ .....                            | 47 |
| $C_{t_0-}^i$           | Ensemble des prêts ayant atteint leur maturité dans l'intervalle $[t^*i, t_0]$ .....                                | 47 |
| $\lambda_1, \lambda_2$ | Paramètres intervenants dans l'intensité du PPP.....  | 47 |
| $b$                    | Observation du nombre de prêts souscrits jusqu'à la date $t_0$ .....  | 48 |
| $b^i$                  | Observation du nombre de prêts souscrits le mois $i$ ...  | 48 |
| $d_j$                  | Observation de la durée du prêt de l'emprunteur $j$ ...   | 48 |
| $i_0$                  | Mois correspondant à la date $t_0$ .....  | 48 |
| $m_j$                  | Observation du montant du prêt de l'emprunteur $j$ ...  | 48 |
| $n^i$                  | Observation du nombre de sinistres déclarés avant la date $t_0$ et concernant les prêts souscrits le mois $i$ ..... | 48 |
| $t_j^i$                | Observation de la date de fin du prêt de l'emprunteur $j$   | 48 |
| $v^i$                  | Observation du nombre de prêts échus avant la date $t_0$ et concernant les prêts souscrits le mois $i$ .....        | 48 |

|                                 |   |     |
|---------------------------------|---|-----|
| $\bar{m}$                       | Estimation de la moyenne des prêts . . . . .  | 50  |
| $\hat{\gamma}$                  | Estimation du coefficient $\gamma$ . . . . .  | 50  |
| $\hat{\lambda}_1$               | Estimation du coefficient $\lambda_1$ . . . . .   | 50  |
| $\hat{\lambda}_2$               | Estimation du coefficient $\lambda_2$ . . . . .   | 50  |
| $\hat{\mu}$                     | Estimation du coefficient $\mu$ . . . . .   | 50  |
| $\hat{\sigma}$                  | Estimation du coefficient $\sigma$ . . . . .  | 50  |
| $\hat{\tau}$                    | Estimation du taux d'emprunt $\tau$ . . . . .   | 50  |
| $\hat{h}$                       | Estimation du coefficient $h$ . . . . .   | 50  |
| $\hat{p}$                       | Estimation du coefficient $p$ . . . . .   | 50  |
| CDS                             | Credit default swap . . . . .   | 64  |
| $\mathbb{E}^{emp}$              | Estimation de l'espérance empirique de la provision . .                                   | 87  |
| $\mathbb{E}^{theo}$             | Estimation de l'espérance théorique de la provision . .                                   | 86  |
| $t_0^*$                         | Date maximale de provisionnement . . . . .  | 88  |
| $S_{\lambda_2}$                 | Sensibilité de l'espérance de la provision face à $\lambda_2$ . . .                       | 97  |
| $S_{\tau}$                      | Sensibilité de l'espérance de la provision face à $\tau$ . . . .                          | 98  |
| $\Psi$                          | Densité d'une v.a. gaussienne centrée et réduite . . . . .                                | 98  |
| $S_{\mu}$                       | Sensibilité de l'espérance de la provision face à $\mu$ . . . .                           | 100 |
| $S_{\sigma}$                    | Sensibilité de l'espérance de la provision face à $\sigma$ . . . .                        | 100 |
| $p(t, d, t^*)$                  | Probabilité que la charge d'un sinistre soit nulle . . . .                                | 103 |
| BS                              | Besoin de solvabilité . . . . .   | 111 |
| $P(h, R)$                       | Provision du mois $i = 0$ calculée à la date 0 . . . . .                                  | 113 |
| $A_h$                           | Ensemble de déclaration des sinistres pour les prêts sous-crits le mois $i = 0$ . . . . . | 114 |
| $BS(h)$                         | Besoin de solvabilité lié au risque de provisionnement                                    | 116 |
| $BS^*(h)$                       | Approximation du besoin de solvabilité . . . . .  | 117 |
| $P(h, R) - \mathbb{E}[P(h, R)]$ | V.a. mesurant le risque de provisionnement . . . . .                                      | 116 |
| $\mathcal{Q}_X(99, 5\%)$        | Quantile de la v.a. $X$ à l'ordre 99, 5% . . . . .  | 116 |
| $C_k$                           | Coefficient intervenant dans la loi asymptotique de la provision . . . . .                | 117 |
| $G$                             | Gaussienne centrée et réduite . . . . .   | 117 |

TABLE DES FIGURES

---

|             |  |
|-------------|--|
| $X^*$       | V.a. intervenant dans la loi asymptotique de la provision<br>117                 |
| $\eta$      | Coefficient intervenant dans la limite de l'espérance de la provision ..... 117  |
| $\kappa$    | Coefficient intervenant dans la loi asymptotique de la provision ..... 117       |
| $\Gamma$    | V.a. de loi gamma et de paramètres $-\frac{2\mu}{\sigma^2}$ ..... 119            |
| $P(h, r)$   | Provision sachant $(R_t)_{t \geq 0}$ ..... 123                                   |
| $X_k(h, r)$ | V.a. intervenant dans la preuve de la loi asymptotique de la provision ..... 123 |
| $d_W$       | Distance de Wasserstein ..... 124  |
| $d_W$       | Distance de Kolmogorow ..... 125   |





Première partie

Introduction générale

---

|   |   |    |
|---|---|----|
| 1 | Généralités et plan de la thèse . . . . . | 3  |
| 2 | Etat de l'art . . . . .                   | 13 |
| 3 | Conclusion et perspectives . . . . .      | 27 |

---





# Généralités et plan de la thèse

---

Dans ce premier chapitre nous présenterons tout d'abord les principales questions auxquelles les compagnies d'assurances doivent faire face. Nous étudierons ensuite la garantie spécifique considérée tout au long de ce document : le cautionnement des prêts habitats. Nous terminerons cette introduction en dressant un plan détaillé de la thèse.

Une traduction de ce chapitre, en anglais, est disponible page [159](#).

## 1.1 Problématique en assurance

Une compagnie d'assurances (nous définirons ce terme plus tard) se doit de protéger un assuré (personne physique ou morale) sur lequel porte un risque. Les risques sont des événements aléatoires et peuvent être classés de différentes manières. Il est possible de distinguer les risques portant sur les personnes physiques (accident, décès, etc.), les risques portant sur les biens matériels (automobile, habitat, etc.) et les risques de responsabilité (civile, professionnelle, etc.). Une autre manière de hiérarchiser les risques est d'opposer les risques liés à la vie de l'assuré des autres risques (appelés risque "non-vie").

La protection par l'assureur de l'assuré en cas de réalisation du risque peut être financière ou matérielle. En contrepartie de cette protection, une prime, ou cotisation (dans le cas d'une société mutuelle), doit être versée à la compagnie d'assurances qui supporte le risque. Cette prime est payée par le souscripteur du contrat. Ainsi lorsque la compagnie d'assurances détermine le montant de la prime, elle ne connaît pas le montant qu'elle devra dépenser suite à l'éventuelle réalisation du risque : ce phénomène est appelé "inversion du cycle de production". Notons que l'aléa sur le risque est double : la réalisation du risque est incertaine et le montant occasionné l'est également. Ceci pose l'une des questions principales à laquelle doivent répondre les assureurs :

**Question 1 : “ Comment déterminer le montant des primes ou des cotisations ? ”**

Il existe plusieurs types de compagnies d'assurances qui se distinguent par leur statut juridique. Nous pouvons distinguer les sociétés d'assurances régies par le code des assurances (société anonyme à but lucratif et société d'assurances mutuelle à but non lucratif), les sociétés mutuelles d'assurances régies, en partie, par le code des mutuelles (but non lucratif et restreint aux assurances de personnes) et les instituts de prévoyance

régis par le code de la sécurité sociale (organisme paritaire à but non lucratif). Ces sociétés d'assurances ont la possibilité de personnaliser leurs tarifs, c'est à dire d'intégrer les caractéristiques personnelles des assurés pour déterminer le montant de la prime dont ils doivent s'acquitter. Ceci introduit la deuxième question à laquelle les assureurs sont confrontés :

**Question 2 : “ Comment détecter les assurés les plus risqués ? ”**

La durée des contrats d'assurances fluctue en fonction des types d'assurances : généralement les contrats d'assurances vie ont une durée plus longue que celle des contrats d'assurances non-vie. Les différentes réglementations auxquelles sont soumises les compagnies d'assurances leur imposent de conserver dans leurs comptes, et ce tout au long de la vie des contrats, le montant nécessaire et suffisant pour faire face à la réalisation des risques assurés ; ce montant étant qualifié de “provision technique” ou “provision statutaire”. Ceci introduit la troisième interrogation à laquelle les assureurs sont confrontés :

**Question 3 : “ Comment calculer le montant des provisions tout au long de la vie des contrats ? ”**

Les montants provisionnés et inscrits dans les comptes des compagnies d'assurances sont investis dans différents produits : actions, obligations, immobilier, etc.. Les produits financiers générés par ces investissements peuvent être extrêmement importants pour certains types d'assurances (compte tenu de la durée d'engagement). Ces investissements sont liés à des engagements auprès d'assurés et doivent être réalisés en respectant certains critères. Le but de ces investissements est double : maximiser le rendement financier et minimiser les pertes. Ceci doit être fait en respectant les règles d'investissement définies par les réglementations en vigueur et par les organes de direction des compagnies. Ceci introduit la quatrième question posée aux assureurs :

**Question 4 : “ Comment investir le montant des primes perçues ? ”**

En plus des réglementations nationales, les compagnies d'assurances seront bientôt soumises à la réglementation Européenne “Solvabilité 2”. Cette réforme (Cf. Section 1.2.2 pour plus de détails) a pour objectif de mieux protéger et mieux informer les souscripteurs d'assurance en évaluant au plus juste la solvabilité des compagnies :

**Question 5 : “ Comment évaluer la solvabilité d'une compagnie d'assurances ? ”**

Passons maintenant à la présentation de la garantie spécifique considérée tout au long de cette thèse.

## 1.2 Le cautionnement des prêts habitats

Après avoir présenté la société CAMCA Assurance, nous introduisons la garantie spécifique étudiée (et proposée par cette compagnie) et concluons par une présentation des différentes provisions que cette compagnie doit constituer.

### 1.2.1 CAMCA Assurance

En 1946, les Caisses régionales du Crédit Agricole ont fondé la mutuelle CAMCA (Caisse d'Assurances Mutuelles du Crédit Agricole) en vue de mutualiser le risque vol ; les garanties de ce type se trouvant difficilement sur le marché de l'assurance au sortir de la guerre.

A la fin des années 1980, CAMCA a étendu son activité à l'assurance para-bancaire (assurance des moyens de paiement, assurance chéquier, etc.) en proposant des garanties liées aux comptes des clients des Caisses régionales. Du fait de son statut de mutuelle, CAMCA ne pouvait proposer de services para-bancaires qu'aux Caisses régionales et non pas directement aux clients de ces Caisses. Une société anonyme n'étant pas soumise à cette contrainte, CAMCA a créé fin des années 90, une société anonyme nommée CAMCA Assurance afin de proposer des services para-bancaires directement aux clients des Caisses régionales. CAMCA Assurance est basée à Luxembourg et ce pour trois raisons majeures :

1. CAMCA disposait à cette époque d'une compagnie de réassurance (dénommée CAMCA Réassurance) déjà implantée à Luxembourg et souhaitait que la compagnie créée (i.e. CAMCA Assurance) puisse s'appuyer sur cette compagnie de réassurance. Il était plus simple de créer la compagnie dans un environnement technique et réglementaire similaire à celui de CAMCA Réassurance.
2. Le Crédit Agricole avait à cette époque une volonté affichée de se développer en Europe et d'être en mesure de proposer des garanties similaires aux autres filiales Européennes. Il était plus logique d'intervenir depuis le Luxembourg sur le mode de la "libre prestation de services" dans les autres pays européens. Rappelons que la libre prestation de services permet à une compagnie d'assurances qui a reçu un agrément dans l'un des pays de l'Union Européenne d'exercer son activité, pour laquelle elle a reçu son agrément, dans n'importe quel autre pays de l'Union sous réserve d'appliquer la réglementation locale du pays où elle prend le risque.
3. La dernière raison était plus pragmatique : au moment de la création de la compagnie en 1997 il fallait environ huit mois sur le plan administratif pour obtenir toutes les autorisations nécessaires auprès du Luxembourg et il en fallait plus du double en France.

Actuellement, la principale activité de CAMCA Assurance, en terme d'encaissement de primes, est la garantie caution (ou cautionnement) des prêts habitats des clients des Caisses régionales du Crédit Agricole (appelées "banques" par la suite). La garantie caution est similaire à une hypothèque notariée : en cas de non paiement définitif des

mensualités d'un prêt de la part d'un emprunteur, CAMCA Assurance se substituera à celui-ci auprès de la banque ayant financé le prêt. Voici une illustration graphique de cette garantie :

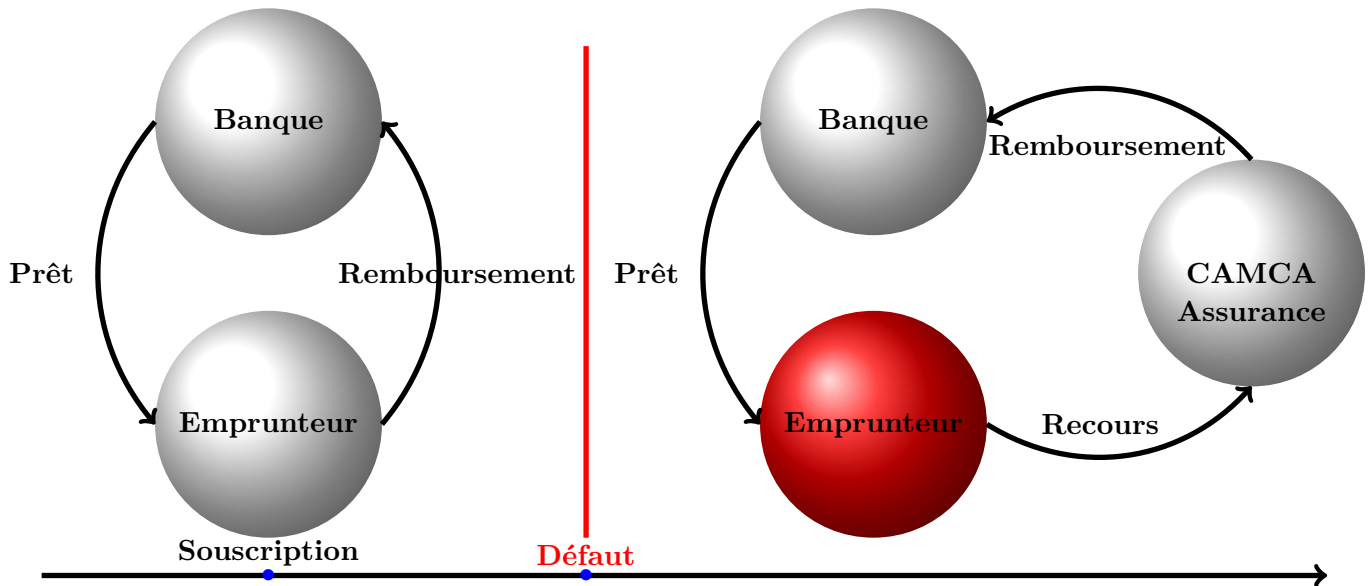


FIGURE 1.1 – Illustration de la garantie proposée par CAMCA Assurance.

### 1.2.2 Garantie caution

Le principe du cautionnement est le suivant. Après paiement à CAMCA Assurance par l'emprunteur, d'un pourcentage sur le montant financé (le montant versé par l'emprunteur est appelé *prime*), CAMCA Assurance accepte de se substituer à celui-ci auprès de la banque en cas de défaut de paiement définitif, c'est à dire lorsque l'emprunteur est dans l'incapacité irréversible de rembourser ses mensualités ; cet état sera par la suite appelé "sinistre". En d'autres termes, CAMCA Assurance remboursera la banque si l'emprunteur est dans l'incapacité permanente de rembourser ses dettes.

Dans cette circonstance, la caution (c'est à dire CAMCA Assurance) a la possibilité d'utiliser tous les moyens légaux pour récupérer tout ou partie de sa créance. Ainsi CAMCA Assurance peut prétendre (ce n'est pas automatique) à la saisie et à la revente du bien financé mais également à toute récupération auprès de l'emprunteur (héritage, retour à meilleure fortune, etc.). CAMCA Assurance a cependant délégué la procédure de recouvrement aux banques finançant les prêts : ce sont ces banques qui récupéreront, en cas de sinistre, l'argent auprès du client et sur le bien financé. Si les montants récupérés, qualifiés de "montant de recours", ne suffisent pas à combler la dette du client alors CAMCA Assurance versera le solde restant dû à la banque : CAMCA Assurance

## 1.2 Le cautionnement des prêts habitats

---

indemniser la banque d'un montant égal à sa perte définitive nette de recours. Notons que les emprunts considérés sont des crédits habitats, i.e. servant au financement d'un bien immobilier (achat, construction, etc.).

L'assurance caution proposée par CAMCA Assurance est différente d'une assurance emprunteur classique (décès, interruption temporaire du temps de travail, invalidité) et elle est également différente d'une assurance crédit ; dans le cas d'une assurance crédit, le souscripteur du contrat est la banque finançant les crédits.

Reformulons ceci selon la terminologie utilisée en assurance :

- CAMCA Assurance accorde une “sureté personnelle” à l'emprunteur : en cas de défaut de paiement de celui-ci, CAMCA Assurance indemniser le créancier de sa perte définitive à l'issue de toute procédure de recouvrement. La garantie offerte par CAMCA Assurance est une caution simple et directe dont la souscription est déléguée aux banques émettrices des prêts.
- Le souscripteur de la garantie est l'emprunteur : c'est cet emprunteur qui paie la prime (unique) à la compagnie d'assurances au début du prêt. C'est également l'emprunteur qui est l'assuré du contrat.
- Le bénéficiaire de la garantie est la banque ayant financé le prêt. Cette banque est également l'intermédiaire permettant la mise en relation des emprunteurs et de la compagnie d'assurances.
- La durée du contrat est égale à la durée du prêt : la garantie est close lorsque le prêt est remboursé, soit par l'emprunteur soit par la compagnie.

Donnons maintenant quelques précisions techniques sur la garantie caution.

Dans certains cas de figure, les banques peuvent, en plus de la garantie fournie par la compagnie d'assurances, mettre en place un autre type de protection. Les deux exemples les plus courants sont les suivants :

1. Lorsque la banque anticipe (via différentes informations) que le client va bientôt se trouver en défaut de paiement ou que la situation de l'emprunteur est plus risquée que lors de la souscription du prêt (divorce, décès du conjoint, etc.), celle-ci peut décider, en plus de la garantie CAMCA Assurance, de prendre une hypothèque notariée (conventionnelle) sur le bien financé. Comme nous l'avons vu dans l'introduction, CAMCA Assurance peut prétendre à la saisie du bien financé ; cette saisie n'est cependant pas systématique en cas de sinistre contrairement à la saisie dans le cadre d'une hypothèque notariée. L'hypothèque notariée, bien que coûteuse, est donc plus “sûre” (au niveau du recouvrement) que la garantie CAMCA Assurance ; les frais, négligeables, liés à l'hypothèque étant à la charge de CAMCA Assurance.
2. Lorsque un emprunteur connaît un ou plusieurs défauts de paiement la banque entame une procédure de recouvrement (à l'amiable ou en contentieux). Cette procédure peut déboucher sur la demande d'une hypothèque judiciaire auprès d'un

juge : le fonctionnement des hypothèques judiciaires est similaire à celui des hypothèques notariées ; les frais, non négligeables, ne sont pas à la charge de CAMCA Assurance.

Nous supposerons que de telles hypothèques sont systématiquement demandées, ce qui implique en particulier qu'en cas de sinistres le bien peut systématiquement être saisi et revendu.

CAMCA Assurance a édité un profil type de clients et de biens financés pour lesquels les banques n'ont pas besoin de demander l'accord de CAMCA Assurance pour que leurs clients puissent bénéficier de la garantie caution. Dans ce cas de figure, l'accord est automatique. Si un client ne respecte pas le profil type édité par CAMCA Assurance (tant sur son profil que sur celui du bien financé), la banque peut demander à CAMCA Assurance, et ce à titre exceptionnel, de prendre en charge la garantie de ce client atypique. Dans le cadre d'un sinistre, si l'application de la garantie CAMCA Assurance s'est effectuée de manière automatique par la banque (sans demande explicite à CAMCA Assurance) et que les profils types sur le client ou sur le bien financé n'ont pas été respectés (fraude ou négligence), alors CAMCA Assurance se réserve le droit de refuser l'indemnisation du sinistre.

Nous supposerons cependant qu'il n'y a aucune fraude ou négligence de la part des banques sur les conditions à respecter pour qu'un client puisse bénéficier de la garantie.

Il est également possible d'interpréter la garantie caution proposée par CAMCA Assurance via les "credit default swap" (Cf. Section 5.1 pour plus de détails) ou via les options d'échange (Cf. Section 5.2 pour plus de détails).

### 1.2.3 Sinistralité et provisionnement

Comme énoncé précédemment, le sinistre pour le type de garantie décrit ci-dessus est le défaut de paiement des emprunteurs. Nous terminons la présentation de cette garantie en expliquant les étapes de déclaration d'un sinistre et la nature des provisions que CAMCA Assurance doit constituer.

En cas de sinistre, le montant que CAMCA Assurance devra rembourser à la banque ayant financé le prêt est appelé "montant du sinistre". La banque a toute liberté pour déclarer qu'un prêt est en sinistre (c'est à dire en défaut définitif de paiement). Ainsi lorsqu'un emprunteur connaît un défaut de paiement, la banque a la possibilité de le déclarer immédiatement en sinistre ou d'attendre que le client réussisse à régulariser sa situation, c'est à dire ne pas le déclarer en sinistre immédiatement. Dans cette situation, le nombre de mensualités impayées peut varier. Notons également qu'un défaut de paiement (non définitif) peut concerner une certaine part de la mensualité : par exemple un incident de paiement de EUR 700 pour une mensualité de EUR 1 500.

Lors de l'indemnisation d'un sinistre, CAMCA Assurance reçoit, de la part de la banque,

## 1.2 Le cautionnement des prêts habitats

le montant de sinistres que le client a contracté, diminué du montant d'une éventuelle revente du bien financé. CAMCA Assurance paie alors ce montant à la banque. Dès ce paiement effectué, CAMCA Assurance n'aura plus besoin de déboursier d'argent pour ce sinistre. Il est cependant possible qu'après ce premier versement, la banque réussisse à récupérer de l'argent auprès de l'emprunteur. Dans ce cas de figure la banque reversera les montants récupérés à CAMCA Assurance. Voici une illustration graphique :



FIGURE 1.2 – Processus de recouvrement.

Cette description des sinistres nous permet de distinguer les deux types de provisions devant être constituées par CAMCA Assurance. A une date  $t_0$ , CAMCA Assurance doit constituer :

1. Une provision pour les sinistres déjà déclarés et indemnisés à la date  $t_0$ . Cette provision consiste en une estimation des recours que les banques effectueront après la date  $t_0$  sur les prêts déjà sinistrés (et indemnisés) à la date  $t_0$ . Cette provision est donc négative. Voici une illustration :

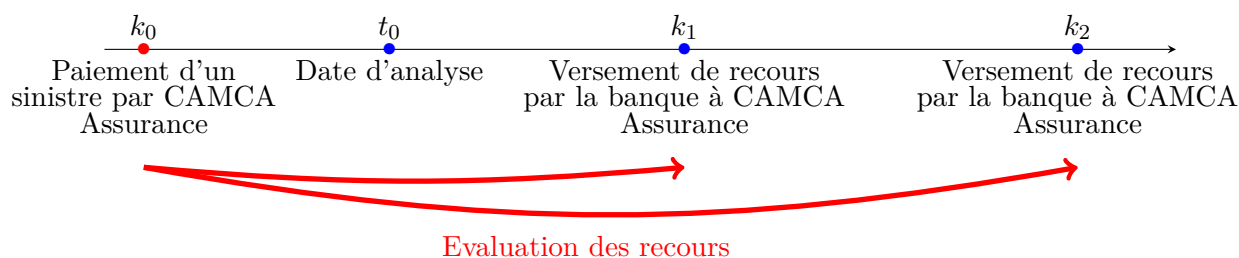


FIGURE 1.3 – Provision (négative) concernant les recours.

2. Une provision pour les sinistres qui seront indemnisés après cette date  $t_0$  et qui concerne les prêts souscrits avant la date  $t_0$ . CAMCA Assurance doit déterminer le montant (inconnu) des sinistres qui seront indemnisés après la date  $t_0$  en tenant compte d'éventuels recours. Voici une illustration :

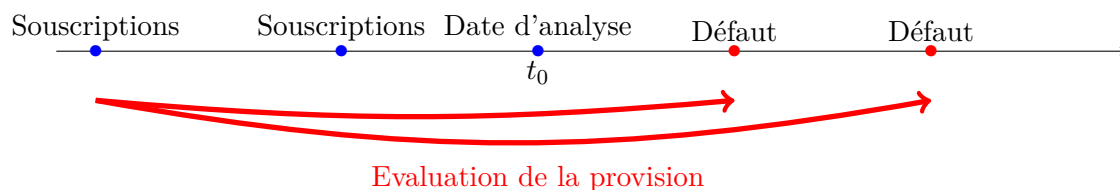


FIGURE 1.4 – Provision (positive).

Les recours reçus après l'indemnisation des sinistres par CAMCA Assurance sont extrêmement rares (moins de 0,01% en montant). Ainsi CAMCA Assurance détermine uniquement le deuxième type de provision.

### 1.3 Plan des Parties *II* et *III*

La deuxième partie de cette thèse sera consacrée à la modélisation des provisions techniques que CAMCA Assurance doit constituer (Cf. Chapitre 4). Nous déterminerons également certains indicateurs de risque (moyenne, variance) ainsi que la loi de la provision (Cf. Section 4.3). Nous concluons ce chapitre en proposant une méthode d'estimation des paramètres présents dans notre modèle (Cf. Section 4.4).

Nous compléterons le Chapitre 4 par des remarques portant notamment sur l'interprétation de la garantie via les "credit default swap" (Cf. Section 5.1), l'investissement de la prime perçue en liant la provision à la notion d'option d'échange (Cf. Section 5.2), une justification des hypothèses de notre modèle (Cf. Section 5.3) et par une application numérique des résultats théoriques obtenus (Cf. Section 5.4).

La troisième partie sera quant à elle consacrée à l'évaluation de la solvabilité de CAMCA Assurance dans le cadre de la réforme Européenne Solvabilité 2. Dans le Chapitre 6, nous discuterons du besoin de solvabilité lié au risque de provisionnement. Nous déterminerons en particulier la loi asymptotique de la provision dans le Théorème 6.2.5 et le besoin de solvabilité associé au risque de provisionnement dans le Théorème 6.2.6. Le Chapitre 6 sera complété par une application numérique dans le Chapitre 7 portant notamment sur le calcul numérique du besoin de solvabilité.

En résumé et pour établir un lien avec les principales problématiques présentées dans la Section 1, nous tenterons de répondre dans cette thèse aux questions 3 et 5, les questions 1, 2 et 4 étant à l'heure actuelle à l'état de projet.

Avant cela nous dresserons dans la prochaine section un rapide état de l'art concernant les deux problématiques traitées dans cette thèse :

1. Le provisionnement,
2. L'évaluation de la solvabilité d'une compagnie d'assurances.







## CHAPITRE 2

# Etat de l'art

---

Dans ce chapitre nous dresserons un rapide état de l'art concernant le provisionnement en assurance non-vie et l'évaluation de la solvabilité dans le cadre de la future réforme Européenne Solvabilité 2.

Une traduction de ce chapitre, en anglais, est disponible page [167](#).

## 2.1 Le provisionnement

Cette section s'inspire en partie des ouvrages [\[23\]](#), [\[43\]](#) et [\[63\]](#).

L'article R-331-1 du code des assurances impose aux assureurs de déterminer la somme d'argent permettant de faire face à leurs engagements techniques ; cette somme est appelée "provision statutaire". En assurance non-vie, les méthodes traditionnellement utilisées pour le calcul de ces provisions sont de deux types : collectives ou individuelles.

### 2.1.1 Les méthodes collectives

Les méthodes collectives se basent sur des données agrégées par génération de contrats (une génération pouvant être un mois, une année, etc.). Ces méthodes utilisent l'évolution passée des sinistres en intégrant parfois des données exogènes à ces sinistres. Elles sont en revanche incapables de prendre en compte les caractéristiques individuelles des assurés. Dans la littérature traitant des problèmes de provisionnement, il existe deux types de méthodes collectives : les méthodes déterministes et les méthodes stochastiques.

Avant de présenter ces méthodes, il est utile d'introduire quelques notations ou concepts.

#### Notation 2.1.1.

1. Les générations de contrats sont notées par l'indice  $i$  :  $i \in \{1, \dots, n\}$  où  $n$  est la dernière génération connue. La génération  $i$  peut représenter une génération de souscriptions de contrats, de survenances de sinistres, d'indemnisations de sinistres, etc.. Pour fixer les idées, nous parlerons dans la suite de génération de souscriptions de contrats et l'unité sera d'un an.
2.  $X_{i,j}$  est le montant des sinistres concernant les contrats souscrits la génération  $i$  et indemnisés l'année  $i + j - 1$ ,  $j \in \{1, \dots, n\}$ . Ainsi  $X_{i,j}$  est une somme de montants

de sinistres individuels.  $C_{i,j}$  est le montant de sinistres concernant les contrats souscrits la génération  $i$  et indemnisés jusqu'à l'année  $i + j - 1$ , i.e. :

$$C_{i,j} := \sum_{k=1}^{k=j} X_{i,k}. \quad (2.1.1)$$

3. A la date  $n$ , la provision de la génération  $i$  est notée  $P_i$  et est définie par :

$$P_i := C_{i,n} - C_{i,n-i+1}. \quad (2.1.2)$$

### Les méthodes déterministes

La méthode la plus connue et utilisée par les praticiens est celle de Chain Ladder qui fut introduite aux alentours des années 1930. Cette méthode déterministe (qui ne repose sur aucune loi de probabilité) se base, pour calculer la provision à une date  $n$ , sur le triangle de développement suivant :

$$\begin{pmatrix} C_{1,1} & C_{1,2} & \cdot & \cdot & C_{1,n} \\ C_{2,1} & C_{2,2} & \cdot & C_{2,n-1} & \\ \cdot & \cdot & \cdot & & \\ C_{n,1} & & & & \end{pmatrix} \quad (2.1.3)$$

La première hypothèse du modèle de Chain Ladder est que tous les contrats ont la même durée et que celle-ci vaut  $n$ . Le but de cette méthode, et plus généralement des méthodes collectives, est d'estimer le montant ultime de sinistres  $C_{i,n}$ , et ainsi la provision  $P_i$ , en complétant le triangle de développement défini ci-dessus, i.e. :

$$\begin{pmatrix} C_{1,1} & C_{1,2} & \cdot & \cdot & C_{1,n} \\ C_{2,1} & C_{2,2} & \cdot & \cdot & C_{2,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ C_{n,1} & C_{n,2} & \cdot & \cdot & C_{n,n} \end{pmatrix} \Rightarrow \begin{pmatrix} P_1 = 0 & & & & \\ P_2 = C_{2,n} - C_{2,n-1} & & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ P_n = C_{n,n} - C_{n,1} & & & & \end{pmatrix} \quad (2.1.4)$$

La méthode de Chain Ladder a été introduite dans le cas où les données sont importantes et lorsque l'évolution du montant de sinistres ne dépend pas de la génération de contrats considérée. Ce qui conduit à introduire la seconde hypothèse : le montant cumulé de sinistres jusqu'à l'année  $i + j$  pour les contrats souscrits l'année  $i$  est proportionnel au montant cumulé de sinistres payés jusqu'à l'année  $i + j - 1$  pour ces mêmes contrats. Le coefficient de proportionnalité, noté  $\lambda_j$ , est déterministe et ne dépend que de l'année de développement  $j$  et non de la génération  $i$  considérée, i.e. :

$$\exists \lambda_j \in \mathbb{R}_+ : C_{i,j+1} = \lambda_j C_{i,j}, \forall i, j : i + j \leq n. \quad (2.1.5)$$

Une fois les paramètres  $(\lambda_j)_{1 \leq j \leq n-1}$  estimés, il est possible de donner une valeur numérique à la provision. La relation (2.1.5) étant valable pour tout  $i \leq n - j$ , le paramètre

$\lambda_j$  est alors déterminé par :

$$\lambda_j = \frac{\sum_{i=1}^{n-j} C_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}}, \quad 1 \leq j \leq n-1, \quad (2.1.6)$$

Comme  $C_{i,j+1} = \lambda_j C_{i,j}$ , le montant ultime de sinistres pour la génération  $i$  vaut :

$$C_{i,n} = C_{i,n+1-i} \prod_{j=1}^{i-1} \lambda_{n-i+j}. \quad (2.1.7)$$

Le montant à provisionner, toujours pour la génération  $i$ , est alors :

$$P_i = C_{i,n} - C_{i,n-i+1} = C_{i,n+1-i} \left[ \prod_{j=1}^{i-1} \lambda_{n-i+j} - 1 \right]. \quad (2.1.8)$$

Appliquons maintenant la méthode de Chain Ladder sur un exemple concret. Considérons le triangle de développement des sinistres suivant ( $n = 3$ ) :

$$\begin{pmatrix} 1 & 4 & 8 \\ 7 & 9 & \\ 2 & & \end{pmatrix} \quad (2.1.9)$$

L'objectif est de déterminer pour chaque génération  $i$  le montant ultime de sinistres, i.e. nous cherchons à déterminer  $C_{1,3}$ ,  $C_{2,3}$  et  $C_{3,3}$ . La première génération étant déjà arrivée à maturité nous connaissons le montant de sinistres ultimes  $C_{1,3}$ . D'après (2.1.6) nous avons :

$$\lambda_1 = \frac{4+9}{1+7} = \frac{13}{8}, \quad (2.1.10)$$

$$\lambda_2 = \frac{8}{4} = 2. \quad (2.1.11)$$

D'après la relation (2.1.8) les montants à provisionner seront donc :

$$P_1 = 0 \text{ (cette génération est déjà arrivée à maturité),}$$

$$P_2 = 9 \times 2 - 9 = 9,$$

$$P_3 = 2 \times 2 \times \frac{13}{8} - 2 = 4,5.$$

Le principal inconvénient de la méthode de Chain Ladder est sa sur-paramétrisation : à la date  $n$  il nous faut estimer  $n-1$  paramètres,  $\lambda_1, \dots, \lambda_{n-1}$ , en se basant sur  $\frac{n(n+1)}{2}$  données. Cette estimation de paramètres (Cf. 2.1.6) confère autant d'importance aux données récentes qu'aux données anciennes. Dans la pratique, les données récentes des

compagnies d'assurances sont souvent plus exhaustives et représentatives de la réalité. La méthode de Chain Ladder pondérée (Cf. [43]) permet de prendre en compte ce phénomène, en accordant plus ou moins d'importance à certaines données (en fonction de leur ancienneté par exemple).

Un autre inconvénient est que l'hypothèse autorégressive (2.1.5) du modèle de Chain Ladder est très restrictive; celle ci peut être relâchée de la manière suivante :

$$C_{i,j+1} = f(C_{i,j}), \forall i, j : i + j \leq n, \quad (2.1.12)$$

où  $f$  est une fonction définie de  $\mathbb{R}_+$  dans  $\mathbb{R}_+$ . Un cas particulier de fonction  $f$  a été proposé par S. Benjamin et L.M. Eagles en 1986 dans le modèle de London Chain :

$$\exists(\lambda_j, \alpha_j) \in (\mathbb{R}_+)^2 : f(C_{i,j}) = \lambda_j C_{i,j} + \alpha_j, \forall i, j : i + j \leq n. \quad (2.1.13)$$

Cette méthode, permettant de prendre en compte le cas où les points  $(C_{i,j}, C_{i,j+1})_{i \geq 1}$  forment une droite ne passant pas par l'origine, augmente le problème de surparamétrisation de la méthode de Chain Ladder. C'est pourquoi E. Straub a proposé en 1988 dans le modèle de London Pivot le choix suivant pour la fonction  $f$  :

$$\exists(\lambda_j, \alpha) \in (\mathbb{R}_+)^2 : f(C_{i,j}) = \lambda_j C_{i,j} + \alpha(\lambda_j - 1), \forall i, j : i + j \leq n. \quad (2.1.14)$$

De nombreuses autres variantes de la méthode de Chain Ladder existent. Toutes ces méthodes, bien que faciles d'implémentation, partagent le même inconvénient : elles ne prennent pas en compte les générations de souscriptions dans le développement futur des sinistres. Pour certaines branches d'assurance, il est naturel de vouloir considérer l'influence des générations par l'intégration d'une information "extérieure" aux sinistres. C'est ce que propose les méthodes Cape Code (H. Bühlmann, 1983, Cf. [17]), Bornhuetter-Ferguson (R.L. Bornhuetter et R.E. Ferguson, 1972, Cf. [14]) ou Benktander (G. Benktander, 1976, Cf. [10]) ; l'information extérieure aux sinistres pouvant être un montant de prime ou un ratio de sinistralité à l'ultime estimé par avis d'expert. D'autres méthodes permettent de prendre en compte l'influence des générations (méthode factorielle des moindres carrés de F. De Vylder, 1978, Cf. [61]) et également l'inflation (méthode de séparation de H.G. Verbeek et G. Taylor, 1972, Cf. [59]).

Toutes ces méthodes étant déterministes, elles ne donnent cependant qu'une estimation ponctuelle de la provision sans fluctuation possible, c'est à dire qu'elles ne fournissent pas d'intervalle de confiance.

Passons maintenant à la présentation des méthodes collectives stochastiques utilisées dans le cadre du provisionnement en assurance non-vie.

### Les méthodes stochastiques

Les méthodes stochastiques utilisées pour le calcul des provisions d'une compagnie d'assurances non-vie sont diverses et variées. Nous proposons ici une classification non ex-

haustive en trois parties : les modèles récurrents, les modèles basés sur des régressions log-normale et les modèles GLM (Generalized Linear Model, Cf. [40]).

La plus connue des méthodes stochastiques récurrentes est la méthode de Mack (T. Mack, 1993, Cf. [34]). Ce modèle est la version stochastique du modèle de Chain Ladder. Les hypothèses retenues dans ce modèle sont les suivantes :

$$(C_{i,1}, \dots, C_{i,n}) \text{ et } (C_{k,1}, \dots, C_{k,n}) \text{ sont deux variables indépendantes } \forall i \neq k, \quad (2.1.15)$$

$$\exists \lambda_j \in \mathbb{R}_+ : \mathbb{E}[C_{i,j+1} | C_{i,1}, \dots, C_{i,j}] = \lambda_j C_{i,j}, \forall i, j : i + j \leq n, \quad (2.1.16)$$

$$\exists \sigma_j \in \mathbb{R}_+ : \mathbb{V}[C_{i,j+1} | C_{i,1}, \dots, C_{i,j}] = \sigma_j C_{i,j}, \forall i, j : i + j \leq n. \quad (2.1.17)$$

Ces trois hypothèses permettent notamment de donner, après estimation des paramètres  $(\lambda_j)_{1 \leq j \leq n-1}$  et  $(\sigma_j)_{1 \leq j \leq n-1}$ , une estimation de la moyenne de la provision et également un intervalle de confiance associé à cette moyenne. De nombreuses extensions au modèle de Mack existent tels que le modèle de Schnauss (1996), le modèle de Munich Chain Ladder (G. Quarg et T. Mack, 2004, Cf. [48]) et le modèle de JAB Chain (B. Verdier et A. Klinger, 2005, Cf. [60]).

Le second type de modèles stochastiques utilisés se base sur une régression log-normale des facteurs de développement  $\frac{C_{i,j+1}}{C_{i,j}}$  (modèle de Hertig, 1985, Cf. [29] et modèle de HFB, 1985, Cf. [30]) ou sur une régression log-normale des variables  $X_{i,j}$  (modèle de S. Christophides, 1990, Cf. [18]).

Une autre alternative est de considérer les modèles GLM. Dans ce type de modèle, les variables  $X_{i,j}$  ont une densité de type exponentiel (normal, gamma, poisson, etc.), i.e. :

$$\exp \left[ \frac{x_{i,j} \theta_{i,j} - v(\theta_{i,j})}{u(\phi)} + w(x_{i,j}, \phi) \right], \quad (2.1.18)$$

où  $u$ ,  $v$  et  $w$  sont des fonctions,  $\theta_{i,j}$  et  $\phi$  sont des paramètres appelés “paramètre naturel” de la famille exponentielle et “paramètre de dispersion”. Les principaux modèles GLM utilisés en assurance non-vie sont ceux de A.E. Renshaw et R.J. Verrall (modèle poissonien, 1998, Cf. [51]) et celui de S.M. Gluck (2000, Cf. [28]).

L’ensemble des méthodes collectives (déterministe et stochastique) présentées jusqu’à maintenant ont plusieurs inconvénients en commun. Le principal inconvénient (déjà évoqué ci-dessus) est leur sur-paramétrisation, d’autre part elles sont incapables de modéliser le montant d’un sinistre concernant un assuré particulier et ainsi d’intégrer les caractéristiques personnelles des assurés. Ces méthodes ne sont pas non plus capables de prendre en compte des garanties avec des contrats de maturité différente et dont aucune génération n’est arrivée à terme (hormis l’utilisation d’un facteur de queue, tail factor en anglais, qui est discutable sur des branches à maturité longue). Un autre inconvénient (présent pour la garantie particulière que nous avons présentée dans le chapitre

précédent) est que le montant estimé de la provision est nul si le montant de sinistres connus à la date de calcul est nul : si  $C_{i,n-i+1} = 0$  alors  $P_i = 0$ . Enfin, l'hypothèse d'indépendance entre les générations n'est pas en pratique souvent vérifiée. C'est pourquoi nous présentons maintenant une alternative à ces méthodes collectives : les méthodes individuelles.

### 2.1.2 Les méthodes individuelles

Dans les modèles individuels, la provision est définie comme une somme, d'un nombre de termes aléatoires, de montant de sinistres individuels concernant un assuré particulier. Le modèle le plus utilisé est le processus de Poisson composé (Cf. [57]) dans lequel la provision de la génération  $i$  à la date  $n$  est notée  $P_i$  et est définie par :

$$P_i := \sum_{j=N_n}^{N_{n+i}} S_j, \quad (2.1.19)$$

où :

- $(N_t)_{t \geq 0}$  est un processus de Poisson et  $N_t$  représente le nombre de sinistres déclarés (ou indemnisés) dans l'intervalle de temps  $[i, t]$ ,
- $(S_j)_{j \geq 1}$  est une suite de variables aléatoires (v.a.) indépendantes et identiquement distribuées (i.i.d.), où  $S_j$  représente le montant du  $j$ -ème sinistres concernant les contrats souscrits la génération  $i$ ,
- Le processus de Poisson  $(N_t)_{t \geq 0}$  est indépendant des montants de sinistres  $(S_j)_{j \geq 1}$ .

Ce type de modèle s'illustre graphiquement de la manière suivante :

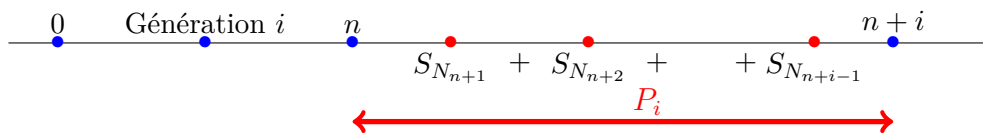


FIGURE 2.1 – Provision individuelle via le processus de Poisson composé.

Les résultats théoriques concernant les processus de Poisson composés sont nombreux : moments (espérance, variance), comportement asymptotique et loi du processus ont été déterminés. Cependant cette modélisation ne permet pas de prendre en compte certains types de dépendances :

- Puisque les v.a.  $(S_j)_{j \geq 1}$  sont indépendantes, il n'est pas possible de considérer d'éventuelles dépendances entre les "assurés".
- De ce fait, la prise en compte de l'influence d'un risque systémique externe (immobilier, récession, etc.) est impossible.



- Il n'est finalement pas permis de considérer un lien entre le processus représentant les déclarations de sinistres et leurs montants car le processus  $(N_t)_{t \geq 0}$  est indépendant des v.a.  $(S_j)_{j \geq 1}$ . En particulier le montant d'un sinistre ne dépend pas de sa date de déclaration (ou d'indemnisation).

Certains modèles permettent de prendre en compte ces dépendances. En ce qui concerne la prise en compte de la date du sinistre dans son montant, nous pouvons citer les "shot noise processes" (Cf. [6], [49] et [54]) dans lesquelles la provision est définie par :

$$P_i := \sum_{j=N_n}^{N_{n+i}} f(T_j, \xi_j), \quad (2.1.20)$$

où  $(N_t)_{t \geq 0}$  est un processus de comptage dont les temps de sauts sont notés  $(T_j)_{j \geq 1}$ ,  $f$  est une fonction mesurable,  $(\xi_j)_{j \geq 1}$  est une suite de v.a. i.i.d. et indépendante du processus de comptage. Ainsi  $f(T_j, \xi_j)$  représente le montant du  $j$ -ème sinistre concernant les contrats souscrits la génération  $i$ . A travers l'expression (2.1.20) il est clair que le montant de chaque sinistre dépend de sa date de déclaration (ou d'indemnisation).

Une autre possibilité pour prendre en compte ce type de dépendance est de modéliser la provision via des  $U$ -statistiques de processus ponctuel de Poisson (Cf. [19]), i.e. :

$$P_i := \sum_{(Y_1, \dots, Y_k) \in \eta_{\neq}^k} f(Y_1, \dots, Y_k), \quad (2.1.21)$$

où  $f$  est une fonction symétrique, l'ensemble de vecteurs aléatoires  $(Y_j)_{j \geq 1}$  forme un processus ponctuel de Poisson et  $\eta_{\neq}^k$  est l'ensemble de tous les  $k$ -tuples points distincts du processus.

Nous pouvons également citer l'article de H. Albrecher et O.J. Boxma (Cf. [2]) et la thèse de R. Biard (Cf. [12]) en ce qui concerne la prise en compte d'une dépendance entre les temps inter-sinistres et leurs montants (applications aux catastrophes naturelles).

En ce qui concerne la modélisation de la dépendance entre les montants de sinistres, nous pouvons mentionner l'utilisation des copules (Cf. [3], [8] et [62]). Il est enfin intéressant de regarder les processus de Poisson marqués (Cf. [33] et [38]).

Les méthodes individuelles présentent également des inconvénients :

1. La technicité de ces méthodes rendent leurs utilisations et explications plus compliquées que les méthodes collectives ; ces méthodes sont donc plus difficiles à mettre en place dans les compagnies d'assurances,
2. Leur temps d'implémentation est supérieur à celui des méthodes collectives,

3. Le nombre de paramètres à estimer peut, parfois, être important,
4. Ces méthodes n'ont pas de caractère générique et sont spécifiques à chaque type de garantie.

La comparaison entre méthodes collectives et individuelles peut se faire, comme le suggère R.J. Verral, de la manière suivante :

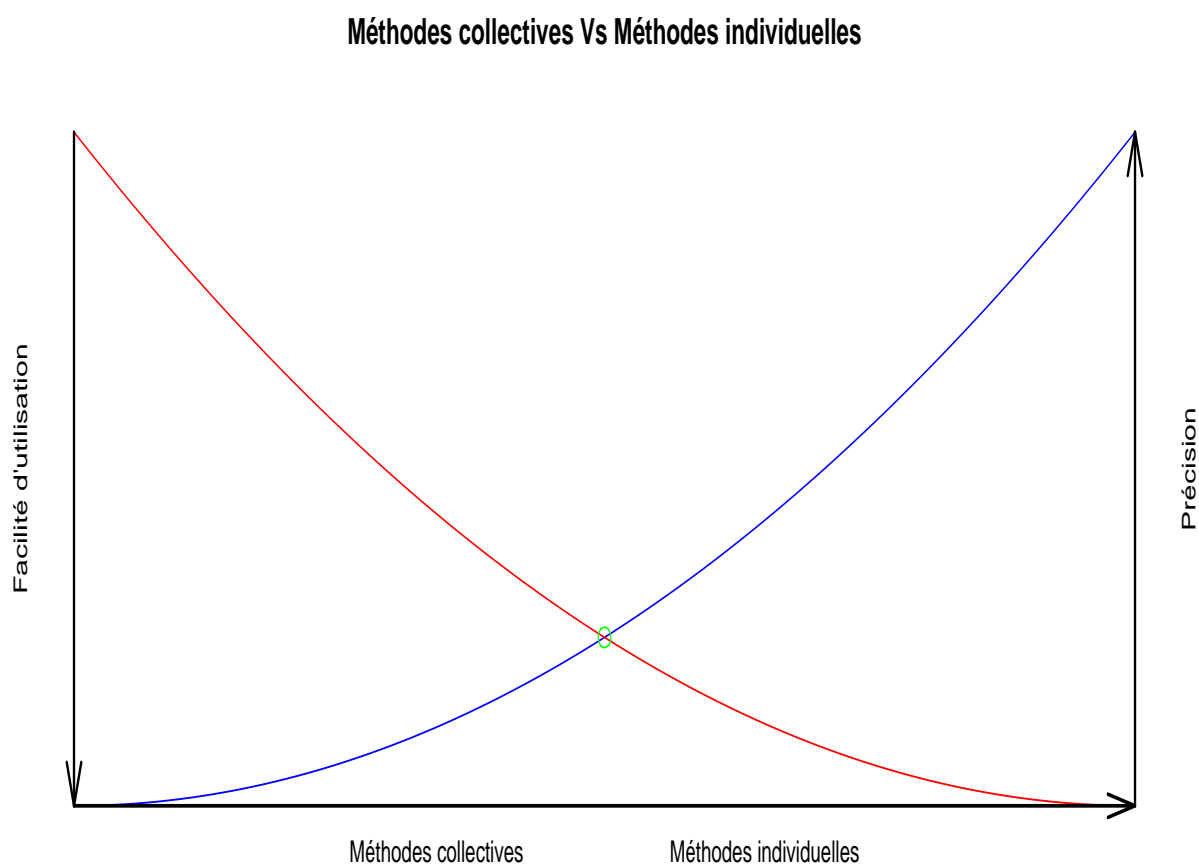


FIGURE 2.2 – Méthodes collectives Vs Méthodes individuelles.

La courbe rouge représente la “facilité d’utilisation” et la courbe bleue représente la “précision”. Le modèle idéal étant celui se situant à l’intersection des courbes bleue et rouge, c’est à dire celui qui maximise la précision et la facilité d’utilisation.

Ceci achève l’état de l’art concernant le provisionnement. Passons maintenant à la problématique de l’évaluation de la solvabilité des compagnies d’assurances.

## 2.2 La réforme Solvabilité 2

### 2.2.1 Le contexte réglementaire

La réforme Européenne Solvabilité 2 sera appliquée le 1er janvier 2016 à l'ensemble des sociétés d'assurances vie, non-vie, de réassurances ainsi qu'aux mutuelles intervenant sur le marché Européen. Cette réforme a pour unique objectif de mieux protéger et de mieux informer les souscripteurs d'assurance. L'uniformisation des règles et contraintes auxquelles sont sujettes les compagnies d'assurances (et mutuelles) devrait permettre d'atteindre cet objectif par l'uniformisation :

- des règles d'évaluation de la solvabilité en prenant en compte les risques spécifiques des compagnies assurances,
- des modes de gouvernance des compagnies,
- des informations transmises au public et aux organes de contrôle.

Le texte fondateur de cette réforme est la Directive [26]. Dans cette Directive la structure générale de la réforme est présentée : il s'agit d'un socle de trois piliers similaires à ceux présents dans la réforme Bâle 2-3 du milieu bancaire (Cf. [27]). Des spécifications et recommandations présentant les points techniques de la réforme sont fournies par l'EIOPA (European Insurance and Occupational Pensions Authority). L'EIOPA est l'organe de contrôle Européen chargé du secteur des assurances : il fait partie des trois autorités Européennes de "Surveillance du Système Européen de Supervision Financière" (avec "l'Autorité Bancaire Européenne" et "l'Autorité Européenne des Marchés Financiers").

### 2.2.2 Les trois piliers

#### Premier pilier :

Le premier pilier de la réforme Solvabilité 2 (Cf. Articles 100 à 127 de la Directive [26]) concerne l'évaluation quantitative du capital de solvabilité requis (CSR) lié aux risques auxquels sont soumises les compagnies d'assurances. Le CSR est défini comme le montant que les compagnies doivent détenir en fond propre pour faire face à la réalisation de leurs risques. Avant de voir une illustration graphique donnons quelques explications sur le bilan d'une compagnie d'assurances :

- L'actif d'une compagnie d'assurances est principalement constitué des "placements" dans lesquels elle a investi les primes perçues : actions, obligations, immobilier, etc.. Il s'agit d'argent que la compagnie possède et est plus ou moins liquide.
- La réassurance est le montant que les assureurs devraient recevoir de la part des réassureurs : il s'agit donc de sommes dues du fait de transfert de primes de risque.
- Les fonds propres sont les capitaux investis par les actionnaires (les "propriétaires") de la compagnie y compris les provisions libres et le résultat à reporter. La compagnie doit cet argent à ses actionnaires.

- Les provisions constituent le montant estimé que la compagnie devra reverser aux bénéficiaires des contrats. Il s'agit donc d'argent dû par la compagnie aux bénéficiaires des contrats d'assurances.

Passons à une illustration graphique du CSR.

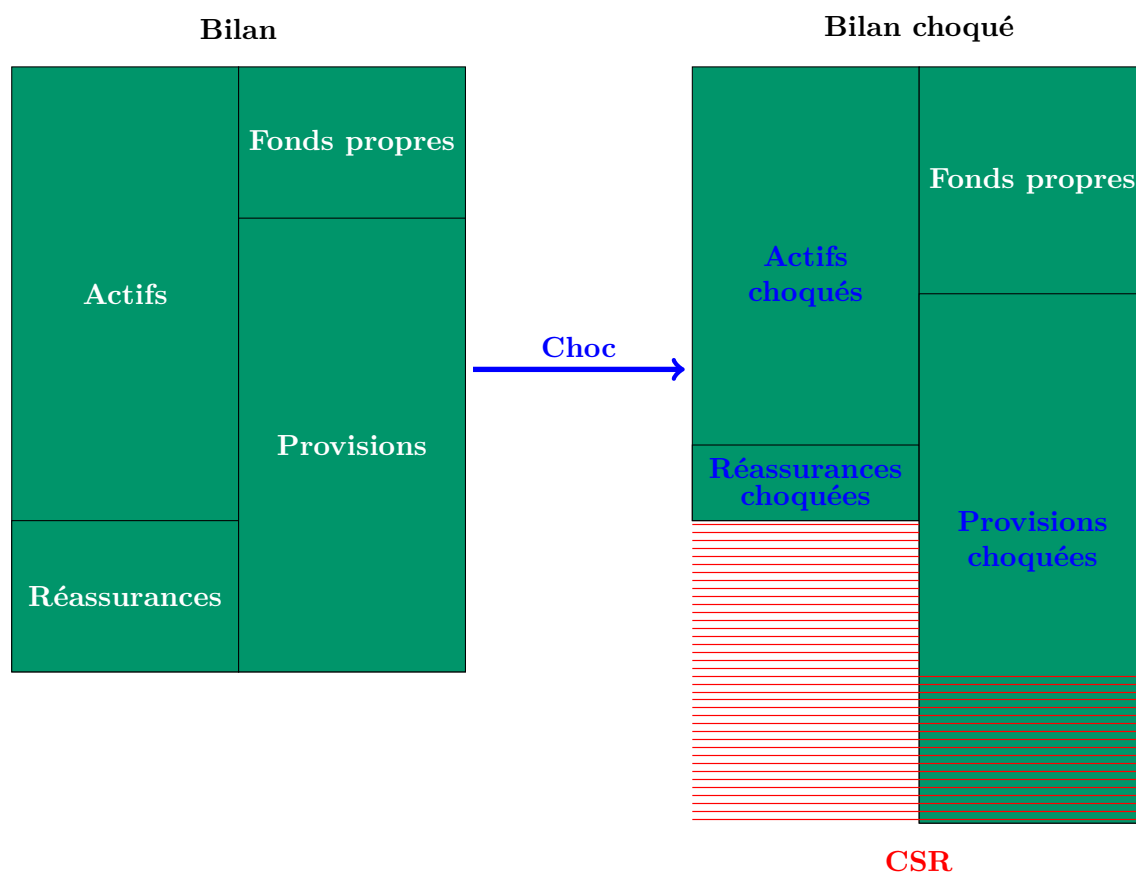


FIGURE 2.3 – Bilan choqué.

Les risques considérés dans l'évaluation du CSR sont imposés par la Directive et dépendent de la nature de la compagnie. Pour les compagnies d'assurances non-vie, les risques à prendre en compte sont au moins les suivants :

- Risque de marché : action, taux, spread, concentration, immobilier, change,
- Risque de défaut des contreparties : réassureurs, banques dépositaires, etc.,
- Risque de souscription : prime et provision (également appelé "prime et réserve"), rachat, catastrophe,

- Risque opérationnel : fraude, défaut du système informatique, etc..

L'EIOPA a proposé une formule standard pour l'évaluation du CSR associé à chacun des risques énoncés ci-dessus. Ce CSR est "calibré" comme une Value At Risk sur un an à 99,5%. En effet l'Article 101 de la Directive [26] spécifie que :

*“Le capital de solvabilité requis correspond à la valeur en risque (Value-at-Risk) des fonds propres de base de l'entreprise d'assurance ou de réassurance, avec un niveau de confiance de 99,5% à l'horizon d'un an.”*

A titre d'exemple, le CSR lié au risque action est calibré comme le montant de la perte sur le portefeuille d'actions et sur un an qui ne sera dépassé que dans 0,5% des cas. L'EIOPA fournit également une matrice de corrélation entre chacun des risques permettant de déterminer un CSR global (tous risques confondus). La Directive laisse la possibilité aux assureurs d'utiliser leur propre formule d'évaluation du CSR, s'ils jugent et démontrent le caractère inadéquat de la formule proposée par l'EIOPA : cette modélisation peut être effectuée sur l'ensemble des risques (modèle interne) ou sur certains risques (modèle interne partiel).

### **Second pilier :**

Le second pilier de la réforme Solvabilité 2 (Cf. Articles 41 à 50 de la Directive [26]) concerne le mode de gouvernance des compagnies d'assurances et également une évaluation interne de leurs risques propres. La Directive impose une certaine structure de gouvernance aux compagnies, mais impose surtout une "procédurisation" de cette gouvernance : le système de gouvernance doit être décrit, commenté et connu des personnes clés de la compagnie. Pour ce faire les compagnies d'assurances doivent déterminer quatre fonctions clés au sein de l'entreprise : la gestion des risques, la conformité, l'audit interne et la fonction actuarielle. Les relations entre ces fonctions clés, le conseil d'administration et les dirigeants de la compagnie doivent être clairement établies et documentées. Les questions de sous-traitance et de communication (aux autorités de contrôle, aux personnels et aux organes de décisions) prennent également une part importante dans le second pilier.

Un autre élément imposé dans ce pilier est l'évaluation par la compagnie elle-même de ses risques propres. Cette évaluation est connue en pratique sous le nom d'ORSA (Own Risk Solvency Assessment) et est décrite par l'Article 45 de la Directive [26] :

*“Dans le cadre de son système de gestion des risques, chaque entreprise d'assurance et de réassurance procède à une évaluation interne des risques et de la solvabilité.”*

Cette évaluation se base sur une cartographie exhaustive des processus présents au sein de la compagnie d'assurances et également sur une cartographie exhaustive des risques supportés par la compagnie. Pour l'ensemble des risques quantifiables, les compagnies d'assurances doivent déterminer le besoin de solvabilité associé : cette évaluation est

similaire à celle imposée dans le premier pilier (CSR) mais ni l'indicateur de risque ni l'horizon d'analyse ne sont spécifiés. Concrètement, les compagnies d'assurances peuvent choisir un autre indicateur de risque que la Value At Risk et un autre horizon d'analyse qu'un an. Les risques considérés dans cet exercice doivent être au minimum ceux du pilier 1. Le dernier point composant l'ORSA est celui des stress tests à appliquer à l'évaluation de la solvabilité dans une vision prospective sur la durée du plan d'affaires. En résumé, le but de l'ORSA est l'évaluation de la solvabilité des compagnies d'assurances par elle-même, limitant ainsi les éventuelles faiblesses présentes dans la formule standard de calcul du CSR.

### Troisième pilier :

Le troisième pilier est quant à lui consacré à la transmission des informations au public et aux autorités de contrôle. Nous ne détaillerons pas plus ce point de la réforme.

### 2.2.3 Les méthodes d'évaluation du capital et du besoin de solvabilité

La seconde partie de cette thèse sera consacrée au calcul du besoin de solvabilité (ORSA) lié au risque de provisionnement. Le risque de provisionnement représente le risque que les provisions calculées par la compagnie d'assurances soient insuffisantes. Il peut être mesuré par la différence entre la provision et son espérance mathématique (puisque c'est cette espérance que l'assureur enregistre dans ses comptes). Ce que nous illustrons graphiquement sur l'exemple suivant (Cf. explications du graphique sur la page suivante) :

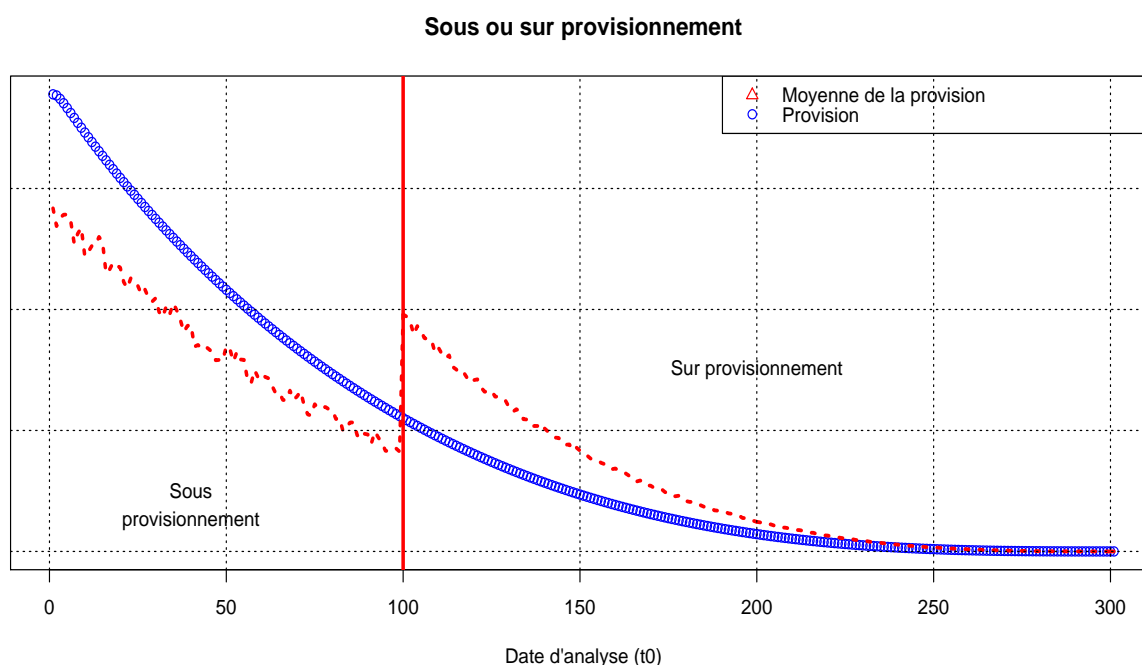


FIGURE 2.4 – Risque de provisionnement.

Dans le précédent graphique :

- La date d'analyse, que nous avons noté  $t_0$ , est en abscisse,
- La moyenne de la provision est représentée par la courbe rouge,
- La "vraie valeur" de la provision est représentée par la courbe bleue.

La différence entre la "vraie valeur" de la provision et son espérance mesure le risque d'un sous ou sur-provisionnement : sur le graphique précédent nous constatons un sous-provisionnement (la moyenne est inférieure à la vraie valeur) avant la date  $t_0 = 100$  et un sur-provisionnement après cette date.

Le CSR lié au risque de provisionnement est alors défini comme la Value At Risk, à 99,5%, sur la variable représentant le risque de provisionnement sur un an. La formule standard proposée par l'EIOPA ne différencie pas le risque de provisionnement du risque de primes qui représente le risque d'une sous-tarifcation. Dans la formule standard ce CSR est approximé par :

$$CSR := 3\sigma V, \quad (2.2.1)$$

où  $V$  est le volume de l'activité incluant un volume de provisions et un volume de primes (acquises, émises) et  $\sigma$  représente la volatilité globale associée au risque de primes et provisions. Nous ne détaillerons pas le calcul des quantités  $V$  et  $\sigma$  qui n'a que peu d'intérêt ici et qui est par ailleurs entièrement fourni dans les spécifications techniques de l'EIOPA (Cf. [1]).

La Directive laisse la possibilité aux compagnies d'assurances de calculer le paramètre  $\sigma$  en utilisant leurs propres données (Undertaking Specific Parameters, USP). L'EIOPA fournit trois méthodes permettant le calcul de ces USP et donc de la volatilité spécifique. Ces méthodes sont détaillées dans le "consultation paper 75" de l'EIOPA et dans les mémoires d'actuariat [32] et [44]. Les méthodes deux et trois se basent sur l'approche proposée par Wuthrich dans [65]. Notons que l'utilisation des USP par les compagnies d'assurances s'inscrit dans le cadre de la formule standard : il ne s'agit pas d'une modélisation interne.

Comme énoncé précédemment, la Directive laisse la possibilité de calculer le CSR lié au risque de provisionnement par une modélisation interne propre à la compagnie d'assurances. Nous pouvons citer par exemple les articles [41] et [46] traitant de ce sujet.

Contrairement au calcul du CSR, la méthode de calcul du besoin de solvabilité, imposé dans le pilier 2 de la Directive, n'est pas spécifiée. Le calcul du besoin de solvabilité peut être interprété comme un modèle interne pour le calcul du CSR dans lequel le choix de la mesure de risque et de l'horizon d'analyse est laissé libre à la compagnie d'assurances. Nous pouvons citer les travaux de [16], [47] et [58] sur ce type de problématique.





# Conclusion et perspectives

---

Une traduction de ce chapitre, en anglais, est disponible page [181](#).

## 3.1 Conclusion

Dans cette thèse nous avons adopté le point de vue d'un assureur devant indemniser une banque à la suite de pertes liées à un défaut de remboursement de ses emprunteurs.

Dans une première partie nous avons modélisé le montant potentiel des pertes liées à ces défauts de paiement pour une période donnée ; ce montant est appelé provision. La quantité clé de notre modèle est le montant d'un défaut. Pour un emprunteur  $j$  et une date de fin de prêt  $T_j$ , ce montant vaut :

$$M_j \times \max [\phi_{T_j}(D_j) - R_{T_j}; 0 ], \quad (3.1.1)$$

où :

- $M_j$  est le montant emprunté et  $D_j$  est la durée du prêt,
- $M_j \phi_{T_j}(D_j)$  est le montant dû par l'emprunteur et dépend des caractéristiques personnelles des emprunteurs et  $\phi$  est une fonction déterministe indépendante de  $j$ ,
- $M_j R_{T_j}$  est le montant de la revente du bien immobilier financé par le prêt. Le coefficient de proportionnalité  $R_{T_j}$  est modélisé par un mouvement Brownien géométrique et représente les fluctuations des prix de l'immobilier. Nous supposons que ce coefficient est identique pour tous les emprunteurs, ainsi il existe une dépendance entre les montants de défauts de chaque emprunteur.

La loi des couples (Date de fin du prêt  $T_j$ , Durée du prêt  $D_j$ ) est modélisée par un processus ponctuel de Poisson ce qui nous permet de définir la provision  $P_h$ , où  $h$  est la durée maximale des contrats considérés, comme une somme d'un nombre aléatoire de montants de défauts individuels :

$$P_h := \sum_{j \geq 1} \mathbb{1}_{\{(T_j, D_j) \in A_h\}} M_j \max [\phi_{T_j}(D_j) - R_{T_j}; 0 ], \quad (3.1.2)$$

où  $A_h$  est l'ensemble de variation des prêts en défaut de paiement. A la fois théorique et appliquée, la formule (3.1.2) permet de calculer, via la prise en compte de plusieurs dépendances, l'espérance et la variance mais aussi de donner un algorithme de simulation de la provision. Le nombre important de données à notre disposition permet d'estimer

les paramètres liés au modèle et de fournir une valeur numérique aux quantiles de la provision.

Dans une deuxième partie nous nous sommes intéressés au besoin de solvabilité associé au risque de provisionnement (problématique imposée par la réforme européenne Solvabilité 2). Le besoin de solvabilité, noté  $BS(h)$ , est défini comme le quantile de la provision centrée, i.e. :

$$BS(h) := \mathcal{Q}_{P_h - \mathbb{E}[P_h]}(99, 5\%), \quad (3.1.3)$$

Pour calculer  $BS(h)$  nous avons étudié le comportement asymptotique de  $P_h$  lorsque  $h \rightarrow +\infty$  et nous avons montré que :

$$\frac{P_h - \mathbb{E}(P_h)}{h\theta_1} \xrightarrow[h \rightarrow +\infty]{\mathcal{L}} G - \int_0^{+\infty} \inf(\theta_2, \theta_3 R_t) dt + \theta_4, \quad (3.1.4)$$

où  $\theta_1, \theta_2, \theta_3$  et  $\theta_4$  sont des paramètres et  $G$  est une variable aléatoire gaussienne centrée, réduite et indépendante de  $(R_t)_{t \geq 0}$ . Par conséquent, nous avons approximé  $BS(h)$  par la quantité  $BS^*(h)$  définie par :

$$BS^*(h) := h\theta_1 \times q_{99,5\%}, \quad (3.1.5)$$

où  $q_{99,5\%}$  est le quantile, à l'ordre 99,5%, de :

$$G - \int_0^{+\infty} \inf(\theta_2, \theta_3 R_t) dt + \theta_4. \quad (3.1.6)$$

La formule précédente est seulement théorique. En effet la loi de  $\int_0^{+\infty} \inf(\theta_2, \theta_3 R_t) dt$  est inconnue. Cependant nous avons évoqué précédemment la possibilité de simuler la provision : ces simulations nous permettent de donner une valeur numérique au besoin de solvabilité.

Notons que s'il n'y avait pas de corrélation entre les montants de chaque sinistre, la loi limite de la provision, convenablement renormalisée, convergerait vers une variable aléatoire gaussienne, centrée et réduite. Le terme additionnel qui s'exprime à l'aide d'une intégrale du mouvement Brownien géométrique prend en compte les corrélations. Il y a donc une "séparation des sources de risques", l'une provenant des caractéristiques individuelles des emprunteurs qui sont indépendants et l'autre provenant du coût "commun" de l'immobilier.

## 3.2 Perspectives

### Implémentation de notre modèle :

La première perspective à l'issue de cette thèse est d'implémenter notre modèle au sein de CAMCA Assurance. Cette implémentation nécessitera :

1. De tester (et “back-tester”) notre modèle sur l’ensemble des données disponibles (nous nous sommes limités dans cette thèse à une unique banque).
2. D’expliquer en détail notre modèle ainsi que les hypothèses que nous avons effectuées.
3. D’implémenter en pratique notre modèle (création d’un package  $\mathbb{R}$ ?) en tenant compte du système informatique de CAMCA Assurance.

#### **Investissement de la prime perçue :**

Nous montrerons dans la Section 5.2 que la garantie proposée par CAMCA Assurance peut s’interpréter via les options d’échange. Cette interprétation suggère la question de l’investissement des primes perçues par la compagnie d’assurances. Pour déterminer une stratégie d’investissement des primes perçues, ou de couverture, il est nécessaire de définir un marché formé d’actifs exogènes à ceux intervenant dans les options d’échange. Lorsque la maturité d’une option d’échange est fixée, il existe des stratégies de couverture de cette option. La difficulté sera ici de prendre en compte l’aléa sur le nombre des défauts, i.e. sur le nombre d’option d’échange, et sur les dates de ces défauts, i.e. sur les dates d’exercice des options d’échange.

#### **Profil de risque :**

Une autre question intéressante est celle du profil de risque des emprunteurs. Nous pourrions déterminer les principales caractéristiques communes des emprunteurs présentant un défaut de paiement et également déterminer les caractéristiques annonçant un défaut futur. L’outil utilisé serait plutôt l’analyse des données.

#### **Solvabilité 2 :**

Enfin nous pourrions nous intéresser à différents aspects de la réforme Solvabilité 2. Dans cette thèse nous nous sommes limités à l’étude du besoin de solvabilité pour une génération donnée. Cependant la méthode asymptotique que nous avons proposée est uniquement théorique. En effet nous voyons à travers la formule (3.1.4) que la loi limite de la provision renormalisée dépend de la variable :

$$\int_0^{+\infty} \inf(\theta_2, \theta_3 R_t) dt. \quad (3.2.1)$$

Il paraît naturel d’étudier la loi de cette intégrale.

Nous pourrions également évaluer la solvabilité liée à l’ensemble des risques évoqués dans la réforme Solvabilité 2 (risque de marché, de contrepartie, etc.).



## Deuxième partie

# Provisionnement face au risque de défaut des emprunteurs

---

|          |   |           |
|----------|---|-----------|
| <b>4</b> | <b>Provisionnement face au risque de défaut des emprunteurs</b> | <b>33</b> |
| 4.1      | Défaut de paiement des emprunteurs                              | 34        |
| 4.2      | Modélisation de la provision                                    | 36        |
| 4.3      | Distribution de la provision                                    | 43        |
| 4.4      | Estimations des paramètres                                      | 46        |
| 4.5      | Preuves   | 50        |
| <b>5</b> | <b>Compléments au modèle de provisionnement</b>                 | <b>65</b> |
| 5.1      | Garantie caution via les “credit default swap”                  | 65        |
| 5.2      | Provision via les options d’échange                             | 67        |
| 5.3      | Justifications des hypothèses                                   | 70        |
| 5.4      | Application numérique   | 76        |

---



# Provisionnement face au risque de défaut des emprunteurs

---

Ce chapitre correspond au contenu d'un article soumis en février 2014 dans la revue "*Insurance : Mathematics and Economics*".

Nous nous intéressons dans ce chapitre au risque de défaut des emprunteurs. Nous nous plaçons du point de vue d'un assureur devant indemniser une banque suite aux pertes liées à un défaut de remboursement de ses emprunteurs.

L'objet de ce chapitre est la modélisation de la provision face au risque de contrepartie des emprunteurs. Contrairement aux méthodes traditionnellement utilisées, notre modèle repose de manière essentielle sur les caractéristiques individuelles des emprunteurs (montant et durée des emprunts). L'apport de notre modèle réside également dans la prise en compte de trois types de dépendance : entre les dates des défauts de paiement et leurs montants, entre le montant de sinistres de chaque emprunteur et celle entre le nombre de sinistres et les montants de sinistres. Nous préciserons les liens entre notre modélisation et les autres modèles considérés dans la littérature.

Un des résultats théoriques importants de ce chapitre est le calcul de la moyenne et de la variance de la provision. Nous fournirons également la loi de la provision conditionnellement au nombre de sinistres. Une autre caractéristique de notre modèle est qu'il est applicable. En effet, une fois donnés les paramètres du modèle, nous obtenons une formule programmable de l'espérance, de la variance ainsi qu'un algorithme de simulation nous permettant de calculer des quantiles. Pour être complet nous proposerons une méthode d'estimation de l'ensemble des paramètres présents dans notre modèle.

**Mots clés :** *Risque de défaut ; Provisionnement stochastique individuel ; Processus ponctuel de Poisson ; Mouvement Brownien géométrique ; Temps de défaut ; Quantile.*

Une traduction de chapitre, en anglais, est disponible page [187](#).

## 4.1 Défaut de paiement des emprunteurs

### 1) Problématique :

Nous considérons dans la suite une compagnie d'assurances devant indemniser une banque à la suite de pertes liées aux défauts de paiement de ses emprunteurs. Les emprunts considérés sont des emprunts immobiliers. Un emprunteur est considéré être dans l'état "sinistre" à une date  $T$  s'il est dans l'incapacité irréversible de rembourser les mensualités de son prêt à partir de la date  $T$ .

La problématique pour la compagnie d'assurances est d'évaluer, à une date  $t_0$ , la charge des sinistres à venir provenant des emprunteurs dont le prêt a commencé avant la date  $t_0$  et qui n'ont pas eu de sinistre. Cette somme est appelée "provision". Il est important de préciser qu'en cas de sinistre la banque peut tenter de récupérer de l'argent directement auprès de l'emprunteur ou également sur le bien financé. Bien entendu le montant réel des futurs sinistres est inconnu à la date  $t_0$  car il dépend d'évènements futurs incertains. Le principal objectif de ce chapitre est de proposer une modélisation stochastique simple mais qui prend en compte les aspects clés du contexte.

La charge d'un sinistre est la quantité clé de notre étude. Pour un client donné et une date de sinistre fixée  $T$ , cette quantité est égale à  $(S_T^1 - S_T^2)_+$ , où  $S_T^1$  est la part non remboursée du prêt à la date  $T$  et  $S_T^2$  est le montant de la revente du bien à la date  $T$  (Cf. Section 4.2.2).  $S_T^1$  est proportionnelle au montant emprunté ; le coefficient de proportionnalité dépend de la durée du prêt et de la date du sinistre.  $S_T^2$  est proportionnel au montant emprunté ; le coefficient de proportionnalité représente l'évolution des prix de l'immobilier, l'apport personnel de l'emprunteur et les coûts liés à la vente. Ce coefficient est modélisé par un mouvement Brownien géométrique. Nous supposons que le processus représentant les fluctuations de l'immobilier est le même pour tous les emprunteurs, donc qu'il existe une dépendance entre chaque montant de sinistres. Nous modéliserons la loi des couples (Date de fin de prêt, Durée initiale du prêt) par un processus ponctuel de Poisson (Cf. Section 4.2.3), ce qui conduira à définir la provision comme la somme des charges de sinistres individuels (Cf. Définition 4.2.6). Pour plus de détails concernant les modèles individuels se référer à [7], [33], [38], [45], [56] et à la Section 2.1.2.

Dans la Section 4.2.4 nous examinerons les liens entre notre modélisation et celles de T. Mack et Cramer - Lundberg. Contrairement à ceux-ci notre modèle considère trois différents types de dépendance (qui sont les principaux apports de notre modèle) :

1. La dépendance entre les montants de sinistres des emprunteurs (car les fluctuations des prix de l'immobilier sont communes pour chaque emprunteur),
2. La dépendance entre les montants de sinistres et les dates de déclaration de ces sinistres,
3. Et celle entre la durée des prêts et les dates de sinistres.



Avec quelques hypothèses supplémentaires, nous calculerons dans la Section 4.3 la moyenne (Cf. Théorème 4.3.2) et la variance (Cf. Théorème 4.3.4) de la provision. Notre modèle permet d'aller plus loin que le calcul de ces deux paramètres. Il est possible de simuler la provision (Cf. Théorème 4.3.6). Nous concluons ce chapitre par l'estimation des paramètres de notre modèle (Cf. Section 4.4). Les nombreuses données individuelles dont CAMCA Assurance dispose nous permettrons de donner une valeur numérique aux paramètres de notre modèle et ainsi obtenir une valeur numérique de l'espérance, de la variance ainsi que du quantile de cette provision (Cf. Section 5.4 du Chapitre 5 pour plus de détails). Les preuves des différents résultats sont données dans la Section 4.5.

### 2) *Etat de l'art :*

Les modèles traditionnellement utilisés en assurance non-vie pour le calcul de ce type de provisions sont des modèles collectifs (Chain Ladder, Mack, Cf. [35] et [64]) dans le sens où ils agrègent les données et ne sont pas capables de calculer la charge d'un unique sinistre. L'inconvénient de ces modèles est double : d'une part l'agrégation des données implique une sur-paramétrisation et d'autre part il n'y a aucune justification pour supposer que les générations de contrat sont indépendantes. D'autres types de modèles existent mais sont très peu utilisés par les assureurs proposant ce type de garantie : il s'agit de modèles individuels issus de la théorie de la ruine, le plus connu étant celui de Cramer - Lundberg (Cf. [52]). Dans ce cadre, la provision est une somme aléatoire de charges de sinistres individuels. L'inconvénient de cette approche est qu'elle ne prend en compte ni la dépendance entre les montants de sinistres, ni la dépendance entre le montant de sinistres et la date du sinistre et ni la dépendance entre le nombre de sinistres et le montant de ces sinistres. Pour une présentation détaillée des méthodes usuelles de provisionnement se référer au Chapitre 2.

### 3) *Conclusion :*

Nous décrivons une méthode alternative de provisionnement adaptée aux contrats d'assurances spécifiques étudiés, reposant sur un modèle stochastique individuel simple qui prend en compte les caractéristiques des emprunteurs et quelques dépendances. A la fois théorique et appliqué, ce modèle donne une formule explicite pour l'espérance et la variance de la provision et permet également de simuler la provision. Cette dernière possibilité autorise le calcul de divers indicateurs de risque (bande de confiance et quantile). Toutefois, certaines des hypothèses de notre modèle sont trop fortes et mériteraient d'être relâchées. Nous avons en particulier supposé que le taux d'emprunt est le même pour tous les emprunteurs et que s'il existe un défaut, il est instantané et irréversible. Nous avons cependant essayé de justifier le choix de l'ensemble de nos hypothèses dans la Section 5.3 du Chapitre 5.

## 4.2 Modélisation de la provision

Nous présenterons les notations et hypothèses dans la Section 4.2.1. Puis, dans la Section 4.2.2, nous modéliserons la charge d'un unique sinistre déclaré à une date fixée. Dans la Section 4.2.3 nous modéliserons les couples (Date de fin du prêt, Durée initiale du prêt) par un processus ponctuel de Poisson (Cf. hypothèse  $(H_7)$ ). Ceci nous permettra d'exprimer la provision comme une somme de charges de sinistres individuels intervenant à des dates aléatoires (Cf. Définition 4.2.6). C'est la définition clé qui va permettre notre analyse quantitative (calcul de moyenne et variance) ainsi que des simulations. Finalement, dans la Section 4.2.4, nous relierons notre approche à celle provenant de la théorie de la ruine et du modèle de Mack.

### 4.2.1 Notations et hypothèses

Dans ce qui suit nous raisonnerons à temps continu. Le temps  $t$  est défini comme une fraction d'année et un mois sera noté  $t^*$ . Il est nécessaire d'introduire certaines notations :

**Notation 4.2.1.** *Nous noterons :*

- $B$  ('B' pour "borrower") l'ensemble des emprunteurs.
- Les emprunteurs par l'indice  $j : j \in B$ ,
- $I_j$  ('I' pour "inception") la date de souscription du prêt de l'emprunteur  $j : I_j \in [0, t_0]$ ,  $t_0$  étant la dernière date considérée. Nous confondrons par la suite  $I_j$  avec la date de déblocage des fonds par la banque et supposons également que l'ensemble des fonds est déblocqué à la même date.
- $D_j$ ,  $j \in B$  ('D' pour "Durée"), la v.a. représentant la durée du prêt de l'emprunteur  $j$ ,
- $M_j$ ,  $j \in B$  ('M' pour "Montant"), la v.a. représentant le montant du prêt de l'emprunteur  $j$ .

Nous associons ainsi à chaque emprunteur  $j$  un couple  $(D_j, M_j)$ .

L'emprunteur  $j$  a souscrit son prêt le mois  $i$  si  $i = \lfloor \frac{I_j}{t^*} \rfloor$ , où  $\lfloor x \rfloor$  est la partie entière de  $x$ . Nous supposons qu'un prêt souscrit le mois  $i$  est souscrit au début du mois  $i$  :

$(H_1)$  : Si l'emprunteur  $j$  a souscrit son prêt à la date  $I_j$  correspondant au mois  $i$  alors nous supposons que  $I_j = t^*i$ .

Cf. la Section 5.3.1 pour une justification de cette hypothèse.

Les v.a. représentant les durées des prêts sont en pratique majorées par une quantité inconnue  $h > 0$  tel que :

$$(H_2) : \begin{cases} D_j \leq h, & j \in B, & (4.2.1) \\ t_0 < I_j + h, & j \in B. & (4.2.2) \end{cases}$$

Cf. la Section 5.3.2 pour une justification de cette hypothèse. Ceci s'illustre de la manière suivante :



FIGURE 4.1 – Majoration de la durée des prêts.

Rappelons que la date 0 représente la première date de souscription d'un prêt.

Nous considérons, dans la suite, des prêts immobiliers dont le remboursement s'effectue de manière mensuelle et constante. Nous supposons que le taux du prêt de chaque emprunteur est le même et qu'il ne varie pas dans le temps. Soit  $\tau$  le taux exponentiel commun à tous les emprunteurs. Nous supposons enfin que les mensualités d'un prêt ne peuvent être remboursées de manière ni anticipée ni partielle. Ainsi nous supposons :

$$(H_3) : \begin{cases} \text{Remboursement mensuel constant,} & (4.2.3) \\ \tau \text{ est le taux exponentiel déterministe,} & (4.2.4) \\ \text{Pas de remboursement anticipé partiel.} & (4.2.5) \end{cases}$$

Lorsqu'un emprunteur connaît un défaut de paiement, a priori non définitif, la banque a deux possibilités : le déclarer immédiatement en sinistre ou attendre que l'emprunteur régularise sa situation, i.e. ne pas le déclarer en sinistre immédiatement. Dans cette situation le nombre de mensualités impayées peut varier. Nous supposons cependant qu'un emprunteur présentant un défaut de paiement est immédiatement déclaré en sinistre par la banque, i.e. que la date du premier défaut de paiement est identique à la date de déclaration du sinistre. Notons également qu'un défaut de paiement peut concerner une certaine part de la mensualité à régler ou son intégralité. Pour simplifier, nous supposons qu'un défaut de paiement concerne l'intégralité d'une mensualité. Ainsi nous supposons :

$$(H_4) : \begin{cases} \text{Emprunteur immédiatement déclaré en sinistre,} & (4.2.6) \\ \text{Un défaut concerne l'intégralité d'une mensualité.} & (4.2.7) \end{cases}$$

Cf. la Section 5.3.4 pour une justification de cette hypothèse.

Comme précisé dans l'introduction, la banque a la possibilité en cas de défaut d'utiliser tous les moyens légaux pour récupérer tout ou partie de sa créance. Les montants ainsi récupérés sont qualifiés de montants de recours et deux cas sont possibles :

1. La banque peut tenter de récupérer et de revendre le bien financé en cas de sinistre : vente automatique si la banque a pris une garantie judiciaire ou probable sinon.
2. La banque peut également espérer récupérer de l'argent auprès de l'emprunteur : héritage, retour à meilleure fortune, etc..

Les recours sur l'emprunteur sont peu fréquents. En effet en cas de "sinistre" la trésorerie de l'emprunteur est souvent quasiment nulle, sinon il ne serait pas en "sinistre". Les

recours proviennent, en grande partie, de la revente du bien. A nouveau pour simplifier, nous supposons que le recours sur l'emprunteur est systématiquement nul et qu'une éventuelle revente du bien financé en cas de sinistre est réalisée à l'instant du sinistre, même si en pratique il faut attendre un certain laps de temps pour encaisser le montant de recours. Nous supposons que le montant de la revente du bien à la date  $T$  est proportionnel au montant du prêt. Le coefficient de proportionnalité dépend de l'évolution des prix de l'immobilier, de l'apport personnel de l'emprunteur et des coûts liés à la revente du bien (frais de notaire, de procédures judiciaires, etc.). Nous supposons pour simplifier que le cours de l'immobilier dépend seulement de l'intervalle de temps entre la date  $T$  et la date de souscription du prêt  $I_j = t^*i$ . Synthétisons l'ensemble des précédentes hypothèses :

$$(H_5) : \begin{cases} \text{Le recours sur l'emprunteur est nul,} \\ \text{Le recours sur le bien financé encaissé à la date } T \text{ vaut :} \\ \gamma M_j R_{T-I_j}, \quad j \in B, \quad I_j \leq T \leq I_j + D_j, \end{cases} \quad (4.2.8)$$

où  $\gamma \in \mathbb{R}$  est un coefficient indépendant des emprunteurs représentant l'apport personnel et les coûts liés à la revente (Cf. (5.3.12)) et  $(R_t)_{t \geq 0}$  représente les fluctuations des prix de l'immobilier. Nous observons à travers (4.2.8) que le coefficient  $R_t$  ne dépend pas des emprunteurs, ainsi il existe une dépendance entre le montant de l'éventuelle revente du bien de chaque emprunteur. Des détails sur cette modélisation sont donnés dans la Section 5.3.5 du Chapitre 5.

Pour les applications numériques nous supposons que  $(R_t)_{t \geq 0}$  est le mouvement Brownien géométrique de paramètres  $\mu$  et  $\sigma$  tel que  $R_0 = 1$ , i.e. :

$$R_t = \exp [\sigma B_t + \mu t ], \quad (4.2.9)$$

où  $(B_t)_{t \geq 0}$  est un mouvement Brownien standard.

## 4.2.2 Modélisation de la charge d'un sinistre

Nous exprimons la charge d'un sinistre déclaré à une date fixée à l'aide des caractéristiques des emprunteurs et de la date de déclaration de ce sinistre. C'est une notion clé de notre approche.

**Définition 4.2.2.** *Considérons un sinistre déclaré à la date  $T$  et concernant l'emprunteur  $j$ . Notons :*

1.  $S_{T-I_j}^{1,j}$  la somme, à la date  $T$ , des capitaux impayés jusqu'à la date d'échéance du prêt, des intérêts impayés jusqu'à la date de déclaration du sinistre et des intérêts de retard,
2.  $S_{T-I_j}^{2,j}$  le montant de la revente du bien à la date  $T$ .

Le montant de sinistres est défini par :

$$(S_{T-I_j}^{1,j} - S_{T-I_j}^{2,j})_+, \quad I_j \leq T \leq I_j + D_j. \quad (4.2.10)$$

Notons que bien que le sinistre intervienne à la date  $T$  nos notations prennent en compte le fait que le prêt soit souscrit à la date  $I_j$  et non à la date 0, c'est pourquoi il existe un "décalage" dans nos notations :  $T - I_j$ .

Il est clair que d'après (4.2.10) le montant d'un sinistre ne peut être négatif. En effet en cas de recours supérieur au montant de retard en capital, en intérêt et en intérêt de retard, la compagnie d'assurances ne versera rien à la banque puisque c'est la banque qui récupère le montant des recours. La charge du sinistre ne représente pas le montant réel du sinistre mais le montant que l'assurance reversera à la banque en cas de sinistre ; ceci explique la présence de la partie positive dans l'expression (4.2.10).

Sous les hypothèses  $(H_1)$  à  $(H_5)$ , les quantités  $S_t^{1,j}$  et  $S_t^{2,j}$  peuvent être exprimées via les caractéristiques personnelles de l'emprunteur  $j$  ( $D_j$  et  $M_j$ ), le coefficient  $\gamma$  (Cf. (5.3.12)) et les fluctuations des prix de l'immobilier  $(R_t)_{t \geq 0}$ .

**Proposition 4.2.3.** *Plaçons nous sous les hypothèses  $(H_1)$  à  $(H_5)$ .*

1. *Nous avons :*

$$S_t^{1,j} := M_j \phi_t(D_j), \quad (4.2.11)$$

$$S_t^{2,j} := \gamma M_j R_t, \quad (4.2.12)$$

où  $\gamma$  et  $(R_t)_{t \geq 0}$  sont introduits dans (4.2.9) et :

$$\phi_t(s) := \begin{cases} \alpha, & \text{si } s \leq t^*, \\ \alpha, & \text{si } t \leq t^* \leq s, \\ \alpha \frac{e^{\tau s} - e^{\tau(t-t^*)}}{e^{\tau s} - 1}, & \text{si } 0 < t^* < t \leq s, \end{cases} \quad (4.2.13)$$

$$\alpha := e^{\tau t^*} (1 + t^* p) - t^* p, \quad (4.2.14)$$

avec  $p \in \mathbb{R}_+$  représentant le coefficient lié aux intérêts de retard (Cf. (4.5.7) pour plus de détails sur ce coefficient).

2. *En combinant (4.2.10), (4.2.11) et (4.2.12) il vient que :*

$$(S_{T-I_j}^{1,j} - S_{T-I_j}^{2,j})_+ := M_j f(T - I_j, D_j, R_{T-I_j}), \quad (4.2.15)$$

où la fonction  $f$  est définie par :

$$f(t, d, u) = (\phi_t(d) - \gamma u)_+. \quad (4.2.16)$$

*Démonstration.* Cf. Section 4.5.1. □

### 4.2.3 Caractéristiques des sinistres et modélisation de la provision

Notons  $T_j^i$  la v.a. représentant la date de fin du prêt de l'emprunteur  $j$  qui a souscrit son prêt le mois  $i$ . Nous déduisons des notations introduites dans la Section 4.2.1 que :

$$t^*i \leq T_j^i \leq t^*i + D_j. \quad (4.2.17)$$

#### Remarque 4.2.4.

1. En pratique aucun sinistre n'est déclaré à la fin d'un prêt, i.e. à la date  $T_j^i = t^*i + D_j$ . D'autre part si un sinistre était déclaré à la fin d'un prêt, le montant du sinistre serait insignifiant. Nous ne prendrons pas en compte d'éventuels sinistres déclarés à la date  $T_j^i = t^*i + D_j$ . Cf. Section 5.3.9 du Chapitre 5 pour une justification de cette hypothèse. De ce fait il est important de distinguer deux situations différentes dans l'inégalité (4.2.17) :
  - $T_j^i = t^*i + D_j$ , si le prêt de l'emprunteur  $j$  est arrivé à maturité sans connaître de défaut de paiement,
  - $t^*i \leq T_j^i < t^*i + D_j$ , si l'emprunteur a connu un défaut avant la maturité.
2. Il est clair que la relation (4.2.17) implique que lorsqu'un défaut a lieu, il existe une dépendance entre la date de défaut et la durée initiale du prêt.

Le fait qu'un prêt peut avoir ou ne pas avoir de défaut conduit à une partition du domaine de variation des couples (Date de fin du prêt, Durée initiale du prêt).

**Définition 4.2.5.** Nous noterons  $A_{t_0-}^i$  (resp.  $A_{t_0+}^i$ ) le domaine de variation des couples (Date de fin du prêt, Durée initiale du prêt) concernant les prêts souscrits le mois  $i$  et ayant connu un défaut dans l'intervalle de temps  $[t^*i, t_0]$  (resp.  $]t_0, t^*i + h[$ ) :

$$A_{t_0-}^i := \{(d, t) \in (\mathbb{R}_+)^2 : t^*i \leq t \leq t_0; t < t^*i + d; d \leq h\}, \quad (4.2.18)$$

$$A_{t_0+}^i := \{(d, t) \in (\mathbb{R}_+)^2 : t_0 < t < t^*i + d; d \leq h\}, \quad (4.2.19)$$

où  $h$  est introduit dans l'hypothèse  $(H_2)$ . Voici une illustration :

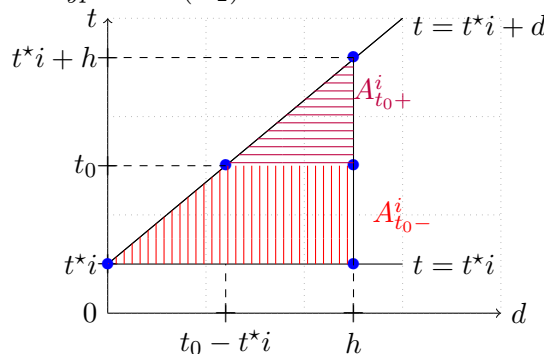


FIGURE 4.2 – Ensembles  $A_{t_0-}^i$  et  $A_{t_0+}^i$ .

Notez qu'à la date  $t_0$  les prêts tels que  $(d, t) \in A_{t_0+}^i$  sont ceux dont la durée est connue mais dont la date de maturité est inconnue tandis que les prêts tels que  $(d, t) \in A_{t_0-}^i$  ont été observés et serviront à notre étalonnage statistique.

Nous nous intéressons à présent à l'ensemble des prêts souscrits le mois  $i$ . Nous faisons l'hypothèse suivante :

( $H_6$ ) : L'ensemble des points  $(T_j^i, D_j)_{j \in B}$  est un processus ponctuel de Poisson (noté PPP dans la suite) d'intensité  $\Lambda_i$ .

Cf. [19] et [21] pour plus de détails sur ce type de processus et Cf. la Section 5.3.6 pour une justification de cette hypothèse.

En supposant que la politique de souscription des prêts ne varie pas d'un mois à l'autre, un sinistre s'explique uniquement par un changement dans la situation personnelle d'un emprunteur, le mois de souscription ne jouant aucun rôle dans la déclaration d'un sinistre. Ainsi nous supposons que :

( $H_7$ ) : Les PPP  $\left( (T_j^i, D_j)_{j \in B} \right)_{i \geq 0}$  sont indépendants. (4.2.20)

et que :

( $H_8$ ) :  $\left\{ \begin{array}{l} (T_j^i, D_j)_{j \in B}, (M_j)_{j \in B} \text{ et } (R_t)_{t \geq 0} \text{ sont indépendants,} \\ \text{Les v.a. } (M_j)_{j \in B} \text{ sont i.i.d.} \end{array} \right.$  (4.2.21)

(4.2.22)

Cf. les Sections 5.3.7 et 5.3.8 pour une justification de ces hypothèses.

Nous allons maintenant modéliser la provision sur l'intervalle de temps  $]t_0, t^*i + h[$  de manière stochastique comme une somme de charges de sinistres individuels.

**Définition 4.2.6.** *Provision de sinistres sur l'intervalle de temps  $]t_0, t^*i + h[$ .*

1. *La charge des sinistres pour tous les prêts souscrits le mois  $i$  avant  $t_0$  (i.e.  $t^*i \leq t_0$ ) et déclarés en sinistre dans l'intervalle de temps  $]t_0, t^*i + h[$  est appelée "provision du mois  $i$ ", notée  $P_{t_0}^i$  ('P' pour "provision") et définie par :*

$$P_{t_0}^i := \sum_{j \geq 1} \mathbb{1}_{\{(T_j^i, D_j) \in A_{t_0+}^i\}} M_j f(T_j^i - t^*i, D_j, R_{T_j^i - t^*i}), \quad (4.2.23)$$

où l'ensemble  $A_{t_0+}^i$  est défini par (4.2.19),  $M_j f(T_j^i - t^*i, D_j, R_{T_j^i - t^*i})$  est défini par (4.2.15) et représente la charge du sinistre au temps  $T_j^i$  concernant l'emprunteur  $j$  ayant souscrit un prêt le mois  $i$ .

2. *La charge des sinistres pour tous les prêts souscrits jusqu'à la date  $t_0$  incluse et déclarés en sinistre dans l'intervalle  $]t_0, t_0 + h[$  est appelée "provision totale", notée  $P_{t_0}$  et définie par :*

$$P_{t_0} := \sum_{i=0}^{\lfloor \frac{t_0}{t^*} \rfloor + 1} P_{t_0}^i. \quad (4.2.24)$$

**Remarque 4.2.7.**

1. L'assureur impose de déterminer la provision totale comme une somme de provisions mensuelles.
2. Comme nous prenons en compte la date des sinistres dans la modélisation de leur montant, nous ne prendrons pas en compte l'actualisation de ces montants dans la provision. Cette prise en compte pourrait être réalisée par l'ajout d'un facteur exponentiel dans la Définition (4.2.23) ; les résultats que nous donnerons dans la section suivante seraient conservés.
3. Il est possible d'interpréter la provision via les options d'échange : Cf. Section 5.2 pour plus de détails. Cette interprétation nous permettra d'évoquer la problématique d'investissement des primes perçues par la compagnie d'assurances.

**4.2.4 Liens avec les modèles existants**

Dans le modèle de Cramer - Lundberg (Cf. [52] pour plus de détails) la provision sur le laps de temps  $]t_0, t_1]$ ,  $0 \leq t_0 < t_1 < +\infty$ , est modélisée par un processus de Poisson composé :

$$\sum_{j \geq 1} Y_j \mathbb{1}_{\{t_0 < T_j \leq t_1\}},$$

où  $T_j$  est la date de déclaration du  $j$ -ème sinistre et  $(Y_j)_{j \geq 1}$  est une suite de v.a. i.i.d. représentant la charge des sinistres. Les v.a.  $Y_j$  sont supposées indépendantes de  $(T_j)_{j \geq 1}$ . Par conséquent le modèle de Cramer - Lundberg ne peut prendre en compte les dépendances entre la date de déclaration et les montants de sinistres, entre les différents montants de sinistres, ni entre le nombre de sinistres et les montants de sinistres.

Dans le modèle de Mack (Cf. [35] et Section 2),  $C_{i,k+1}$  est le montant de sinistres, pour les prêts souscrits le mois  $i$ , indemnisés entre les instants  $i$  et  $i + k + 1$ . Par conséquent, d'après la Définition 4.2.6 cette quantité peut être exprimée par :

$$C_{i,k+1} = C_{i,k} + \sum_{j \geq 1} \mathbb{1}_{\{(T_j^i, D_j) \in A_{(i+k+1)+}^i \setminus A_{(i+k)+}^i\}} M_j f(T_j^i - t^*i, D_j, R_{T_j^i - t^*i}). \quad (4.2.25)$$

Notons maintenant  $\mathcal{F}_{i,k}$  "l'information disponible" à la date  $i + k$ , i.e. :

$$\mathcal{F}_{i,k} := \sigma \left[ (T_j^i, T_j^i \leq it^* + k)_{j \in B}; (R_t)_{t \leq k}; (D_j)_{j \in B: I_j = t^*i}; (M_j)_{j \in B: I_j = t^*i} \right]. \quad (4.2.26)$$

Supposons que (4.2.9) soit vérifiée, puisque  $(T_j^i, D_j)_{j \in B}$  est un PPP et que conditionnellement à  $\mathcal{F}_{i,k}$  :

$$(R_{s+it^*+k}; 0 \leq s \leq 1) \stackrel{(d)}{=} (R_{it^*+k} R'_s; 0 \leq s \leq 1), \quad (4.2.27)$$

où  $R'$  est une copie indépendante de  $R$ , il vient que :

$$\mathbb{E} \left[ C_{i,k+1} | \mathcal{F}_{i,k} \right] = C_{i,k} + \beta_{i,k}(R_k), \quad (4.2.28)$$



où :

$$\beta_{i,k}(x) := \mathbb{E} \left[ \sum_{j \geq 1} \mathbb{1}_{\{(T_j^i, D_j) \in A_{(i+k+1)+}^i \setminus A_{(i+k)+}^i\}} M_j f(T_j^i - t^*i, D_j, x R_{T_j^i - t^*i}) \right]. \quad (4.2.29)$$

Notre modélisation est donc complètement différente de celle proposée par Mack. En effet dans le modèle de Mack l'hypothèse principale est :

$$\mathbb{E} [C_{i,k+1} | \mathcal{F}_{i,k}] = f_k C_{i,k}, \quad (4.2.30)$$

où  $f_k$  est un coefficient déterministe ne dépendant pas de la génération  $i$ . Notre modèle, d'un point de vue collectif, est un modèle additif où le coefficient  $\beta_{i,k}$  (Cf. (4.2.29)) prend en compte à la fois le mois de souscription  $i$  et la date de développement  $k$ .

### 4.3 Distribution de la provision

Nous commencerons par calculer l'espérance et la variance de la provision puis nous nous intéresserons à la fonction de répartition de cette variable.

#### 4.3.1 Espérance et variance de la provision

**Notation 4.3.1.** *Il est commode de noter dans la suite :*

$$F_1(t, x, \mu, \sigma) := x \Phi \left[ \frac{\ln(x) - \mu t}{\sigma \sqrt{t}} \right] - e^{(\mu + \frac{\sigma^2}{2})t} \Phi \left[ \frac{\ln(x) - (\mu + \sigma^2)t}{\sigma \sqrt{t}} \right], \quad (4.3.1)$$

où  $x \in \mathbb{R}_+$ ,  $\Phi$  est la fonction de répartition d'une v.a. gaussienne centrée et réduite et  $\mu, \sigma$  sont les paramètres du mouvement Brownien géométrique défini par (4.2.9).

**Théorème 4.3.2.** *Plaçons nous sous les hypothèses  $(H_1)$  à  $(H_8)$  et supposons que  $(R_t)_{t \geq 0}$  est le mouvement Brownien géométrique de paramètres  $\mu$  et  $\sigma$  défini par (4.2.9). Il vient alors que :*

$$\mathbb{E} [P_{t_0}^i] = \gamma \mathbb{E} [M_1] \int_{A_{t_0+}^i} F_1 \left[ t - t^*i, \frac{\phi_{t-t^*i}(d)}{\gamma}, \mu, \sigma \right] d\Lambda_i(t, d), \quad (4.3.2)$$

$$\mathbb{E} [P_{t_0}] = \gamma \mathbb{E} [M_1] \sum_{i=0}^{\lfloor \frac{t_0}{i^*} \rfloor + 1} \int_{A_{t_0+}^i} F_1 \left[ t - t^*i, \frac{\phi_{t-t^*i}(d)}{\gamma}, \mu, \sigma \right] d\Lambda_i(t, d), \quad (4.3.3)$$

où  $F_1$  et  $\phi_t(d)$  sont les fonctions définies par (4.3.1) et (4.2.13) et  $\mathbb{E} [M_1]$  est l'espérance du montant des prêts.

La preuve est donnée dans la Section 4.5.2. Les formules (4.3.2) et (4.3.3) ne sont pas uniquement théoriques. Quand la mesure  $\Lambda_i$  est la mesure de Lebesgue et quand les paramètres prennent des valeurs fixes, il est possible de calculer une valeur numérique de la valeur moyenne de la provision, voir la Section 5.4.2 du Chapitre 5.

Passons maintenant au calcul de la variance de la provision.

**Notation 4.3.3.** Nous noterons dans la suite :

$$\tilde{d}(x, y, z, t) := \frac{\ln(x) - yt}{z\sqrt{t}}, \quad (4.3.4)$$

$$F_2(t, d) := 2\frac{\phi_t(d)}{\gamma}F_1\left[t, \frac{\phi_t(d)}{\gamma}, \mu, \sigma\right] - F_1\left[t, \left(\frac{\phi_t(d)}{\gamma}\right)^2, 2\mu, 2\sigma\right], \quad (4.3.5)$$

$$F_3(t, d, t', d', x) := \frac{e^{\mu t + \sigma\sqrt{t}x - \frac{x^2}{2}}}{\sqrt{2\pi}} \left( \frac{\phi_t(d)}{\gamma} - e^{\mu t + \sigma\sqrt{t}x} \right) F_1\left[t' - t, \frac{\phi_{t'}(d')}{\gamma e^{\mu t + \sigma\sqrt{t}x}}, \mu, \sigma\right], \quad (4.3.6)$$

où rappelons que les fonctions  $\phi_t(d)$  et  $F_1$  sont respectivement définies par (4.2.13) et (4.3.1).

**Théorème 4.3.4.** Plaçons nous sous les hypothèses  $(H_1)$  à  $(H_8)$  et supposons que  $(R_t)_{t \geq 0}$  est le mouvement Brownien géométrique de paramètres  $\mu$  et  $\sigma$  défini par (4.2.9). La variance de la provision du mois  $i$  vaut :

$$\begin{aligned} \mathbb{V}\left[P_{t_0}^i\right] &= \gamma^2 \mathbb{E}\left[(M_1)^2\right] \int_{A_{t_0+}^i} F_2(t - t^*i, d) d\Lambda_i(t, d) + \gamma^2 \left[\mathbb{E}(M_1)\right]^2 \\ &\times \left[ \int_{(A_{t_0+}^i)^2} \left\{ \int_{-\infty}^{\tilde{d}\left(\frac{\phi_{t-t^*i}(d)}{\gamma}, \mu, \sigma, t-t^*i\right)} F_3\left[t - t^*i, d, t' - t^*i, d', x\right] dx \right\} \right. \\ &\left. d\Lambda_i(t, d) d\Lambda_i(t', d') - \left\{ \int_{A_{t_0+}^i} F_1\left[t - t^*i, \frac{\phi_{t-t^*i}(d)}{\gamma}, \mu, \sigma\right] d\Lambda_i(t, d) \right\}^2 \right], \quad (4.3.7) \end{aligned}$$

où  $\mathbb{E}\left[(M_1)^2\right]$  est le moment d'ordre 2 des montants de prêts.

La preuve de ce résultat est donnée dans la Section 4.5.3. La variance de la provision totale peut également être calculée (Cf. Section 4.5.6). Comme pour l'espérance, la formule (4.3.7) n'est pas uniquement théorique. Quand la mesure  $\Lambda_i$  est la mesure de Lebesgue et quand les paramètres sont fixés, il est possible de calculer une valeur numérique de la variance de la provision.

La formule (4.3.7) n'est pas triviale. Ceci est dû au fait que les montants de sinistres ne sont pas indépendants du fait de la présence du processus  $(R_t)_{t \geq 0}$ . Il faut ainsi tenir compte de la corrélation entre deux montants de sinistres.

**Remarque 4.3.5.** Il est possible d'obtenir une formule plus générale pour la variance sans supposer que  $(R_t)_{t \geq 0}$  est le mouvement Brownien géométrique défini par (4.2.9). En effet nous avons :

$$\begin{aligned} \mathbb{V}\left[P_{t_0}^i\right] &= \mathbb{V}\left[N_i(A_{t_0+}^i)\right] \mathbb{E}\left[\left(M_1 f(T_1^i - t^*i, D_1, R_{T_1^i - t^*i})\right)^2\right] \\ &+ \left[\mathbb{E}\left[N_i(A_{t_0+}^i)\right]\right]^2 \left\{ C^{i,i} - \left[\mathbb{E}\left[M_1 f(T_1^i - t^*i, D_1, R_{T_1^i - t^*i})\right]\right]^2 \right\}, \quad (4.3.8) \end{aligned}$$

où le couple  $(T_1^i, D_1)$  désigne un point générique du PPP  $(T_j^i, D_j)_{j \in B}$  et :

$$N_i(A) := \sum_{j \geq 1} \mathbb{1}_{\{(T_j^i, D_j) \in A\}}, \quad (4.3.9)$$

$$C^{i, i'} := \left[ \mathbb{E}[M_1] \right]^2 \mathbb{E} \left[ g \left( T_1^i - t^* i, D_1, T_2^{i'} - t^* i', D_2 \right) \right], \quad (4.3.10)$$

$$g(t, d, t', d') := \mathbb{E} \left[ \left( \phi_t(d) - \gamma R_t \right)_+ \left( \phi_{t'}(d') - \gamma R_{t'} \right)_+ \right]. \quad (4.3.11)$$

où le couple  $(T_2^i, D_2)$  représente un autre point générique du PPP  $(T_j^i, D_j)_{j \in B}$  et indépendant de  $(T_1^i, D_1)$ . Le résultat (4.3.8) (démontré dans la Section 4.5.3) nous permet d'identifier deux sources de risques. Le premier terme de cette expression représente la variabilité du nombre de sinistres autour du nombre moyen de sinistres. Le second terme de l'expression (4.3.8) est, comme dans le cas classique des processus de Poisson composés (lorsque les montants de sinistres sont indépendants), égal à la variance des montants de sinistres. Donc, le second terme représente la variabilité des montants de sinistres autour de leur moyenne. Dans le cas où  $(R_t)_{t \geq 0}$  est le mouvement Brownien géométrique défini par (4.2.9) nous pouvons calculer explicitement la fonction  $g$  et en déduire ainsi une expression de la variance. Nous pourrions remplacer le mouvement Brownien géométrique par un autre processus. Toutefois pour avoir des formules explicites il faut que les quantités  $\mathbb{E}[(\phi_t(s) - \gamma R_t)_+]$  et  $g(t, d, t', d')$  soient calculables.

### 4.3.2 Loi de la provision $P_{t_0}^i$

Il est commode d'ordonner les temps de fin de prêt de la manière suivante :

$$T_{(1)}^i < T_{(2)}^i < \dots < T_{(n)}^i < \dots \quad (4.3.12)$$

La durée du prêt associée à la date  $T_{(j)}^i$  est notée  $D_{(j)}$ . Il vient alors que :

**Théorème 4.3.6.** *Plaçons nous sous les hypothèses  $(H_1)$  à  $(H_8)$ .*

1. Rappelons que :

$$\mathcal{L} \left[ N_i(A_{t_0+}^i) \right] = \mathcal{P} \left[ \Lambda_i(A_{t_0+}^i) \right].$$

2. La loi conditionnelle de  $P_{t_0}^i$  sachant  $N_i(A_{t_0+}^i) = n$  est donnée par :

$$\mathcal{L} \left[ P_{t_0}^i | N_i(A_{t_0+}^i) = n \right] = \mathcal{L} \left[ u_n \left( (T_{(j)}^i)_{1 \leq j \leq n}, (D_{(j)})_{1 \leq j \leq n}, (M_j)_{1 \leq j \leq n}, (G_j)_{1 \leq j \leq n} \right) \right], \quad (4.3.13)$$

où :

$$\begin{aligned} & u_n \left( (t_{(j)}^i)_{1 \leq j \leq n}, (d_{(j)})_{1 \leq j \leq n}, (m_j)_{1 \leq j \leq n}, (g_j)_{1 \leq j \leq n} \right) \\ & := \sum_{j=1}^n m_j \left\{ \phi_{t_{(j)}^i - t^* i}(d_{(j)}) - \gamma \prod_{k=1}^j r_k \right\}_+, \end{aligned} \quad (4.3.14)$$

$$r_k = \exp \left[ \sigma \sqrt{t_{(k)}^i - t_{(k-1)}^i} g_k + \mu(t_{(k)}^i - t_{(k-1)}^i) \right], \text{ avec la convention : } t_{(0)}^i = 0, \quad (4.3.15)$$

$(G_k)_{k \geq 1}$  est une suite de gaussienne i.i.d. réduite, centrée et indépendante de  $\left( (T_{(j)}^i, D_{(j)})_{1 \leq j \leq n}, (M_j)_{1 \leq j \leq n} \right)$ .

Cf. la Section 4.5.4 pour la preuve de ce résultat.

A l'évidence la loi de la provision n'est pas explicite. Toutefois elle s'exprime comme une fonction compliquée mais dépendante de v.a. "faciles" à simuler. Le Théorème 4.3.6 nous permet de réaliser des simulations de Monte Carlo. Pour cela, il nous faut connaître la loi des montants  $M_j$ . Nous disposons, pour les prêts souscrits le mois  $i = 0$ , de  $b^i$  observations sur le montant des prêts. L'ordre de grandeur de  $b^i$  étant de 300, il est raisonnable de remplacer la f.d.r. de la loi de  $M_1$  par sa f.d.r. empirique. Notons  $(m_j)_{1 \leq j \leq b^i}$  les observations des montants de prêts et soit  $F_M$  la fonction de répartition empirique de cet ensemble, i.e. :

$$F_M(m) := \frac{1}{b^i} \sum_{j=1}^{b^i} \mathbb{1}_{\{m_j \leq m\}}.$$

Nous simulerons alors les montants de prêts par inversion de la fonction de répartition empirique  $F_M$ , c'est à dire :

- Simuler une variable de loi uniforme sur  $[0, 1]$  et nous noterons  $\zeta$  la réalisation,
- Calculer ensuite :  $m_{\lfloor b^i \zeta \rfloor + 1}$ .

Ceci nous permet d'obtenir l'algorithme de simulation suivant (nous supposons pour ce faire que  $\Lambda_i$  est connu) :

Cet algorithme nous permettra, dès lors que l'intensité du PPP est spécifiée et que les paramètres prennent des valeurs fixes, de simuler la provision et d'en déduire divers indicateurs de risques.

## 4.4 Estimations des paramètres

Le modèle stochastique présenté précédemment a été créé afin de pouvoir calculer explicitement la provision avec les données fournies par CAMCA Assurance. L'analyse théorique présentée dans la Section 4.3 (Cf. Théorèmes 4.3.2, 4.3.4 et 4.3.6) permet de calculer numériquement la moyenne, la variance et les quantiles de la provision dès que :

1. La mesure  $\Lambda_i$  est spécifiée,
2. Les paramètres du modèle sont estimés.

Le nombre important d'emprunteurs pour lesquels nous connaissons le montant, la durée du prêt et éventuellement la date du défaut nous permettra d'estimer les paramètres du modèle (Cf. Section 4.4.2). Commençons tout d'abord par spécifier  $\Lambda_i$ .

---

**Algorithme 4.1** Simulation d'un échantillon de taille  $l$  de la variable  $P_{t_0}^i$ .

---

- 1: **for**  $k = 1$  to  $l$  **do**
- 2: Simuler une v.a. de loi  $\mathcal{P}\left(\Lambda_i(A_{t_0+}^i)\right)$ . Nous noterons  $n$  sa réalisation.
- 3: Simuler  $n$  v.a. i.i.d. de loi uniforme sur  $A_{t_0+}^i$ . Nous noterons  $(t_j, d_j)_{1 \leq j \leq n}$  les réalisations. Ordonner les temps  $(t_j)_{1 \leq j \leq n} : t_{(1)} \leq \dots \leq t_{(n)}$ . La durée associée au temps  $t_{(j)}$  est notée  $d_{(j)}$ .
- 4: Les variables  $(M_j)_{1 \leq j \leq n}$  sont simulées par inversion de la f.d.r. empirique. Nous noterons  $(\tilde{m}_j)_{1 \leq j \leq n}$  les réalisations.
- 5: Simuler  $n$  v.a. i.i.d. de loi normale centrée et réduite. Nous noterons  $(g_j)_{1 \leq j \leq n}$  les réalisations.
- 6: Calculer :

$$p_k^i = \sum_{j=1}^n \tilde{m}_j \left\{ \phi_{t_{(j)} - t^*i}(d_{(j)}) - \gamma \prod_{k=1}^j r_k \right\}_+, \quad (4.3.16)$$

où  $\phi$  et  $r_k$  sont respectivement définis par (4.2.13) et (4.3.15).

- 7: **end for**
- 

#### 4.4.1 Choix de l'intensité $\Lambda_i$

Rappelons que la partition,  $A_{t_0-}^i$  et  $A_{t_0+}^i$  (Cf. (4.2.18) et (4.2.19)), du domaine de variation des couples (Date de fin du prêt, Durée du prêt), concerne les prêts ayant connus un défaut.

**Définition 4.4.1.** Soit  $C_{t_0-}^i$  (resp.  $C_{t_0+}^i$ ) le "domaine de variation" des couples (Date de fin du prêt, Durée initiale du prêt) concernant les prêts souscrits le mois  $i$  et ayant atteint leur maturité dans l'intervalle de temps  $[t^*i, t_0]$  (resp.  $]t_0, t^*i + h]$ ) :

$$C_{t_0-}^i := \left\{ (d, t) \in (\mathbb{R}_+)^2 : t = t^*i + d; t \leq t_0; d \leq h \right\}, \quad (4.4.1)$$

$$C_{t_0+}^i := \left\{ (d, t) \in (\mathbb{R}_+)^2 : t = t^*i + d; t > t_0; d \leq h \right\}. \quad (4.4.2)$$

Voici une illustration graphique de ces ensembles :

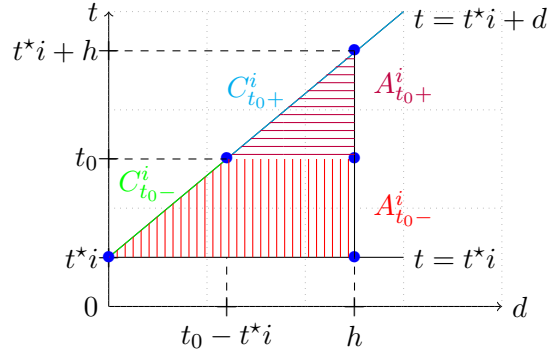
Passons maintenant au choix de l'intensité  $\Lambda_i$ . Nous supposons que :

$$\Lambda_i := \lambda_1 \text{Leb}_{|C_{t_0-}^i \cup C_{t_0+}^i} + \lambda_2 \text{Leb}_{|A_{t_0-}^i \cup A_{t_0+}^i}, \quad (4.4.3)$$

où :

- $\lambda_1$  et  $\lambda_2$  sont deux paramètres réels et strictement positifs,
- $\text{Leb}_{|C_{t_0-}^i \cup C_{t_0+}^i}$  (resp.  $\text{Leb}_{|A_{t_0-}^i \cup A_{t_0+}^i}$ ) est la restriction de la mesure de Lebesgue à l'ensemble  $C_{t_0-}^i \cup C_{t_0+}^i$  (resp.  $A_{t_0-}^i \cup A_{t_0+}^i$ ).

Cf. Section 5.3.6 pour une justification de cette hypothèse.


 FIGURE 4.3 – Ensembles  $C_{t_0-}^i$  et  $C_{t_0+}^i$ .

#### 4.4.2 Estimation des paramètres

##### Liste des paramètres :

Commençons par lister les paramètres de notre modèle :

1. Le montant moyen des prêts :  $\mathbb{E}(M_1)$ ,
2. Le taux mensuel des prêts :  $\tau$  défini par (4.2.4),
3. Le coefficient  $\gamma$  représentant l'apport personnel et les coûts liés à la revente, Cf. la formule (4.2.8),
4. Le coefficient de pénalité,  $p$ , concernant les intérêts de retard, Cf. (4.5.7),
5. Les coefficients  $\mu$  et  $\sigma$  du mouvement Brownien géométrique définis par (4.2.9),
6. Les paramètres  $\lambda_1$  et  $\lambda_2$  liés à l'intensité  $\Lambda_i$  du PPP, Cf. (4.4.3),
7. Le paramètre  $h$  représentant la durée maximale des prêts, Cf. l'expression (4.2.1).

##### Données disponibles :

A la date  $t_0$  nous disposons de trois types de données : les données communes à chaque emprunteur, les données liées à chaque sinistre et les données exogènes à la compagnie d'assurances. Les deux premiers types de données sont privés car fournis par la compagnie d'assurances tandis que le dernier est public. Pour chaque emprunteur  $j$  ayant souscrit un prêt le mois  $i$  tel que  $it^* \leq t_0$ , nous disposons de la durée  $d_j$ , du montant  $m_j$  et de la date de fin du prêt  $t_j^i$ . Nous disposons également :

- Du nombre  $b^i$  de prêts souscrits le mois  $i$ . Le nombre total de prêts souscrits jusqu'à la date  $t_0$  sera noté  $b$  et défini par :

$$b := \sum_{i=0}^{i_0} b^i, \quad (4.4.4)$$

où :

$$i_0 := \lfloor \frac{t_0}{t^*} \rfloor \in \mathbb{N}^*. \quad (4.4.5)$$

- Du nombre  $n^i$  de sinistres déclarés avant la date  $t_0$  pour les prêts souscrits le mois  $i$ ,
- Du nombre  $v^i$  de prêts arrivés à maturité à la date  $t_0$  sans connaitre de défaut.

Ainsi le nombre de “prêts en vie” à la date  $t_0$ , c’est à dire ni sinistrés ni échus, est donné par  $b^i - n^i - v^i$ .

Les données externes à la compagnie d’assurances proviennent de l’INSEE (Institut National de la Statistique et des Etudes Economiques) et du courtier en assurance “Empruntis”. L’INSEE fournit l’indice des prix des logements anciens et neufs en France métropolitaine : cet indice est noté  $u_t$  à la date  $t$ . Le courtier “Empruntis” publie chaque année des statistiques sur les emprunts immobiliers, en particulier le taux d’apport personnel moyen (Cf. coefficient  $\gamma$  défini par (4.2.12)) et le taux actuariel moyen.

Cf. Section 5.4.1 pour une présentation de quelques statistiques descriptives sur les données considérées.

#### Estimation des paramètres :

Le montant emprunté moyen, i.e.  $\mathbb{E}(M_1)$ , sera estimé par la moyenne empirique. Le taux exponentiel sera estimé à l’aide du taux actuariel proposé par Empruntis (taux actuariel estimé à 5%). Le coefficient  $\gamma$  (Cf. Section 5.3.5) sera estimé à l’aide des données d’Empruntis (coefficient  $a$ , Cf. Section 5.3.5, représentant l’apport personnel est estimé à 30%) et à l’aide d’avis d’experts (coefficient  $c$ , Cf. (5.3.11), représentant les coûts liés à la vente est estimé à 10%). Le coefficient  $p$  lié aux intérêts de retard sera déterminé par avis d’expert à 2 (Cf. expression (4.5.7)). Enfin les coefficients  $\mu$  et  $\sigma$  liés au mouvement Brownien géométrique seront estimés via les données de l’INSEE (ou par avis d’expert dans le cas de scénarios de stress).

Passons maintenant à l’estimation des paramètres  $\lambda_1$ ,  $\lambda_2$  et de la durée maximale  $h$ . Notons  $\hat{\lambda}_1$ ,  $\hat{\lambda}_2$  et  $\hat{h}$  les estimateurs obtenus par maximum de vraisemblance.

**Théorème 4.4.2.** *Plaçons nous sous les hypothèses  $(H_1)$  à  $(H_8)$ . Alors  $x = \hat{\lambda}_1$ ,  $y = \hat{\lambda}_2$  et  $z = \hat{h}$  sont solutions du système non linéaire suivant :*

$$\sum_{i=0}^{i_0} \left[ \frac{v^i}{x} + \frac{(b^i - n^i - v^i)\sqrt{2}(z - t_0 + t^*i)}{x\sqrt{2}(z - t_0 + t^*i) + \frac{y(z-t_0+t^*i)^2}{2}} \right] = \sqrt{2}(i_0 + 1)z, \quad (4.4.6)$$

$$(S) : \sum_{i=0}^{i_0} \left[ \frac{n^i}{y} + \frac{(b^i - n^i - v^i)(z - t_0 + t^*i)^2}{2x\sqrt{2}(z - t_0 + t^*i) + y(z - t_0 + t^*i)^2} \right] = \frac{(i_0 + 1)z^2}{2}, \quad (4.4.7)$$

$$\begin{aligned} & \sum_{i=0}^{i_0} \left[ \frac{n^i(t_0 - t^*i)}{(t_0 - t^*i)\left(z - \frac{1}{2}(t_0 - t^*i)\right)} + \frac{(b^i - n^i - v^i)\left(\sqrt{2}x + y(z - t_0 + t^*i)\right)}{x\sqrt{2}(z - t_0 + t^*i) + \frac{y(z-t_0+t^*i)^2}{2}} \right] \\ & = (x\sqrt{2} + yz)(i_0 + 1), \end{aligned} \quad (4.4.8)$$

La preuve de ce résultat est donnée dans la Section 4.5.5. Notons que ce système peut être résolu numériquement avec le logiciel  $\mathbb{R}$  et en particulier avec la fonction “nleqslv” du package portant le même nom (Cf. Section 5.4.5 pour plus de détails sur cette fonction).

Synthétisons l’estimation de l’ensemble des paramètres :

| Paramètre                 | Notation                                    | Estimation   | Source                     |
|---------------------------|---|--|----------------------------|
| $\mathbb{E}(M_1)$         | $\bar{m}$                                   | $\frac{1}{b} \sum_{j=1}^b m_j$                             | Assureur                   |
| $\tau$                    | $\hat{\tau}$                                | $\ln\left(1 + \frac{5}{100} \frac{1}{12}\right)$           | Empruntis                  |
| $\gamma$                  | $\hat{\gamma}$                              | 1, 17  | Avis d’expert et Empruntis |
| $p$                       | $\hat{p}$                                   | 2  | Avis d’expert              |
| $\mu$                     | $\hat{\mu}$                                 | $\frac{\sum_{i=1}^{i_0} x_i}{i_0 t^*}$                     | INSEE                      |
| $\sigma$                  | $\hat{\sigma}$                              | $\frac{\sum_{i=1}^{i_0} (x_i - \hat{\mu} t^*)^2}{i_0 t^*}$ | INSEE                      |
| $\lambda_1, \lambda_2, h$ | $\hat{\lambda}_1, \hat{\lambda}_2, \hat{h}$ | Maximum de vraisemblance                                   | Assureur                   |

FIGURE 4.4 – Synthèse des estimations.

où :

$$x_t = \ln(u_t) - \ln(u_{t-1}). \quad (4.4.9)$$

Nous verrons dans la Section 5.4 du Chapitre 5 comment utiliser ces estimations pour obtenir une valeur numérique de la moyenne, de la variance et des quantiles de la provision.

Passons maintenant aux preuves des résultats obtenus dans les Sections 4.2 et 4.3.

## 4.5 Preuves

### 4.5.1 Preuve de la Proposition 4.2.3

Considérons un emprunteur  $j$  ayant souscrit un prêt à la date  $I_j$  d’un montant  $M_j$  pour une durée  $D_j$ . Soit  $\tilde{\tau}$  le rendement mensuel constant. Par conséquent :

$$(1 + \tilde{\tau})^{t^*} = e^{\tilde{\tau} t^*}. \quad (4.5.1)$$

Ainsi d’après [20] pages 172 à 178 il vient que :



1. Le capital remboursé par l'emprunteur  $j$  à la date  $t$  vaut :

$$C_t^j = \begin{cases} (e^{\tau t^*} - 1)M_j \left( \frac{e^{\tau(t-I_j-t^*)}}{e^{\tau D_j} - 1} \right), & \text{si } I_j < t \leq I_j + D_j, \\ 0, & \text{si } t = I_j. \end{cases} \quad (4.5.2)$$

2. L'intérêt remboursé par l'emprunteur  $j$  à la date  $t$  vaut :

$$U_t^j = \begin{cases} (e^{\tau t^*} - 1)M_j \left( \frac{e^{\tau D_j} - e^{\tau(t-I_j-t^*)}}{e^{\tau D_j} - 1} \right), & \text{si } I_j < t \leq I_j + D_j, \\ 0, & \text{si } t = I_j. \end{cases} \quad (4.5.3)$$

3. La mensualité versée par l'emprunteur  $j$  à la date  $t$  vaut :

$$ME_t^j = \begin{cases} (e^{\tau t^*} - 1)M_j \left( \frac{e^{\tau D_j}}{e^{\tau D_j} - 1} \right), & \text{si } I_j < t \leq I_j + D_j, \\ 0, & \text{si } t = I_j. \end{cases} \quad (4.5.4)$$

Supposons maintenant que cet emprunteur présente un défaut de paiement à une date  $T$  fixée et déterministe. L'hypothèse  $(H_4)$  implique que la banque le déclare immédiatement en sinistre. La Définition 4.2.2 et les formules (4.5.1), (4.5.2), (4.5.3) et (4.5.4) nous permettent de déterminer les différents flux intervenant dans la charge d'un sinistre.

1. Le capital impayé de la date  $T$  jusqu'à la date d'échéance du prêt,  $I_j + D_j$ , vaut :

$$\sum_{t=T}^{I_j+D_j} C_t^j = \begin{cases} M_j \left( \frac{e^{\tau D_j} - e^{\tau(T-I_j-t^*)}}{e^{\tau D_j} - 1} \right), & \text{si } I_j < T \leq I_j + D_j, \\ M_j, & \text{si } T = I_j. \end{cases} \quad (4.5.5)$$

2. Les intérêts pour l'unique mensualité impayée de la date  $T$  valent :

$$U_T^j = \begin{cases} (e^{\tau t^*} - 1)M_j \left( \frac{e^{\tau D_j} - e^{\tau(T-I_j-t^*)}}{e^{\tau D_j} - 1} \right), & \text{si } I_j < T \leq I_j + D_j, \\ 0, & \text{si } T = I_j. \end{cases} \quad (4.5.6)$$

3. Tout retard dans le paiement d'une mensualité entraîne le paiement d'intérêts de retard proportionnels au délai de retard. Cet intérêt de retard est égal à l'intérêt classique multiplié par un coefficient. Le coefficient multiplicatif lié aux intérêts de retard est constant par rapport au temps et identique pour tous les emprunteurs. Il sera noté  $p$  (' $p'$  pour "pénalité"). Les intérêts de retard pour l'unique mensualité impayée valent :

$$t^*pU_T^j = \begin{cases} t^*p(e^{\tau t^*} - 1)M_j \left( \frac{e^{\tau D_j} - e^{\tau(T-I_j-t^*)}}{e^{\tau D_j} - 1} \right), & \text{si } I_j < T \leq I_j + D_j, \\ 0, & \text{si } T = I_j, \end{cases} \quad (4.5.7)$$

L'intérêt de retard est multiplié par le laps de temps où la mensualité est restée impayée : dans ce cas nous multiplions par  $t^*$  car la durée de l'intérêt impayé jusqu'à la date de déclaration du sinistre est de  $t^*$  (de la date  $T - t^*$  à la date  $T$ ) d'après l'hypothèse  $(H_4)$ .

4. Le montant de recours encaissé à la date  $T$  vaut :

$$\gamma M_j R_{T-I_j}, \quad (4.5.8)$$

où  $\gamma$  est défini par l'expression (5.3.12). Cf. l'hypothèse ( $H_5$ ) et la Section 5.3.5 pour plus de détails.

La formule (4.2.11) s'obtient en combinant (4.5.5), (4.5.6) et (4.5.7). La formule (4.2.12) s'obtient avec (4.5.8). Notons que nous avons posé  $\phi_t(s) = \alpha$  si  $t \leq t^*$  ou  $s \leq t^*$  pour assurer la continuité de la fonction  $\phi.(s)$ .

### 4.5.2 Preuve du Théorème 4.3.2

La preuve du Théorème 4.3.2 se base sur le lemme suivant :

**Lemme 4.5.1.** *Soient  $K \in \mathbb{R}_+$ ,  $(\alpha, \beta) \in \mathbb{R}^2$  et  $(B_t)_{t \geq 0}$  un mouvement Brownien standard. Il vient que :*

$$\mathbb{E} \left[ \left( K - \exp[\alpha B_t + \beta t] \right)_+ \right] = F_1(t, K, \beta, \alpha), \quad (4.5.9)$$

où la fonction  $F_1$  est définie par (4.3.1).

*Preuve du Lemme 4.5.1.* Un calcul similaire à celui effectué dans [31] page 303 donne (4.5.9). □

Revenons à la preuve du Théorème 4.3.2. Plaçons nous sous les hypothèses ( $H_1$ ) à ( $H_8$ ) et supposons que  $(R_t)_{t \geq 0}$  est le mouvement Brownien géométrique défini par (4.2.9). Puisque  $(T_j^i, D_j)_{j \in B}$  est un PPP, sachant que  $N_i(A_{t_0+}^i) = k$  les couples de v.a.  $(T_j^i, D_j)_{1 \leq j \leq k}$  sont i.i.d. de loi  $\frac{1}{\Lambda_i(A_{t_0+}^i)} \Lambda_i |_{A_{t_0+}^i}$ , où  $N_i(A_{t_0+}^i)$  est défini par (4.3.9). Il vient alors :

$$\begin{aligned} \mathbb{E} \left[ P_{t_0}^i \right] &= \sum_{k \geq 0} \mathbb{P} \left[ N_i(A_{t_0+}^i) = k \right] \mathbb{E} \left[ \sum_{j=1}^k M_j f(T_j^i - t^*i, D_j, R_{T_j^i - t^*i}) \right] \\ &= \mathbb{E} \left[ N_i(A_{t_0+}^i) \right] \mathbb{E} \left[ M_1 f(T_1^i - t^*i, D_1, R_{T_1^i - t^*i}) \right] \end{aligned} \quad (4.5.10)$$

$$= \mathbb{E} \left[ N_i(A_{t_0+}^i) \right] \mathbb{E} \left[ M_1 \right] \int_{A_{t_0+}^i} \mathbb{E} \left[ \left\{ \phi_{t-t^*i}(d) - \gamma R_{t-t^*i} \right\}_+ \right] \frac{d\Lambda_i(t, d)}{\Lambda_i(A_{t_0+}^i)}, \quad (4.5.11)$$

avec  $\phi_t(d)$  définie par (4.2.13). En combinant (4.5.11) et (4.5.9) il vient alors que :

$$\mathbb{E} \left[ P_{t_0}^i \right] = \gamma \mathbb{E} \left[ M_1 \right] \int_{A_{t_0+}^i} F_1 \left( t - t^*i, \frac{\phi_{t-t^*i}(d)}{\gamma}, \mu, \sigma \right) d\Lambda_i(t, d). \quad (4.5.12)$$

Ce qui démontre la formule (4.3.2) du Théorème. La formule (4.3.3) se déduit alors facilement.

### 4.5.3 Preuve du Théorème 4.3.4

Plaçons nous sous les hypothèses  $(H_1)$  à  $(H_8)$ . Donnons pour commencer un lemme technique qui nous servira par la suite :

#### Lemme 4.5.2.

1. Pour tout  $j \neq l$  nous avons :

$$\mathbb{E}\left[M_j f(T_j^i - t^*i, D_j, R_{T_j^i - t^*i}) M_l f(T_l^i - t^*i, D_l, R_{T_l^i - t^*i})\right] = C^{i,i}, \quad (4.5.13)$$

où  $C^{i,i}$  est défini par (4.3.10). Et pour  $i \neq i'$  nous avons :

$$\mathbb{E}\left[M_j f(T_j^i - t^*i, D_j, R_{T_j^i - t^*i}) M_l f(T_l^{i'} - t^*i', D_l, R_{T_l^{i'} - t^*i'})\right] = C^{i,i'}. \quad (4.5.14)$$

2. Soit  $K \in \mathbb{R}_+$ . Considérons  $(\alpha, \beta) \in \mathbb{R}^2$  et  $(B_t)_{t \geq 0}$  un mouvement Brownien standard. Alors :

$$\mathbb{E}\left[\left\{\left(K - \exp[\alpha B_t + \beta t]\right)_+\right\}^2\right] = 2KF_1(t, K, \beta, \alpha) - F_1(t, K^2, 2\beta, 2\alpha), \quad (4.5.15)$$

où la fonction  $F_1$  est définie par (4.3.1).

3. Nous avons :

$$\begin{aligned} \mathbb{V}\left[P_{t_0}^i\right] &= \mathbb{V}\left[N_i(A_{t_0+}^i)\right] \mathbb{E}\left[\left(M_1 f(T_1^i - t^*i, D_1, R_{T_1^i - t^*i})\right)^2\right] \\ &\quad + \left[\mathbb{E}\left[N_i(A_{t_0+}^i)\right]\right]^2 \left\{C^{i,i} - \left[\mathbb{E}\left[M_1 f(T_1^i - t^*i, D_1, R_{T_1^i - t^*i})\right]\right]^2\right\}. \end{aligned} \quad (4.5.16)$$

4. Si  $(R_t)_{t \geq 0}$  est le mouvement Brownien géométrique défini par (4.2.9) alors :

$$\mathbb{E}\left[\left(M_1 f(T_1^i - t^*i, D_1, R_{T_1^i - t^*i})\right)^2\right] = \gamma^2 \mathbb{E}\left[(M_1)^2\right] \int_{A_{t_0+}^i} F_2(t - t^*i, d) \frac{d\Lambda_i(t, d)}{\Lambda_i(A_{t_0+}^i)}, \quad (4.5.17)$$

$$\begin{aligned} C^{i,i'} &= \gamma^2 \left[\mathbb{E}\left[M_1\right]\right]^2 \int_{A_{t_0+}^i \times A_{t_0+}^{i'}} \left\{ \int_{-\infty}^{\tilde{d}(\frac{\phi_{t-t^*i}(d)}{\gamma}, \mu, \sigma, t-t^*i)} F_3(t - t^*i, d, t' - t^*i', d', x) dx \right\} \\ &\quad \frac{d\Lambda_i(t, d) d\Lambda_{i'}(t', d')}{\Lambda_i(A_{t_0+}^i) \Lambda_{i'}(A_{t_0+}^{i'})}, \end{aligned} \quad (4.5.18)$$

où les fonctions  $\tilde{d}$ ,  $F_2$  et  $F_3$  sont définies par (4.3.4), (4.3.5) et (4.3.6) et  $\mathbb{E}\left[(M_1)^2\right]$  est le moment d'ordre 2 des montants de prêts.

La preuve de ce lemme sera fournie à la fin de cette section. Revenons maintenant à la preuve du Théorème 4.3.4. Nous cherchons à calculer la variance de la provision du mois  $i$ . D'après les expressions (4.5.16), (4.3.2), (4.5.17) et (4.5.18) il vient que :

$$\mathbb{V}\left[P_{t_0}^i\right] = \gamma^2 \mathbb{E}\left[(M_1)^2\right] \int_{A_{t_0+}^i} F_2(t - t^*i, d) d\Lambda_i(t, d) + \gamma^2 \left[\mathbb{E}\left[M_1\right]\right]^2$$

$$\times \left[ \int_{(A_{t_0+}^i)^2} \left\{ \int_{-\infty}^{\tilde{d}(\frac{\phi_{t-t^*i}(d)}{\gamma}, \mu, \sigma, t-t^*i)} F_3(t-t^*i, d, t'-t^*i, d', x) dx \right\} d\Lambda_i(t, d) d\Lambda_i(t', d') - \left\{ \int_{A_{t_0+}^i} F_1\left(t-t^*i, \frac{\phi_{t-t^*i}(d)}{\gamma}, \mu, \sigma\right) d\Lambda_i(t, d) \right\}^2 \right].$$

Ce qui démontre la formule (4.3.7). Revenons pour conclure à la preuve du Lemme 4.5.2.

*Preuve du Lemme 4.5.2.*

1. En combinant (4.2.15) et (4.3.11), il vient que  $\mathbb{E}\left[M_j f(T_j^i - t^*i, D_j, R_{T_j^i - t^*i}) M_l f(T_l^i - t^*i, D_l, R_{T_l^i - t^*i})\right]$  vaut :

$$\begin{aligned} & \mathbb{E}\left[M_j M_l \left\{ \phi_{T_j^i - t^*i}(D_j) - \gamma R_{T_j^i - t^*i} \right\}_+ \left\{ \phi_{T_l^i - t^*i}(D_l) - \gamma R_{T_l^i - t^*i} \right\}_+\right] \\ &= \left[\mathbb{E}[M_1]\right]^2 \mathbb{E}\left[\left\{ \phi_{T_j^i - t^*i}(D_j) - \gamma R_{T_j^i - t^*i} \right\}_+ \left\{ \phi_{T_l^i - t^*i}(D_l) - \gamma R_{T_l^i - t^*i} \right\}_+\right] \\ &= \left[\mathbb{E}[M_1]\right]^2 \mathbb{E}\left[g\left(T_j^i - t^*i, D_j, T_l^i - t^*i, D_l\right)\right] \\ &= \left[\mathbb{E}[M_1]\right]^2 \mathbb{E}\left[g\left(T_1^i - t^*i, D_1, T_2^i - t^*i, D_2\right)\right] \\ &= C^{i,i}, \end{aligned}$$

car puisque  $(T_j^i, D_j)_{j \in B}$  est un PPP alors les v.a.  $(T_j^i, D_j)_{j \geq 1}$  sont i.i.d.. Ce qui démontre la formule (4.5.13) du lemme.

De la même manière  $\mathbb{E}\left[M_j f(T_j^i - t^*i, D_j, R_{T_j^i - t^*i}) M_l f(T_l^{i'} - t^*i', D_l, R_{T_l^{i'} - t^*i'})\right]$  vaut :

$$\begin{aligned} & \mathbb{E}\left[M_j M_l \left\{ \phi_{T_j^i - t^*i}(D_j) - \gamma R_{T_j^i - t^*i} \right\}_+ \left\{ \phi_{T_l^{i'} - t^*i'}(D_l) - \gamma R_{T_l^{i'} - t^*i'} \right\}_+\right] \\ &= \left[\mathbb{E}[M_1]\right]^2 \mathbb{E}\left[\left\{ \phi_{T_j^i - t^*i}(D_j) - \gamma R_{T_j^i - t^*i} \right\}_+ \left\{ \phi_{T_l^{i'} - t^*i'}(D_l) - \gamma R_{T_l^{i'} - t^*i'} \right\}_+\right] \\ &= \left[\mathbb{E}[M_1]\right]^2 \mathbb{E}\left[g\left(T_j^i - t^*i, D_j, T_l^{i'} - t^*i', D_l\right)\right] \\ &= \left[\mathbb{E}[M_1]\right]^2 \mathbb{E}\left[g\left(T_1^i - t^*i, D_1, T_2^{i'} - t^*i', D_2\right)\right] \\ &= C^{i,i'}. \end{aligned}$$

Ce qui démontre la formule (4.5.14).

2. Si  $x \geq 0$  alors :

$$\left((K - x)_+\right)^2 = 2K(K - x)_+ - (K^2 - x^2)_+.$$

En appliquant le Lemme 4.5.1 et la dernière observation nous pouvons conclure que :

$$\begin{aligned} & \mathbb{E}\left[\left\{\left(K - \exp[\alpha B_t + \beta t]\right)_+\right\}^2\right] \\ &= 2K\mathbb{E}\left[\left(K - \exp[\alpha B_t + \beta t]\right)_+\right] - \mathbb{E}\left[\left(K^2 - \exp[2\alpha B_t + 2\beta t]\right)_+\right], \\ &= 2KF_1(t, K, \beta, \alpha) - F_1(t, K^2, 2\beta, 2\alpha). \end{aligned}$$

Ce qui démontre la formule (4.5.15).

3. Calculons tout d'abord le moment d'ordre 2 de la provision du mois  $i$ . Nous en déduirons ensuite la variance de la provision du mois  $i$ . Nous avons :

$$\begin{aligned} \mathbb{E}\left[(P_{t_0}^i)^2\right] &= \sum_{k \geq 1} \mathbb{P}\left[N_i(A_{t_0+}^i) = k\right] \mathbb{E}\left[\left\{\sum_{j=1}^k M_j f(T_j^i - t^*i, D_j, R_{T_j^i - t^*i})\right\}^2\right] \\ &= A + B, \end{aligned} \tag{4.5.19}$$

où  $N_i$  est défini par (4.3.9) et :

$$A = \sum_{k \geq 1} \mathbb{P}\left[N_i(A_{t_0+}^i) = k\right] \mathbb{E}\left[\sum_{j=1}^k (M_j)^2 \{f(T_j^i - t^*i, D_j, R_{T_j^i - t^*i})\}^2\right],$$

et

$$\begin{aligned} B &= \sum_{k \geq 1} \mathbb{P}\left[N_i(A_{t_0+}^i) = k\right] \mathbb{E}\left[2 \sum_{1 \leq j < l \leq k} M_j f(T_j^i - t^*i, D_j, R_{T_j^i - t^*i}) \right. \\ &\quad \left. \times M_l f(T_l^i - t^*i, D_l, R_{T_l^i - t^*i})\right]. \end{aligned}$$

Commençons par calculer  $A$ . Une analyse similaire à celle effectuée dans la preuve du Théorème 4.3.2, montre que :

$$\begin{aligned} A &= \sum_{k \geq 1} \mathbb{P}\left[N_i(A_{t_0+}^i) = k\right] \sum_{j=1}^k \mathbb{E}\left[\left(M_1 f(T_1^i - t^*i, D_1, R_{T_1^i - t^*i})\right)^2\right] \\ &= \sum_{k \geq 1} \mathbb{P}\left[N_i(A_{t_0+}^i) = k\right] k \mathbb{E}\left[\left(M_1 f(T_1^i - t^*i, D_1, R_{T_1^i - t^*i})\right)^2\right] \\ &= \mathbb{E}\left[N_i(A_{t_0+}^i)\right] \mathbb{E}\left[\left(M_1 f(T_1^i - t^*i, D_1, R_{T_1^i - t^*i})\right)^2\right]. \end{aligned} \tag{4.5.20}$$

Calculons maintenant  $B$ . D'après (4.5.13) il vient que :

$$\begin{aligned} B &= \sum_{k \geq 1} \mathbb{P}\left[N_i(A_{t_0+}^i) = k\right] \times 2 \sum_{1 \leq j < l \leq k} C^{i,i} \\ &= \sum_{k \geq 1} \mathbb{P}\left[N_i(A_{t_0+}^i) = k\right] k(k-1) C^{i,i} \end{aligned}$$

$$\begin{aligned}
 &= \mathbb{E} \left[ (N_i(A_{t_0+}^i))^2 \right] C^{i,i} - \mathbb{E} \left[ N_i(A_{t_0+}^i) \right] C^{i,i} \\
 &= \left[ \mathbb{E} \left[ N_i(A_{t_0+}^i) \right] \right]^2 C^{i,i},
 \end{aligned} \tag{4.5.21}$$

où  $C^{i,i}$  est défini par (4.3.10). En combinant (4.5.19), (4.5.20) et (4.5.21) il vient :

$$\mathbb{E} \left[ (P_{t_0}^i)^2 \right] = \mathbb{E} \left[ N_i(A_{t_0+}^i) \right] \mathbb{E} \left[ \left( M_1 f(T_1^i - t^*i, D_1, R_{T_1^i - t^*i}) \right)^2 \right] + \left[ \mathbb{E} \left[ N_i(A_{t_0+}^i) \right] \right]^2 C^{i,i}. \tag{4.5.22}$$

Revenons maintenant au calcul de la variance de la provision du mois  $i$ . D'après (4.5.10) et (4.5.22) il vient que :

$$\begin{aligned}
 \mathbb{V} \left[ P_{t_0}^i \right] &= \mathbb{E} \left[ (P_{t_0}^i)^2 \right] - \left[ \mathbb{E} \left[ P_{t_0}^i \right] \right]^2 \\
 &= \mathbb{V} \left[ N_i(A_{t_0+}^i) \right] \mathbb{E} \left[ \left( M_1 f(T_1^i - t^*i, D_1, R_{T_1^i - t^*i}) \right)^2 \right] \\
 &\quad + \left[ \mathbb{E} \left[ N_i(A_{t_0+}^i) \right] \right]^2 \left\{ C^{i,i} - \left[ \mathbb{E} \left[ M_1 f(T_1^i - t^*i, D_1, R_{T_1^i - t^*i}) \right] \right]^2 \right\}.
 \end{aligned}$$

Ceci démontre la formule (4.5.16).

4. D'après l'hypothèse ( $H_8$ ), (4.2.15) et (4.2.9) il vient que :

$$\begin{aligned}
 \mathbb{E} \left[ \left( M_1 f(T_1^i - t^*i, D_1, R_{T_1^i - t^*i}) \right)^2 \right] &= \mathbb{E} \left[ (M_1)^2 \right] \times \int_{A_{t_0+}^i} \mathbb{E} \left[ \left\{ (\phi_{t-t^*i}(d) - \gamma R_{t-t^*i})_+ \right\}^2 \right] \\
 &\quad \frac{d\Lambda_i(t, d)}{\Lambda_i(A_{t_0+}^i)}.
 \end{aligned}$$

En appliquant la formule (4.5.15) nous obtenons :

$$\mathbb{E} \left[ \left( M_1 f(T_1^i - t^*i, D_1, R_{T_1^i - t^*i}) \right)^2 \right] = \gamma^2 \mathbb{E} \left[ (M_1)^2 \right] \int_{A_{t_0+}^i} F_2(t - t^*i, d) \frac{d\Lambda_i(t, d)}{\Lambda_i(A_{t_0+}^i)},$$

où  $F_2$  est définie par (4.3.5). Ce qui démontre la formule (4.5.17).

Nous calculons maintenant le coefficient  $C^{i,i'}$  qui est défini par (4.3.10). Nous voulons tout d'abord évaluer  $g(t, d, t', d')$ , pour  $t \neq t'$ , où  $g$  est définie par (4.3.11). Supposons par exemple  $t' > t$ . Soient  $0 < t_1 < \dots < t_n$ . D'après la définition de  $R$ , conditionnellement à  $R_{t_1} = x_1, \dots, R_{t_{n-1}} = x_{n-1}$ ,  $R_{t_n}$  a même loi que  $x_{n-1} R_{t_n - t_{n-1}}$ , c'est à dire :

$$\mathcal{L} \left[ R_{t_n} | R_{t_1} = x_1, \dots, R_{t_{n-1}} = x_{n-1} \right] = \mathcal{L} \left[ x_{n-1} R_{t_n - t_{n-1}} \right]. \tag{4.5.23}$$

En particulier,

$$\mathcal{L} \left( R_t, R_{t'} \right) = \mathcal{L} \left( R_t, R'_{t'-t} \right), \tag{4.5.24}$$

où :

$$R'_{t'-t} := \exp [\sigma(B_{t'} - B_t) + \mu(t' - t)]. \quad (4.5.25)$$

Alors pour  $t' > t$  :

$$\begin{aligned} g(t, d, t', d') &= \gamma^2 \mathbb{E} \left[ \left\{ \frac{\phi_t(d)}{\gamma} - R_t \right\}_+ \left\{ \frac{\phi_{t'}(d')}{\gamma} - R_{t'} \right\}_+ \right] \\ &= \gamma^2 \mathbb{E} \left[ \left\{ \frac{\phi_t(d)}{\gamma} - R_t \right\}_+ \left\{ \frac{\phi_{t'}(d')}{\gamma} - R_t R'_{t'-t} \right\}_+ \right] \\ &= \gamma^2 \mathbb{E} \left[ \left\{ \frac{\phi_t(d)}{\gamma} - R_t \right\}_+ \tilde{g}(R_t) \right], \end{aligned} \quad (4.5.26)$$

où :

$$\tilde{g}(z) := z \mathbb{E} \left[ \left\{ \frac{\phi_{t'}(d')}{z\gamma} - R_{t'-t} \right\}_+ \right].$$

D'après (4.5.9) il vient que :

$$\tilde{g}(z) = z F_1 \left( t' - t, \frac{\phi_{t'}(d')}{z\gamma}, \mu, \sigma \right). \quad (4.5.27)$$

Revenons au calcul de  $g$ . En combinant les expressions (4.3.4), (4.5.26) et (4.5.27) nous obtenons :

$$\begin{aligned} g(t, d, t', d') &= \gamma^2 \mathbb{E} \left[ \left\{ \frac{\phi_t(d)}{\gamma} - R_t \right\}_+ R_t F_1 \left( t' - t, \frac{\phi_{t'}(d')}{R_t \gamma}, \mu, \sigma \right) \right] \\ &= \gamma^2 \int_{-\infty}^{\tilde{d}(\frac{\phi_t(d)}{\gamma}, \mu, \sigma, t)} \frac{e^{\mu t + \sigma \sqrt{t} x - \frac{x^2}{2}}}{\sqrt{2\pi}} \left( \frac{\phi_t(d)}{\gamma} - e^{\mu t + \sigma \sqrt{t} x} \right) \\ &\quad \times F_1 \left( t' - t, \frac{\phi_{t'}(d')}{\gamma e^{\mu t + \sigma \sqrt{t} x}}, \mu, \sigma \right) dx \\ &= \int_{-\infty}^{\tilde{d}(\frac{\phi_t(d)}{\gamma}, \mu, \sigma, t)} \gamma^2 F_3(t, d, t', d', x) dx, \end{aligned} \quad (4.5.28)$$

où la fonction  $F_3$  est définie par (4.3.6). D'après les formules (4.3.10) et (4.5.28) il vient que :

$$\begin{aligned} C^{i, i'} &= \gamma^2 \left[ \mathbb{E}[M_1] \right]^2 \int_{A_{t_0+}^i \times A_{t_0+}^{i'}} \left\{ \int_{-\infty}^{\tilde{d}(\frac{\phi_{t-t^*i}(d)}{\gamma}, \mu, \sigma, t-t^*i)} F_3(t-t^*i, d, t'-t^*i, d', x) dx \right\} \\ &\quad \frac{d\Lambda_i(t, d) d\Lambda_{i'}(t', d')}{\Lambda_i(A_{t_0+}^i) \Lambda_{i'}(A_{t_0+}^{i'})}. \end{aligned}$$

Ce qui démontre la formule (4.5.18) et achève la preuve du Lemme 4.5.2.

□

#### 4.5.4 Preuve du Théorème 4.3.6

Pour simplifier la preuve nous considérons un mois  $i$  particulier et nous noterons :

$$T_j^i = T_j.$$

Commençons par ranger par ordre croissant les temps de défaut  $(T_j)_{1 \leq j \leq n}$  :

$$T_{(1)} = T_{j_1} < \dots < T_{(n)} = T_{j_n}, \quad (4.5.29)$$

où  $(j_1, \dots, j_n) \in B^n : j_k \neq j_l \forall k \neq l \in \{1, \dots, n\}$ . Par convention nous supposons :  $T_{(0)} = 0$ .

La durée (resp. le montant) du prêt associée au sinistre ayant eu lieu à la date  $T_{(j)}$  est notée  $D_{(j)}$  (resp.  $M_{(j)}$ ). Ainsi par exemple  $D_{(1)} = D_{j_1}$ .

D'après la Définition 4.2.6 et la Proposition 4.2.3, la provision du mois  $i$  se réécrit :

$$P_{t_0}^i = \sum_{j \geq 1} \mathbb{1}_{\{(T_{(j)}, D_{(j)}) \in A_{t_0+}^i\}} M_{(j)} \left\{ \phi_{T_{(j)} - t^* i}(D_{(j)}) - \gamma R_{T_{(j)} - t^* i} \right\}_+.$$

D'après l'identité en loi (4.5.23) nous avons :

$$\begin{aligned} \mathcal{L} \left[ \left( R_{T_{(1)}}, \dots, R_{T_{(n)}} \right) \right] &= \mathcal{L} \left[ \left( \exp \left( \sigma \sqrt{T_{(1)}} G_1 + \mu T_{(1)} \right), \dots, \right. \right. \\ &\quad \left. \left. \prod_{k=1}^n \exp \left( \sigma \sqrt{T_{(k)} - T_{(k-1)}} G_k + \mu (T_{(k)} - T_{(k-1)}) \right) \right) \right], \end{aligned} \quad (4.5.30)$$

où  $(G_j)_{1 \leq j \leq n}$  est une suite de gaussienne i.i.d. centrée et réduite et indépendante de  $((T_j, D_j)_{1 \leq j \leq n}, (M_j)_{1 \leq j \leq n})$ .

Ainsi la loi de  $P_{t_0}^i$  conditionnellement à  $N_i(A_{t_0+}^i) = n$  est :

$$\mathcal{L} \left[ P_{t_0}^i | N_i(A_{t_0+}^i) = n \right] = \mathcal{L} \left[ u_n \left( (T_{(j)})_{1 \leq j \leq n}, (D_{(j)})_{1 \leq j \leq n}, (M_{(j)})_{1 \leq j \leq n}, (G_j)_{1 \leq j \leq n} \right) \right],$$

où :

$$u_n \left( (t_{(j)})_{1 \leq j \leq n}, (d_{(j)})_{1 \leq j \leq n}, (m_{(j)})_{1 \leq j \leq n}, (g_j)_{1 \leq j \leq n} \right) := \sum_{i=j}^n m_{(j)} \left\{ \phi_{t_{(j)} - t^* i}(d_{(j)}) - \gamma \prod_{k=1}^j r_k \right\}_+,$$

$r_k = \exp \left[ \sigma \sqrt{t_{(k)} - t_{(k-1)}} g_k + \mu (t_{(k)} - t_{(k-1)}) \right]$ , avec la convention :  $t_{(0)} = 0$ ,

$(G_k)_{k \geq 1}$  est une suite de gaussienne i.i.d. et indépendante de  $((T_j, D_j)_{1 \leq j \leq n}, (M_j)_{1 \leq j \leq n})$ .

D'autre part, les indices aléatoires  $j_1, \dots, j_n$  définis par (4.5.29) ne dépendent que de  $T_1, \dots, T_n$ . Or  $(M_j)_{1 \leq j \leq n}$  est indépendant de  $(T_j)_{1 \leq j \leq n}$ , ainsi :

$$\mathcal{L} \left[ (M_{j_1}, \dots, M_{j_n}) \right] = \mathcal{L} \left[ (M_1, \dots, M_n) \right]. \quad (4.5.31)$$

Donc :

$$\mathcal{L} \left[ P_{t_0}^i | N_i(A_{t_0+}^i) = n \right] = \mathcal{L} \left[ u_n \left( (T_{(j)})_{1 \leq j \leq n}, (D_{(j)})_{1 \leq j \leq n}, (M_j)_{1 \leq j \leq n}, (G_j)_{1 \leq j \leq n} \right) \right].$$



### 4.5.5 Preuve du Théorème 4.4.2

Avant de commencer la preuve du Théorème 4.4.2 nous énonçons un lemme technique.

#### Lemme 4.5.3.

1. Considérons la suite  $(Z_i, Z'_i, Z''_i)_{0 \leq i \leq i_0}$  de v.a. vérifiant les conditions suivantes :

(a) Les v.a.  $\{Z_i, Z'_i, Z''_i, 0 \leq i \leq i_0\}$  sont indépendantes,

(b) Les lois marginales sont :

$$\mathcal{L}(Z_i) = \mathcal{P}(f_i(\theta)), \quad \mathcal{L}(Z'_i) = \mathcal{P}(f'_i(\theta)), \quad \mathcal{L}(Z''_i) = \mathcal{P}(f''_i(\theta)), \quad (4.5.32)$$

où  $f_i, f'_i$  et  $f''_i$  sont de classes  $C^1(\mathbb{R}^3)$  et  $\theta = (\theta_1, \theta_2, \theta_3) \in \mathbb{R}^3$ .

Soient également  $(z_i, z'_i, z''_i)_{0 \leq i \leq i_0}$  une réalisation de la suite  $(Z_i, Z'_i, Z''_i)_{0 \leq i \leq i_0}$ . Les trois équations de vraisemblance sont les suivantes :

$$\sum_{i=0}^{i_0} \left[ z_i \frac{\partial f_i(\theta)}{\partial \theta_k} + z'_i \frac{\partial f'_i(\theta)}{\partial \theta_k} + z''_i \frac{\partial f''_i(\theta)}{\partial \theta_k} \right] - \sum_{i=0}^{i_0} \frac{\partial}{\partial \theta_k} \left[ f_i(\theta) + f'_i(\theta) + f''_i(\theta) \right] = 0, \quad k = 1, 2, 3. \quad (4.5.33)$$

2. Soient  $Leb_1$  et  $Leb_2$  les mesures de Lebesgue sur  $\mathbb{R}$  et  $\mathbb{R}^2$ . Alors :

$$Leb_1(C_{t_0-}^i) = \sqrt{2}(t_0 - t^*i), \quad (4.5.34)$$

$$Leb_2(A_{t_0-}^i) = (t_0 - t^*i) \left( h - \frac{1}{2}(t_0 - t^*i) \right), \quad (4.5.35)$$

$$Leb_2(A_{t_0+}^i) = \frac{1}{2}(h - t_0 + t^*i)^2, \quad (4.5.36)$$

$$Leb_1(C_{t_0+}^i) = \sqrt{2}(h - t_0 + t^*i), \quad (4.5.37)$$

où  $C_{t_0-}^i, A_{t_0-}^i, A_{t_0+}^i$  et  $C_{t_0+}^i$  sont définis par (4.4.1), (4.2.18), (4.2.19) et (4.4.2).

La preuve de ce lemme sera donnée à la fin de cette section. Revenons maintenant à la preuve du Théorème 4.4.2. Posons :

$$\begin{aligned} Z_i &= N_i(C_{t_0-}^i) \text{ et } z_i = v^i, \\ Z'_i &= N_i(A_{t_0-}^i) \text{ et } z'_i = n^i, \\ Z''_i &= N_i(A_{t_0+}^i \cup C_{t_0+}^i) \text{ et } z''_i = b^i - n^i - v^i, \\ \theta &= (\theta_1, \theta_2, \theta_3) = (\lambda_1, \lambda_2, h). \end{aligned}$$

En utilisant (4.3.9), (4.4.3) et le Lemme 4.5.3 nous en déduisons que (4.5.32) a bien lieu et nous avons :

$$\begin{aligned} f_i(\theta) &= \lambda_1 Leb_1(C_{t_0-}^i) = \lambda_1 \sqrt{2}(t_0 - t^*i), \\ f'_i(\theta) &= \lambda_2 Leb_2(A_{t_0-}^i) = \lambda_2 (t_0 - t^*i) \left( h - \frac{1}{2}(t_0 - t^*i) \right), \end{aligned}$$

$$f_i''(\theta) = \lambda_1 Leb_1(C_{t_0+}^i) + \lambda_2 Leb_2(A_{t_0+}^i) = \lambda_1 \sqrt{2}(h - t_0 + t^*i) + \frac{\lambda_2}{2}(h - t_0 + t^*i)^2,$$

$$f_i(\theta) + f_i'(\theta) + f_i''(\theta) = \lambda_1 \sqrt{2}h + \frac{\lambda_2}{2}h^2.$$

Les variables  $Z_i = N_i(C_{t_0-}^i)$ ,  $Z_i' = N_i(A_{t_0-}^i)$  et  $Z_i'' = N_i(A_{t_0+}^i \cup C_{t_0+}^i)$  étant indépendantes nous pouvons appliquer le Lemme 4.5.3. Les équations de vraisemblance, où  $z_i, z_i'$  et  $z_i''$  sont remplacées par  $v^i, n^i$  et  $b^i - n^i - v^i$  (Cf. la Section 4.4.2), conduisent au système suivant :

$$\sum_{i=0}^{i_0} \left[ v^i \frac{\frac{\partial \lambda_1 Leb_1(C_{t_0-}^i)}{\partial \theta_k}}{\lambda_1 Leb_1(C_{t_0-}^i)} + n^i \frac{\frac{\partial \lambda_2 Leb_2(A_{t_0-}^i)}{\partial \theta_k}}{\lambda_2 Leb_2(A_{t_0-}^i)} + (b^i - n^i - v^i) \frac{\frac{\partial (\lambda_1 Leb_1(C_{t_0+}^i) + \lambda_2 Leb_2(A_{t_0+}^i))}{\partial \theta_k}}{\lambda_1 Leb_1(C_{t_0+}^i) + \lambda_2 Leb_2(A_{t_0+}^i)} \right]$$

$$- \sum_{i=0}^{i_0} \frac{\partial}{\partial \theta_k} \left[ \lambda_1 Leb_1(C_{t_0-}^i) + \lambda_2 Leb_2(A_{t_0-}^i) + \lambda_1 Leb_1(C_{t_0+}^i) + \lambda_2 Leb_2(A_{t_0+}^i) \right] = 0,$$

où  $k = 1, 2, 3$ . En utilisant le Lemme 4.5.3 les équations de vraisemblance se réécrivent :

$$\sum_{i=0}^{i_0} \left[ \frac{v^i}{\lambda_1} + \frac{(b^i - n^i - v^i) \sqrt{2}(h - t_0 + t^*i)}{\lambda_1 \sqrt{2}(h - t_0 + t^*i) + \frac{\lambda_2(h - t_0 + t^*i)^2}{2}} \right] = \sqrt{2}(i_0 + 1)h,$$

$$(S) : \sum_{i=0}^{i_0} \left[ \frac{n^i}{\lambda_2} + \frac{(b^i - n^i - v^i)(h - t_0 + t^*i)^2}{2\lambda_1 \sqrt{2}(h - t_0 + t^*i) + \lambda_2(h - t_0 + t^*i)^2} \right] = \frac{(i_0 + 1)(h)^2}{2},$$

$$\sum_{i=0}^{i_0} \left[ \frac{n^i(t_0 - t^*i)}{(t_0 - t^*i)(h - \frac{1}{2}(t_0 - t^*i))} + \frac{(b^i - n^i - v^i)(\sqrt{2}\lambda_1 + \lambda_2(h - t_0 + t^*i))}{\lambda_1 \sqrt{2}(h - t_0 + t^*i) + \frac{\lambda_2(h - t_0 + t^*i)^2}{2}} \right]$$

$$= (\lambda_1 \sqrt{2} + \lambda_2 h)(i_0 + 1).$$

Nous noterons  $(\hat{\lambda}_1, \hat{\lambda}_2, \hat{h})$  la solution de ce système qui correspond aux estimateurs de maximum de vraisemblance des paramètres  $\lambda_1, \lambda_2$  et  $h$ . Ce qui achève la preuve du Théorème 4.4.2.

Revenons pour conclure à la preuve du Lemme 4.5.3.

*Preuve du Lemme 4.5.3.*

1. La vraisemblance de la suite  $(Z_i, Z_i', Z_i'')_{0 \leq i \leq i_0}$  est notée  $L_{i_0}$ . Par indépendance des variables  $\{Z_i, Z_i', Z_i'', 0 \leq i \leq i_0\}$  il vient que :

$$L_{i_0} = \frac{\prod_{i=0}^{i_0} [f_i(\theta)]^{z_i} [f_i'(\theta)]^{z_i'} [f_i''(\theta)]^{z_i''}}{\prod_{i=0}^{i_0} z_i! z_i'! z_i''!} \exp \left[ - \sum_{i=0}^{i_0} (f_i(\theta) + f_i'(\theta) + f_i''(\theta)) \right],$$

où  $z_i, z'_i, z''_i \in \mathbb{N}^3$ . D'où :

$$\begin{aligned} \ln(L_{i_0}) &= -\ln \left( \prod_{i=0}^{i_0} z_i! z'_i! z''_i! \right) + \sum_{i=0}^{i_0} \left[ z_i \ln [f_i(\theta)] + z'_i \ln [f'_i(\theta)] + z''_i \ln [f''_i(\theta)] \right] \\ &\quad - \sum_{i=0}^{i_0} \left[ f_i(\theta) + f'_i(\theta) + f''_i(\theta) \right]. \end{aligned} \quad (4.5.38)$$

Les équations de vraisemblance sont alors :

$$\sum_{i=0}^{i_0} \left[ z_i \frac{\partial f_i(\theta)}{\partial \theta_k} + z'_i \frac{\partial f'_i(\theta)}{\partial \theta_k} + z''_i \frac{\partial f''_i(\theta)}{\partial \theta_k} \right] - \sum_{i=0}^{i_0} \frac{\partial}{\partial \theta_k} \left[ f_i(\theta) + f'_i(\theta) + f''_i(\theta) \right] = 0, \quad k = 1, 2, 3.$$

2. La preuve est une conséquence directe des définitions des ensembles  $C_{t_0-}^i$ ,  $A_{t_0-}^i$ ,  $A_{t_0+}^i$ , et  $C_{t_0+}^i$ .

□

#### 4.5.6 Variance de la provision totale

Nous cherchons à calculer la variance de la provision totale. Nous avons :

$$\mathbb{V}[P_{t_0}] = \sum_{i=0}^{\lfloor \frac{t_0}{t^*} \rfloor + 1} \mathbb{V}[P_{t_0}^i] + 2 \sum_{0 \leq i < i' \leq \lfloor \frac{t_0}{t^*} \rfloor + 1} \left\{ \mathbb{E}[P_{t_0}^i P_{t_0}^{i'}] - \mathbb{E}[P_{t_0}^i] \mathbb{E}[P_{t_0}^{i'}] \right\}. \quad (4.5.39)$$

Calculons  $\mathbb{E}[P_{t_0}^i P_{t_0}^{i'}]$  pour commencer.

$$\begin{aligned} \mathbb{E}[P_{t_0}^i P_{t_0}^{i'}] &= \mathbb{E} \left\{ \left[ \sum_{j \geq 1} \mathbb{1}_{\{(T_j^i, D_j) \in A_{t_0+}^i\}} M_j f(T_j^i - t^* i, D_j, R_{T_j^i - t^* i}) \right] \right. \\ &\quad \times \left. \left[ \sum_{l \geq 1} \mathbb{1}_{\{(T_l^{i'}, D_l) \in A_{t_0+}^{i'}\}} M_l f(T_l^{i'} - t^* i', D_l, R_{T_l^{i'} - t^* i'}) \right] \right\} \\ &= \sum_{k \geq 0, h \geq 0} \left\{ \mathbb{P}[N_i(A_{t_0+}^i) = k] \mathbb{P}[N_{i'}(A_{t_0+}^{i'}) = h] \right\} \\ &\quad \times \sum_{1 \leq j \leq k, 1 \leq l \leq h} \mathbb{E} \left[ M_j f(T_j^i - t^* i, D_j, R_{T_j^i - t^* i}) M_l f(T_l^{i'} - t^* i', D_l, R_{T_l^{i'} - t^* i'}) \right] \\ &= \sum_{k \geq 0, h \geq 0} \left\{ \mathbb{P}[N_i(A_{t_0+}^i) = k] \mathbb{P}[N_{i'}(A_{t_0+}^{i'}) = h] \right\} C^{i, i'} k h \\ &= C^{i, i'} \mathbb{E}[N_i(A_{t_0+}^i)] \mathbb{E}[N_{i'}(A_{t_0+}^{i'})]. \end{aligned} \quad (4.5.40)$$

En combinant les expressions (4.5.39), (4.5.40), (4.5.18) et (4.3.2) il vient que :

$$\mathbb{V}[P_{t_0}] = \sum_{i=0}^{\lfloor \frac{t_0}{t^*} \rfloor + 1} \mathbb{V}[P_{t_0}^i] + 2 \sum_{0 \leq i < i' \leq \lfloor \frac{t_0}{t^*} \rfloor + 1} \gamma^2 [\mathbb{E}[M_1]]^2 \left\{ \int_{A_{t_0+}^i \times A_{t_0+}^{i'}} \int_{-\infty}^{\tilde{d}(\frac{\phi_{t-t^*i}(d)}{\gamma}, \mu, \sigma, t-t^*i)} \right\}$$

$$\begin{aligned} & F_3\left(t - t^*i, d, t' - t^*i', d', x\right) dx \} d\Lambda_i(t, d) d\Lambda_{i'}(t', d') \\ & - \int_{A_{t_0+}^i} F_1\left(t - t^*i, \frac{\phi_{t-t^*i}(d)}{\gamma}, \mu, \sigma\right) d\Lambda_i(t, d) \\ & \times \int_{A_{t_0+}^{i'}} F_1\left(t - t^*i', \frac{\phi_{t-t^*i'}(d)}{\gamma}, \mu, \sigma\right) d\Lambda_{i'}(t, d) \}, \end{aligned}$$

où  $\mathbb{V}\left[P_{t_0}^i\right]$  est obtenue via la formule (4.3.7).





# Compléments au modèle de provisionnement

---

Dans ce chapitre nous donnerons quelques compléments aux résultats obtenus dans le Chapitre 4. Nous commencerons par expliquer le lien entre la garantie proposée par CAMCA Assurance et les “credit default swap”. Puis nous montrerons qu’il est possible d’exprimer la provision comme une somme d’un nombre aléatoire de payoff d’options d’échange intervenant à des dates d’exercice aléatoire ; ceci nous permettra de discuter de la problématique d’investissement des primes perçues par CAMCA Assurance. Ensuite nous justifierons, dans la mesure du possible, les hypothèses du modèle de provisionnement. Nous proposerons une application numérique dans laquelle nous fournirons des statistiques décrivant les prêts considérés, une analyse quantitative des résultats obtenus dans le chapitre précédent (moyenne théorique, moyenne sur simulations, etc.) ainsi qu’une analyse de la sensibilité des résultats face à une variation de certains paramètres. Nous concluons ce chapitre par une présentation de la méthode numérique utilisée pour résoudre le système d’équations non linéaires présenté dans le Théorème 4.4.2.

Une traduction de ce chapitre, en anglais, est disponible page 217.

## 5.1 Garantie caution via les “credit default swap”

Nous montrons dans cette section le lien entre les “credit default swap” (CDS) et la garantie proposée par CAMCA Assurance.

Un CDS est un “contrat de protection financière” dans lequel :

1. L’acheteur de la protection est une entité (personne, entreprise, etc.) notée A,
2. Le vendeur de la protection est noté B,
3. L’entité de référence sur lequel porte le risque est notée C.

L’acheteur A se doit de payer une prime (en général trimestriellement) au vendeur B pour bénéficier de sa protection contre un évènement touchant C. En cas de réalisation de l’évènement (défini dans le contrat) sur l’entité C, alors B indemniserà A du préjudice subi (également défini dans le contrat).

Les évènements touchant l’entité C et couverts par le contrat peuvent être : un défaut de paiement, une faillite, une dégradation de la note, etc..

Notons que le souscripteur *A* du contrat est également son bénéficiaire.

Le but des CDS est de transférer le risque que supporte l'entité *A* vers l'entité *B* si l'entité *A* a un lien direct avec *C*. Cependant l'entreprise *A* peut avoir aucun lien avec l'entité *C*, dans ce cas de figure le CDS est un produit spéculatif et ne correspond plus à une protection. Voici une illustration :

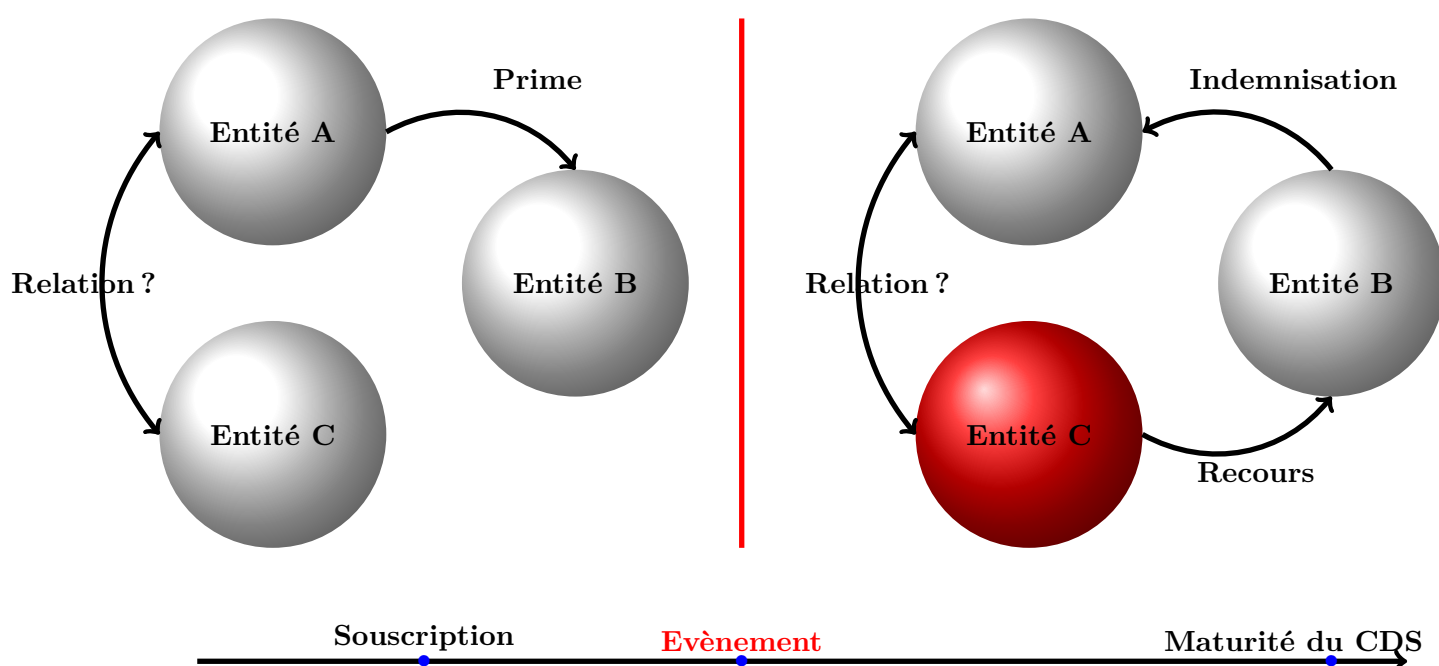


FIGURE 5.1 – Les “credits defaults swap”.

La précédente figure ressemble à la Figure 1.2.1 présentée lors de la description de la garantie proposée par CAMCA Assurance. Interprétons maintenant cette garantie via les CDS et expliquons les éventuelles différences. Dans la garantie proposée par la compagnie d'assurances :

- L'entité *A* est la banque,
- L'entité *B* est la compagnie d'assurances,
- L'entité *C* est l'emprunteur.

La garantie proposée par CAMCA Assurance est similaire à un CDS avec quelques différences :

1. L'entité *C* (sur lequel porte le risque) paie la prime du CDS.
2. Le souscripteur est l'entité *C* (l'emprunteur) et le bénéficiaire est l'entité *A* (la banque).



3. La prime est payée au début du contrat en un unique versement.
4. L'événement affectant l'entité C est la déclaration en sinistre du prêt par l'entité A : l'entité A décide du déclenchement de l'évènement concernant l'entité C entraînant un éventuel paiement de la part de l'entité B.

Hormis les quelques différences énoncées ci-dessus la garantie proposée par CAMCA Assurance est donc similaire à un CDS.

Voyons maintenant comment sont calculées les primes des CDS. Considérons pour ce faire un CDS de nominal  $N$ , de maturité  $D$  et soient  $t_0 = 0 < t_1 < \dots < t_n = D$  les dates de versement de la prime  $s$ . Soient enfin  $T$  le temps, aléatoire, de défaut et  $R$  le taux aléatoire de recouvrement (i.e. la part récupérée sur l'entité C en cas d'évènement déclenchant le CDS). Illustrons ceci graphiquement :



FIGURE 5.2 – Illustration des versements dans un “credit default swap”.

Pour calculer la prime d'un CDS il faut (Cf. [4] et [37] pour plus de détails) :

1. Déterminer la moyenne du taux de recouvrement  $\mathbb{E}[R]$ ,
2. Déterminer la f.d.r. des instants de défauts,
3. Et enfin estimer les taux sans risque.

Ce type de modélisation est différent du notre puisque dans ce modèle le nombre de défauts n'est pas aléatoire : une provision peut être constituée sur chaque contrat.

## 5.2 Provision via les options d'échange

Dans cette section nous interprétons la charge d'un sinistre déclaré à une date aléatoire  $T_j^i$  et concernant un prêt de caractéristiques  $(D_j, M_j)$  comme le payoff d'une option d'échange. Plus précisément nous voulons interpréter la formule (4.2.10) comme le payoff d'une option d'échange de maturité aléatoire  $T_j^i$  (Cf. [31] et [36] pour une présentation détaillée de ces options), i.e. :

$$(S_{T-I_j}^{1,j} - S_{T-I_j}^{2,j})_+, \quad I_j \leq T \leq I_j + D_j. \quad (5.2.1)$$

Par conséquent, d'après (4.2.23), la provision de sinistres pourra être vue comme une somme d'un nombre aléatoire de payoff d'options d'échange intervenant à des dates aléatoires.

Rappelons le principe des options d'échange. Une option européenne d'échange est une option permettant à son détenteur d'échanger, à une date  $T$  (appelée maturité), un actif

risqué de prix  $S_T^2$  contre un autre actif risqué de prix  $S_T^1$ . Le prix de cette option est appelé “prime”. Le payoff d’une telle option est défini par :

$$(S_T^1 - S_T^2)_+.$$

En effet à la maturité de l’option son détenteur a deux possibilités :

1. Si  $S_T^1 > S_T^2$ , alors le détenteur achète, à l’instant  $T$ , un actif de prix  $S_T^2$  et l’échange contre un actif de prix  $S_T^1$  qu’il revend immédiatement. Son gain est donc (sans considérer la prime versée) :

$$S_T^1 - S_T^2 = (S_T^1 - S_T^2)_+.$$

2. Si  $S_T^1 \leq S_T^2$  alors le détenteur de l’option ne fait rien. Son gain est donc :

$$0 = (S_T^1 - S_T^2)_+.$$

Montrons maintenant à travers un exemple le lien entre la garantie proposée par la compagnie d’assurances et les options d’échanges.

**Exemple 5.2.1.** Supposons qu’une personne désire emprunter auprès d’une banque une certaine somme le mois  $i = 0$ . Pour se décharger de son risque de contrepartie, l’emprunteur souscrit à la garantie proposée par la compagnie d’assurances, laquelle se substituera à l’emprunteur en cas de défaut de celui-ci. Considérons une option d’échange de maturité aléatoire correspondant à la date du défaut de l’emprunteur et considérons les actifs suivants :

- L’actif de prix  $(S_t^{1,j})_{t \geq 0}$  représentant la somme due par l’emprunteur  $j$  à la date  $t$ ,
- L’actif de prix  $(S_t^{2,j})_{t \geq 0}$  représentant le montant de la revente à la date  $t$  du bien financé par le prêt.

La société d’assurance est vue comme le vendeur de l’option d’échange de l’actif de prix  $S_t^{2,j}$  contre l’actif de prix  $S_t^{1,j}$ . L’acheteur de l’option est l’emprunteur. La date, aléatoire, d’exercice de l’option est la date de déclaration du sinistre  $T_j^0$ . Ainsi à la date 0 :

- L’emprunteur paie la prime pour détenir l’option d’échange. Grâce à la somme emprunté  $M_j$  l’emprunteur détient, compte tenu de son apport personnel, un bien d’une valeur  $(1 + a)M_j$ . Dans les options d’échange classiques aucun n’achat n’est effectué par l’acheteur à la date initiale.
- La compagnie d’assurances reçoit la prime de l’option.

A la date d’exercice, qui correspond à la date du sinistre, deux choix s’offrent à l’emprunteur (détenteur de l’option) :

1. Si le montant des recours (sur le bien financé) est supérieur ou égal au montant que l’emprunteur doit à la banque, alors l’emprunteur n’exerce pas son option d’échange : le bien est vendu au prix  $S_{T_j^0}^{2,j}$ . L’emprunteur peut alors rembourser la banque qui recouvre ainsi sa créance. La charge du sinistre, c’est à dire le payoff de l’option, est nulle pour la compagnie d’assurances.

2. Si le montant des recours est inférieur au montant que l'emprunteur doit à la banque, alors l'emprunteur exerce son option d'échange : il échange son actif de prix  $S_{T_j^0}^{2,j}$  contre l'actif de prix  $S_{T_j^0}^{1,j}$  qu'il reversera ensuite à la banque (sous forme de cash). La charge du sinistre est alors de  $S_{T_j^0}^{1,j} - S_{T_j^0}^{2,j}$  pour la compagnie d'assurances. En effet la compagnie d'assurances achètera à la date  $T_j^0$  un actif de prix  $S_{T_j^0}^{1,j}$  et l'échangera à l'emprunteur contre un actif de prix  $S_{T_j^0}^{2,j}$  que la compagnie vendra immédiatement.

Insistons sur le fait que le payoff de l'option ne représente pas le "gain" en cas de sinistre pour l'emprunteur mais uniquement le montant que la banque recevra de la part de l'assureur.

Le développement précédent montre que la charge d'un sinistre déclaré à la date  $T_j^i$  est le payoff d'une option d'échange de date d'exercice  $T_j^i$ . Ainsi la provision de sinistres concernant l'ensemble des prêts souscrits le mois  $i$  qui seront déclarés en sinistre dans l'intervalle de temps  $]t_0, t^*i + h[$  s'interprète clairement comme une somme d'un nombre Poissonien de payoff d'options d'échange entre les actifs de prix  $S_t^{2,j}$  et  $S_t^{1,j}$  à la date d'exercice égale au temps aléatoire de défaut  $T_j^i$ . Pour plus de détails sur la modélisation du risque de défaut via les options se référer à [22], [42] et [55].

Cette interprétation via les options d'échange suggère la question de la couverture, i.e. la question de l'investissement des primes perçues par la compagnie d'assurances. Pour un emprunteur  $j \in B$  notons  $p_j$  le montant de la prime perçue par la compagnie d'assurances. Nous avons vu que le contrat d'assurance proposé par la compagnie peut s'interpréter comme une option d'échange entre deux actifs dont le montant de la prime est  $p_j$ . La date d'exercice  $T_j^i$  de l'option est cependant aléatoire. D'autre part le nombre de contrats considérés, et ainsi le nombre d'options d'échange, est aléatoire et vaut :

$$\sum_{i=0}^{i_0} N_i (C_{t_0+}^i \cup C_{t_0-}^i \cup A_{t_0+}^i \cup A_{t_0-}^i), \quad (5.2.2)$$

où  $N_i$ ,  $C_{t_0-}^i$ ,  $A_{t_0-}^i$ ,  $A_{t_0+}^i$  et  $C_{t_0+}^i$  sont définis par (4.3.9), (4.4.1), (4.2.18), (4.2.19) et (4.4.2).

Ainsi pour déterminer une stratégie d'investissement des primes perçues par la compagnie d'assurances il nous faudrait définir un marché dépendant d'actifs exogènes (autres que les actifs de prix  $(S_t^1)_{t \geq 0}$  et  $(S_t^2)_{t \geq 0}$ ) et tenir compte de deux sources d'aléas :

1. Un aléa sur les dates d'exercices des options,
2. Et un aléa sur le nombre de ces options.

## 5.3 Justifications des hypothèses

Dans cette section nous essaierons de justifier l'ensemble des hypothèses réalisées dans le Chapitre 4.

### 5.3.1 Hypothèse 1

Rappelons la première hypothèse du modèle :

( $H_1$ ) : Si l'emprunteur  $j$  a souscrit son prêt à la date  $I_j$  correspondant au mois  $i$  alors nous supposons que  $I_j = t^*i$ .

Ainsi nous supposons que les prêts sont souscrits au début de chaque mois. Cette hypothèse n'a pas de réelle justification. Les prêts sont émis de manière continue mais un mois est une période courte au regard de la durée des prêts. D'autre part CAMCA Assurance ne dispose que du mois d'émission des prêts (et de l'année) et non de la date précise.

### 5.3.2 Hypothèse 2

Les v.a. représentant les durées des prêts sont en pratique majorées par une quantité inconnue  $h > 0$  tel que :

$$(H_2) : \begin{cases} D_j \leq h, & j \in B, \\ t_0 < I_j + h, & j \in B. \end{cases} \quad (5.3.1)$$

$$(5.3.2)$$

Voici une illustration :



Comme précisé dans la Section 1.2.2, CAMCA Assurance a édité un profil type de clients et de biens financés pour lesquels la banque n'a pas besoin de demander l'accord de la compagnie pour que leurs clients puissent bénéficier de la garantie caution. Dans ce profil type la compagnie a instauré une durée maximale concernant les prêts. Ainsi il est naturel de supposer que les v.a. représentant les durées des prêts sont majorées par une quantité déterministe.

Cependant si le client ne respecte pas le profil type édité par la compagnie d'assurances, la banque peut demander à la compagnie de prendre en charge la garantie de ce client atypique. La durée du prêt, dans ce cas de figure, pourrait dépasser la durée maximale inscrite dans le profil type. D'autre part les prêts peuvent être renégociés au cours de leur vie et la maturité peut ainsi changer : nous ne traitons pas ce cas de figure dans notre modèle au vu du faible nombre de prêts que cela représente (0,43% des prêts ont été restructurés). C'est pourquoi nous supposons que les v.a. représentant les durées des prêts sont majorées par une quantité déterministe mais que cette quantité est inconnue.

### 5.3.3 Hypothèse 3

Rappelons la troisième hypothèse du modèle :

$$(H_3) : \begin{cases} \text{Remboursement mensuel constant,} & (5.3.3) \\ \tau \text{ est le taux exponentiel déterministe,} & (5.3.4) \\ \text{Pas de remboursement anticipé partiel.} & (5.3.5) \end{cases}$$

L'hypothèse (5.3.3) est vérifiée en pratique car le nombre de prêts à taux variable assurés par CAMCA Assurance est infime.

L'hypothèse (5.3.4) n'est quant à elle pas vérifiée : le taux varie clairement en fonction des emprunteurs. L'impact sur le calcul des provisions du taux d'emprunt sera analysé dans la Section 5.4.3.

L'hypothèse (5.3.5) n'est pas satisfaite en pratique. Toutefois sous (5.3.5) le montant de provisions calculées est plus important que dans le cas d'une hypothèse contraire et assure à la compagnie une plus grande provision en augmentant ainsi la couverture du risque.

### 5.3.4 Hypothèse 4

Rappelons la quatrième hypothèse du modèle :

$$(H_4) : \begin{cases} \text{Emprunteur immédiatement déclaré en sinistre,} & (5.3.6) \\ \text{Un défaut concerne l'intégralité d'une mensualité.} & (5.3.7) \end{cases}$$

L'hypothèse (5.3.7) est vérifiée en pratique.

En revanche l'hypothèse (5.3.6) n'est clairement pas vérifiée en pratique : le délai moyen de déclaration d'un prêt en sinistre par la banque, suite à un défaut de paiement, est de 24 mois. En effet la banque a toute liberté pour déclarer le prêt en sinistre. Voyons un exemple (nous noterons 'P' pour paiement, 'D' pour défaut de paiement et 'DS' pour déclaration du sinistre) :

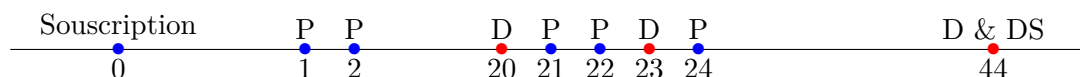


FIGURE 5.3 – Déclaration d'un sinistre.

Dans l'exemple ci-dessus, la banque a attendu trois défauts de paiement pour déclarer le prêt en sinistre et envisager ainsi un éventuel remboursement de la part de la compagnie d'assurances.

Rappelons que le montant dû par l'emprunteur en cas de sinistre est constitué du capital non remboursé et des intérêts de retard. Dans l'exemple précédent, les intérêts de retard seront plus importants que ceux que nous considérons dans notre modèle : les intérêts des mensualités 20, 23 et 44 sont impayés tandis que dans notre modèle le sinistre serait déclaré immédiatement lors de la mensualité 20 et seul l'intérêt de la mensualité 20 serait impayé. Ainsi, dans notre modèle, nous sous-estimons le montant du sinistre en intérêt de retard. Cependant nous constatons dans la pratique que la part des intérêts de retard est peu importante, voir négligeable, par rapport au capital restant dû de l'emprunteur comme l'illustre ce graphique :

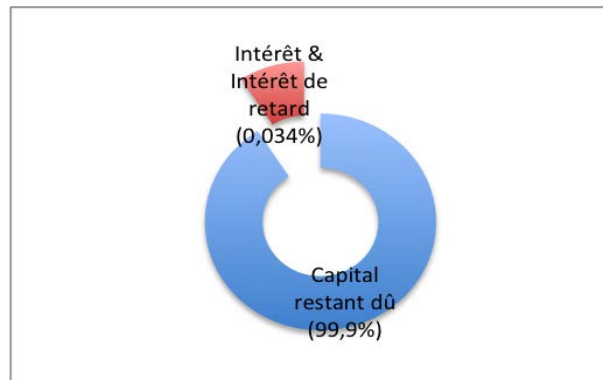


FIGURE 5.4 – Capital restant dû Vs Intérêts et intérêts de retard.

Dans notre modèle l'emprunteur est déclaré en sinistre lors de la mensualité 20 et dans ce cas de figure les mensualités 21, 22, ..., 43 et 44 sont impayées. Dans l'exemple ci-dessus l'emprunteur a remboursé les mensualités 21, 22 et 24 après son premier défaut à la date 20. Ceci implique que nous sur-estimons le montant de retard en capital dans notre modèle.

Il est enfin possible que l'emprunteur rembourse, entre la date du premier défaut et la date de déclaration du sinistre par la banque, une partie du capital non payé ce qui aurait pour effet de diminuer le capital que l'emprunteur doit à la banque. Reprenons notre précédent exemple en notant ' $R_k$ ' tout éventuel remboursement par l'emprunteur d'une mensualité  $k$  non payée :

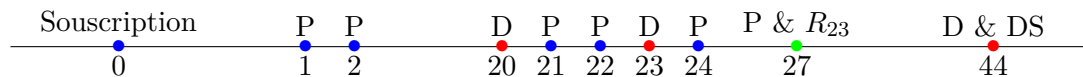


FIGURE 5.5 – Remboursement d'une mensualité non payée.

A la date 27 l'emprunteur paie la mensualité (capital et intérêt) correspondante à cette date et en plus rembourse la mensualité 23. Dans ce cas de figure notre modèle sur-estime le montant de retard en capital et en intérêt.

### 5.3.5 Hypothèse 5

La cinquième hypothèse est la suivante :

$$(H_5) : \begin{cases} \text{Le recours sur l'emprunteur est nul,} \\ \text{Le recours sur le bien financé encaissé à la date } T \text{ vaut :} \\ \gamma M_j R_{T-I_j}, j \in B, I_j \leq T \leq I_j + D_j, \end{cases} \quad (5.3.8)$$

Les recours sur l'emprunteur sont peu fréquents. En effet en cas de "sinistre" la trésorerie de l'emprunteur est souvent quasiment nulle, sinon il ne serait pas en "sinistre". C'est pourquoi nous supposons que les recours sur l'emprunteur sont nuls.

Nous supposons que la revente sur le bien financé est systématique bien qu'en réalité la banque ne réussisse à saisir le bien financé que dans 70% des cas : il s'agit d'une faiblesse de notre modèle.

A l'instant initial, nous supposons que la valeur du bien financé est égale au montant emprunté augmenté de l'apport personnel de l'emprunteur. Il paraît raisonnable de supposer que l'apport personnel de l'emprunteur est proportionnel au montant financé puisque ce montant est, en partie, conditionnel à l'apport personnel. Nous supposons que le coefficient de proportionnalité est constant par rapport au temps et identique pour tous les emprunteurs. Nous noterons alors :

$$a \in \mathbb{R}_+ \text{ le coefficient représentant l'apport personnel d'un emprunteur.} \quad (5.3.9)$$

Une possibilité que nous n'avons pas pris en compte est que le coefficient d'apport personnel soit aléatoire et / ou varie en fonction des emprunteurs.

Nous supposons d'autre part que le montant du bien à l'instant  $T$  dépend uniquement du cours de l'immobilier ; les éventuelles dégradations ou améliorations du bien ne sont pas prises en compte. Nous supposons en particulier que la valeur du bien à l'instant  $T$  est égale au montant du bien à la date initiale multiplié par une v.a.  $R_{I_j, T}$  ('R' pour "Real estate") représentant la variation des prix de l'immobilier entre les dates  $I_j$  et  $T$ . De nouveau, pour simplifier, nous supposerons que le coefficient  $R_{I_j, T}$  dépend uniquement de l'intervalle de temps entre la date initiale du prêt  $I_j$  et la date  $T$ , i.e. :

$$R_{I_j, T} = R_{T-I_j}, \quad (5.3.10)$$

avec en particulier  $R_0 = 1$ .

En cas de sinistre, la vente du bien financé engendre un coût (procédures judiciaires, etc.). Ces frais sont supposés proportionnels au montant du bien à l'instant du sinistre. Le coefficient de proportionnalité est supposé être identique pour tous les emprunteurs. Nous noterons alors :

$$c \in [0, 1] \text{ le coefficient représentant les coûts liés à la vente du bien.} \quad (5.3.11)$$

Ainsi, le montant de l'éventuelle revente du bien en cas de sinistre à la date  $T$  vaut :

$$(1+a)M_j R_{T-I_j} - c(1+a)M_j R_{T-I_j} = \gamma M_j R_{T-I_j},$$

où :

$$\gamma := (1-c)(1+a). \quad (5.3.12)$$

### 5.3.6 Hypothèse 6

Dans la sixième hypothèse nous avons supposé que :

( $H_6$ ) : L'ensemble des points  $(T_j^i, D_j)_{j \in B}$  est un PPP (Cf. [19] et [21]) d'intensité  $\Lambda_i$  :

$$\forall i : \Lambda_i := \lambda_1 \text{Leb}_{|C_{t_0-}^i \cup C_{t_0+}^i} + \lambda_2 \text{Leb}_{|A_{t_0-}^i \cup A_{t_0+}^i}, \quad (5.3.13)$$

où :

- $\lambda_1$  et  $\lambda_2$  sont deux paramètres réels et strictement positifs,
- $\text{Leb}_{|C_{t_0-}^i \cup C_{t_0+}^i}$  (resp.  $\text{Leb}_{|A_{t_0-}^i \cup A_{t_0+}^i}$ ) est la restriction de la mesure de Lebesgue à l'ensemble  $C_{t_0-}^i \cup C_{t_0+}^i$  (resp.  $A_{t_0-}^i \cup A_{t_0+}^i$ ) défini par (4.4.2) et (4.4.1) (resp. (4.2.19) et (4.2.18)).

Il s'agit d'une hypothèse classique qui n'a pas de réelle justification mais qui paraît raisonnable.

### 5.3.7 Hypothèse 7

Rappelons la septième hypothèse du modèle :

$$(H_7) : \text{Les PPP } \left( (T_j^i, D_j)_{j \in B} \right)_{i \geq 0} \text{ sont indépendants.} \quad (5.3.14)$$

Il n'y a, à priori, aucun lien entre les emprunteurs. Si un tel lien existe il est marginal. C'est pourquoi nous faisons cette hypothèse.

### 5.3.8 Hypothèse 8

La huitième hypothèse est la suivante :

$$(H_8) : \begin{cases} (T_j^i, D_j)_{j \in B}, (M_j)_{j \in B} \text{ et } (R_t)_{t \geq 0} \text{ sont indépendants,} & (5.3.15) \\ \text{Les v.a. } (M_j)_{j \in B} \text{ sont i.i.d..} & (5.3.16) \end{cases}$$

$R_t$  représente la variation de la valeur du bien financé due à la variation des prix de l'immobilier. Nous pourrions penser que, par exemple, une forte variation des prix de l'immobilier pourrait signifier, ou traduire, un risque plus important de déclaration des sinistres : un effondrement du marché de l'immobilier pourrait être synonyme d'une crise plus large touchant toute la population (crise systémique). Nous supposons qu'un tel



lien, difficilement mesurable, n'existe pas et donc que le PPP est indépendant de la variation de l'immobilier, i.e. :

$$(T_j^i, D_j)_{j \in B} \text{ et } (R_t)_{t \geq 0} \text{ sont indépendants.} \quad (5.3.17)$$

Nous pourrions également penser que le cours de l'immobilier varie de façon différente selon la valeur des biens considérés. Par exemple il pourrait exister :

- une forte variation de l'immobilier pour les biens à faible valeur,
- et une faible variation de l'immobilier pour les biens à forte valeur.

Cette différence de variation serait significative sur une gamme de biens très large. Chez l'assureur en question le montant des biens est relativement "concentré" (Cf. Section 5.4.1), c'est pourquoi nous supposons que :

$$(M_j)_{j \in B} \text{ et } (R_t)_{t \geq 0} \text{ sont indépendants.} \quad (5.3.18)$$

D'autre part le montant emprunté ne joue aucun rôle sur la déclaration d'un sinistre : un sinistre est causé par un changement dans la situation personnelle d'un emprunteur et non dans les caractéristiques du prêt, c'est pourquoi nous supposons que :

$$(T_j^i, D_j)_{j \in B} \text{ et } (M_j)_{j \in B} \text{ sont indépendants.} \quad (5.3.19)$$

L'hypothèse (5.3.16) se justifie par le fait que les montants de prêts sont relativement homogènes du fait du profil sur le bien financé établi par CAMCA Assurance. D'autre part l'indépendance entre les montants de prêts se justifie par le fait qu'il n'y a, à priori, aucun lien entre les emprunteurs.

### 5.3.9 Complément à la Remarque 4.2.4

Dans cette section nous supposons  $i = 0$ .

Dans la Remarque 4.2.4 nous avons expliqué que les sinistres intervenant lors du dernier remboursement du prêt n'étaient pas pris en compte dans notre modèle. Si un défaut de paiement avait lieu lors de la dernière échéance alors son montant serait très faible.

Nous avons vu que la fonction  $f$  (Cf. (4.2.16)) intervenant dans le montant d'un sinistre dépend de l'unité de temps  $t^*$ . Il vient alors que la limite de la charge d'un sinistre déclaré à une date aléatoire  $T = D$  pour un prêt d'une durée  $D$  et d'un montant  $M$  est nulle lorsque l'unité de temps  $t^*$  devient très petite :

$$\begin{aligned} \lim_{t^* \rightarrow 0} M f(D, D, R_D) &= \lim_{t^* \rightarrow 0} M \left[ \alpha \frac{e^{\tau D} - e^{\tau(D-t^*)}}{e^{\tau D} - 1} - \gamma R_D \right]_+ \\ &= 0, \end{aligned} \quad (5.3.20)$$

où la fonction  $f$  est définie par (4.2.16).

Passons maintenant à une application numérique des résultats théoriques obtenus dans le Chapitre 4.

## 5.4 Application numérique

Dans la Section 5.4.1 nous donnerons des statistiques descriptives sur les prêts considérés. Dans la Section 5.4.2 nous comparerons les résultats obtenus par simulation à l'aide de l'Algorithme 4.1 de ceux obtenus grâce aux formules analytiques présentes dans les Théorèmes 4.3.2 et 4.3.4. Nous étudierons également l'évolution des provisions en fonction de la date d'analyse et dans une hypothèse de run-off. Dans la Section 5.4.3 nous analyserons la sensibilité de la moyenne de la provision face à une variation de certains paramètres et dans la Section 5.4.4 nous étudierons la probabilité que la charge d'un sinistre soit nulle. Enfin nous expliquerons dans la Section 5.4.5, la méthode numérique utilisée pour résoudre le système d'équations présenté dans le Théorème 4.4.2.

Afin d'alléger la présentation, l'unité des générations sera l'année, et non le mois comme dans le chapitre précédent. Dans toute cette section les données considérées sont celles récoltées par la compagnie d'assurances du mois de janvier 2000 au mois de décembre 2013 pour une banque particulière. Le nombre de contrats considérés s'élève à 45 126 répartis par génération de la manière suivante :

| Génération | 00  | 01  | 02  | 03  | 04  | 05    | 06    |
|------------|-----|-----|-----|-----|-----|-------|-------|
| Nombre     | 361 | 409 | 429 | 330 | 394 | 1 734 | 3 041 |
| En %       | 0,8 | 0,9 | 0,9 | 0,7 | 0,9 | 3,8   | 6,7   |

| Génération | 07    | 08    | 09    | 10    | 11    | 12    | 13    |
|------------|-------|-------|-------|-------|-------|-------|-------|
| Nombre     | 6 722 | 6 048 | 3 545 | 5 763 | 6 356 | 4 581 | 5 413 |
| En %       | 14,9  | 13,4  | 7,8   | 12,8  | 14,1  | 10,1  | 12    |

FIGURE 5.6 – Nombre de prêts par génération.

avec la convention : 00 signifie génération 2000, 01 signifie génération 2001, etc..

### 5.4.1 Statistique descriptive

Nous présenterons quelques statistiques décrivant les données suivantes :

- Montant et durée des prêts,
- Capital restant dû et intérêts,
- Nombre de sinistres,
- Prêts échus.

#### Montant et durée des prêts

|                         | Moyenne | Ecart-type | Minimum | Maximum | Médiane |
|-------------------------|---------|------------|---------|---------|---------|
| Montant des prêts (EUR) | 65 378  | 60 093     | 150     | 700 000 | 50 000  |
| Durée (mois)            | 201     | 69,9       | 12      | 360     | 180     |

FIGURE 5.7 – Statistiques descriptives sur le portefeuille considéré.

#### 5.4 Application numérique

---

Toutes générations de contrats confondues, le montant moyen des prêts considérés, au 31/12/2013, est de EUR 65 378 pour une durée moyenne de financement de 201 mois. L'étendue concernant les montants et les durées est relativement importante (EUR 699 850 pour les montants et 348 mois pour les durées). Ceci est dû :

- aux demandes de prêts atypiques ne respectant pas le profil type de bien, Cf. Section 1,
- aux refinancement de certains prêts (durée et montant faible),
- ou enfin aux prêts relais.

Les écarts types (EUR 60 093 pour les montants et 70 mois pour les durées) nous indiquent cependant que le non respect du profil de bien est plus généralement lié à un montant supérieur au plafond qu'à une durée de financement atypique car les écarts types des montants de prêts sont plus importants que ceux de la durée (en pourcentage).

Observons maintenant la distribution du montant des prêts :

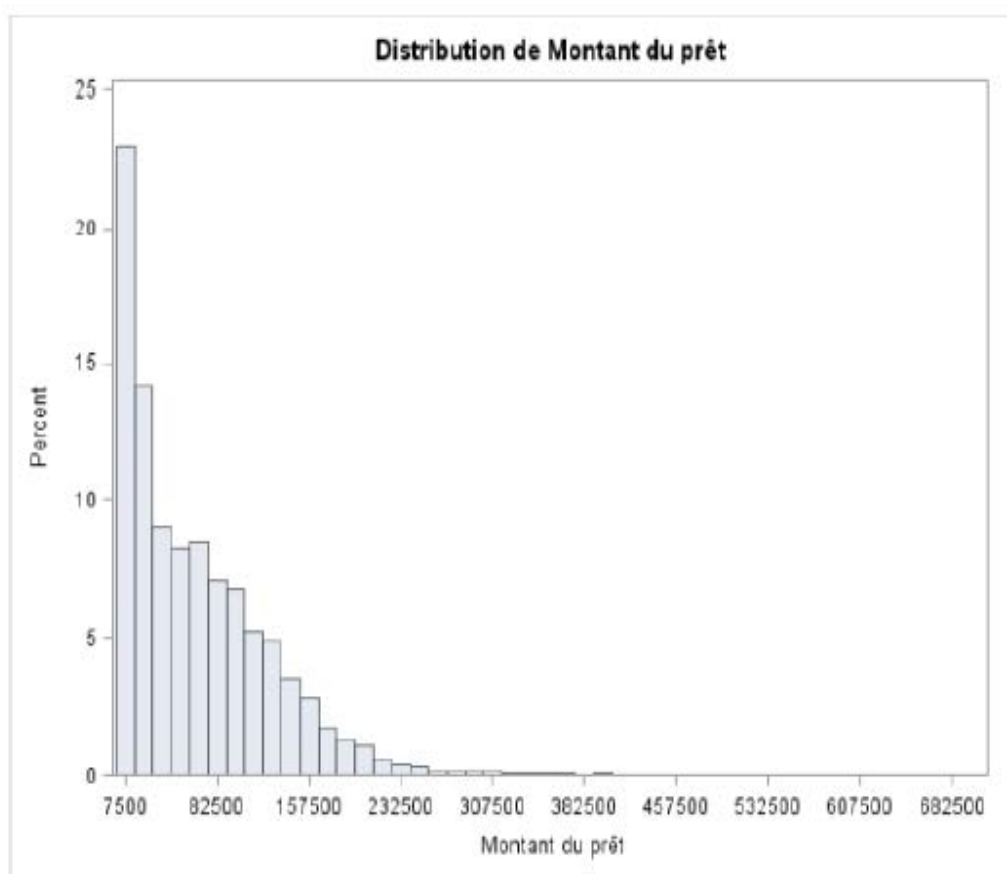


FIGURE 5.8 – Histogramme du montant des prêts.

L'histogramme précédent est déterminé par classe de montant de prêts. Nous avons choisi un découpage en  $k$  classes, où  $k$  vaut (méthode de Sturges) :

$$k := 1 + 3,3 \times \log(n), \quad (5.4.1)$$

avec  $n$  le nombre de données considérées ; soit environ 36 classes dans notre cas de figure. Nous constatons que la distribution est uni-modale. Observons maintenant l'histogramme de la durée des prêts :

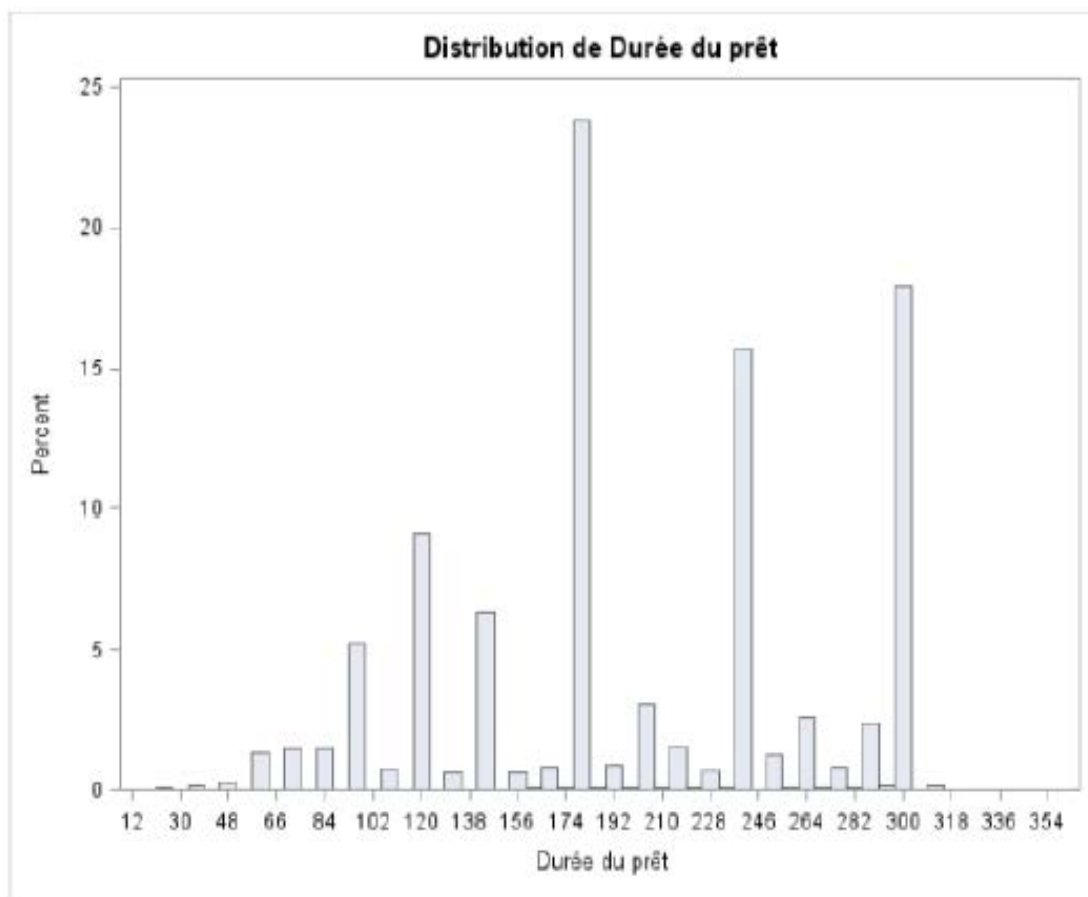


FIGURE 5.9 – Histogramme de la durée des prêts.

Nous constatons que la distribution empirique de la durée des prêts est multi-modale et que les durées sont assez éparpillées autour de la médiane (180 mois).

Nous allons maintenant analyser des boxplots (ou boîtes à moustache, ou diagrammes en boîte, ou boîtes de Tukey) dont nous rappelons l'interprétation :

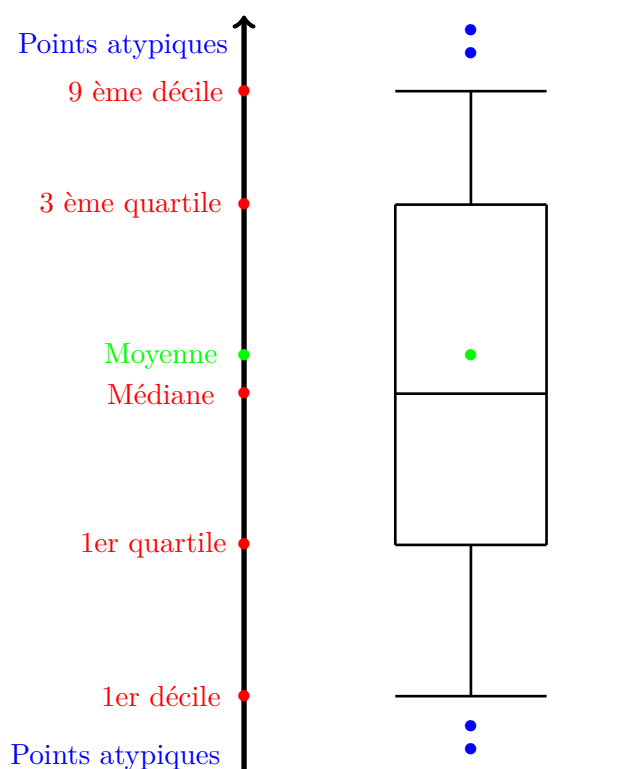


FIGURE 5.10 – Explication boxplot.

Observons maintenant les boxplots du montant et de la durée des prêts.

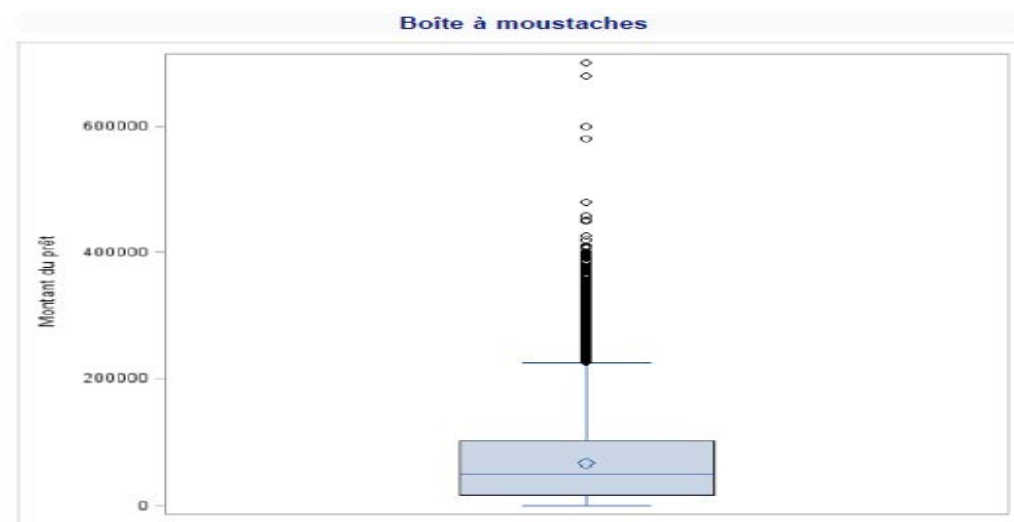


FIGURE 5.11 – Boxplot du montant des prêts.

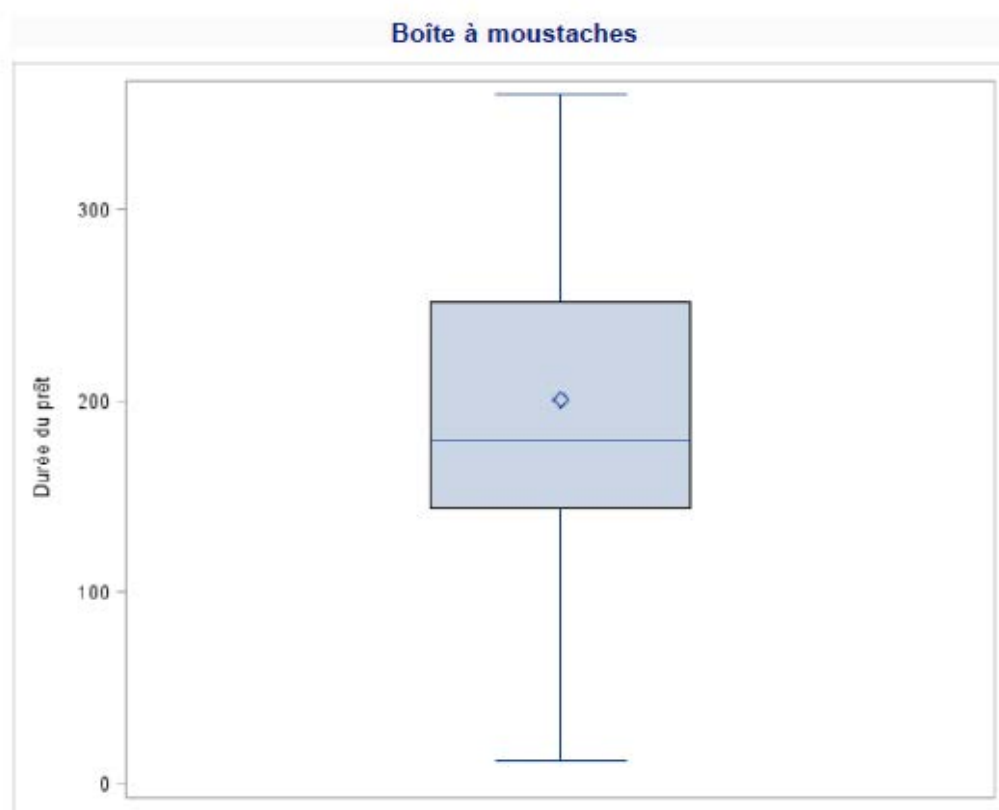


FIGURE 5.12 – Boxplot de la durée des prêts.

Analysons maintenant les données par génération :

| Génération                     | 00  | 01  | 02  | 03 | 04  | 05  | 06  |
|--------------------------------|-----|-----|-----|----|-----|-----|-----|
| Mnt génération / Mnt total (%) | 0,1 | 0,1 | 0,1 | 0  | 0,5 | 3,6 | 6,7 |

| Génération                     | 07   | 08 | 09  | 10   | 11   | 12   | 13   |
|--------------------------------|------|----|-----|------|------|------|------|
| Mnt génération / Mnt total (%) | 16,1 | 12 | 6,2 | 14,2 | 14,9 | 11,2 | 14,3 |

FIGURE 5.13 – Poids (en montant de prêts) par génération.

où “Mnt génération” signifie “Montant des prêts par génération de contrats” et “Mnt total” signifie “Montant des prêts toutes générations de contrats confondues”. Ce tableau nous indique que les générations 2006 à 2013 représentent à elles seules près de 95% du montant total de prêts cautionnés.

Voyons pour conclure les boxplots par génération de contrats :

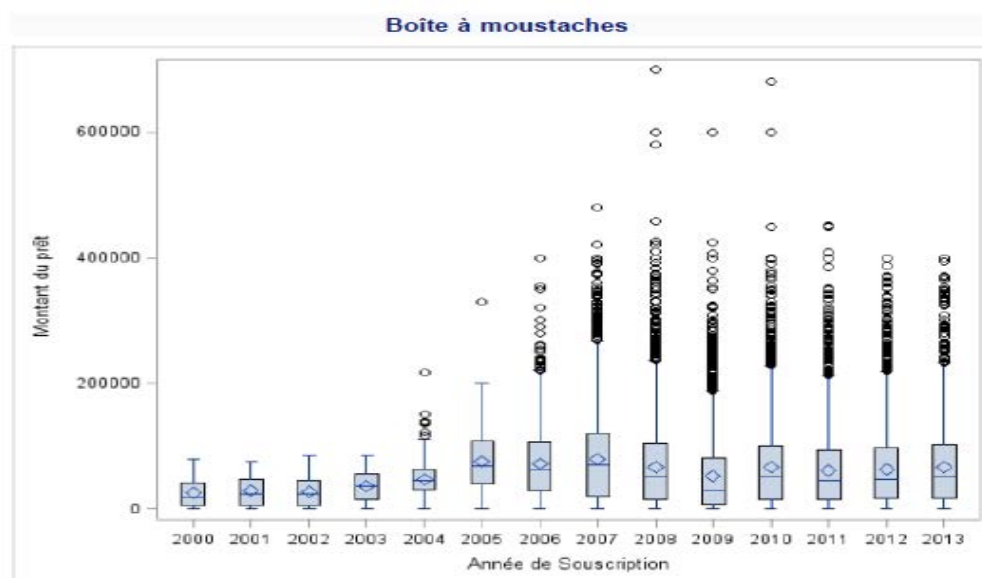


FIGURE 5.14 – Boxplot du montant des prêts par génération.

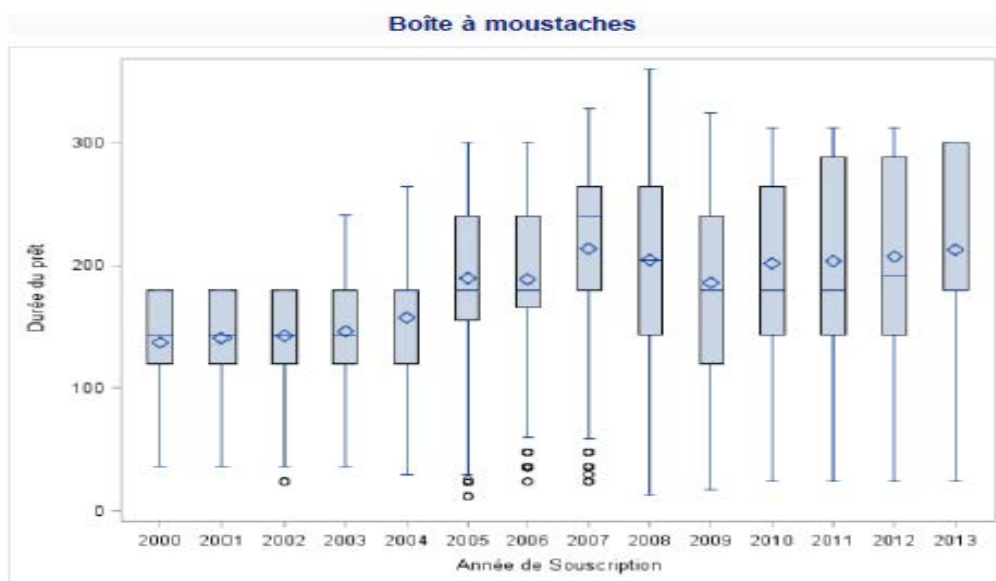


FIGURE 5.15 – Boxplot de la durée des prêts par génération.

Le montant moyen des prêts tend à rester constant sur les dernières générations. Une légère augmentation de la durée moyenne d'endettement est toutefois à constater. Nous observons que le nombre de prêts atypiques en montant (synonyme de dérogation) n'évolue pas sur les générations de 2006 à 2013. Cela traduit le fait que les demandes de dérogations effectuées par la banque auprès de la compagnie d'assurances, pour obtenir un

montant emprunté supérieur au plafond, reste stable au fil du temps : une augmentation pourrait générer une augmentation des risques. Le boxplot de la durée des prêts confirme que les durées sont moins concentrées autour de la moyenne que les montants.

### Capital restant dû et intérêt de retard

Le montant d'intérêt de retard représente 0,034% du montant total de capital restant dû et d'intérêt de retard. Nous nous concentrerons donc sur l'analyse du capital restant dû. Le montant moyen de capital restant dû est de EUR 53 945 (pour un montant moyen de prêt de EUR 65 378). L'évolution, par génération, des boxplots est la suivante :

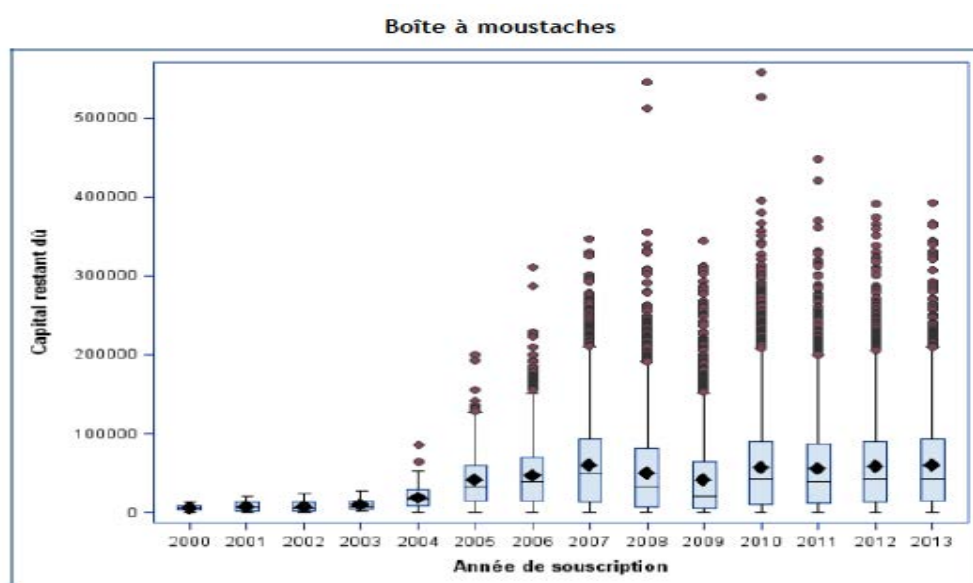


FIGURE 5.16 – Boxplot du capital restant dû par génération.

Plusieurs constatations peuvent être effectuées :

- Le capital restant dû représente le capital sous risque (intérêts de retard omis) : le montant maximal de provisions ne peut, a priori, dépasser le capital restant dû.
- Le montant moyen de capital restant dû augmente plus la génération est récente.

Comparons maintenant le montant du capital restant dû au montant de prêt : où “CRD génération” signifie “capital restant dû par génération de contrats” et rappelons que “Mnt génération” signifie “Montant des prêts par génération de contrats”. Comme précédemment, le ratio capital restant dû sur montant des prêts augmente plus la génération est récente. D’autre part les générations anciennes (antérieures à 2006) ont un ratio inférieur à 50% : en cas d’un éventuel sinistre, la revente du bien pourrait couvrir la totalité du montant dû par l’emprunteur. Observons pour conclure le ratio  $\text{CRD génération } i / \text{CRD toutes générations confondues}$  :



## 5.4 Application numérique

| Génération                          | 00   | 01 | 02   | 03   | 04   | 05 | 06   |
|-------------------------------------|------|----|------|------|------|----|------|
| CRD génération / Mnt génération (%) | 15,6 | 24 | 23,4 | 25,2 | 38,6 | 53 | 63,1 |

| Génération                          | 07   | 08   | 09   | 10   | 11   | 12   | 13   |
|-------------------------------------|------|------|------|------|------|------|------|
| CRD génération / Mnt génération (%) | 74,1 | 77,2 | 80,2 | 86,7 | 90,5 | 93,8 | 90,2 |

FIGURE 5.17 – Ratio CRD sur montant des prêts.

| Génération                     | 00 | 01 | 02 | 03 | 04  | 05  | 06  |
|--------------------------------|----|----|----|----|-----|-----|-----|
| CRD génération / CRD total (%) | 0  | 0  | 0  | 0  | 0,2 | 2,3 | 5,2 |

| Génération                     | 07   | 08   | 09  | 10   | 11   | 12   | 13   |
|--------------------------------|------|------|-----|------|------|------|------|
| CRD génération / CRD total (%) | 14,6 | 11,3 | 6,1 | 15,1 | 16,5 | 12,8 | 15,7 |

FIGURE 5.18 – Poids (en CRD) par génération.

où “CRD total” signifie “capital restant dû toutes générations de contrats confondues”.  
Voici une illustration graphique de ce tableau :

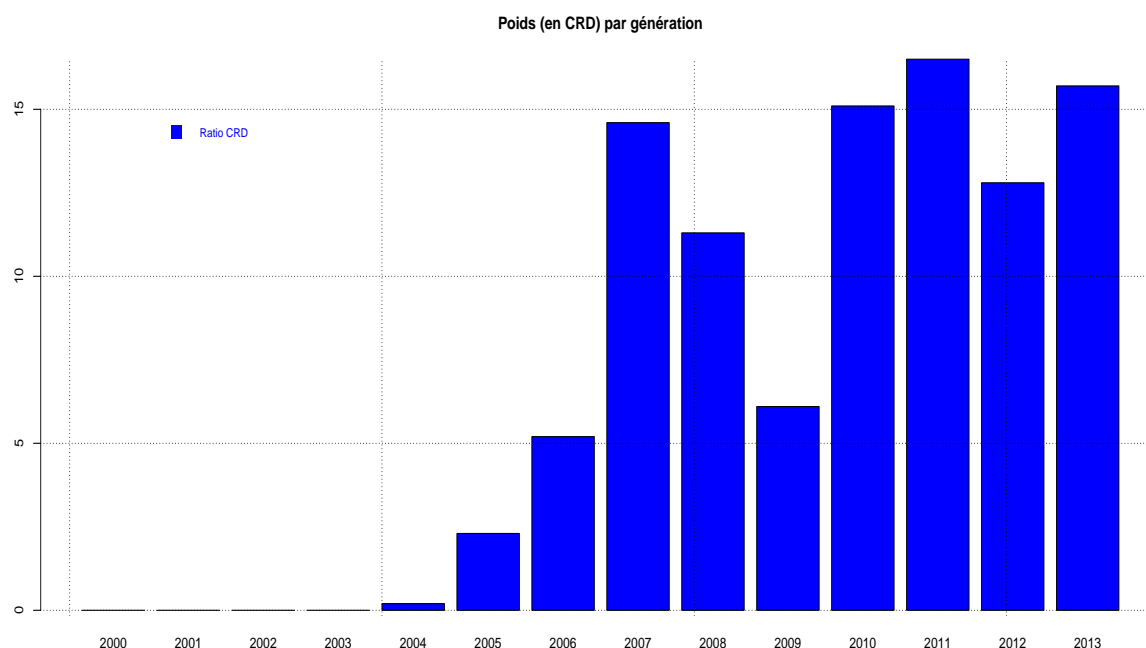


FIGURE 5.19 – Poids (en CRD) par génération - Graphique.

Notons que les générations 2006 à 2014 représentent plus de 95% du capital restant dû de toutes les générations : ces générations représentent la majorité des risques.

## Sinistres

Par mesure de confidentialité nous ne fournissons pas les données en terme de montant mais uniquement en terme de nombre.

Voici la répartition, par génération de contrats, des 330 sinistres enregistrés du 1er janvier 2000 au 31 décembre 2013 :

| Génération                                   | 00   | 01  | 02  | 03   | 04  | 05   | 06   |
|--|------|-----|-----|------|-----|------|------|
| Nb sin. génération                           | 41   | 30  | 32  | 36   | 15  | 64   | 37   |
| Nb sin. génération / Nb sin. total (%)       | 12,4 | 9,1 | 9,7 | 11   | 4,5 | 19,4 | 11,2 |
| Nb sin. génération / Nb prêts génération (%) | 11,3 | 7,3 | 7,4 | 10,9 | 3,8 | 3,7  | 1,2  |

| Génération                                   | 07   | 08  | 09  | 10  | 11  | 12  | 13 |
|--|------|-----|-----|-----|-----|-----|----|
| Nb sin. génération                           | 35   | 13  | 21  | 2   | 2   | 2   | 0  |
| Nb sin. génération / Nb sin. total (%)       | 10,6 | 3,9 | 6,4 | 0,6 | 0,6 | 0,6 | 0  |
| Nb sin. génération / Nb prêts génération (%) | 0,5  | 0,2 | 0,6 | 0   | 0   | 0   | 0  |

FIGURE 5.20 – Nombre de sinistres.

où “Nb sin. génération” signifie “nombre de sinistres par génération de contrats”, “Nb sin. total” signifie “nombre de sinistres toutes générations confondues de contrats” et “Nb prêts génération” signifie “nombre de prêts par génération de contrats”.

Un sinistre déclaré dans les premières années de la vie d’un prêt n’entraîne pas, en général, de remboursement de la part de la compagnie d’assurances, cf. Sections 5.3.9 et 5.4.4. Cette observation est confirmée par les précédentes données : la génération 2013 n’a pas encore nécessité d’indemnisation de la part de la compagnie d’assurances et la compagnie est intervenue deux fois pour chacune des générations 2010 à 2012.

La troisième ligne du précédent tableau indique que les générations 2000, 2003, 2005, 2006 et 2007 sont les générations qui ont le plus grand nombre de sinistres : elles représentent en effet près de 65% du nombre total de sinistres indemnisés mais pour un poids de 27% du nombre total de prêts.

La dernière ligne du tableau précédent nous fournit le ratio, par génération de contrats, du nombre de sinistres par le nombre de prêts initiaux ; ce ratio est appelé taux de sinistralité. Nous constatons logiquement que les générations les plus anciennes ont des taux de sinistralité plus élevés que les générations récentes.

## Prêts échus

Le nombre de prêts échus de manière naturelle, i.e. sans défaut de paiement définitif, se décompose par génération de la manière suivante :

#### 5.4 Application numérique

---

| Génération                                    | 00  | 01  | 02  | 03  | 04  | 05  | 06  |
|---|-----|-----|-----|-----|-----|-----|-----|
| Nb échus génération                           | 279 | 292 | 269 | 191 | 147 | 483 | 770 |
| Nb échus génération / Nb prêts génération (%) | 77  | 71  | 63  | 58  | 37  | 28  | 25  |

| Génération                                    | 07   | 08   | 09  | 10  | 11  | 12  | 13  |
|---|------|------|-----|-----|-----|-----|-----|
| Nb échus génération                           | 1695 | 1422 | 523 | 439 | 299 | 133 | 15  |
| Nb échus génération / Nb prêts génération (%) | 25   | 23   | 15  | 8   | 5   | 3   | 0,3 |

FIGURE 5.21 – Prêts échus.

où “Nb échus génération” signifie “nombre de prêts échus de manière naturelle par génération de contrats”. Comme attendu le nombre de prêts échus augmente plus la génération est ancienne.

### 5.4.2 Espérance, intervalle de confiance et nombre de sinistres estimés

Comme nous l'avons expliqué dans le Chapitre 4 :

1. Nous pouvons obtenir une valeur numérique de l'espérance et de la variance de la provision avec les formules (4.3.2) et (4.3.7) des Théorèmes 4.3.2 et 4.3.4.
2. Nous pouvons simuler la provision par la méthode de Monte Carlo proposée dans l'Algorithme 4.1 de la Section 4.3.2 et obtenir ainsi une valeur numérique approchée des quantiles, de la moyenne et de la variance avec l'échantillon simulé.

Ces deux méthodes, l'une théorique et l'autre basée sur des simulations, devraient fournir des résultats numériques proches pour la moyenne et la variance. Au préalable nous devons disposer d'une valeur numérique des paramètres de notre modèle. Pour l'obtenir nous nous baserons sur la méthode d'estimation proposée dans la Section 4.4 appliquée au portefeuille de contrats présenté dans la section précédente. Ceci conduit aux estimations suivantes :

| Paramètre         | Notation          | Valeur |
|-------------------|-------------------|--------|
| $\mathbb{E}(M_1)$ | $\bar{m}$         | 65 378 |
| $\tau$            | $\hat{\tau}$      | 0,07   |
| $\gamma$          | $\hat{\gamma}$    | 1,17   |
| $p$               | $\hat{p}$         | 2      |
| $\mu$             | $\hat{\mu}$       | -0,07  |
| $\sigma$          | $\hat{\sigma}$    | 0,1958 |
| $\lambda_1$       | $\hat{\lambda}_1$ | 94,89  |
| $\lambda_2$       | $\hat{\lambda}_2$ | 0,46   |
| $h$ (année)       | $\hat{h}$         | 23,26  |

FIGURE 5.22 – Synthèse estimation paramètres.

Notez que conformément à ce que nous avons expliqué dans la Section 4.4, les paramètres liés au mouvement brownien géométrique ont été fixés par avis d'expert. En effet ces paramètres représentent l'évolution de l'immobilier sur une période pouvant attendre 30 ans : il apparaît plus judicieux de fixer ces paramètres que de les estimer sur des données. Le choix effectué correspond à une baisse de 30% des prix de l'immobilier sur une période de 10 ans. Ce scénario correspond, en partie, au benchmark utilisé par la Banque Centrale Européenne dans le cadre des stress tests réalisés dans le milieu bancaire ; cet exercice est dénommé "Asset Quality Review" ou AQR.

Nous présenterons plusieurs résultats concernant la provision :

1. Comparaison de l'espérance théorique et de l'espérance obtenue par simulation,
2. Evolution de l'espérance en fonction de la date d'analyse pour une génération fixée,
3. Evolution de l'espérance en fonction des générations (à une date d'analyse fixée),
4. Evolution de l'espérance en fonction de la date d'analyse et sous hypothèse run-off,
5. Comparaison avec la méthode de provisionnement utilisée par la compagnie.

### Espérance théorique Vs espérance simulée

Nous comparons ici pour la génération 2000 (i.e.  $i = 0$ ) et en faisant évoluer la date d'analyse  $t_0$  :

1. L'espérance de la provision obtenue en remplaçant la valeur des paramètres par leur "estimation" dans la formule (4.3.2) du Théorème 4.3.2 qui sera notée :

$$\mathbb{E}^{theo}[P_{t_0}^0] := \hat{\lambda}_2 \bar{m} \times \int_{\hat{A}_{t_0+}^0} \left\{ \hat{\phi}_t(s) \Phi \left[ \frac{\ln \left[ \frac{\hat{\phi}_t(s)}{\hat{\gamma}} \right] - \hat{\mu}t}{\hat{\sigma}\sqrt{t}} \right] - \hat{\gamma} e^{(\hat{\mu} + \frac{\hat{\sigma}^2}{2})t} \Phi \left[ \frac{\ln \left[ \frac{\hat{\phi}_t(s)}{\hat{\gamma}} \right] - (\hat{\mu} + \hat{\sigma}^2)t}{\hat{\sigma}\sqrt{t}} \right] \right\} ds dt, \quad (5.4.2)$$

où :

$$\hat{A}_{t_0+}^0 := \left\{ (s, t) \in (\mathbb{R}_+)^2 : t_0 < t < s; s \leq \hat{h} \right\}, \quad (5.4.3)$$

$$\hat{\phi}_t(s) := \begin{cases} \left( e^{\hat{\tau}t^*} (1 + t^*\hat{p}) - t^*\hat{p} \right) \times \frac{e^{\hat{\tau}s} - e^{\hat{\tau}(t-t^*)}}{e^{\hat{\tau}s} - 1}, & \text{si } t^* < t \leq s, \\ \left( e^{\hat{\tau}t^*} (1 + t^*\hat{p}) - t^*\hat{p} \right), & \text{si } t \leq t^* \text{ ou } s \leq t^*. \end{cases} \quad (5.4.4)$$

2. Et l'espérance de la provision obtenue en simulant la provision à l'aide de l'algorithme suivant :

---

**Algorithme 5.1** Simulation d'un échantillon de taille  $l$  de la variable  $P_{t_0}^0$ .

---

- 1: **for**  $k = 1$  to  $l$  **do**
- 2: Simuler une v.a. de loi  $\mathcal{P}(\hat{\lambda}_2 \text{Leb}(\hat{A}_{t_0+}^0))$ . Nous noterons  $n$  sa réalisation.
- 3: Simuler  $n$  v.a. i.i.d. de loi uniforme sur  $\hat{A}_{t_0+}^0$ . Nous noterons  $(t_j, d_j)_{1 \leq j \leq n}$  les réalisations. Ordonner les temps  $(t_j)_{1 \leq j \leq n} : t_{(1)} \leq \dots \leq t_{(n)}$ . Nous noterons  $d_{(j)}$  la durée associée à la date  $t_{(j)}$ .
- 4: Les variables  $(M_j)_{1 \leq j \leq n}$  sont simulées par inversion de la f.d.r. empirique. Nous noterons  $(\tilde{m}_j)_{1 \leq j \leq n}$  les réalisations.
- 5: Simuler  $n$  v.a.r. i.i.d. de loi normale centrée et réduite. Nous noterons  $(g_j)_{1 \leq j \leq n}$  les réalisations.
- 6: Calculer :

$$p_k^0 = \sum_{j=1}^n \tilde{m}_j \left\{ \hat{\phi}_{t_{(j)}}(d_{(j)}) - \hat{\gamma} \prod_{k=1}^j \hat{r}_k \right\}_+, \quad (5.4.5)$$

où  $\hat{\phi}$  est défini par (5.4.4) et :

$$\hat{r}_k := \exp \left[ \hat{\sigma} \sqrt{t_{(k)} - t_{(k-1)}} g_k + \hat{\mu}(t_{(k)} - t_{(k-1)}) \right]. \quad (5.4.6)$$

- 7: **end for**
-

Notons que la taille de l'échantillon est prise égale à  $l = 5\,000$ . L'espérance obtenue sur l'échantillon simulée sera notée  $\mathbb{E}^{emp}(P_{t_0}^0)$  et définie par :

$$\mathbb{E}^{emp}(P_{t_0}^0) := \frac{1}{l} \sum_{k=1}^l p_k^0. \quad (5.4.7)$$

Traçons maintenant les courbes  $[t_0, \mathbb{E}^{theo}(P_{t_0}^0)]$  (Cf. (5.4.2)) et  $[t_0, \mathbb{E}^{emp}(P_{t_0}^0)]$  (Cf. (5.4.7)) pour  $t_0 \in \{0, \dots, \lfloor \hat{h} \rfloor\}$  (Cf. Figure 5.22) :

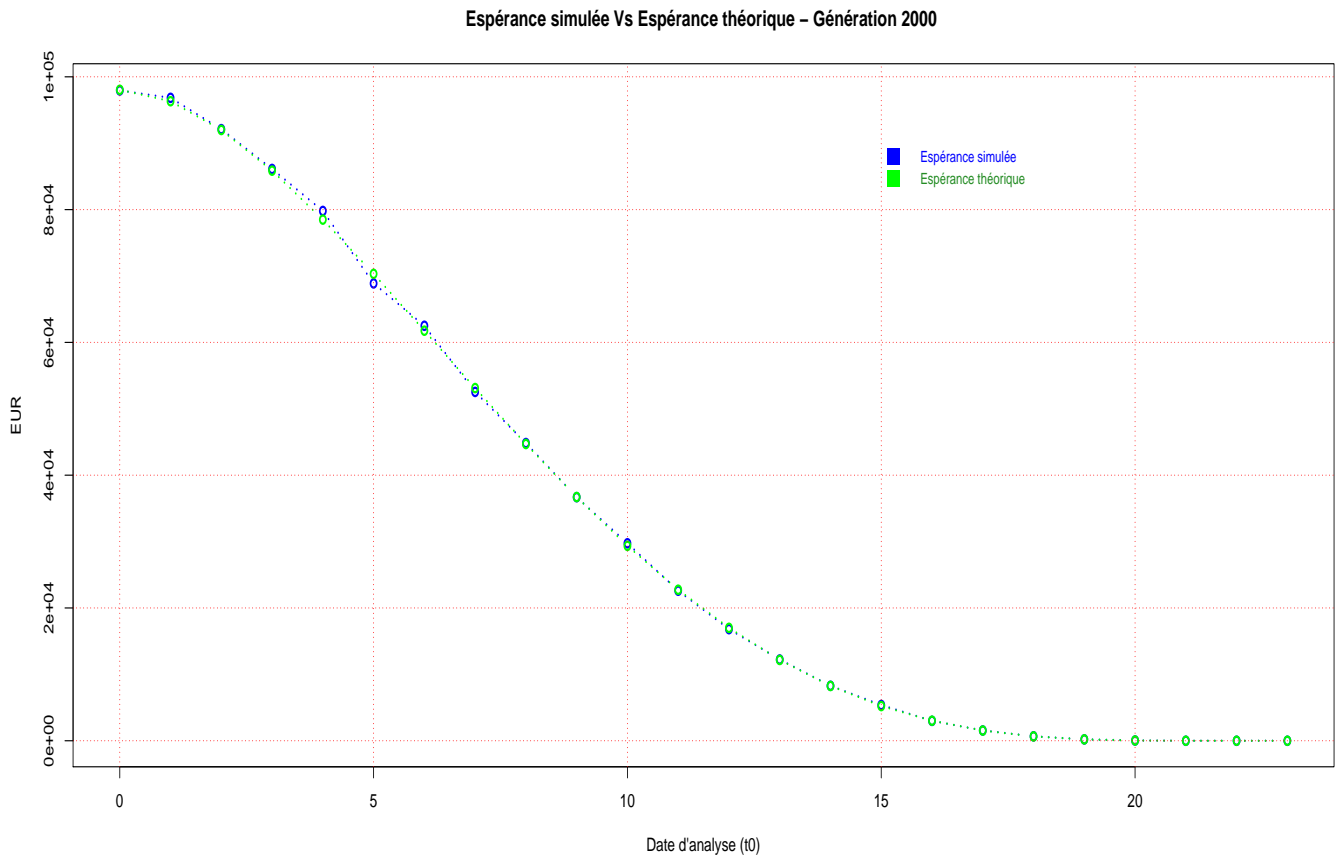


FIGURE 5.23 – Espérance théorique Vs espérance simulée.

De la Figure 5.23 nous pouvons tirer les conséquences suivantes :

1. L'espérance théorique et l'espérance obtenue par simulation sont extrêmement proches et ce quelque soit la date d'analyse.
2. Comme attendu l'espérance de la provision est décroissante en fonction de la date d'analyse mais cette décroissance n'est pas linéaire.
3. Enfin, dès 2018 l'espérance de la provision est quasiment nulle. Ceci nous permettrait de déterminer numériquement la date  $t_0^*$  (dans cet exemple  $t_0^* = 18$  années) à

## 5.4 Application numérique

partir de laquelle l'espérance de la provision devient négligeable, i.e. :

$$t_0^* := \inf\{t_0 \geq 0 : \mathbb{E}[P_{t_0}^0] \leq \beta\}, \quad (5.4.8)$$

où  $\beta$  serait fixé très petit. Ceci fournirait à la compagnie d'assurances la durée maximale de provisionnement par génération, i.e. la durée pendant laquelle la compagnie devrait enregistrer une provision dans ses comptes.

Notons que grâce aux simulations de Monte Carlo il est possible d'obtenir un intervalle de confiance de l'espérance (au seuil de 99,5% dans l'exemple ci-dessous) :

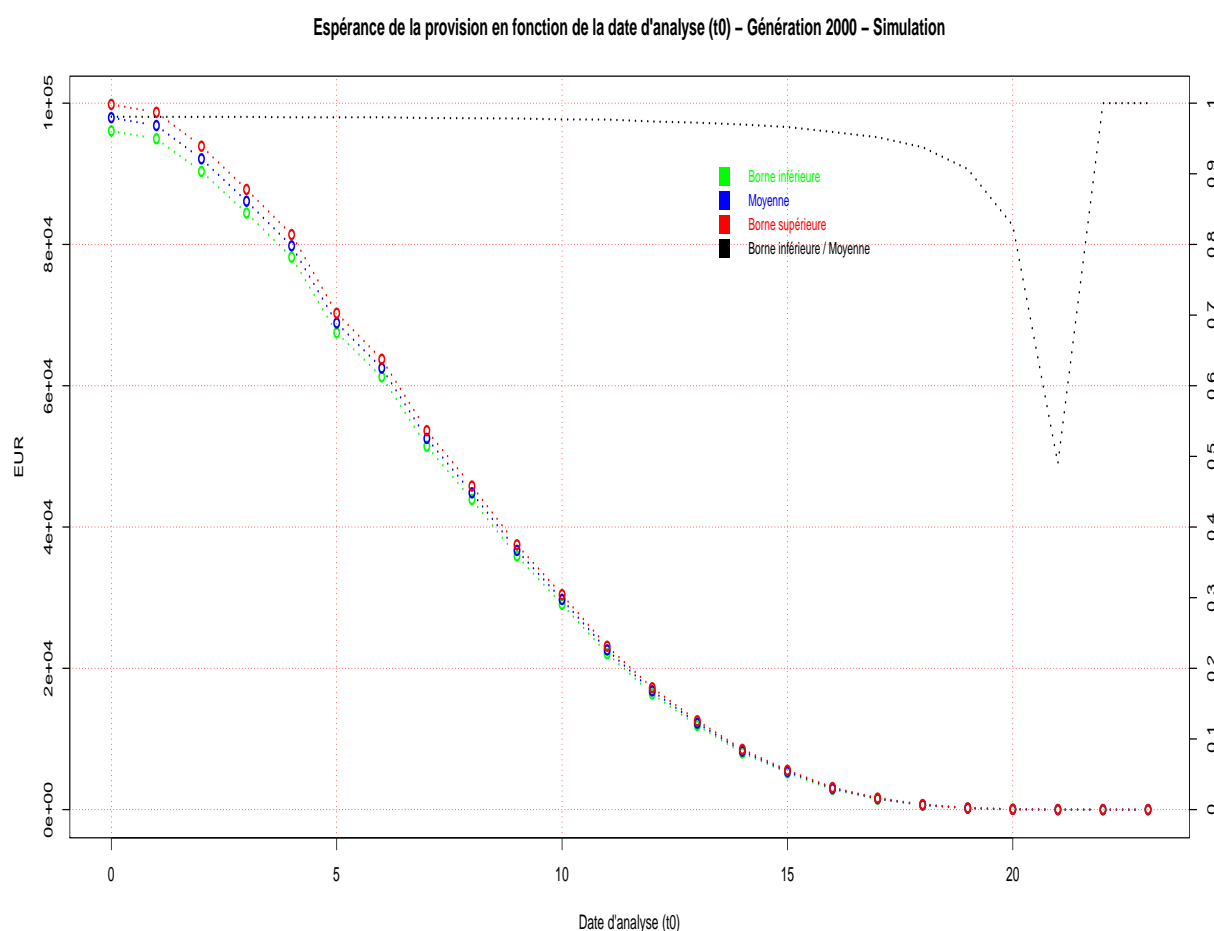


FIGURE 5.24 – Espérance et intervalle de confiance en fonction de la date d'analyse.

Vu la simplicité de l'algorithme utilisé et le nombre de simulations effectués, les intervalles de confiance sont très précis, comme le montre la courbe noire représentant l'écart relatif entre la borne inférieure de l'intervalle de confiance et l'espérance de la provision.

### Evolution de l'espérance en fonction des générations

Ici nous fixons la date d'analyse au 31 décembre 2013 et nous faisons varier les générations de contrats. Nous présentons uniquement les résultats obtenus via les simulations, i.e. nous traçons la courbe  $[i, \mathbb{E}^{emp}(P_{t_0}^i)]$  (Cf. (5.4.7)) pour  $i \in \{0, \dots, 13\}$ .

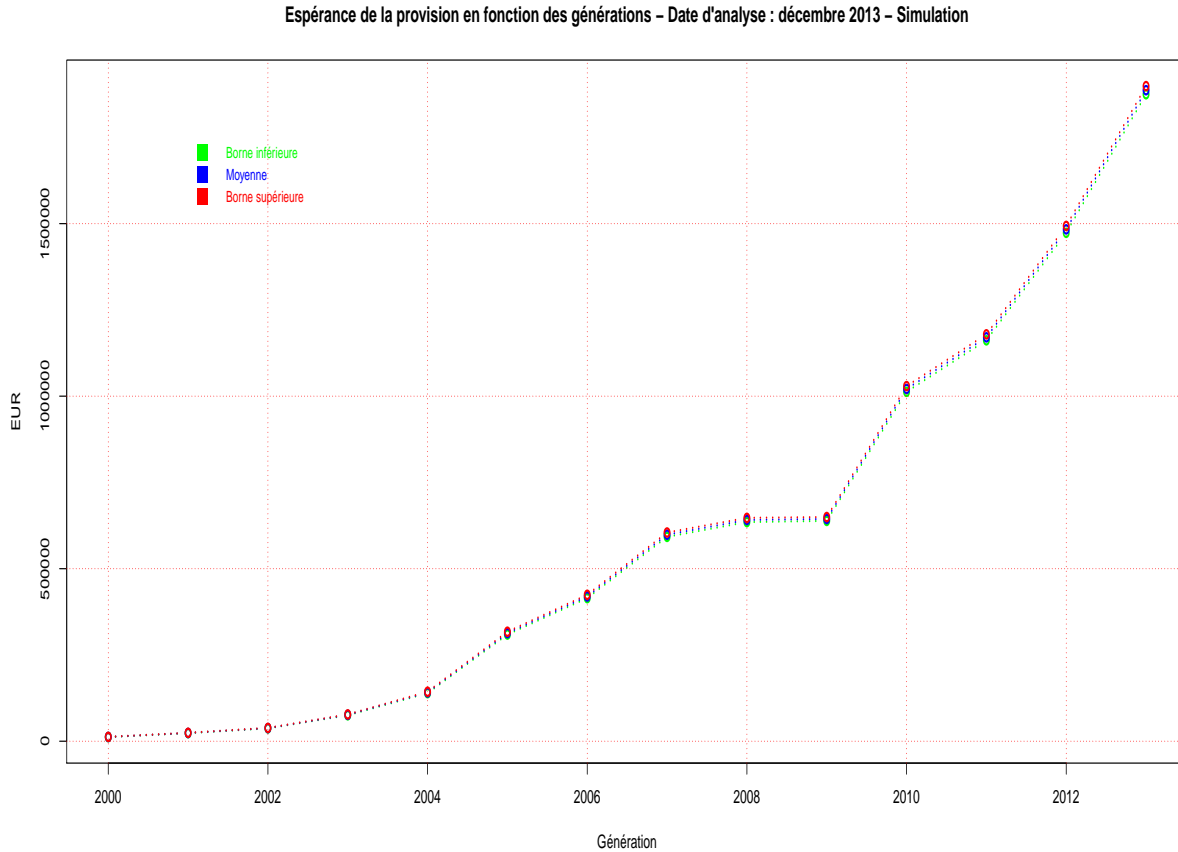


FIGURE 5.25 – Espérance de la provision en fonction des générations.

Comme attendu, la moyenne de la provision croît plus la génération est récente.

Il est intéressant de comparer le ratio  $\mathbb{E}^{emp}(P_{t_0}^i) / \mathbb{E}^{emp}(P_{t_0})$  au ratio préalablement introduit “CRD génération  $i$  sur CRD toutes générations” (Cf. page 82), où  $\mathbb{E}^{emp}(P_{t_0})$  est définie par :

$$\mathbb{E}^{emp}(P_{t_0}) := \sum_{i=0}^{\lfloor \frac{t_0}{t^*} \rfloor + 1} \mathbb{E}^{emp}(P_{t_0}^i). \quad (5.4.9)$$

En effet le pourcentage de provisions que représente chaque génération doit être sensiblement le même que celui du capital restant dû, qui rappelons le représente le capital sous risque. Nous cherchons à vérifier que la provision est “correctement répartie” entre



## 5.4 Application numérique

chaque génération. Notons cependant qu'en raison de la possibilité de revente du bien financé en cas de sinistre, il n'est pas illogique que pour les générations les plus anciennes le ratio de provisions soit inférieur au ratio de capital restant dû. Voici les ratios :

| Génération                                  | 00  | 01  | 02  | 03  | 04  | 05  | 06  |
|---|-----|-----|-----|-----|-----|-----|-----|
| CRD génération / CRD total (%)              | 0   | 0   | 0   | 0   | 0,2 | 2,3 | 5,2 |
| $E^{emp}(P_{t_0}^i) / E^{emp}(P_{t_0})$ (%) | 0,4 | 0,7 | 1,1 | 1,6 | 2,3 | 3,2 | 4,3 |

| Génération                                  | 07   | 08   | 09  | 10   | 11   | 12   | 13   |
|---|------|------|-----|------|------|------|------|
| CRD génération / CRD total (%)              | 14,6 | 11,3 | 6,1 | 15,1 | 16,5 | 12,8 | 15,7 |
| $E^{emp}(P_{t_0}^i) / E^{emp}(P_{t_0})$ (%) | 5,6  | 7,1  | 9   | 11,7 | 14   | 17,5 | 21,4 |

FIGURE 5.26 – Ratio CRD Vs Ratio Espérance (%).

où rappelons que “CRD génération” signifie “capital restant dû par génération de contrats” et “CRD total” signifie “capital restant dû toutes générations de contrats confondues”. Voici une illustration graphique de ce tableau :

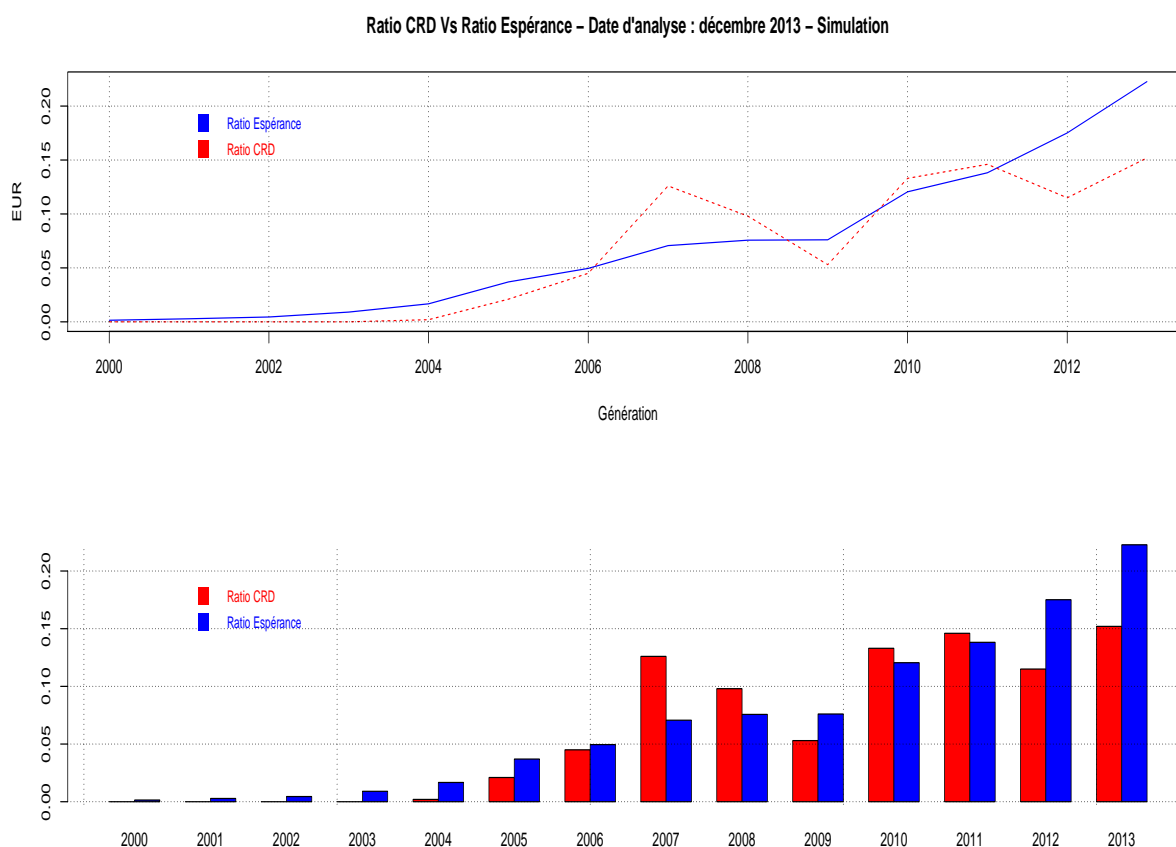


FIGURE 5.27 – Ratio CRD Vs Ratio espérance (graphiques).

Nous constatons que le ratio de la provision suit “grossièrement” la même répartition par génération que le ratio du capital restant dû. Nous constatons cependant que pour les générations 2007, 2008, 2010 et 2011 le ratio des CRD dépasse celui des provisions : cela illustre une faiblesse du modèle.

L’apriori des praticiens est que les générations antérieures à 2008 sont les plus risquées. La crise de 2008 aurait modifiée la politique de souscription des Caisses Régionales du Crédit Agricole : les contrats assurés par CAMCA Assurance depuis cette date seraient vraisemblablement moins risqués. Il est cependant encore trop tôt pour juger si cet apriori est confirmé : les générations postérieures à 2008 n’ont pas encore atteint une maturité suffisante pour être comparées aux générations antérieures. Quoiqu’il en soit notre modélisation de l’intensité des PPP ne permet pas de prendre en compte ce phénomène car nous avons supposé que les paramètres intervenant dans l’intensité des PPP sont indépendants des générations de contrats ; une solution possible pour corriger ce problème serait de pondérer les générations de prêts par un poids déterminé par avis d’expert.

Nous constatons également que le ratio de la provision est supérieur à celui du CRD pour les générations 2012 et 2013 : ceci intéresse la compagnie d’assurances car elle préfère s’assurer une marge de sécurité dans les premières années de vie d’une génération.

Observons pour conclure le nombre de sinistres prédits par le modèle ainsi que le ratio “nombre de sinistres total / nombre de prêts initial”. Nous négligerons les générations 2000 à 2005 dans l’analyse suivante qui ne sont pas significatives.

| Génération                                    | 06   | 07   | 08 | 09   | 10   | 11 | 12   | 13   |
|---|------|------|----|------|------|----|------|------|
| Nb sin. prévus génération                     | 39,1 | 43,9 | 49 | 54,4 | 60,1 | 66 | 72,2 | 78,7 |
| Nb sin. total génération / Nb prêts total (%) | 2,5  | 1,2  | 1  | 2,1  | 1    | 1  | 1,6  | 1,4  |

FIGURE 5.28 – Nombre de sinistres prévus.

où “Nb sin. prévus génération” signifie “moyenne du nombre de sinistres prévus par génération de contrats” et “Nb sin. total génération” signifie “nombre de sinistres indemnisés plus moyenne du nombre de sinistres prévus par génération de contrats”.

Nous constatons plusieurs éléments :

1. Le taux de la génération 2013 est entièrement imputable au modèle puisqu’il n’y a pas encore eu de sinistres indemnisés par la compagnie à la date de calcul,
2. Les générations antérieures à 2008 n’apparaissent pas plus risquées que les autres générations contrairement à l’apriori des praticiens,
3. L’intensité des PPP ne prend pas en compte les générations, donc le nombre de sinistres prévus ne tient pas compte du nombre de prêts souscrits par génération : cela illustre une faiblesse du modèle.

### Evolution de l'espérance en fonction de la date d'analyse et sous l'hypothèse d'un run-off

Nous nous intéressons maintenant à l'évolution de l'espérance de la provision toute générations confondues, en fonction de la date d'analyse  $t_0$  mais sous l'hypothèse d'un run-off au 31 décembre 2013, c'est à dire sans supposer de nouvelles souscriptions de prêts à compter de cette date. Nous traçons pour ce faire la courbe  $[t_0, \mathbb{E}^{emp}(P_{t_0}^{Runoff})]$  pour  $t_0 \in \{0, \dots, \lfloor \hat{h} + 13 \rfloor\}$  où  $P_{t_0}^{Runoff}$  est ici définie par :

$$P_{t_0}^{Runoff} := \sum_{i=0}^{\min(13, \lfloor \frac{t_0}{\tau^*} \rfloor)} P_{t_0}^i, \quad (5.4.10)$$

avec  $P_{t_0}^i$  définie par (4.2.23). Voici une illustration :

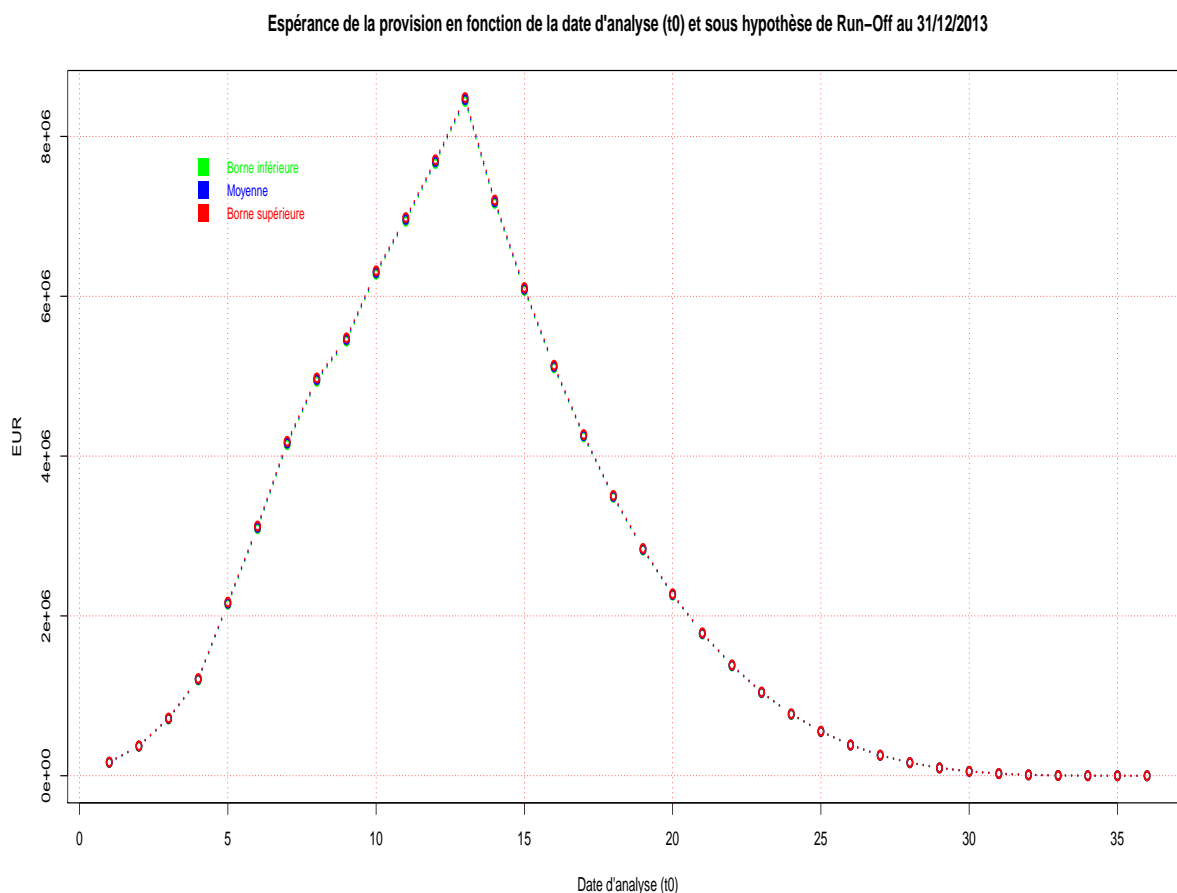


FIGURE 5.29 – Espérance de la provision sous hypothèse de run-off.

Nous constatons que l'espérance de la provision est quasiment nulle dès l'année 2031 sous l'hypothèse d'un run-off fin 2013.

## Comparaison avec la méthode actuelle de provisionnement

La compagnie d'assurances détermine deux types de provisions concernant la garantie caution des prêts habitats : une provision individuelle (sur chaque contrat) pour les contrats risqués et une provision collective pour les contrats jugés non risqués ; la compagnie est en effet capable, au moyen de différentes informations, d'isoler les contrats les plus risqués.

La provision individuelle, déterminée contrat par contrat, est calculée comme "l'expected loss bilan", ou ELB. La méthodologie utilisée est celle de la réforme bancaire Bâle 2, c'est à dire que l'ELB est égale à la probabilité de défaut du contrat multiplié par la perte au moment du défaut ("loss given default") multiplié par l'exposition au moment du défaut ("exposition at default"). Ces provisions sont fournies par la banque émettrice des prêts. Nous ne rentrerons pas plus en détails sur le calcul de cette provision car nous ne disposons pas de toutes les informations nécessaires.

Le deuxième type de provision est une provision collective du type "Coût moyen \* Nombre de sinistres", où :

- Le coût moyen est déterminé par génération de contrats comme le capital restant dû moyen (en incluant les intérêts de retard),
- Le nombre de sinistres est déterminé via une pseudo méthode de Chain Ladder.

Ce sont ces deux provisions que CAMCA Assurance inscrit dans ses comptes ; il s'agit de provisions statutaires.

La future réforme Solvabilité 2 (Cf. Section 2.1.2 pour plus de détails) imposera aux compagnies d'assurances, début 2016, de calculer les provisions techniques en mode "Best Estimate" (BE dans la suite). Le BE des provisions techniques est défini dans les spécifications techniques fournies par l'EIOPA de la manière suivante (Cf. [1]) :

*"The best estimate should correspond to the probability weighted average of future cash-flows taking account of the time value of money".*

Schématiquement, le BE des provisions correspond à une actualisation des provisions techniques statutaires en ne prenant plus en considération d'éventuelle marge de sécurité prudentielle : ainsi le BE des provisions est en général inférieur aux provisions statutaires. CAMCA Assurance procède, depuis 2010, à un calcul de BE de ces provisions : ce calcul consiste en une réadaptation de la méthode de calcul des provisions statutaires que nous ne détaillerons pas ici.

Comparons, au 31 décembre 2013, le montant des provisions statutaires calculées par la compagnie, le montant du BE de ces provisions et le montant calculé via notre modèle (mais cette fois-ci actualisé) : les deux premières provisions sont fournies par la compagnie. Nous présentons les résultats par génération de contrats. Dans le graphique

## 5.4 Application numérique

ci-dessous les bâtons rouges représentent les provisions statutaires calculées par CAMCA Assurance, les bâtons bleus représentent le BE de ces provisions, les bâtons verts représentent les provisions calculées via notre méthode et l'axe des abscisses représente les générations de contrats. Nous négligerons les générations 2000 à 2005.

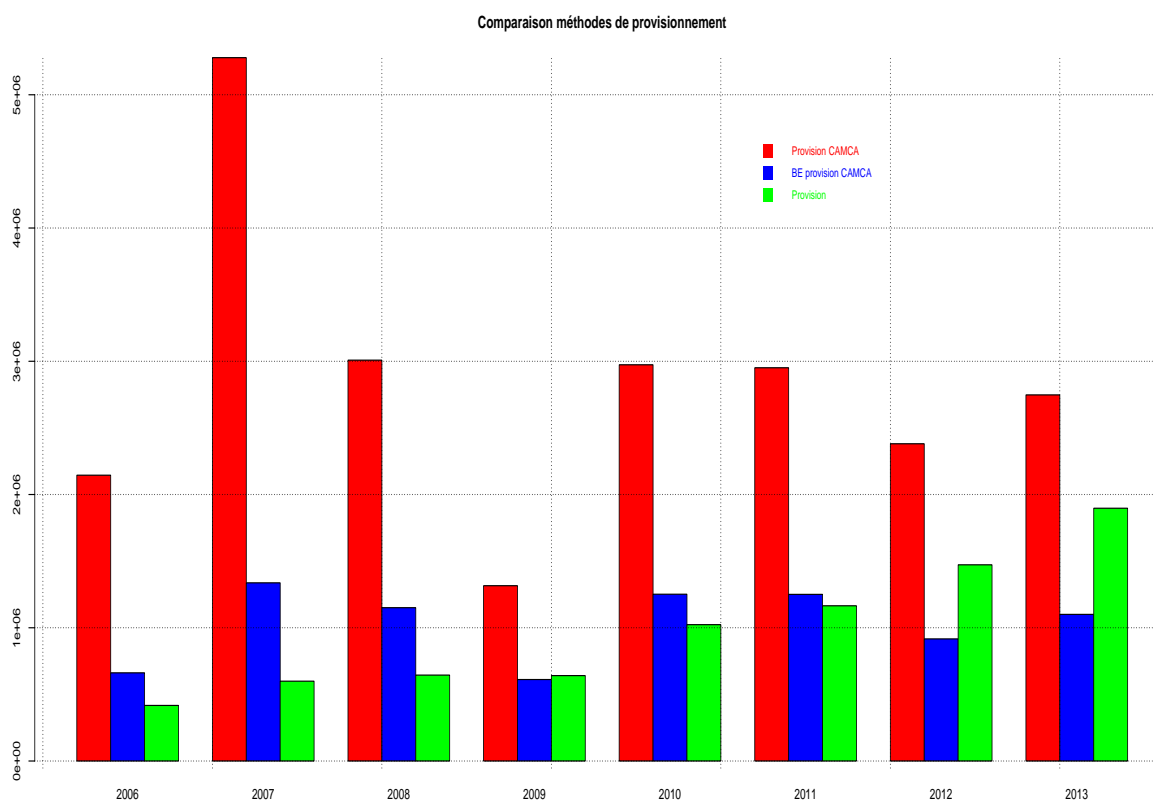


FIGURE 5.30 – Comparaison des méthodes de provisionnement.

Nous pouvons tirer de la Figure 12.29 les conséquences suivantes :

1. Comme énoncé précédemment les provisions statutaires sont largement supérieures aux BE des provisions.
2. Nous constatons également que les provisions statutaires et le BE des provisions prennent en compte le caractère, soit disant, plus risqué des générations 2008 et antérieures.
3. Les provisions de notre modèle et le BE des provisions sont relativement proches pour les générations 2009 à 2011, mais divergent quelque peu pour les autres.

Comparons maintenant les ratios “Provision génération  $i$  sur Provisions toutes générations confondues” au ratio “CRD génération  $i$  sur CRD toutes générations confondues”. Cette comparaison nous indiquera dans quelle mesure les provisions sont “correctement” réparties par génération de contrats.

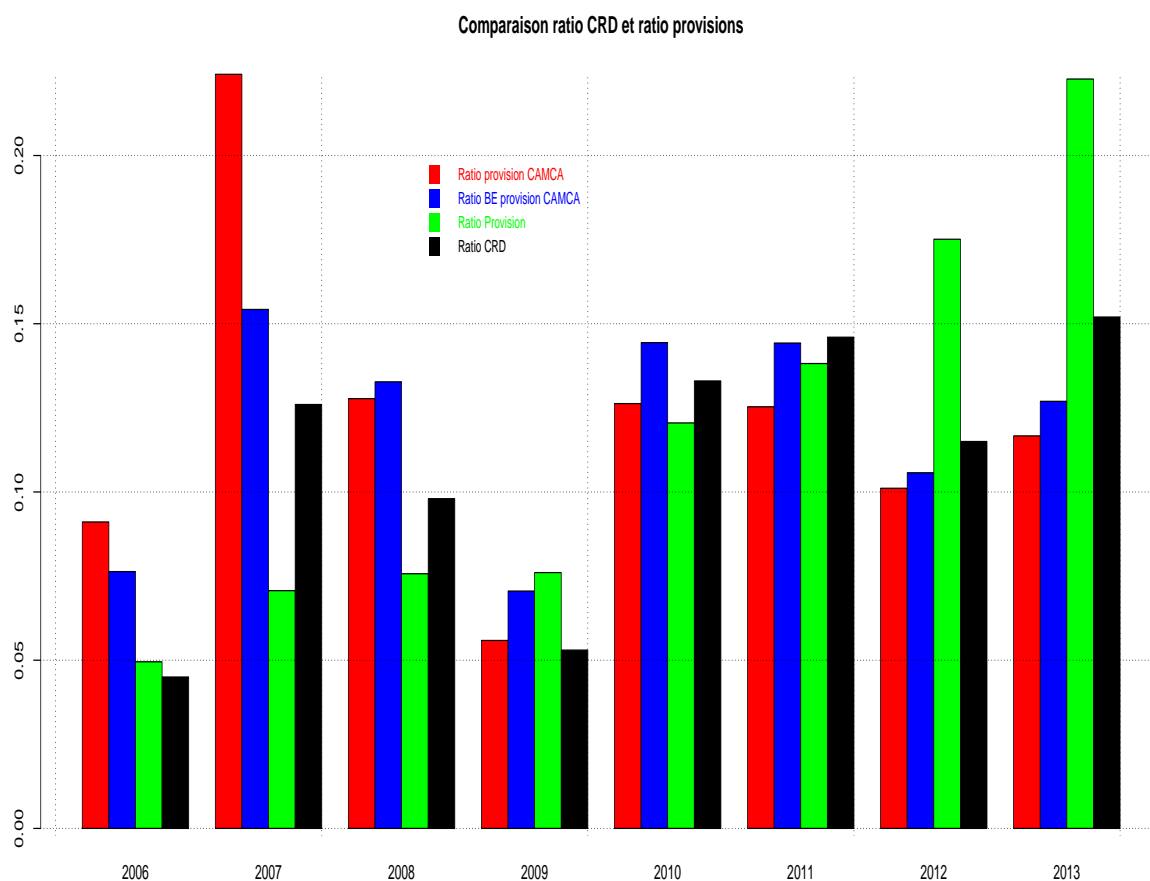


FIGURE 5.31 – Ratio CRD Vs Ratio espérance en fonction des méthodes de provisionnement.

Quelques observations :

1. Comme évoqué précédemment nous constatons que seul notre modèle sur-provisionne les deux dernières générations (i.e. 2012 et 2013) : ceci intéresse la compagnie d'assurances.
2. En revanche les provisions statutaires et le BE de ces provisions prennent en compte, contrairement à notre modélisation, le caractère soit disant plus risqué des générations 2008 et antérieures.
3. Pour les générations 2009 à 2011 les ratios des différentes provisions et du CRD sont semblables.

Comparons pour conclure les ratios de sinistralité de ces différentes méthodes de provisionnement, c'est à dire que nous comparons le ratio "Provisions + montants des sinistres

indemnisés divisé par primes perçues”.

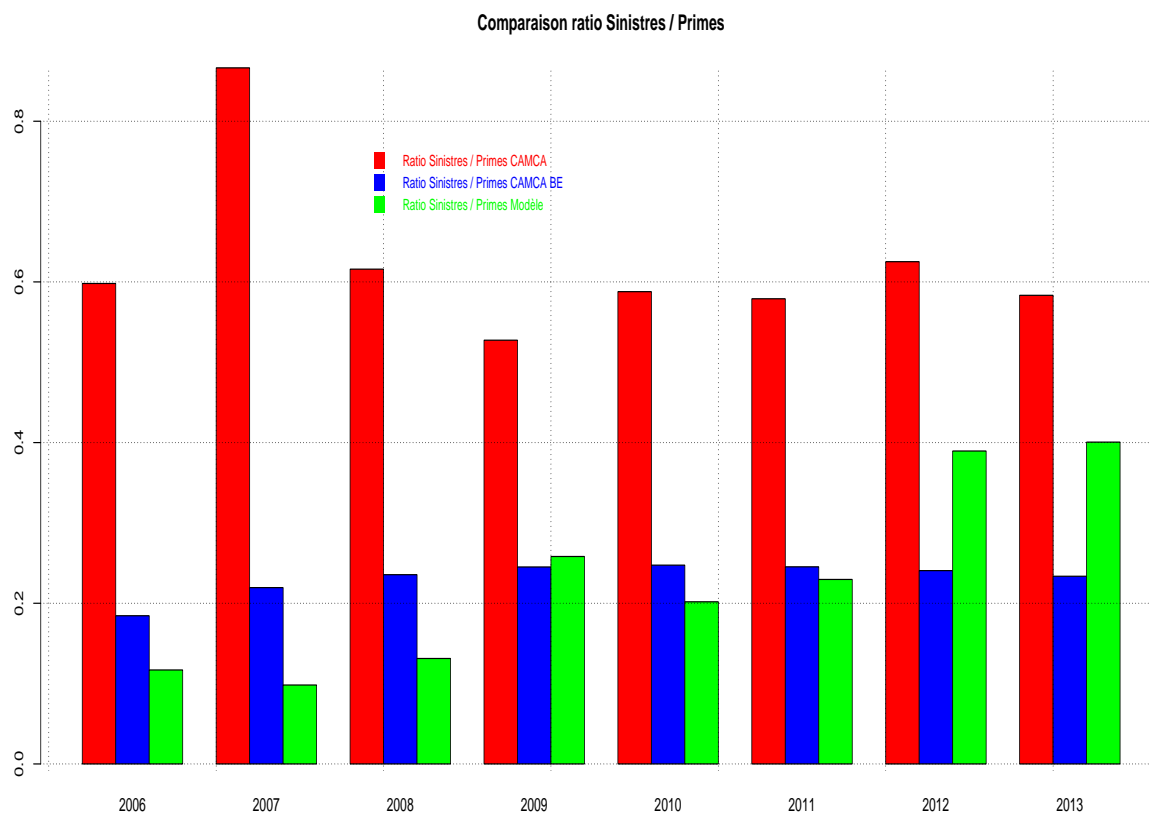


FIGURE 5.32 – Ratio Sinistres / Primes en fonction des méthodes de provisionnement.

Quelques observations :

1. Les taux de sinistralité déterminés via les provisions statutaires dépassent largement les autres taux de sinistralité.
2. Les taux de sinistralité des BE des provisions sont quasi constants pour chaque génération.
3. Les taux de sinistralité déterminé via notre modèle sont beaucoup plus importants pour les générations 2012 et 2013 et ressemblent à ceux calculés via le BE des provisions pour les autres générations.

Notre modèle laisse penser que les générations 2008 et antérieures ne seraient pas plus risquées que les autres ; cela infirmerait l’apriori des praticiens. Le taux de sinistralité connu à la date de calcul renforce cette conclusion.

### 5.4.3 Sensibilité aux paramètres

Dans cette section nous considérons la génération  $i = 0$ . D'après le Théorème 4.3.2 nous avons :

$$\mathbb{E}[P_{t_0}^0] = \lambda_2 \gamma \mathbb{E}[M_1] \int_{A_{t_0+}^0} F_1 \left[ t, \frac{\phi_t(s, \tau)}{\gamma}, \mu, \sigma \right] dt ds. \quad (5.4.11)$$

où  $F_1$  est définie par (4.3.1),  $\mathbb{E}[M_1]$  est l'espérance du montant des prêts et :

$$\phi_t(s, \tau) := \begin{cases} \left[ e^{\tau t^*} (1 + t^* p) - t^* p \right] \frac{e^{\tau s} - e^{\tau(t-t^*)}}{e^{\tau s} - 1}, & \text{si } t^* < t \leq s, \\ \left[ e^{\tau t^*} (1 + t^* p) - t^* p \right], & \text{si } t \leq t^* \text{ ou } s \leq t^*, \end{cases} \quad (5.4.12)$$

Nous noterons :

$$g(\lambda_2, \tau, \mu, \sigma) := \mathbb{E}[P_{t_0}^0], \quad 0 \leq t_0 \leq h.$$

Nous voulons mesurer la sensibilité de l'espérance de la provision suite à une variation des différents paramètres la constituant et indépendamment les uns des autres, c'est à dire mesurer la variation de  $\mathbb{E}(P_{t_0}^0)$  suite à une variation :

1. de l'intensité du PPP :  $\lambda_2$ ,
2. du taux d'emprunt :  $\tau$ ,
3. des paramètres liés à l'évolution des prix de l'immobilier :  $\mu$  et  $\sigma$ .

Nous comparerons également l'évolution de ces différentes sensibilités en fonction de la date d'analyse  $t_0$ .

#### Sensibilité par rapport à l'intensité du PPP

Nous cherchons ici à quantifier l'impact sur l'espérance de la provision d'une variation de l'intensité du PPP : c'est à dire l'impact sur  $\mathbb{E}(P_{t_0}^0)$  d'une variation du nombre moyen de sinistre par unité de temps en supposant que tous les autres paramètres sont constants. Cette sensibilité nous donnera une indication sur la variation de l'espérance de la provision suite à une variation de la sinistralité en terme de fréquence.

**Définition 5.4.1.** *La sensibilité de l'espérance de la provision face une variation de l'intensité du PPP est notée  $S_{\lambda_2}$  ( $S$  pour sensibilité) et est définie par :*

$$S_{\lambda_2} := \frac{\partial g(\lambda_2, \tau, \mu, \sigma)}{\partial \lambda_2}. \quad (5.4.13)$$

Il paraît clair que  $\mathbb{E}[P_{t_0}^0]$  est proportionnelle à  $\lambda_2$ , ainsi :

$$S_{\lambda_2} = \gamma \mathbb{E}[M_1] \int_{A_{t_0+}^0} F_1 \left[ t, \frac{\phi_t(s, \tau)}{\gamma}, \mu, \sigma \right] dt ds. \quad (5.4.14)$$

Nous constatons que  $S_{\lambda_2}$  n'est pas fonction de  $\lambda_2$ , c'est à dire que la variation de  $\mathbb{E}(P_{t_0}^0)$  suite à une variation de  $\lambda_2$  est identique pour toute valeur de  $\lambda_2$ .



### Sensibilité par rapport au taux d'emprunt

Nous voulons ici déterminer comment fluctue l'espérance de la provision pour une fluctuation du taux d'emprunt  $\tau$ .

**Définition 5.4.2.** La sensibilité de l'espérance de la provision face à une variation du taux d'emprunt est notée  $S_\tau$  ( $S$  pour sensibilité) et est définie par :

$$S_\tau := \frac{\partial g(\lambda_2, \tau, \mu, \sigma)}{\partial \tau}. \quad (5.4.15)$$

Soit  $\Psi$  la densité d'une loi normale centrée et réduite. D'après la formule (5.4.11) il vient que :

$$\begin{aligned} S_\tau = & \lambda_2 \mathbb{E}[M_1] \int_0^h \left\{ \int_0^s \left[ \frac{\partial}{\partial \tau} \phi_t(s, \tau) \Phi \left[ \tilde{d} \left( \frac{\phi_t(s, \tau)}{\gamma}, \mu, \sigma, t \right) \right] \right. \right. \\ & \left. \left. + \frac{\frac{\partial}{\partial \tau} \phi_t(s, \tau)}{\sigma \sqrt{t}} \left( \Psi \left[ \tilde{d} \left( \frac{\phi_t(s, \tau)}{\gamma}, \mu, \sigma, t \right) \right] - \frac{\gamma e^{(\mu + \frac{\sigma^2}{2})t}}{\phi_t(s, \tau)} \Psi \left[ \tilde{d} \left( \frac{\phi_t(s, \tau)}{\gamma}, \mu + \sigma^2, \sigma, t \right) \right] \right) \right] dt \right\} ds, \end{aligned} \quad (5.4.16)$$

où :

$$\frac{\partial}{\partial \tau} \phi_t(s, \tau) = \begin{cases} t^*(1 + t^*p) \frac{e^{\tau s} - e^{\tau(t-t^*)}}{e^{\tau s} - 1} + [e^{\tau t^*}(1 + t^*p) - t^*p] \\ \quad \times \frac{[se^{\tau s} - (t - t^*)e^{\tau(t-t^*)}](e^{\tau s} - 1) - se^{\tau s}[e^{\tau s} - e^{\tau(t-t^*)}]}{(e^{\tau s} - 1)^2}, & \text{si } t^* < t \leq s, \\ t^* e^{\tau t^*} (1 + t^*p), & \text{si } t \leq t^* \text{ ou } s \leq t^*, \end{cases} \quad (5.4.17)$$

où la fonction  $\phi$  est définie par (5.4.12).

Traçons la courbe  $(\tau, S_\tau)$  pour  $\tau \in [\frac{\hat{\tau}}{2}; 2\hat{\tau}]$  et pour différentes valeurs de  $t_0$  :

- La courbe verte représente l'évolution de la sensibilité en fonction de  $\tau$  pour la date d'analyse  $t_0 = 0$ ,
- La courbe bleue représente l'évolution de la sensibilité en fonction de  $\tau$  pour la date d'analyse  $t_0 = 5$ ,
- La courbe rouge représente l'évolution de la sensibilité en fonction de  $\tau$  pour la date d'analyse  $t_0 = 10$ ,
- La courbe noire représente l'évolution de la sensibilité en fonction de  $\tau$  pour la date d'analyse  $t_0 = 15$ .

La variable  $\tau$  est en abscisse et la valeur de la sensibilité en ordonnée. Enfin la courbe verticale jaune représentera la valeur des différentes sensibilités pour la valeur de  $\tau$  estimée dans la Section 5.4.2, i.e.  $\tau = 0,07$ .

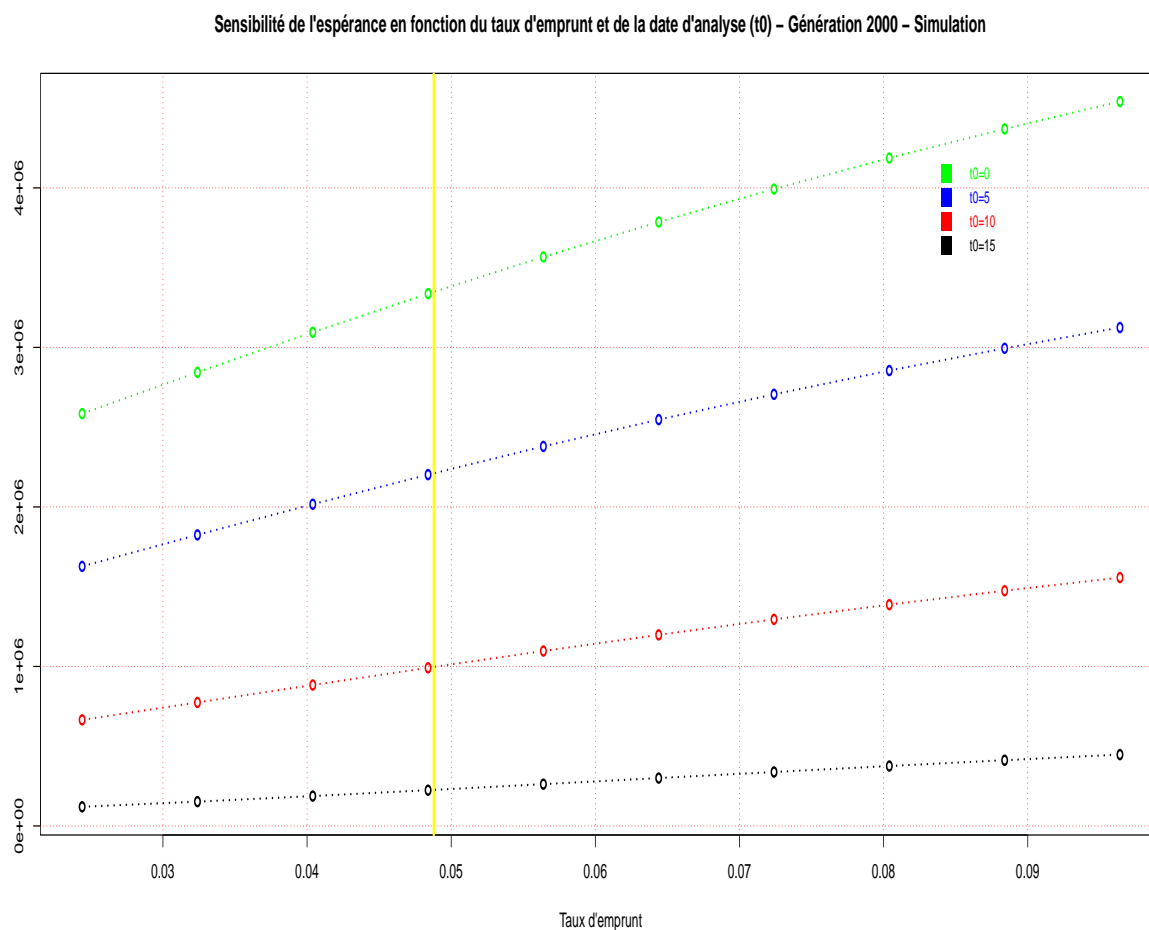


FIGURE 5.33 – Sensibilité de l'espérance face au taux d'intérêt.

Quelques constatations :

1. Comme attendu, la sensibilité de l'espérance par rapport à  $\tau$  est positive, c'est à dire que l'espérance de la provision est une fonction croissante par rapport à  $\tau$ .
2. Nous constatons également que la sensibilité est positive, c'est à dire que l'espérance de la provision est une fonction convexe par rapport à  $\tau$  ; il s'agit du moins d'une "hypothèse" suggérée par la figure.
3. D'autre part l'influence du paramètre  $\tau$  diminue en fonction de la date d'analyse  $t_0$ .

### Sensibilité par rapport aux paramètres liés à la variation des prix de l'immobilier

Nous cherchons ici à quantifier l'impact sur l'espérance de la provision d'une variation des paramètres  $\mu$  et  $\sigma$  en supposant que tous les autres paramètres sont constants. Ces

deux sensibilités nous donnerons une indication de la variation de  $\mathbb{E}(P_{t_0}^0)$  suite à une variation des prix de l'immobilier, i.e. suite à une variation de la sinistralité en terme de montant. Définissons ces deux sensibilités :

**Définition 5.4.3.** *La sensibilité de l'espérance de la provision suite à une variation de  $\mu$  (respectivement  $\sigma$ ) est notée  $S_\mu$  (respectivement  $S_\sigma$ ) et est définie par :*

$$S_\mu := \frac{\partial g(\lambda_2, \tau, \mu, \sigma)}{\partial \mu}, \quad (5.4.18)$$

$$S_\sigma := \frac{\partial g(\lambda_2, \tau, \mu, \sigma)}{\partial \sigma}. \quad (5.4.19)$$

D'après la formule (5.4.11) il vient que :

$$\begin{aligned} S_\mu = & -\lambda_2 \mathbb{E}[M_1] \int_0^h \left\{ \int_0^s \frac{\sqrt{t} \phi_t(s, \tau)}{\sigma} \Psi \left[ \tilde{d} \left( \frac{\phi_t(s, \tau)}{\gamma}, \mu, \sigma, t \right) \right] \right. \\ & \left. + \gamma e^{(\mu + \frac{\sigma^2}{2})t} \left[ t \Phi \left[ \tilde{d} \left( \frac{\phi_t(s, \tau)}{\gamma}, \mu + \sigma^2, \sigma, t \right) \right] - \frac{\sqrt{t}}{\sigma} \Psi \left[ \tilde{d} \left( \frac{\phi_t(s, \tau)}{\gamma}, \mu + \sigma^2, \sigma, t \right) \right] \right] dt \right\} ds, \end{aligned} \quad (5.4.20)$$

$$\begin{aligned} S_\sigma = & -\lambda_2 \mathbb{E}[M_1] \int_0^h \left\{ \int_0^s \frac{\phi_t(s, \tau)}{\sigma} \tilde{d} \left( \frac{\phi_t(s, \tau)}{\gamma}, \mu, \sigma, t \right) \Psi \left[ \tilde{d} \left( \frac{\phi_t(s, \tau)}{\gamma}, \mu, \sigma, t \right) \right] \right. \\ & \left. + \gamma e^{(\mu + \frac{\sigma^2}{2})t} \left[ \sigma t \Phi \left[ \tilde{d} \left( \frac{\phi_t(s, \tau)}{\gamma}, \mu + \sigma^2, \sigma, t \right) \right] - \left( 2\sqrt{t} + \frac{\tilde{d} \left( \frac{\phi_t(s, \tau)}{\gamma}, \mu + \sigma^2, \sigma, t \right)}{\sigma} \right) \right. \right. \\ & \left. \left. \times \Psi \left[ \tilde{d} \left( \frac{\phi_t(s, \tau)}{\gamma}, \mu + \sigma^2, \sigma, t \right) \right] \right] dt \right\} ds, \end{aligned} \quad (5.4.21)$$

où rappelons que  $\Psi$  est la densité de la loi normale centrée et réduite.

Nous traçons ici les courbes  $(\mu, S_\mu)$  et  $(\sigma, S_\sigma)$  pour  $\mu \in [-1; 1]$ ,  $\sigma \in [0; 1]$  et pour différentes valeurs de  $t_0$  :

- Les courbes vertes représentent l'évolution de la sensibilité en fonction de  $\mu$  et de  $\sigma$  pour la date d'analyse  $t_0 = 0$ ,
- Les courbes bleues représentent l'évolution de la sensibilité en fonction de  $\mu$  et de  $\sigma$  pour la date d'analyse  $t_0 = 5$ ,
- Les courbes rouges représentent l'évolution de la sensibilité en fonction de  $\mu$  et de  $\sigma$  pour la date d'analyse  $t_0 = 10$ ,
- Les courbes noires représentent l'évolution de la sensibilité en fonction de  $\mu$  et de  $\sigma$  pour la date d'analyse  $t_0 = 15$ .

Le paramètre  $\mu$  (resp.  $\sigma$ ) est en abscisse et la sensibilité en ordonnée. Enfin la courbe verticale jaune représentera la valeur des différentes sensibilités pour les valeurs de  $\mu$  et  $\sigma$  estimées dans la Section 5.4.2, i.e.  $\mu = -0,07$  et  $\sigma = 0,1958$ .

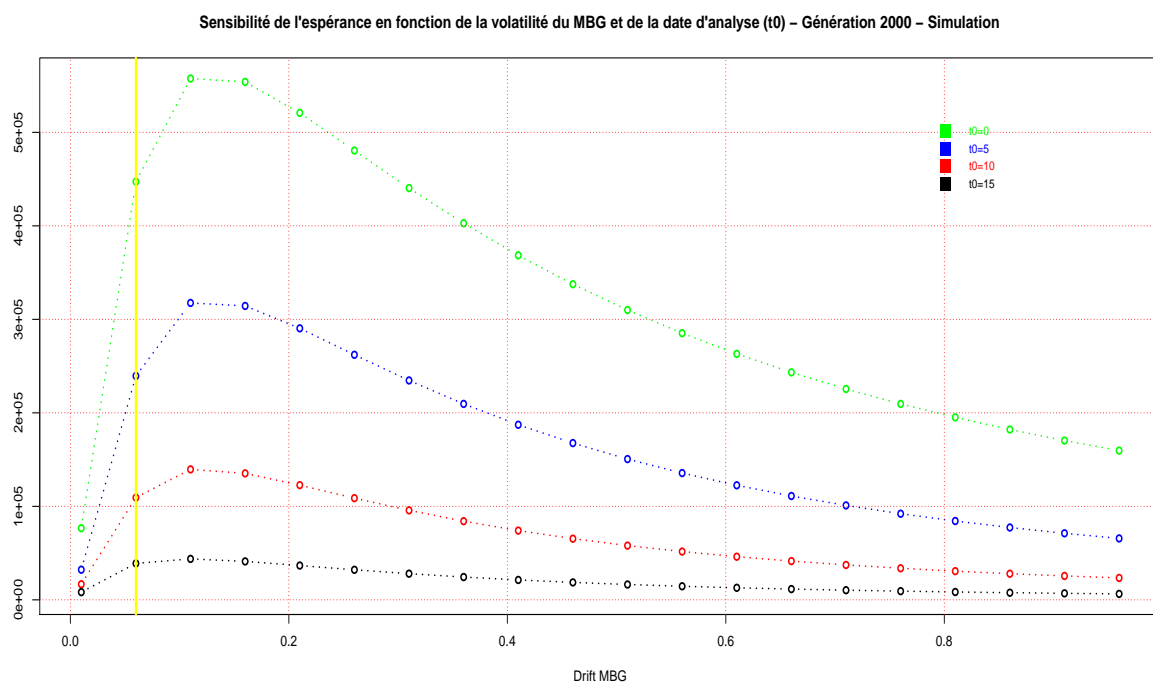
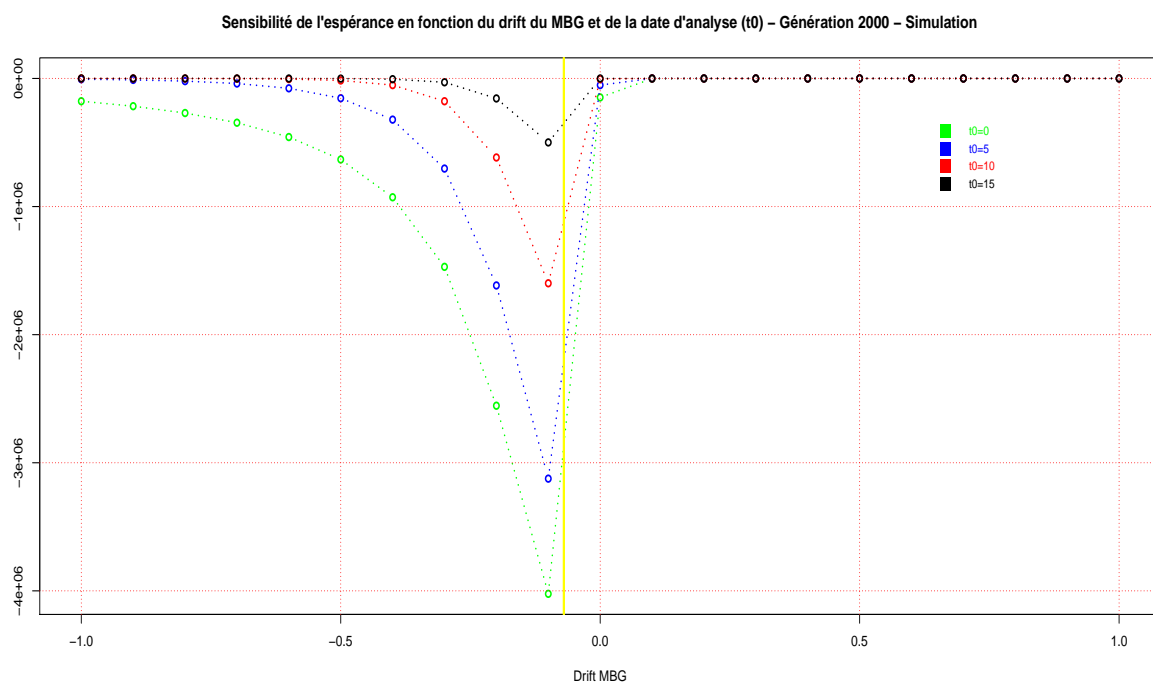


FIGURE 5.34 – Sensibilité de l'espérance face aux paramètres  $\mu$  et  $\sigma$ .

De la Figure 5.34 nous pouvons en déduire :

1. Comme attendu, la sensibilité de l'espérance de la provision par rapport au paramètre  $\mu$  est négative. Cela signifie que l'espérance de la provision décroît lorsque  $\mu$  croît. En effet le paramètre  $\mu$  représente la tendance des prix de l'immobilier : lorsque ces prix augmentent, le recours sur le bien financé est plus important et par conséquent le montant du sinistre est plus faible. Nous observons également que l'espérance est une fonction concave de  $\mu$ . D'autre part l'influence du paramètre  $\mu$  diminue en fonction de la date d'analyse  $t_0$ .
2. La sensibilité de l'espérance de la provision par rapport au paramètre  $\sigma$  est positive. Cela signifie que l'espérance de la provision croît quand  $\sigma$  croît. Cette sensibilité diminue en fonction de la date d'analyse  $t_0$ . Nous observons également que l'espérance est une fonction convexe de  $\sigma$ .
3. A l'heure actuelle nous ne pouvons expliquer la rupture observée pour  $\mu \sim -0,1$  dans la Figure 5.34.

### Comparaison des différentes sensibilités

Nous voulons ici comparer l'impact des différents paramètres sur l'espérance de la provision en fonction de la date d'analyse  $t_0$ . Nous traçons pour cela les courbes  $(t_0, S_k)$  où :  $k \in \{\hat{\lambda}_2, \hat{\tau}, \hat{\mu}, \hat{\sigma}\}$ . Ces courbes représentent l'évolution en fonction de la date  $t_0$  des différentes sensibilités.

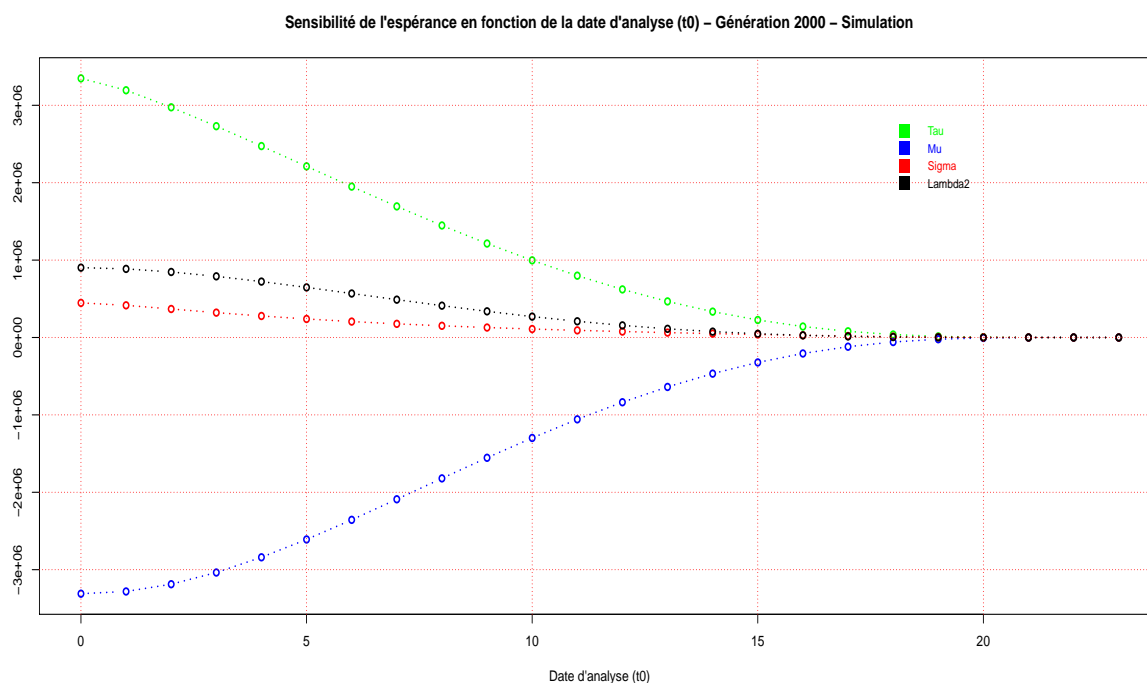


FIGURE 5.35 – Comparaison des différents sensibilités en fonction de  $t_0$ .

Nous constatons différents éléments :

1. Seul le paramètre  $\mu$  permet de diminuer l'espérance de la provision. L'espérance de la provision augmente lorsque les paramètres  $\tau$ ,  $\sigma$  et  $\lambda_2$  augmentent.
2. Nous constatons également que le taux d'emprunt  $\tau$  a le plus d'impact sur l'espérance de la provision.
3. D'autre part et compte tenu de nos estimations, la variabilité du mouvement Brownien géométrique, i.e. la variabilité de la variation des prix de l'immobilier, n'a que peu d'influence sur l'espérance de la provision.

#### 5.4.4 Probabilité de nullité de la charge d'un sinistre

Il est intéressant de regarder la probabilité, conditionnellement à  $[(T, D) = (t, d)]$ , que la charge d'un sinistre soit nulle. Notons  $p(t, d, t^*)$  cette probabilité, i.e. :

$$p(t, d, t^*) := \mathbb{P}\left[Mf(T, D, R_T) = 0 \mid (D, T) = (d, t)\right], \quad (5.4.22)$$

où la fonction  $f$  est définie par (4.2.16) (page 39). Notons que la probabilité  $p$  dépend de l'unité de temps  $t^*$  car la fonction  $f$  dépend implicitement de  $t^*$ . Nous analyserons plus loin l'évolution de cette probabilité en fonction des valeurs du ratio  $\frac{t}{d}$ .

Par définition de la fonction  $f$  il vient que :

$$\begin{aligned} Mf(t, d, R_t) = 0 &\Leftrightarrow M[\phi_t(d) - R_t\gamma]_+ = 0 \\ &\Leftrightarrow \phi_t(d) - R_t\gamma \leq 0 \\ &\Leftrightarrow R_t \geq \frac{\phi_t(d)}{\gamma} \\ &\Leftrightarrow B_t \geq \frac{\ln\left[\frac{\phi_t(d)}{\gamma}\right] - \mu t}{\sigma t}. \end{aligned} \quad (5.4.23)$$

Par définition d'un mouvement Brownien il vient que :

$$p(t, d, t^*) = 1 - \Phi\left[\frac{\ln\left[\frac{\phi_t(d)}{\gamma}\right] - \mu t}{\sigma\sqrt{t}}\right], \quad (5.4.24)$$

où  $\phi$  est définie par (4.2.13) et  $\Phi$  est la f.d.r. d'une v.a. gaussienne standard.

Nous souhaitons maintenant voir comment évolue la probabilité  $p(t, d, t^*)$  en fonction de différentes valeurs du ratio :

$$\frac{t}{d} =: \beta, \quad (5.4.25)$$

i.e. en fonction des différents étapes de la vie d'un prêt.

Pour ce faire nous traçons la courbe  $\left(\beta, \min_d(p(\beta d, d, t^*))\right)$  pour le choix de paramètres expliqué dans la Section 5.4.2 :

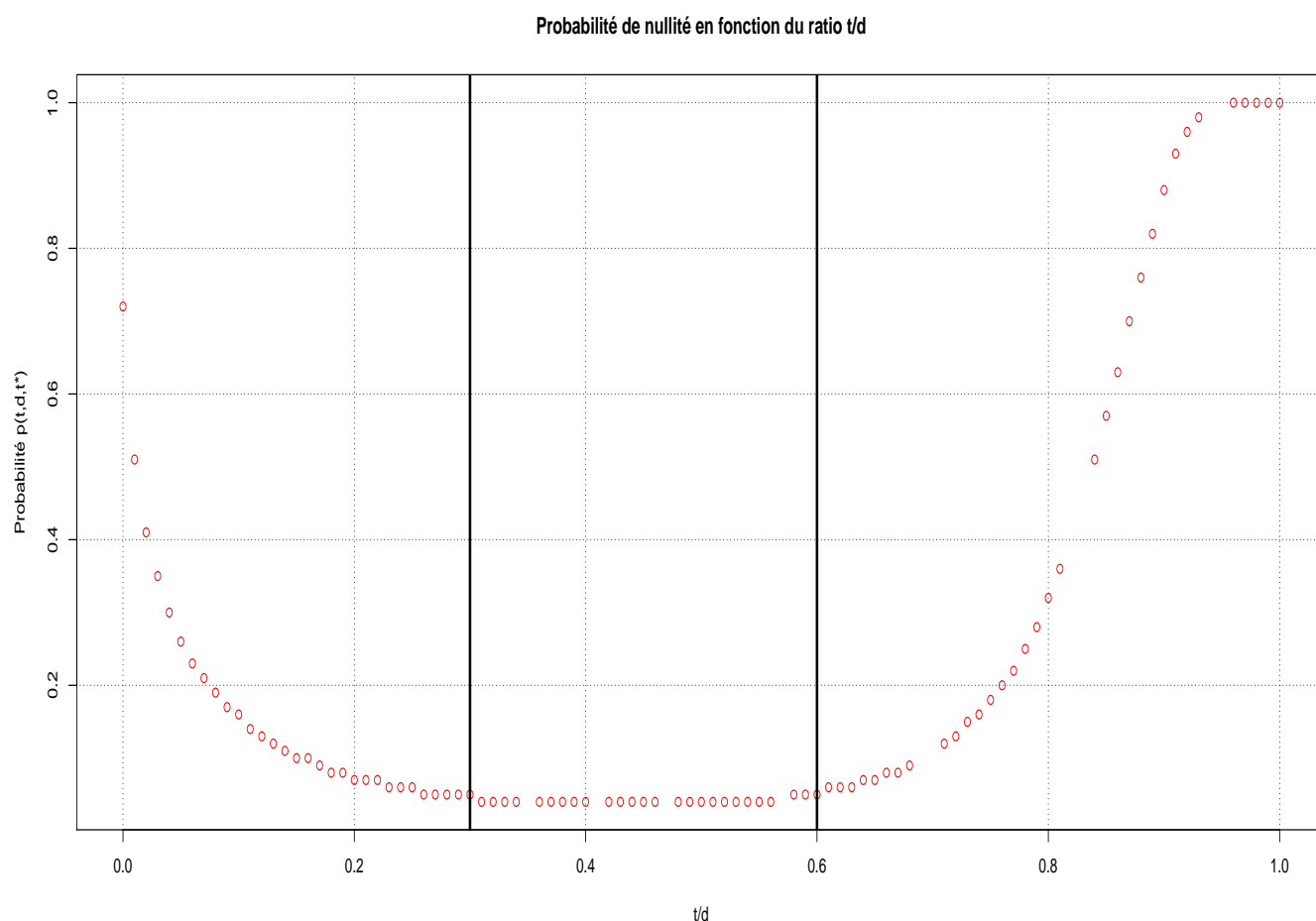


FIGURE 5.36 – Probabilité de nullité de la charge d'un sinistre.

Le graphique précédent nous renseigne sur deux éléments importants.

1. Si un sinistre est déclaré au début de la vie d'un prêt alors le montant de ce sinistre a de fortes chances d'être nul. Cette observation confirme l'a priori des praticiens. L'explication avancée par les praticiens est que si un sinistre est déclaré au début de la vie du prêt alors le montant dû par l'emprunteur est inférieur au montant de la revente du bien. Expliquons de manière plus précise ce point. Considérons un sinistre déclaré à la date 0, alors :

$$\lim_{t^* \rightarrow 0} Mf(0, D, R_0) = M[\alpha - (1+a)(1-c)]_+, \quad (5.4.26)$$

où rappelons que le coefficient  $a$  représente l'apport personnel des emprunteurs et le coefficient  $c$  représente les coûts liés à la revente. Donc si  $\frac{1}{c} - \frac{1}{a} \geq \alpha$  alors le montant du sinistre déclaré à la date 0 est nul. Compte tenu de nos estimations

des paramètres  $a$  et  $c$  il vient que :

$$(1 + a)(1 - c) = 1,17 \text{ et } \alpha = 1,15, \quad (5.4.27)$$

donc :

$$\lim_{t^* \rightarrow 0} Mf(0, D, R_0) = 0. \quad (5.4.28)$$

2. D'autre part nous constatons que la probabilité  $p(t, d, t^*)$  :

- Décroit au début de la vie des prêts,
- Reste proche de 0 entre les ratios 0,3 et 0,6,
- Augmente à partir du ratio 0,6 pour atteindre une probabilité proche de 1 pour le ratio 1.

Cette observation est également confirmée par les praticiens. Une des raisons avancées est qu'au début de la vie des prêts l'apport personnel suffit à annuler le montant du sinistre (c'est pourquoi  $p(t, d, t^*)$  est élevé au début de la vie des prêts). D'autre part plus un prêt avance dans le temps, moins le montant restant à rembourser est important, c'est pourquoi  $p(t, d, t^*)$  augmente à partir du ratio 0,6.

#### 5.4.5 Fonction “nleqslv” du logiciel $\mathbb{R}$

Dans le Théorème 4.4.2 de la Section 4.4 nous avons montré que  $\lambda_1, \lambda_2$  et  $h$  pouvaient être estimés en résolvant un système de trois équations non linéaires de la forme :

$$(S) : \begin{cases} x \in \mathbb{R}^N, \\ g(x) = 0_{\mathbb{R}^N}, \end{cases} \quad (5.4.29)$$

$$(5.4.30)$$

où  $g \in C(\mathbb{R}^N, \mathbb{R}^N)$ .

Les deux principales méthodes numériques de résolution des systèmes non linéaires sont les méthodes de point fixe et les méthodes de Newton. La plus simple est la première. Dans cette méthode la fonction  $f$  est définie par :

$$\begin{aligned} f : (E, d) &\longrightarrow (E, d) \\ x &\longrightarrow x + g(x), \end{aligned}$$

où  $E$  est un espace métrique muni d'une distance  $d$ . Nous remarquons que si  $f(x) = x$  alors  $g(x) = 0$  (et inversement). Si la fonction  $f$  est contractante, i.e. si :

$$\exists k \in ]0; 1[ : d(f(x), f(y)) \leq kd(x, y), \quad \forall x, y \in E.$$

alors il est possible de démontrer que :

$$\exists ! \bar{x} \in E : f(\bar{x}) = \bar{x}.$$

De plus si  $x_0 \in E$  et  $x_{n+1} = f(x_n)$  alors la suite  $(x_n)_{n \geq 1}$  converge vers  $\bar{x}$ .



Si la fonction  $f$  n'est pas contractante il est possible d'utiliser la méthode du point fixe de relaxation (Cf. [5] pour plus de détails) mais la vitesse de convergence est linéaire.

Les méthodes de Newton répondent à cet inconvénient. L'algorithme de résolution dans cette méthode est le suivant :

---

**Algorithme 5.2** Algorithme de Newton.

---

- 1: **for**  $n \geq 1$  **do**
  - 2:   Initialisation de  $x_0 \in E$ .
  - 3:   Calculer :  $Dg(x_n)$ .
  - 4:   Inverser :  $Dg(x_n)(x_{n+1} - x_n) = g(x_n)$ .
  - 5: **end for**
- 

où  $Dg(x) \in \mathbb{M}_N(\mathbb{R})$  est la matrice représentant la différentielle de  $g$  en  $x$ .

La convergence de cet algorithme est quadratique. Cependant à chaque itération il faut :

1. Calculer  $Dg(x_n)$ ,
2. Résoudre le système linéaire  $Dg(x_n)(x_{n+1} - x_n) = g(x_n)$ .

Le coût de calcul est donc très important. Pour pallier à cet inconvénient plusieurs solutions existent telle que la méthode du "Faux quasi-Newton", la méthode de "Newton incomplet", la méthode de la sécante (dans le cas  $N = 1$ ) et les méthodes du type "quasi-Newton" (généralisation de la méthode de la sécante au cas  $N > 1$ ) avec en particulier la méthode de Broyden. Cette dernière méthode est basée sur l'algorithme suivant :


---


**Algorithme 5.3** Algorithme de Broyden.

---

- 1: **for**  $n \geq 1$  **do**
  - 2:   Initialisation de  $x_0 \in E$ ,  $x_1 \in E$  et  $B_0 \in \mathbb{M}_N(\mathbb{R})$ .
  - 3:   Calculer :  $\delta_n = x_n - x_{n-1}$  et  $y_n = g(x_n - x_{n-1})$ .
  - 4:   Calculer :  $B_n = B_{n-1} + \frac{y_n - B_{n-1}\delta_n}{\delta_n^t \delta_n} (\delta_n) \delta_n^t$ .
  - 5:   Résoudre :  $B_n(x_{n+1} - x_n) = -g(x_n)$ .
  - 6: **end for**
- 

Il s'agit d'un algorithme à deux pas dont la vitesse de convergence n'est plus quadratique.

La fonction "nleqslv" du package "nleqslv" sous le logiciel  permet de résoudre les systèmes non linéaires du type  $(S)$  en utilisant deux méthodes :

1. La méthode de Newton : si les dérivés de  $g$  ne sont pas renseignés,  procède à leur calcul en utilisant des méthodes numériques.
2. La méthode de Broyden.

Nous avons utilisé la méthode de Broyden pour la résolution du système présenté dans le Théorème 4.4.2.



## Troisième partie

# Besoin de solvabilité lié au risque de provisionnement

---

|          |   |            |
|----------|---|------------|
| <b>6</b> | <b>Besoin de solvabilité lié au risque de provisionnement dans un contexte d'ORSA</b> . . . . . | <b>111</b> |
| 6.1      | Solvabilité 2 . . . . .   | 111        |
| 6.2      | Besoin de solvabilité lié au risque de provisionnement . . . . .                                | 114        |
| 6.3      | Preuves techniques . . . . .  | 130        |
| <b>7</b> | <b>Compléments numériques au calcul du besoin de solvabilité</b> . . . . .                      | <b>139</b> |
| 7.1      | Espérances asymptotique et théorique . . . . .  | 139        |
| 7.2      | Distribution empirique de la provision . . . . .  | 142        |
| 7.3      | Besoin de solvabilité empirique . . . . .   | 147        |

---



# Besoin de solvabilité lié au risque de provisionnement dans un contexte d'ORSA

---

Nous nous plaçons du point de vue d'un assureur devant indemniser une banque à la suite de pertes liées à un défaut de remboursement de ses emprunteurs.

Le principal objectif de ce chapitre est de calculer le besoin de solvabilité associé au risque de provisionnement dans un contexte d'ORSA (Own Risk Solvency Assessment) imposé par la réforme Solvabilité 2 (Cf. Section 2.1.2). Pour ce faire nous nous baserons sur la modélisation individuelle de la provision introduite dans le Chapitre 4. Notre méthode, contrairement aux autres existantes, tient compte, dans l'évaluation du besoin de solvabilité, des caractéristiques des emprunteurs et de la dépendance entre les montants de sinistres via un risque systémique (immobilier). D'un point de vue mathématique nous étudierons la convergence en loi d'une somme, d'un nombre de terme aléatoire, de variables dépendantes et de lois différentes.

**Mots clés :** *Solvabilité 2 ; ORSA ; Besoin de solvabilité ; Risque de provisionnement ; Quantile ; Convergence en loi ; Perpetual integral functional of Brownian motion.*

Un traduction de ce chapitre, en anglais, est disponible page 261.

## 6.1 Solvabilité 2

### 1) Introduction :

Comme expliqué dans la Section 2.1.2, la réforme Solvabilité 2 a pour objectif de mieux protéger et de mieux informer les souscripteurs d'assurance. Cette réforme, similaire à la réforme engagée dans le secteur bancaire Bâle 2/3 (Cf. [27]), entend, entre autres choses, déterminer la perte associée aux différents risques auxquels sont soumises les sociétés d'assurances. L'Article 45 de la Directive [26] précise en effet que :

*“Dans le cadre de son système de gestion des risques, chaque entreprise d'assurance et de réassurance procède à une évaluation interne des risques et de la solvabilité. Cette évaluation porte au moins sur les éléments suivants :*

1. *Le besoin global de solvabilité, compte tenu du profil de risque spécifique, des limites approuvées de tolérance au risque et de la stratégie commerciale de l'entreprise ;*
2. *...”*

Cette évaluation est connue en pratique sous le nom de “dispositif ORSA” (Own Risk Solvency Assesment) et intervient au sein du second pilier de la réforme Solvabilité 2.

Nous nous concentrerons dans ce chapitre sur l'évaluation du risque lié à un mauvais provisionnement que nous appellerons, conformément à l'Article 45, *Besoin de Solvabilité* (BS).

Nous considérons le cas particulier d'une compagnie d'assurances devant indemniser une banque à la suite de pertes liées aux défauts de paiement de ses emprunteurs et pour simplifier les notations nous évaluerons le BS à la date de souscription des contrats.

2) *Méthodes existantes et plan du chapitre :*

Le calcul du BS nécessite de définir un indicateur de risque, i.e. une mesure de risque, ainsi qu'un horizon sur lequel il s'appliquera : ces deux éléments ne sont pas précisés par la Directive. Le calcul du BS diffère du calcul du capital de solvabilité requis (CSR) puisque dans le calcul du CSR l'indicateur de risque et l'horizon sont définis. En effet l'Article 101 de la Directive [26] précise que :

*“Le capital de solvabilité requis correspond à la valeur en risque (Value-at-Risk) [...] avec un niveau de confiance de 99,5% à l'horizon d'un an.”*

Le BS a pour objectif de pallier aux défauts présents dans le calcul du CSR. Comme précisé dans [41], la durée d'un an dans le calcul du CSR semble mal adaptée pour les assurances proposant des garanties sur le long terme : évaluer la solvabilité sur un an pour des compagnies d'assurances ayant des engagements sur plusieurs décennies semble inapproprié. Contrairement à [58], où le BS est défini comme le maximum des pertes pour chaque exercice futur, nous définirons le BS sur la durée maximale des contrats considérés (Cf. Définition 6.2.2 et Remarque 6.2.3). La mesure de risque considérée sera la Value At Risk à 99,5% (Cf. [24] pour une description des mesures de risques usuelles). A l'instar de [16], et pour simplifier les calculs, nous ne considérons pas d'actualisation ni de marge de risque dans ce chapitre (Cf. [47] pour la prise en compte de ces éléments).

Pour évaluer le BS nous utiliserons la modélisation individuelle de la provision introduite dans le Chapitre 4 et rappelée dans la Section 6.2.1. Contrairement à [46], cette modélisation individuelle de la provision permet de prendre en compte la dépendance entre les individus via un facteur de risque systémique (immobilier) et également de prendre en compte la dépendance entre les montants de sinistres et le processus de comptage. Cf. la Section 2.1.2, [8], [13], [15], [62] et références incluses pour plus de détails sur la prise en compte de ce type de dépendance.

Nous verrons que pour évaluer le BS, l'étape clé est de déterminer le quantile de la provision, i.e. d'inverser la fonction de répartition (f.d.r.) de la provision. Déterminer la f.d.r. théorique de la provision étant impossible, nous approximerons cette f.d.r. avec

l'étude asymptotique de la provision lorsque la durée maximale des prêts devient grande (Cf. Théorème 6.2.5). Nous donnerons la preuve de ce résultat, avec quelques résultats intermédiaires intéressants, dans la Section 6.2.3. Ce résultat de convergence nous permettra de donner une approximation du BS dans le Théorème 6.2.6.

Afin d'alléger la présentation des résultats, les notations du Chapitre 4 seront légèrement modifiées. Les preuves techniques seront données dans la Section 6.3.

3) *Résultat principal du chapitre :*

La v.a. mesurant le risque de provisionnement est définie comme la différence entre la provision et sa moyenne (puisque c'est cette espérance que la compagnie d'assurances enregistre dans ses comptes). Donc à la date 0 le BS, concernant les prêts souscrits le mois  $i = 0$ , est défini par :

$$BS(h) := \mathcal{Q}_{P(h) - \mathbb{E}[P(h, R)]}(99, 5\%), \quad (6.1.1)$$

où  $h$  est la durée maximale des prêts, la v.a.  $P(h, R)$  représente la provision calculée à la date 0 et concernant les prêts souscrits le mois 0,  $\mathbb{E}[P(h, R)]$  est la moyenne de  $P(h, R)$  et  $\mathcal{Q}_{P(h, R) - \mathbb{E}[P(h, R)]}$  est le quantile de  $P(h, R) - \mathbb{E}[P(h, R)]$ . Rappelons que (Cf. (6.2.1)) :

$$P(h, R) := \sum_{j \geq 1} \mathbb{1}_{\{(T_j, D_j) \in A_h\}} M_j f(T_j, D_j, R_{T_j}), \quad (6.1.2)$$

avec  $(T_j, D_j)_{j \geq 1}$  un PPP représentant l'ensemble de points (Date de fin du prêt, Durée du prêt),  $A_h$  l'ensemble où un défaut apparait possible (Cf. (6.2.3)),  $(M_j)_{j \geq 1}$  une suite de v.a. i.i.d. représentant les montants empruntés, la fonction  $f$  définie par (6.2.4), et  $R := (R_t, t \geq 0)$  représentant les fluctuations des prix de l'immobilier entre les dates 0 et  $t$ . Le processus  $(R_t)_{t \geq 0}$  est le mouvement Brownien géométrique défini par (6.2.7). Nous supposons aussi que les v.a.  $(T_j, D_j)_{j \geq 1}$ ,  $(M_j)_{j \geq 1}$  et le processus  $(R_t)_{t \geq 0}$  sont indépendants.

Nous voyons à travers la formule (6.1.1) que le calcul de  $BS(h)$  nécessite l'inversion de la fonction de répartition de  $P(h, R) - \mathbb{E}[P(h, R)]$ . En pratique  $h = 276$  mois, ce qui est considéré comme une période "longue". D'un point de vue mathématique, nous considérons que nous sommes en régime asymptotique, i.e.  $h \rightarrow +\infty$ . Dans le Théorème 6.2.5 nous montrerons que la provision, correctement renormalisée, converge en loi vers un mélange de Gaussienne et d'une intégrale d'une fonction d'un mouvement Brownien géométrique, i.e. :

$$\frac{P(h, R) - \mathbb{E}[P(h, R)]}{h\theta_1} \xrightarrow[h \rightarrow +\infty]{\mathcal{L}} G - \int_0^{+\infty} \inf(\theta_2, \theta_3 R_t) dt + \theta_4, \quad (6.1.3)$$

où  $\theta_1, \theta_2, \theta_3$  et  $\theta_4$  sont des constantes qui s'expriment à l'aide des paramètres de notre modèle et  $G$  est une variable aléatoire gaussienne centrée, réduite et indépendante de

$(R_t)_{t \geq 0}$ . Par conséquent, nous approximerons  $BS(h)$  par la quantité  $BS^*(h)$  (Cf. Théorème 6.2.6 pour plus de détails) définie par :

$$BS^*(h) := h\theta_1 \mathcal{Q}(99,5\%), \quad (6.1.4)$$

où  $\mathcal{Q}(99,5\%)$  est le quantile, à l'ordre 99,5%, de :

$$G - \int_0^{+\infty} \inf(\theta_2, \theta_3 R_t) dt + \theta_4. \quad (6.1.5)$$

La formule (6.1.4) est seulement théorique. En effet la loi de  $\int_0^{+\infty} \inf(\theta_2, \theta_3 R_t) dt$  est inconnue. Nous discuterons cependant dans la Section 7.3 de la possibilité de simuler la provision et d'en déduire une valeur numérique du besoin de solvabilité.

Notons que s'il n'y avait pas de corrélation entre les montants de chaque sinistre, le membre de gauche de l'expression (6.1.3) convergerait vers une variable aléatoire gaussienne centrée et réduite. Le terme additionnel qui s'exprime à l'aide d'une intégrale du mouvement Brownien géométrique prend en compte les corrélations. Il y a donc une "séparation des sources de risques", l'une provenant des caractéristiques individuelles des emprunteurs qui sont indépendants et l'autre provenant du coût de l'immobilier qui affecte tous les emprunteurs.

## 6.2 Besoin de solvabilité lié au risque de provisionnement

Dans ce qui suit nous raisonnons à temps continu. Le temps  $t$  est défini comme une fraction d'année et  $t^*$  représente un mois.

Dans la Section 6.2.1 nous commencerons par rappeler le contexte du provisionnement introduit dans le Chapitre 4. Nous expliquerons les principales hypothèses et notations liées à ce modèle. La Section 6.2.2 sera consacrée au calcul du BS, i.e. à l'obtention de la formule (6.1.4). La preuve du résultat (6.1.3) figurera dans la Section 6.2.3.

### 6.2.1 Provisionnement face au risque de défaut

Une compagnie d'assurances doit indemniser une banque suite aux défauts de paiement de ses emprunteurs. Les emprunts considérés sont des emprunts immobiliers. Le principal objectif pour la compagnie d'assurances est d'évaluer, à la date  $t_0 = 0$ , les montants de sinistres venant des emprunteurs ayant souscrits un prêt le mois 0. Ce montant s'appelle une "provision". Il est important de noter qu'en cas de défaut la banque peut tenter de revendre le bien financé par le prêt.

Dans le Chapitre 4 la provision calculée au temps 0 et concernant les prêts souscrits à la date 0 est définie par (nous modifions légèrement les notations) :

$$P(h, R) := \sum_{j \geq 1} \mathbb{1}_{\{(T_j, D_j) \in A_h\}} M_j f(T_j, D_j, R_{T_j}), \quad (6.2.1)$$



où :

- $D_j$  (resp.  $M_j$ ) est la durée (resp. le montant) du prêt souscrit le mois 0 par l'emprunteur  $j$ . La durée maximale  $h$  (inconnue) des prêts est telle que :

$$D_j \leq h, \quad j \geq 1. \quad (6.2.2)$$

- $T_j$  représente la date de fin du prêt de l'emprunteur  $j$ . Il est important de distinguer deux possibilités :

1.  $T_j = D_j$  signifie que l'emprunteur a remboursé son prêt normalement jusqu'à la maturité,
2.  $T_j < D_j$  signifie que l'emprunteur a connu un défaut de paiement avant la maturité.

$R := (R_t, t \geq 0)$  représente les fluctuations des prix de l'immobilier entre les dates 0 et  $t$ . L'ensemble  $A_h$  est défini par :

$$A_h := \{(s, t) \in (\mathbb{R}_+)^2 : 0 \leq t < s; s \leq h\}. \quad (6.2.3)$$

Voici une illustration :

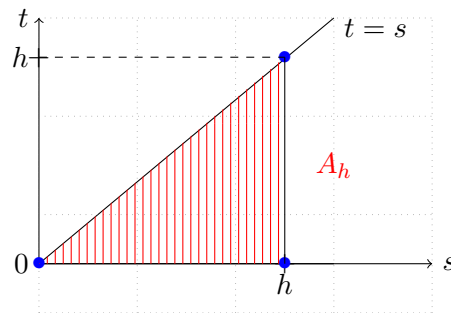


FIGURE 6.1 – Ensemble  $A_h$ .

- $M_j f(T_j, D_j, R_{T_j})$  est le montant que la compagnie d'assurances reversera à la banque en cas de défaut de paiement de l'emprunteur  $j$  à la date  $T_j$ . La fonction  $f$  est définie par :

$$f(t, s, u) := [\phi_t(s) - \gamma u]_+, \quad (6.2.4)$$

où  $0 \leq t \leq s, u > 0, \gamma \in ]0, 1[$  et :

$$\phi_t(s) := \begin{cases} \alpha, & \text{si } t \leq s \leq t^*, \\ \alpha, & \text{si } t \leq t^* \leq s, \\ \alpha \frac{e^{\tau s} - e^{\tau(t-t^*)}}{e^{\tau s} - 1}, & \text{si } 0 < t^* < t \leq s, \end{cases} \quad (6.2.5)$$

$$\alpha := e^{\tau t^*} (1 + t^* p) - t^* p, \quad (6.2.6)$$

avec  $p \in \mathbb{R}_+$  représentant le coefficient de pénalité dû aux intérêts de retard, et  $\tau \in \mathbb{R}_+$  le taux exponentiel des emprunts.

Notons que d'après (6.2.1) et (6.2.4),  $M_j \phi_{T_j}(D_j)$  représente le montant dû par l'emprunteur  $j$  à la date  $T_j$  et  $\gamma M_j R_{T_j}$  est le montant de la revente du bien à la date  $T_j$ . Le montant d'un sinistre ne peut être négatif. Quand le montant des recours est supérieur au montant dû par l'emprunteur, alors la compagnie d'assurances n'indemniserait pas la banque (car en réalité c'est la banque qui collecte le montant des recours). Le montant de sinistres ne représente pas le montant réel de la perte pour la banque mais uniquement le montant qu'elle recevra de la part de la compagnie d'assurances.

Rappelons les principales hypothèses effectuées dans le Chapitre 4 :

**Hypothèse 6.2.1.**

(H<sub>5</sub>) :  $(R_t)_{t \geq 0}$  est le mouvement Brownien géométrique de paramètres  $\mu$  et  $\sigma$  tel que  $R_0 = 1$ , i.e. :

$$R_t := \exp[\sigma B_t + \mu t], \tag{6.2.7}$$

où  $(B_t)_{t \geq 0}$  est un mouvement Brownien standard partant de 0.

(H<sub>6</sub>) : L'ensemble des points  $(T_j, D_j)_{j \geq 1}$  forment un PPP d'intensité  $\Lambda$  :

$$\Lambda := \lambda_1 \text{Leb}|_{C_h} + \lambda_2 \text{Leb}|_{A_h}, \tag{6.2.8}$$

où  $(\lambda_1, \lambda_2) \in (\mathbb{R}_+)^2$ ,  $C_h := \{(t, t) \in (\mathbb{R}_+)^2 : t \leq h\}$ , et  $\text{Leb}|_{C_h}$  (resp.  $\text{Leb}|_{A_h}$ ) est la mesure de Lebesgue restreinte à  $C_h$  (resp.  $A_h$ ).

(H<sub>8</sub>) : Les v.a.  $(T_j, D_j)_{j \geq 1}$ ,  $(M_j)_{j \geq 1}$  et le processus  $(R_t)_{t \geq 0}$  sont indépendants et les v.a.  $(M_j)_{j \geq 1}$  sont i.i.d..

Nous supposons d'autre part que :

$$(H_9) : \quad \mathbb{E}[(M_1)^4] < +\infty \text{ et } M_1 > 0 \text{ p.s..} \tag{6.2.9}$$

Passons maintenant au calcul du besoin de solvabilité.

**6.2.2 Calcul du besoin de solvabilité**

Comme expliqué dans la Section 6.1, le BS correspond au montant que les compagnies d'assurances doivent détenir en fonds propre pour faire face à un sous-provisionnement sur la durée maximale des prêts considérés.

Dans cette section, nous définirons tout d'abord la v.a. mesurant le risque de provisionnement (Cf. Définition 6.2.2). Dans la Remarque 6.2.3 nous évoquerons le modèle proposé par Wüthrich et al. dans [65]. Dans cet article, Wüthrich et al. s'intéressent à deux indicateurs de risque : la variance et l'erreur conditionnelle moyenne. Ces deux indicateurs sont également étudiés dans [16]. Cependant nous utiliserons, comme dans [41] et [47], la Value At Risk comme indicateur de risque pour définir le BS. Ensuite

nous donnerons le principal résultat théorique de ce chapitre (Cf. Théorème 6.2.5), i.e. la convergence en loi de :

$$\frac{P(h, R) - \mathbb{E}[P(h, R)]}{h}, \quad (6.2.10)$$

lorsque  $h \rightarrow +\infty$ . Finalement ce résultat nous permettra d'évaluer théoriquement le BS dans le Théorème 6.2.6. Nous concluons cette section avec quelques remarques portant notamment sur le lien avec les autres modèles ou encore sur le rôle du paramètre  $\mu$ , Cf. Remarques 6.2.7, 6.2.8, 6.2.9, 6.2.10 et 6.2.11.

**Définition 6.2.2.**

1. La v.a. mesurant le risque de provisionnement du mois 0 est définie par :

$$P(h, R) - \mathbb{E}[P(h, R)], \quad (6.2.11)$$

où  $P(h, R)$  est définie par (6.2.1).

2. Le BS associé au risque de provisionnement du mois 0, dénoté  $BS(h)$ , est le quantile d'ordre 99,5% (ou Value At Risk) de la v.a.  $P(h, R) - \mathbb{E}[P(h, R)]$ , i.e. :

$$BS(h) := \mathcal{Q}_{P(h, R) - \mathbb{E}[P(h, R)]}(99,5\%), \quad (6.2.12)$$

où  $\mathcal{Q}_X(99,5\%)$  est le quantile de la v.a.  $X$  à l'ordre 99,5%, i.e. :

$$\mathbb{P}\left[X \leq \mathcal{Q}_X(99,5\%)\right] = 99,5\%. \quad (6.2.13)$$

**Remarque 6.2.3.**

1. Remarquons que si  $c \in \mathbb{R}$  et  $Y$  est une v.a. alors :

$$\mathcal{Q}_{Y+c}(99,5\%) = \mathcal{Q}_Y(99,5\%) + c. \quad (6.2.14)$$

2. Dans la Section 2.3 de [65], la v.a. mesurant le risque de provisionnement sur un an, et calculée à la date 0, est notée  $CDR_1$  (Claim Development Result) et définie par (nous utilisons nos notations) :

$$\begin{aligned} CDR_1 := & \mathbb{E}[P(h, R)|\mathcal{F}_0] - \left\{ \sum_{j \geq 1} \mathbb{1}_{\{(T_j, D_j) \in A_h, T_j \leq 1\}} M_j f(T_j, D_j, R_{T_j}) \right. \\ & \left. + \mathbb{E}\left[ \sum_{j \geq 1} \mathbb{1}_{\{(T_j, D_j) \in A_h, T_j > 1\}} M_j f(T_j, D_j, R_{T_j}) \middle| \mathcal{F}_1 \right] \right\}. \end{aligned} \quad (6.2.15)$$

Les sommes  $\sum_{j \geq 1} \mathbb{1}_{\{(T_j, D_j) \in A_h, T_j \leq 1\}} M_j f(T_j, D_j, R_{T_j})$  et  $\sum_{j \geq 1} \mathbb{1}_{\{(T_j, D_j) \in A_h, T_j > 1\}} M_j f(T_j, D_j, R_{T_j})$  sont les provisions pour les périodes de temps  $[0, 1]$  et  $]1, h]$ ,  $\mathcal{F}_0$  et  $\mathcal{F}_1$  représentent "l'information disponible" aux dates 0 et 1, i.e. :

$$\mathcal{F}_t := \sigma\left[(D_j)_{j \geq 1}, (M_j)_{j \geq 1}, (R_s)_{s \leq t}, (T_j \wedge t)_{j \geq 1}\right]. \quad (6.2.16)$$

Comme mentionné dans leur article, Wuthrich et al. s'intéresse à une analyse du risque à "court terme" comme imposé dans le premier pilier de la réforme Solvabilité 2. Nous nous sommes cependant intéressé à une analyse des risques sur le "long terme", dans une approche ORSA. Si nous considérons la Définition (13.2.15) proposée par Wuthrich et al. non sur un an mais sur la durée maximale des prêts,  $h$ , nous obtenons :

$$CDR_h := - \left[ P(h, R) - \mathbb{E}[P(h, R) | \mathcal{F}_0] \right]. \quad (6.2.17)$$

Donc notre définition de la v.a. mesurant le risque de provisionnement est similaire à celle proposée par Wuthrich et al. avec quelques ajustements :

- (a) Nous ne prenons pas en compte l'information disponible à la date  $t_0$ ,
- (b) Nous utilisons une modélisation individuelle de la provision tandis que Wuthrich et al. utilisent une méthode collective pour déterminer la provision (Chain Ladder).
- (c) L'horizon d'analyse est la durée maximale des prêts.

Notons que  $\mathbb{E}[P(h, R)]$  est connu (Cf. Théorème 4.3.2 du Chapitre 4), la loi de  $P(h, R)$  est inconnue mais elle peut être simulée. Dans le Théorème 6.2.5 nous montrerons que la quantité définie par (6.2.10) converge en loi lorsque  $\mu < 0$ . Par conséquent nous approximerons  $BS(h)$  par la quantité  $BS^*(h)$  définie par :

$$BS^*(h) := hQ_X(99, 5\%), \quad (6.2.18)$$

où  $X$  est la limite de  $\frac{P(h, R) - \mathbb{E}[P(h, R)]}{h}$  lorsque  $h \rightarrow +\infty$  et  $Q_X$  est défini par (6.2.13).

Pour calculer  $BS^*(h)$  il est utile d'adopter les notations suivantes :

**Notation 6.2.4.** Nous noterons :

$$G \text{ une v.a. Gaussienne centrée, réduite et indépendante de } (R_t)_{t \geq 0}, \quad (6.2.19)$$

$$C_k := \lambda_2 \frac{\alpha^k}{2} \mathbb{E}[(M_1)^k], k \geq 1, \quad (6.2.20)$$

$$X^* := G - \frac{2C_1}{\sqrt{C_2}} \left[ \frac{e^{-\tau t^*}}{\tau} + \int_0^{+\infty} \inf\left(1, \frac{\gamma}{\alpha} R_t\right) dt \right], \quad (6.2.21)$$

$$\kappa := \frac{\lambda_2 \mathbb{E}[M_1]}{\sqrt{C_2}} \int_0^{+\infty} \left( \alpha \Phi \left[ \frac{\mu t - \ln \left[ \frac{\alpha}{\gamma} \right]}{\sigma \sqrt{t}} \right] + \gamma e^{(\mu + \frac{\sigma^2}{2})t} \Phi \left[ \frac{\ln \left[ \frac{\alpha}{\gamma} \right] - (\mu + \sigma^2)t}{\sigma \sqrt{t}} \right] \right) dt, \quad (6.2.22)$$

$$\eta := \lambda_2 \mathbb{E}[M_1] \int_0^{+\infty} \left\{ \alpha \Phi \left[ \frac{\ln \left[ \frac{\alpha}{\gamma} \right] - \mu t}{\sigma \sqrt{t}} \right] - \gamma e^{(\mu + \frac{\sigma^2}{2})t} \Phi \left[ \frac{\ln \left[ \frac{\alpha}{\gamma} \right] - (\mu + \sigma^2)t}{\sigma \sqrt{t}} \right] \right\} dt, \quad (6.2.23)$$

où  $\Phi$  est la f.d.r. de  $G$ ,  $\mathbb{E}[(M_1)^k]$  est le moment d'ordre  $k$  de  $M_1$ ,  $\lambda_2$ ,  $\alpha$ ,  $\tau$ ,  $\gamma$ ,  $\mu$  et  $\sigma$  sont définis par (6.2.8), (6.2.6), (6.2.5), (6.2.4), (6.2.7) et (6.2.7).

**Théorème 6.2.5.** *Plaçons nous sous les hypothèses  $(H_1)$  à  $(H_9)$ .*

1. *Le comportement asymptotique de la moyenne de la provision est le suivant :*

$$\lim_{h \rightarrow +\infty} \frac{\mathbb{E}[P(h, R)]}{h^2} = \lambda_2 \alpha \mathbb{E}[M_1] \left( \frac{1}{2} \mathbb{1}_{\{\mu < 0\}} + \frac{1}{4} \mathbb{1}_{\{\mu = 0\}} \right). \quad (6.2.24)$$

*D'autre part, quand  $\mu > 0$  nous avons :*

$$\lim_{h \rightarrow +\infty} \frac{\mathbb{E}[P(h, R)]}{h} = \eta, \quad (6.2.25)$$

*où  $\eta$  est défini par (6.2.23).*

2. *Si  $\mu < 0$  alors la provision re-normalisée converge en loi :*

$$\frac{P(h, R) - h^2 C_1}{h \sqrt{C_2}} \xrightarrow[h \rightarrow +\infty]{\mathcal{L}} X^*, \quad (6.2.26)$$

$$\frac{P(h, R) - \mathbb{E}[P(h, R)]}{h \sqrt{C_2}} \xrightarrow[h \rightarrow +\infty]{\mathcal{L}} X^* + \kappa, \quad (6.2.27)$$

*où  $C_k$ ,  $X^*$  et  $\kappa$  sont définis par (6.2.20), (6.2.21) et (6.2.22).*

La preuve du Théorème 6.2.5 est donnée dans la Section 6.2.3 : ceci nous permettra d'isoler certains résultats intermédiaires intéressants. Nous donnerons quelques commentaires relatifs au Théorème 6.2.5, Cf. Remarques 6.2.8, 6.2.9, 6.2.10 et 6.2.11, mais nous commençons tout d'abord par donner la principale conséquence du Théorème 6.2.5, i.e. le calcul de  $BS^*(h)$ .

**Théorème 6.2.6.** *Plaçons nous sous les hypothèses  $(H_1)$  à  $(H_9)$ . Si  $\mu < 0$  alors :*

$$BS^*(h) = h \left[ \kappa + \mathcal{Q}_{X^*}(99, 5\%) \right] \sqrt{\frac{\lambda_2 \alpha^2 \mathbb{E}[(M_1)^2]}{2}}, \quad (6.2.28)$$

*où  $\mathcal{Q}_{X^*}(99, 5\%)$  est définie par (6.2.13) et la constante  $\kappa$  est définie par (6.2.22).*

*Démonstration.* Il s'agit d'une conséquence directe de la définition de  $BS^*(h)$ , Cf. (6.2.18), du Théorème 6.2.5, Cf. (6.2.27), et de (6.2.14).  $\square$

**Remarque 6.2.7.**

1. *Dans le Chapitre 4 une méthode d'estimation des paramètres avec des données réelles a été proposée. Il ressort que la valeur estimée de  $\mu$  est strictement négative, c'est pourquoi nous avons uniquement considéré ce cas dans le théorème précédent.*
2. *La formule (6.2.28) est uniquement théorique. En effet, la loi de :*

$$\int_0^{+\infty} \inf \left( 1, \frac{\gamma}{\alpha} R_t \right) dt, \quad (6.2.29)$$

*est inconnue. Cependant nous avons vu dans la Section 5.4.2 qu'il était possible de simuler la provision : nous discuterons dans la Section 7.3 de la possibilité de calculer le Besoin de Solvabilité via ces simulations.*

Faisons maintenant quelques remarques concernant le Théorème 6.2.5.

**Remarque 6.2.8.**

Il est possible de donner une autre expression de la v.a.  $X^*$  définie par (6.2.21). En remarquant que :

$$\inf(1, x) = x - (x - 1)_+, \quad (6.2.30)$$

il vient que :

$$\int_0^{+\infty} \inf(1, \frac{\gamma}{\alpha} R_t) dt = \int_0^{+\infty} \frac{\gamma}{\alpha} R_t dt - \int_0^{+\infty} [\frac{\gamma}{\alpha} R_t - 1]_+ dt. \quad (6.2.31)$$

Dans [25] il est démontré que :

$$\int_0^{+\infty} R_t dt \stackrel{(d)}{=} \frac{2}{\sigma^2 \Gamma}, \quad (6.2.32)$$

où  $\Gamma$  est une v.a. de loi Gamma de paramètre  $-\frac{2\mu}{\sigma^2}$ . Ainsi en combinant (6.2.32), (6.2.31) et (6.2.21) il vient que :

$$X^* \stackrel{(d)}{=} G - \frac{2C_1}{\sqrt{C_2}} \left[ \frac{e^{-\tau t^*}}{\tau} + \frac{2\gamma}{\alpha \sigma^2 \Gamma} - \int_0^{+\infty} [\frac{\gamma}{\alpha} R_t - 1]_+ dt \right]. \quad (6.2.33)$$

Toutefois la v.a.  $\Gamma$  présente dans (6.2.33) n'est pas indépendante de  $\int_0^{+\infty} [\frac{\gamma}{\alpha} R_t - 1]_+ dt$ . Les intégrales du type (6.2.32) sont appelées "perpetual integral functionals of Brownian motion" (Cf. [11], [25], [53], [66] et [67] pour plus de détails).

**Remarque 6.2.9.**

Le signe de  $\mu$  joue un rôle essentiel dans l'étude asymptotique de la provision.

1. Si  $\mu < 0$  alors  $R_h \xrightarrow{h \rightarrow +\infty} 0$ . Cela signifie que nous ne considérons pas de revente du bien et cela permet à la compagnie d'assurances d'être plus prudente. Il s'agit du cas le plus intéressant pour la compagnie d'assurances.
2. Si  $\mu > 0$  alors  $R_h \xrightarrow{h \rightarrow +\infty} +\infty$ .  $\mu > 0$  signifie que nous considérons une augmentation "exponentielle" du marché de l'immobilier. Ce cas de figure n'a pas été approfondi car il n'a que peu d'intérêt en pratique.
3.  $\mu = 0$  signifie qu'il y a seulement des fluctuations sur la valeur du bien financé. Dans ce cas de figure nous n'avons pas été capables de déterminer la loi asymptotique de  $P(h, R) - \mathbb{E}[P(h, R)]$ .

**Remarque 6.2.10.**

1. Le facteur de re-normalisation  $h\sqrt{C_2}$  (Cf. (6.2.26) et (6.2.27)) n'est pas égal à l'écart type de la provision  $P(h, R)$  comme le cas du théorème central limite classique.

2. Notons que s'il n'y avait pas de corrélation entre les montants de chaque sinistre, la v.a.  $X^*$  (Cf. (6.2.21)) serait égale à une variable aléatoire gaussienne centrée et réduite. Le terme additionnel dans la v.a.  $X^*$  qui s'exprime à l'aide d'une intégrale du mouvement Brownien :

$$\frac{2C_1}{\sqrt{C_2}} \int_0^{+\infty} \inf \left( 1, \frac{\gamma}{\alpha} R_t \right) dt, \quad (6.2.34)$$

prend en compte les corrélations entre les montant de sinistres. Le terme :

$$\frac{2C_1}{\sqrt{C_2}} \frac{e^{-\tau t^*}}{\tau}, \quad (6.2.35)$$

représente quant à lui l'erreur commise lors de la re-normalisation. Comme expliqué dans l'introduction il y a donc une "séparation des sources de risque", l'une provenant des caractéristiques individuelles des emprunteurs qui sont indépendants et l'autre provenant du coût de l'immobilier commun à tous les emprunteurs.

**Remarque 6.2.11.** Comparaison avec d'autres "modèles avec dépendance".

Rappelons deux modèles de PPP avec dépendance :

1. Les  $U$ -statistics de PPP sont définies par :

$$\sum_{(X_1, \dots, X_k) \in \eta_{\neq}^k} f(X_1, \dots, X_k), \quad (6.2.36)$$

où  $f$  est une fonction symétrique, l'ensemble de points  $(X_i)_{i \geq 1}$  est un PPP et  $\eta_{\neq}^k$  est l'ensemble de tous les  $k$ -tuples points distincts du processus.

2. Les processus "shot noise" sont définis par :

$$\sum_{j=1}^{N_h} f(T_j, \xi_j), \quad 0 < t < +\infty,$$

où  $f$  est une fonction mesurable explicite,  $(T_j)_{j \geq 1}$  sont les temps de sauts du processus de comptage  $(N_s)_{s \geq 0}$ ,  $(\xi_j)_{j \geq 1}$  est une collection de v.a. i.i.d. et indépendantes de  $(N_s)_{s \geq 0}$ .

Pour de tels modèles il est possible de montrer un théorème central limite (Cf. [49] et [50]). Cependant ces modèles ne permettent pas de prendre en considération le processus  $(R_t)_{t \geq 0}$ .

### 6.2.3 Preuve du Théorème 6.2.5

**Preuve de l'item 1 du Théorème 6.2.5 :**

Nous prouvons tout d'abord que la moyenne de la provision converge vers une constante qui dépend de  $\mu$  quand  $h \rightarrow +\infty$ . Rappelons que  $i = 0$  et que :

$$\Lambda(dt ds) = \lambda_2 dt ds, \quad (6.2.37)$$

où  $\Lambda$  et  $\lambda_2$  sont définis par (6.2.8). En utilisant le Théorème 4.3.2 du Chapitre 4 et les changements de variables  $s = hy$  et  $t = hx$  nous obtenons :

$$\mathbb{E}[P(h, R)] = \lambda_2 \mathbb{E}[M_1] \int_0^h ds \left[ \int_0^s \{g_1(t, s) - g_2(t, s)\} dt \right] \quad (6.2.38)$$

$$= h^2 \lambda_2 \mathbb{E}[M_1] \int_0^1 dy \left[ \int_0^y \{g_1(hx, hy) - g_2(hx, hy)\} dx \right], \quad (6.2.39)$$

où :

$$g_1(u, v) := \phi_u(v) \Phi \left[ \frac{\ln \left[ \frac{\phi_u(v)}{\gamma} \right] - \mu u}{\sigma \sqrt{u}} \right], \quad u < v, \quad (6.2.40)$$

$$g_2(u, v) := \gamma e^{(\mu + \frac{\sigma^2}{2})u} \Phi \left[ \frac{\ln \left[ \frac{\phi_u(v)}{\gamma} \right] - (\mu + \sigma^2)u}{\sigma \sqrt{u}} \right], \quad u < v, \quad (6.2.41)$$

et  $\Phi$  désigne la f.d.r. de la v.a. gaussienne centrée, réduite et la fonction  $\phi$  est définie par (6.2.5). Notons que la condition  $0 < x < y$  implique :

$$\lim_{h \rightarrow +\infty} \phi_{hx}(hy) = \alpha, \quad (6.2.42)$$

avec  $\alpha$  défini par (6.2.6). Nous avons donc :

$$\lim_{h \rightarrow +\infty} g_1(hx, hy) = \alpha \left( \mathbb{1}_{\{\mu < 0\}} + \frac{1}{2} \mathbb{1}_{\{\mu = 0\}} \right). \quad (6.2.43)$$

Pour étudier le comportement asymptotique de  $g_2(hx, hy)$  nous devons distinguer deux cas :

1. Si  $\mu + \sigma^2 \leq 0$ , alors  $\mu + \frac{\sigma^2}{2} < 0$ . Puisque  $\Phi(z) \leq 1$  il vient que :

$$\lim_{h \rightarrow +\infty} g_2(hx, hy) = 0. \quad (6.2.44)$$

2. Si  $\mu + \sigma^2 > 0$ , rappelons que :

$$\Phi(z) \sim \frac{1}{\sqrt{2\pi}|z|} e^{-\frac{z^2}{2}}, \quad z \rightarrow -\infty. \quad (6.2.45)$$

Donc lorsque  $h \rightarrow +\infty$  nous avons successivement :

$$\begin{aligned} g_2(hx, hy) &\sim e^{(\mu + \frac{\sigma^2}{2})hx} \frac{\gamma}{\sqrt{2\pi}} \frac{\sigma \sqrt{hx}}{|\ln \left[ \frac{\phi_{hx}(hy)}{\gamma} \right] - (\mu + \sigma^2)hx|} \\ &\quad \times \exp \left[ -\frac{1}{2\sigma^2 hx} \left\{ \ln \left[ \frac{\phi_{hx}(hy)}{\gamma} \right] - (\mu + \sigma^2)hx \right\}^2 \right] \\ &\sim \frac{\sigma \gamma}{\sqrt{2\pi}(\mu + \sigma^2)\sqrt{hx}} \exp \left[ \frac{(\mu + \sigma^2) \ln \left[ \frac{\phi_{hx}(hy)}{\gamma} \right]}{\sigma^2} - \frac{\left( \ln \left[ \frac{\phi_{hx}(hy)}{\gamma} \right] \right)^2}{2\sigma^2 hx} \right] \end{aligned}$$



$$\begin{aligned}
 & + \left[ \left( \mu + \frac{\sigma^2}{2} \right) - \frac{(\mu + \sigma^2)^2}{2\sigma^2} \right] hx \Big] \\
 \sim & \frac{\sigma\gamma}{\sqrt{2\pi}(\mu + \sigma^2)\sqrt{hx}} \exp \left[ \frac{(\mu + \sigma^2) \ln(\frac{\alpha}{\gamma})}{\sigma^2} \right] \exp \left[ -\frac{\mu^2}{2\sigma^2} hx \right].
 \end{aligned}$$

Ainsi :

$$\lim_{h \rightarrow +\infty} g_2(hx, hy) = 0. \quad (6.2.46)$$

En combinant (6.2.44) et (6.2.46) nous obtenons :

$$\lim_{h \rightarrow +\infty} g_2(hx, hy) = 0. \quad (6.2.47)$$

Une application du Lemme 4.5.1 conduit à :

$$\mathbb{E}[f(hx, hy, R_{hx})] = g_1(hx, hy) - g_2(hx, hy), \quad (6.2.48)$$

il vient alors d'après la formule (6.3.6) du Lemme 6.3.2 que :

$$0 \leq g_1(hx, hy) - g_2(hx, hy) \leq \alpha, \forall h > 0, \forall 0 < x < y. \quad (6.2.49)$$

Par conséquent en utilisant (6.2.39), (6.2.43), (6.2.47), (6.2.49) et le théorème de convergence dominé nous avons :

$$\lim_{h \rightarrow +\infty} \frac{\mathbb{E}[P(h, R)]}{h^2} = \lambda_2 \alpha \mathbb{E}[M_1] \left( \frac{1}{2} \mathbb{1}_{\{\mu < 0\}} + \frac{1}{4} \mathbb{1}_{\{\mu = 0\}} \right). \quad (6.2.50)$$

Ceci établit la formule (6.2.24).

Nous voulons maintenant étudier la limite de  $\frac{\mathbb{E}[P(h, R)]}{h}$  quand  $\mu > 0$  et  $h \rightarrow +\infty$ . Enonçons un lemme qui nous servira par la suite :

**Lemme 6.2.12.** Soient  $a \in \mathbb{R}$ ,  $b \in \mathbb{R}$  et  $c > 0$ . Alors :

$$\int_0^{+\infty} \Phi \left[ \frac{a - bt}{\sqrt{t}} \right] dt < +\infty, \quad b > 0, \quad (6.2.51)$$

$$\int_0^{+\infty} e^{(b + \frac{c^2}{2})t} \Phi \left[ \frac{a - (b + c^2)t}{c\sqrt{t}} \right] dt < +\infty, \quad (6.2.52)$$

où rappelons que  $\Phi$  est la f.d.r. d'une v.a. gaussienne centrée et réduite.

Ce lemme sera démontré dans la Section 6.3.1. Revenons à l'étude de la limite de  $\frac{\mathbb{E}[P(h, R)]}{h}$ . En utilisant (6.2.38) et le changement de variable  $s = hy$  nous obtenons :

$$\begin{aligned}
 \Gamma_h := \frac{\mathbb{E}[P(h, R)]}{h\lambda_2\mathbb{E}[M_1]} &= \frac{1}{h} \int_0^h ds \left[ \int_0^s \{g_1(t, s) - g_2(t, s)\} dt \right] \\
 &\stackrel{s=hy}{=} \int_0^1 dy \left[ \int_0^{hy} \{g_1(t, hy) - g_2(t, hy)\} dt \right], \quad (6.2.53)
 \end{aligned}$$

où les fonctions  $g_1$  et  $g_2$  sont définies par (6.2.40) et (6.2.41) respectivement. Puisque  $t \leq hy$  il vient que :

$$\lim_{h \rightarrow +\infty} \phi_t(hx) = \alpha, \quad t > 0, x > 0. \quad (6.2.54)$$

Ainsi :

$$\lim_{h \rightarrow +\infty} g_1(t, hy) = \alpha \Phi \left[ \frac{\ln \left[ \frac{\alpha}{\gamma} \right] - \mu t}{\sigma \sqrt{t}} \right], \quad (6.2.55)$$

$$\lim_{h \rightarrow +\infty} g_2(t, hy) = \gamma e^{(\mu + \frac{\sigma^2}{2})t} \Phi \left[ \frac{\ln \left[ \frac{\alpha}{\gamma} \right] - (\mu + \sigma^2)t}{\sigma \sqrt{t}} \right]. \quad (6.2.56)$$

D'après (6.3.6) nous avons :

$$\phi_t(hy) \leq \alpha. \quad (6.2.57)$$

Ainsi :

$$0 \leq g_1(t, hy) \leq \alpha \Phi \left[ \frac{\ln \left( \frac{\alpha}{\gamma} \right) - \mu t}{\sigma \sqrt{t}} \right], \quad (6.2.58)$$

$$0 \leq g_2(t, hy) \leq \gamma e^{(\mu + \frac{\sigma^2}{2})t} \Phi \left[ \frac{\ln \left( \frac{\alpha}{\gamma} \right) - (\mu + \sigma^2)t}{\sigma \sqrt{t}} \right]. \quad (6.2.59)$$

En combinant (6.2.53), (6.2.55), (6.2.56), (6.2.58), (6.2.59), (6.2.51) et (6.2.52) il vient que :

$$\lim_{h \rightarrow +\infty} \Gamma_h = \int_0^{+\infty} \left\{ \alpha \Phi \left[ \frac{\ln \left[ \frac{\alpha}{\gamma} \right] - \mu t}{\sigma \sqrt{t}} \right] - \gamma e^{(\mu + \frac{\sigma^2}{2})t} \Phi \left[ \frac{\ln \left[ \frac{\alpha}{\gamma} \right] - (\mu + \sigma^2)t}{\sigma \sqrt{t}} \right] \right\} dt.$$

Ce qui prouve (6.2.25).

### Preuve de la formule (6.2.26) du Théorème 6.2.5 :

Introduisons quelques notations.

#### Notation 6.2.13.

$$X_k(h, r) := \lambda_2 \mathbb{E} \left[ (M_1)^k \right] \int_{A_h} [f(t, s, r_t)]^k dt ds, k \geq 1, \quad (6.2.60)$$

$$P(h, r) := \sum_{j \geq 1} \mathbf{1}_{\{(T_j, D_j) \in A_h\}} M_j f(T_j, D_j, r_{T_j}), \quad (6.2.61)$$

$$\begin{aligned} r. : \mathbb{R}_+ &\longrightarrow \mathbb{R}_+ \\ t &\longmapsto r_t, \end{aligned} \quad (6.2.62)$$

$$r := (r_t, t \geq 0), \quad (6.2.63)$$

où  $A_h$  et  $f$  sont définis par (6.2.3) et (6.2.4) respectivement. Notons que  $X_k(h, R)$  est une v.a.  $\sigma[R_t, t \leq h]$  mesurable.

Nous étudierons la convergence en loi de la provision sachant  $(R_t = r_t, t \geq 0)$ . Ce qui signifie que nous allons considérer le comportement asymptotique de  $P(h, r)$  définie par (6.2.61). Nous allons montrer (Cf. Proposition 6.2.14) que cette variable, convenablement renormalisée, converge en distribution vers une variable gaussienne lorsque  $h \rightarrow +\infty$ . Dans un second temps nous étudierons le comportement asymptotique des coefficients de renormalisation, i.e. de  $\frac{X_k(h, R)}{h^2}$ , en fonction de la valeur de  $\mu$  et nous prouverons que  $\frac{X_k(h, R)}{h^2}$  converge en probabilité vers une constante quand  $\mu < 0$ , Cf. (6.2.82). Dans un dernier temps nous prouverons que  $\left(\frac{X_1(h, R)}{h^2} - C_1\right)h$  converge presque sûrement vers l'intégrale d'une fonction d'un mouvement Brownien géométrique lorsque  $\mu < 0$ , Cf. (6.2.86).

Ces résultats nous permettront de prouver que la provision convenablement renormalisée converge en loi vers un mélange d'une variable aléatoire gaussienne et de l'intégrale d'une fonction d'un mouvement Brownien géométrique.

Commençons par l'étude asymptotique de la provision  $P(h, r)$ . Pour ce faire nous avons besoin d'introduire la distance de Wasserstein, notée  $d_W$ , définie pour deux v.a.  $X$  et  $Y$  par :

$$d_W(X, Y) := \sup_{\xi \in Lip(1)} \left| \mathbb{E}[\xi(X)] - \mathbb{E}[\xi(Y)] \right|, \quad (6.2.64)$$

où  $Lip(1)$  est l'ensemble de toutes les fonctions  $\xi : \mathbb{R} \rightarrow \mathbb{R}$  qui sont 1-lipschitziennes, i.e. vérifiant :

$$|\xi(x) - \xi(y)| \leq |x - y|, \quad \forall x, y \in \mathbb{R}. \quad (6.2.65)$$

**Proposition 6.2.14.** *Plaçons nous sous les hypothèses  $(H_1)$  à  $(H_9)$ .*

1. *Nous avons :*

$$d_W \left[ \frac{P(h, r) - \mathbb{E}[P(h, r)]}{\sqrt{\mathbb{V}[P(h, r)]}}, G \right] \leq \frac{\sqrt{X_4(h, r)}}{X_2(h, r)}, \quad (6.2.66)$$

où  $X_k(h, r)$  est définie par (6.2.60).

2. *D'autre part :*

$$\mathbb{E}[P(h, r)] = X_1(h, r), \quad (6.2.67)$$

$$\mathbb{V}[P(h, r)] = X_2(h, r). \quad (6.2.68)$$

3. *Si la fonction  $t \rightarrow r_t$  vérifie :*

$$\lim_{t \rightarrow +\infty} r_t = 0, \quad (6.2.69)$$

*alors :*

$$\lim_{h \rightarrow +\infty} \frac{\sqrt{X_4(h, r)}}{X_2(h, r)} = 0. \quad (6.2.70)$$

4. Ainsi :

$$\frac{P(h, r) - \mathbb{E}[P(h, r)]}{\sqrt{\mathbb{V}[P(h, r)]}} \xrightarrow[h \rightarrow +\infty]{\mathcal{L}} G. \quad (6.2.71)$$

La preuve de cette proposition sera faite dans la Section 6.3.2.

**Remarque 6.2.15.** La distance de Kolmogorow entre deux v.a. réelles  $X$  et  $Y$  est notée  $d_{Kol}$  et définie par :

$$d_{Kol}(X, Y) := \sup_{x \in \mathbb{R}} |\mathbb{P}[X \leq x] - \mathbb{P}[Y \leq x]|. \quad (6.2.72)$$

Dans [39] il est rappelé que si  $G$  est une variable gaussienne centrée et réduite alors :

$$d_{Kol}(X, G) \leq 2\sqrt{d_W(X, G)}. \quad (6.2.73)$$

En combinant (6.2.66) et (6.2.73) il vient que :

$$\left| \mathbb{P}\left[\frac{P(h, r) - \mathbb{E}[P(h, r)]}{\sqrt{\mathbb{V}[P(h, r)]}} \leq x\right] - \mathbb{P}[G \leq x] \right| \leq 2\sqrt{\frac{\sqrt{X_4(h, r)}}{X_2(h, r)}}, \quad x > 0. \quad (6.2.74)$$

Et d'après (6.3.13) il suit que :

$$\left| \mathbb{P}\left[\frac{P(h, r) - \mathbb{E}[P(h, r)]}{\sqrt{\mathbb{V}[P(h, r)]}} \leq x\right] - \mathbb{P}[G \leq x] \right| \leq 2 \left[ \frac{\alpha^2 \mathbb{E}[(M_1)^4]}{\lambda_2 [\mathbb{E}[(M_1)^2]]^2 \int_{A_h} [f(t, s, r_t)]^2 dt ds} \right]^{\frac{1}{4}}, \quad (6.2.75)$$

D'après (6.3.22) il suit que :

$$\begin{aligned} \int_{A_h} [f(t, s, r_t)]^2 dt ds &\geq h \int_0^1 dy \left[ \int_0^{+\infty} \mathbf{1}_{t \leq hy} \left( [\alpha(1 - e^{-\tau t^*}) - r_t]_+ \right)^2 dt \right] \\ &\geq h \int_0^{+\infty} \mathbf{1}_{\frac{t}{h} \leq 1} \left( [\alpha(1 - e^{-\tau t^*}) - r_t]_+ \right)^2 \left(1 - \frac{t}{h}\right) dt. \end{aligned} \quad (6.2.76)$$

D'après (6.2.69) :  $\exists \tilde{t} : \forall t \geq \tilde{t} : r_t < \alpha(1 - e^{-\tau t^*})$ . Donc :

$$\int_{A_h} [f(t, s, r_t)]^2 dt ds \geq h \int_{\tilde{t}}^{+\infty} \mathbf{1}_{\frac{t}{h} \leq 1} \left( [\alpha(1 - e^{-\tau t^*}) - r_t]_+ \right)^2 \left(1 - \frac{t}{h}\right) dt. \quad (6.2.77)$$

Soit  $\tilde{t}'$  tel que :  $\forall t > \tilde{t}' : r_t < \frac{\alpha}{2}(1 - e^{-\tau t^*})$  et :

$$\tilde{C} := \frac{\alpha}{2}(1 - e^{-\tau t^*}). \quad (6.2.78)$$

Pour  $h > 2\tilde{t}'$  il suit que :

$$\int_{A_h} [f(t, s, r_t)]^2 dt ds \geq h\tilde{C} \int_{\tilde{t}'}^h \left(1 - \frac{t}{h}\right) dt.$$

$$\begin{aligned} &\geq h^2 \tilde{C} \int_{\frac{\tilde{v}}{h}}^1 (1-u) du \\ &\geq h^2 \frac{\tilde{C}}{2}. \end{aligned} \quad (6.2.79)$$

En combinant (6.2.80) et (6.2.79) il suit que :

$$\left| \mathbb{P} \left[ \frac{P(h, r) - \mathbb{E}[P(h, r)]}{\sqrt{\mathbb{V}[P(h, r)]}} \leq x \right] - \mathbb{P}[G \leq x] \right| \leq \frac{C}{\sqrt{h}}, x > 0, \quad (6.2.80)$$

où :

$$C = 2 \sqrt{\frac{\alpha}{\mathbb{E}[(M_1)^2]}} \left( \frac{2\mathbb{E}[(M_1)^4]}{\lambda_2 \tilde{C}} \right)^{\frac{1}{4}}. \quad (6.2.81)$$

Le résultat (6.2.80) nous donne une vitesse de convergence de la provision renormalisée vers une variable gaussienne lorsque le processus  $(R_t)_{t \geq 0}$  est supposé connu.

Etudions maintenant le comportement asymptotique de  $X_k(h, R)$  en fonction de la valeur de  $\mu$  :

**Proposition 6.2.16.**

Plaçons nous sous les hypothèses  $(H_1)$  à  $(H_8)$ . Soit également  $k \geq 1$ .

1. Lorsque  $\mu < 0$  il vient que :

$$\frac{X_k(h, R)}{h^2} \xrightarrow[h \rightarrow +\infty]{\mathcal{P}} C_k, \quad (6.2.82)$$

où le coefficient  $C_k$  est défini par (6.2.20). Lorsque  $\mu = 0$  alors :

$$\frac{X_k(h, R)}{h^2} \xrightarrow[h \rightarrow +\infty]{\mathcal{L}} 2C_k \int_0^1 (1-t) \mathbb{1}_{\{B_t < 0\}} dt, \quad (6.2.83)$$

où  $(B_t)_{t \geq 0}$  est le mouvement Brownien associé à  $(R_t)_{t \geq 0}$ , Cf. (6.2.7).

2. D'autre part si  $\mu < 0$  nous avons :

$$\left( \frac{X_1(h, R)}{h^2} - C_1 \right) h \xrightarrow[h \rightarrow +\infty]{p.s.} -2C_1 \left[ \frac{e^{-\tau t^*}}{\tau} + \int_0^{+\infty} \inf \left( 1, \frac{\gamma}{\alpha} R_t \right) dt \right], \quad (6.2.84)$$

où  $(R_t)_{t \geq 0}$  est le mouvement Brownien géométrique défini par (6.2.7).

La preuve de cette proposition sera effectuée dans la Section 6.3.3.

**Remarque 6.2.17.**

1. Remarquons que si  $\mu > 0$  alors :

$$\int_0^{+\infty} \inf \left( 1, \frac{\gamma}{\alpha} R_t \right) dt = +\infty. \quad (6.2.85)$$

Donc lorsque  $\mu > 0$  le résultat (6.2.84) devient :

$$\left( \frac{X_1(h, R)}{h^2} - C_1 \right) h \xrightarrow[h \rightarrow +\infty]{p.s.} +\infty. \quad (6.2.86)$$

Dans le cas  $\mu > 0$  le facteur de re-normalisation n'est pas suffisant pour assurer une limite finie.

2. Le résultat (6.2.84) implique en particulier que (6.2.82) a lieu pour  $k = 1$  et la convergence a lieu presque sûrement.

Notons que nous ne pouvons pas déduire des Propositions 6.2.14 et 6.2.16, la convergence de  $\frac{P(h, R) - h^2 C_1}{h\sqrt{C_2}}$ , nous avons besoin d'un résultat supplémentaire :

**Proposition 6.2.18.**

Plaçons nous sous les hypothèses  $(H_1)$  à  $(H_9)$  et supposons que  $\mu < 0$ . Il vient alors que :

$$\left[ \frac{P(h, R) - X_1(h, R)}{\sqrt{X_2(h, R)}}, \left( \frac{X_1(h, R)}{h^2} - C_1 \right) h \right] \xrightarrow[h \rightarrow +\infty]{\mathcal{L}} \left[ G, -2C_1 \left[ \frac{e^{-\tau t^*}}{\tau} + \int_0^{+\infty} \inf \left( 1, \frac{\gamma}{\alpha} R_t \right) dt \right] \right]. \quad (6.2.87)$$

La preuve de cette proposition sera faite dans la Section 6.3.4.

Revenons à la preuve du résultat (6.2.26). Nous avons :

$$\frac{P(h, R) - h^2 C_1}{h\sqrt{C_2}} = \frac{P(h, R) - X_1(h, R)}{\sqrt{X_2(h, R)}} \sqrt{\frac{X_2(h, R)}{h^2 C_2}} + \left( \frac{X_1(h, R)}{h^2} - C_1 \right) \frac{h}{\sqrt{C_2}}. \quad (6.2.88)$$

La Proposition 6.2.18 combinée à (6.2.82) avec  $k = 2$  nous permet de déduire :

$$\frac{P(h, R) - h^2 C_1}{h\sqrt{C_2}} \xrightarrow[h \rightarrow +\infty]{\mathcal{L}} X^*,$$

où  $X^*$  est définie par (6.2.21). Ceci établit la formule (6.2.26).

**Preuve de la formule (6.2.27) du Théorème 6.2.5 :**

Enonçons tout d'abord un lemme technique qui nous servira par la suite :

**Lemme 6.2.19.**

$$\lim_{h \rightarrow +\infty} \int_0^1 dy \int_0^{hy} [\phi_t(hy) - \alpha] dt = -\alpha \frac{e^{-\tau t^*}}{\tau}. \quad (6.2.89)$$

Ce lemme sera prouvé dans la Section 6.3.5. Concentrons nous maintenant sur la preuve de la formule (6.2.27). Nous avons :

$$\frac{P(h, R) - \mathbb{E}[P(h, R)]}{h\sqrt{C_2}} = \frac{P(h, R) - h^2 C_1}{h\sqrt{C_2}} + \frac{h}{\sqrt{C_2}} \left[ C_1 - \frac{\mathbb{E}[P(h, R)]}{h^2} \right]. \quad (6.2.90)$$

D'après (6.2.38) et en utilisant le changement de variable  $s = hy$  il vient que :

$$h \left[ C_1 - \frac{\mathbb{E}[P(h, R)]}{h^2} \right] = \lambda_2 \mathbb{E}[M_1] \int_0^1 dy \int_0^{hy} [\Gamma_1(t, hy) + \Gamma_2(t, hy)] dt, \quad (6.2.91)$$

où :

$$\Gamma_1(t, hy) := \alpha - \phi_t(hy) \Phi \left[ \frac{\ln \left[ \frac{\phi_t(hy)}{\gamma} \right] - \mu t}{\sigma \sqrt{t}} \right], \quad (6.2.92)$$

$$\Gamma_2(t, hy) := \gamma e^{(\mu + \frac{\sigma^2}{2})t} \Phi \left[ \frac{\ln \left[ \frac{\phi_t(hy)}{\gamma} \right] - (\mu + \sigma^2)t}{\sigma \sqrt{t}} \right]. \quad (6.2.93)$$

D'après (6.2.54) il vient que :

$$\begin{aligned} \lim_{h \rightarrow +\infty} \Gamma_1(t, hy) &= \alpha - \alpha \Phi \left[ \frac{\ln \left[ \frac{\alpha}{\gamma} \right] - \mu t}{\sigma \sqrt{t}} \right] \\ &= \alpha \Phi \left[ \frac{\mu t - \ln \left[ \frac{\alpha}{\gamma} \right]}{\sigma \sqrt{t}} \right], \end{aligned} \quad (6.2.94)$$

$$\lim_{h \rightarrow +\infty} \Gamma_2(t, hy) = \gamma e^{(\mu + \frac{\sigma^2}{2})t} \Phi \left[ \frac{\ln \left[ \frac{\alpha}{\gamma} \right] - (\mu + \sigma^2)t}{\sigma \sqrt{t}} \right]. \quad (6.2.95)$$

D'après (6.3.6) et (6.3.21) nous avons :

$$\alpha(1 - e^{-\tau t^*}) \leq \phi_t(hy) \leq \alpha. \quad (6.2.96)$$

Ainsi :

$$\begin{aligned} \Gamma_1(t, hy) &= \alpha - \phi_t(hy) \left( 1 - \Phi \left[ \frac{\mu t - \ln \left[ \frac{\phi_t(hy)}{\gamma} \right]}{\sigma \sqrt{t}} \right] \right) \\ &= \alpha - \phi_t(hy) + \phi_t(hy) \Phi \left[ \frac{\mu t - \ln \left[ \frac{\phi_t(hy)}{\gamma} \right]}{\sigma \sqrt{t}} \right] \\ &\leq \alpha - \phi_t(hy) + \alpha \Phi \left[ \frac{\mu t - \ln \left[ \frac{\alpha(1 - e^{-\tau t^*})}{\gamma} \right]}{\sigma \sqrt{t}} \right], \end{aligned} \quad (6.2.97)$$

$$\Gamma_2(t, hy) \leq \gamma e^{(\mu + \frac{\sigma^2}{2})t} \Phi \left[ \frac{\ln \left[ \frac{\alpha}{\gamma} \right] - (\mu + \sigma^2)t}{\sigma \sqrt{t}} \right]. \quad (6.2.98)$$

En combinant (6.2.91), (6.2.92), (6.2.93), (6.2.94), (6.2.95), (6.2.97), (6.2.98), (6.2.51), (6.2.52) et (6.2.89) il vient que :

$$\lim_{h \rightarrow +\infty} h \left[ C_1 - \frac{\mathbb{E}[P(h, R)]}{h^2} \right] = \lambda_2 \mathbb{E}[M_1] \int_0^{+\infty} \left( \alpha \Phi \left[ \frac{\mu t - \ln \left[ \frac{\alpha}{\gamma} \right]}{\sigma \sqrt{t}} \right] \right)$$

$$+ \gamma e^{(\mu + \frac{\sigma^2}{2})t} \Phi \left[ \frac{\ln \left[ \frac{\alpha}{\gamma} \right] - (\mu + \sigma^2)t}{\sigma\sqrt{t}} \right] dt. \quad (6.2.99)$$

Donc en combinant (6.2.90), (6.2.99) et (6.2.26) nous obtenons la formule (6.2.27).

## 6.3 Preuves techniques

### 6.3.1 Preuve du Lemme 6.2.12

Rappelons que :

$$\Phi(z) \sim \frac{1}{\sqrt{2\pi}|z|} e^{-\frac{z^2}{2}}, \quad z \rightarrow -\infty. \quad (6.3.1)$$

La formule (6.2.51) est une conséquence directe de (6.3.1). Prouvons maintenant la formule (6.2.52). Nous distinguons deux cas :

1. Si  $b + c^2 \leq 0$  alors  $b + \frac{c^2}{2} < 0$  et la formule (6.2.52) est vérifiée car  $\Phi(x) \leq 1$  pour tout  $x \in \mathbb{R}$ .
2. Si  $b + c^2 > 0$  nous utilisons que :

$$\Phi(z) \leq \frac{1}{\sqrt{2\pi}|z|} e^{-\frac{z^2}{2}}, \quad z < 0. \quad (6.3.2)$$

Donc :

$$\begin{aligned} e^{(b + \frac{c^2}{2})t} \Phi \left[ \frac{a - (b + c^2)t}{c\sqrt{t}} \right] &\leq \frac{c\sqrt{t}}{\sqrt{2\pi}|a - (b + c^2)t|} \exp \left[ \left( b + \frac{c^2}{2} \right)t - \frac{1}{2tc^2} [a - (b + c^2)t]^2 \right] \\ &\leq \frac{c\sqrt{t}}{\sqrt{2\pi}|a - (b + c^2)t|} \exp \left[ -\frac{b^2}{2c^2}t + \frac{a}{c^2}(b + c^2) - \frac{a^2}{2tc^2} \right]. \end{aligned} \quad (6.3.3)$$

Ce qui prouve (6.2.52).

### 6.3.2 Preuve de la Proposition 6.2.14

**Notation 6.3.1.** *Il est utile d'adopter les notations suivantes :*

$$f_k := f(T_k, D_k, r_{T_k}), \quad k \geq 1, \quad (6.3.4)$$

$$N(A) := \sum_{j \geq 1} \mathbf{1}_{\{(T_j, D_j) \in A\}}, \quad (6.3.5)$$

où  $f$  est définie par (6.2.4),  $A$  est un sous ensemble de  $\{(s, t), 0 \leq t \leq s\}$  et  $(r_t, t \geq 0)$  est une fonction continue à valeurs dans  $]0, +\infty[$ .

La preuve de la Proposition 6.2.14 est basée sur le lemme suivant :

#### Lemme 6.3.2.



1. Pour tout  $t, s, u \geq 0 : t < s$  nous avons :

$$f(t, s, u) \leq \phi_t(s) \leq \alpha, \quad (6.3.6)$$

où  $f$  et  $\alpha$  sont définies par (6.2.4) et (6.2.6).

2. Nous avons :

$$\mathbb{E}[P(h, r)] = X_1(h, r), \quad (6.3.7)$$

$$\mathbb{V}[P(h, r)] = X_2(h, r). \quad (6.3.8)$$

3. Si :

$$\lim_{t \rightarrow +\infty} r_t = 0, \quad (6.3.9)$$

alors :

$$\lim_{h \rightarrow +\infty} \int_{A_h} [f(t, s, r_t)]^2 dt ds = +\infty. \quad (6.3.10)$$

La preuve de ce lemme sera effectuée à la fin de cette section. Nous démontrons maintenant la Proposition 6.2.14. Les formules (6.2.67) et (6.2.68) sont prouvées par (6.3.7) et (6.3.8). D'après le Théorème 4.1 de [50] nous avons :

$$d_W \left[ \frac{P(h, r) - \mathbb{E}[P(h, r)]}{\mathbb{V}[P(h, r)]}, G \right] \leq \frac{\sqrt{X_4(h, r)}}{\mathbb{V}[P(h, r)]}, \quad (6.3.11)$$

où  $d_W$  est la distance de Wasserstein définie par (6.2.64) : ceci combiné à (6.3.8) démontre la formule (6.2.66). Montrons maintenant que le terme de droite dans l'inéquation (6.3.11) converge vers 0 lorsque  $h \rightarrow +\infty$ . D'après (6.2.60), (6.3.8) et (6.3.6) il vient que :

$$\begin{aligned} \frac{X_4(h, r)}{[\mathbb{V}[P(h, r)]]^2} &= \frac{X_4(h, r)}{[X_2(h, r)]^2} \\ &= \frac{\mathbb{E}[(M_1)^4] \int_{A_h} [f(t, s, r_t)]^4 dt ds}{\lambda_2 [\mathbb{E}[(M_1)^2] \int_{A_h} [f(t, s, r_t)]^2 dt ds]^2}. \end{aligned}$$

Puisque :

$$[f(t, s, r_t)]^4 \leq \alpha^2 [f(t, s, r_t)]^2, \quad (6.3.12)$$

il suit que :

$$\frac{X_4(h, r)}{[\mathbb{V}[P(h, r)]]^2} \leq \frac{\alpha^2 \mathbb{E}[(M_1)^4]}{\lambda_2 [\mathbb{E}[(M_1)^2]]^2 \int_{A_h} [f(t, s, r_t)]^2 dt ds}. \quad (6.3.13)$$

D'après (6.3.10), (6.3.13) et l'hypothèse  $(H_9)$  il suit que :

$$\frac{\sqrt{X_4(h, r)}}{\mathbb{V}[P(h, r)]} \xrightarrow{h \rightarrow +\infty} 0. \quad (6.3.14)$$

Ce qui démontre la formule (6.2.70). En combinant (6.3.14) et (6.3.11) il suit que (6.2.71) est vérifiée. Ceci achève la preuve de la Proposition 6.2.14. Revenons maintenant à la preuve du Lemme 6.3.2.

*Preuve du Lemme 6.3.2.*

1. D'après (6.2.4), puisque  $u \geq 0$  et  $\gamma \in ]0, 1[$  il vient que :

$$f(t, s, u) \leq \phi_t(s), \quad (6.3.15)$$

où  $f$  et  $\phi$  sont définies par (6.2.4) et (6.2.5). Par définition de  $\phi_t(s)$  nous déduisons de (6.3.15) que :

$$f(t, s, u) \leq \alpha, \text{ si } t \leq t^*. \quad (6.3.16)$$

Lorsque  $s > t^*$  la fonction  $t \rightarrow \phi_t(s)$  est décroissante sur  $]t^*, s]$  et :

$$\lim_{t \rightarrow t^*} \phi_t(s) = \alpha.$$

Nous en déduisons que :

$$f(t, s, u) \leq \alpha, \text{ si } t > t^*. \quad (6.3.17)$$

Ce qui prouve (6.3.6).

2. Rappelons que  $N(A_h) \sim P(\Lambda(A_h))$  et sachant que  $N(A_h) = n$  les v.a.  $(T_1, D_1), \dots, (T_n, D_n)$  sont i.i.d. de loi  $\frac{1}{\Lambda(A_h)} \Lambda|_{A_h}$ . Il vient alors que :

$$\begin{aligned} \mathbb{E}[P(h, r)] &= \mathbb{E}\left[\sum_{j \geq 1} M_j f_j \mathbf{1}_{\{(T_j, D_j) \in A_h\}}\right] \\ &= X_1(h, r). \end{aligned}$$

Ce qui prouve la formule (6.3.7). Calculons maintenant la variance de  $P(h, r)$  :

$$\begin{aligned} \mathbb{V}[P(h, r)] &= \mathbb{E}\left[\left(\sum_{j \geq 1} M_j f_j \mathbf{1}_{\{(T_j, D_j) \in A_h\}}\right)^2\right] - [X_1(h, r)]^2 \\ &= \sum_{n \geq 0} \mathbb{P}[N(A_h) = n] \left\{ \mathbb{E}\left[\sum_{j=1}^n (M_j f_j)^2\right] + \mathbb{E}\left[\sum_{1 \leq j \neq l \leq n} M_j f_j M_l f_l\right] \right\} \\ &\quad - [X_1(h, r)]^2 \\ &= \sum_{n \geq 0} \mathbb{P}[N(A_h) = n] \left\{ n \mathbb{E}[(M_1)^2] \mathbb{E}[(f_1)^2] + n(n-1) \left(\mathbb{E}[M_1] \mathbb{E}[f_1]\right)^2 \right\} \\ &\quad - [X_1(h, r)]^2. \end{aligned} \quad (6.3.18)$$

D'autre part :

$$\mathbb{E}[(f_1)^k] = \frac{\lambda_2}{\Lambda(A_h)} \int_{A_h} [f(t, s, r_t)]^k dt ds$$

$$\begin{aligned}
 &= \frac{\lambda_2}{\mathbb{E}[N(A_h)]} \int_{A_h} [f(t, s, r_t)]^k dt ds \\
 &= \frac{X_k(h, r)}{\mathbb{E}[(M_1)^k]} \frac{1}{\mathbb{E}[N(A_h)]}.
 \end{aligned} \tag{6.3.19}$$

Ainsi :

$$\begin{aligned}
 \mathbb{V}[P(h, r)] &= \sum_{n \geq 0} \mathbb{P}[N(A_h) = n] \left\{ n \frac{X_2(h, r)}{\mathbb{E}[N(A_h)]} + n(n-1) \left[ \frac{X_1(h, r)}{\mathbb{E}[N(A_h)]} \right]^2 \right\} \\
 &\quad - [X_1(h, r)]^2 \\
 &= X_2(h, r) + \left[ \mathbb{E}[(N(A_h))^2] - \mathbb{E}[N(A_h)] \right] \left[ \frac{X_1(h, r)}{\mathbb{E}[N(A_h)]} \right]^2 - [X_1(h, r)]^2.
 \end{aligned}$$

Puisque :

$$\mathbb{V}[N(A_h)] = \mathbb{E}[N(A_h)] = \mathbb{E}[(N(A_h))^2] - (\mathbb{E}[N(A_h)])^2,$$

nous en déduisons que :

$$\mathbb{V}[P(h, r)] = X_2(h, r).$$

Ceci établit la formule (6.3.8).

3. Nous avons :

$$\begin{aligned}
 \int_{A_h} [f(t, s, r_t)]^2 dt ds &= \int_0^h ds \left[ \int_0^s [f(t, s, r_t)]^2 dt \right] \\
 &\stackrel{s=hy}{=} h \int_0^1 dy \left[ \int_0^{+\infty} \mathbb{1}_{t \leq hy} \left( [\phi_t(hy) - r_t]_+ \right)^2 dt \right].
 \end{aligned} \tag{6.3.20}$$

Lorsque  $s > t^*$  la fonction  $t \rightarrow \phi_t(s)$  est décroissante sur  $]t^*, s]$ , nous en déduisons :

$$\phi_s(s) \geq \alpha(1 - e^{-\tau t^*}) > 0, \quad 0 \leq t \leq s. \tag{6.3.21}$$

Donc :

$$\int_{A_h} [f(t, s, r_t)]^2 dt ds \geq h \int_0^1 dy \left[ \int_0^{+\infty} \mathbb{1}_{t \leq hy} \left( [\alpha(1 - e^{-\tau t^*}) - r_t]_+ \right)^2 dt \right]. \tag{6.3.22}$$

D'après (6.3.9) nous avons :

$$\int_0^{+\infty} [(a - r_t)_+]^2 dt = +\infty, \quad a > 0.$$

Ainsi :

$$\lim_{h \rightarrow +\infty} \int_{A_h} [f(t, s, r_t)]^2 dt ds = +\infty.$$

Ceci achève la preuve du Lemme 6.3.2.

□

### 6.3.3 Preuve de la Proposition 6.2.16

1. Il est clair que la preuve des formules (6.2.82) et (6.2.83) de la Proposition 6.2.16 est une conséquence directe du lemme suivant :

**Lemme 6.3.3.** *Soit  $k \geq 1$ .*

(a) *Pour tout  $h > 0$  :*

$$X_k(h, R) \stackrel{(d)}{=} \tilde{X}_k(h, R), \quad (6.3.23)$$

où la fonction  $X_k(h, R)$  est définie par (6.2.60) et :

$$\tilde{X}_k(h, R) := h^2 \lambda_2 \mathbb{E}[(M_1)^k] \int_0^1 dy \left[ \int_0^y [f(hx, hy, \exp[\mu hx + \sigma \sqrt{h} B_x])]^k dx \right]. \quad (6.3.24)$$

(b) *Si  $\mu < 0$  alors :*

$$\lim_{h \rightarrow +\infty} \frac{\tilde{X}_k(h, R)}{h^2} = C_k, \quad p.s., \quad (6.3.25)$$

où la constante  $C_k$  est définie par (6.2.20).

(c) *Si  $\mu = 0$  alors :*

$$\lim_{h \rightarrow +\infty} \frac{\tilde{X}_k(h, R)}{h^2} = 2C_k \int_0^1 (1-t) \mathbb{1}_{\{B_t < 0\}} dt, \quad p.s.. \quad (6.3.26)$$

La preuve de ce lemme sera effectuée à la fin de cette section.

2. Maintenant nous voulons prouver le résultat (6.2.84). Introduisons la notation suivante :

$$\tilde{C}_1(h) := \frac{\lambda_2}{h} \mathbb{E}[M_1] \int_0^1 du \int_0^{hu} \phi_t(hu) dt. \quad (6.3.27)$$

La preuve est basée sur le lemme suivant :

**Lemme 6.3.4.**

(a) *Si  $\mu < 0$  il vient que :*

$$\left[ \frac{X_1(h, R)}{h^2} - \tilde{C}_1(h) \right] h \xrightarrow[h \rightarrow +\infty]{p.s.} -\lambda_2 \mathbb{E}[M_1] \int_0^{+\infty} \inf(\alpha, \gamma R_t) dt. \quad (6.3.28)$$

(b) *D'autre part :*

$$\lim_{h \rightarrow +\infty} [\tilde{C}_1(h) - C_1] h = -2 \frac{e^{-\tau t^*}}{\tau} C_1. \quad (6.3.29)$$

La preuve de ce lemme sera effectuée à la fin de cette section. Maintenant nous pouvons prouver la formule (6.2.84). Puisque :

$$\left( \frac{X_1(h, R)}{h^2} - C_1 \right) h = \left[ \frac{X_1(h, R)}{h^2} - \tilde{C}_1(h) \right] h + [\tilde{C}_1(h) - C_1] h, \quad (6.3.30)$$

et en utilisant (6.3.28) et (6.3.29) nous en déduisons la formule (6.2.84).

Revenons à la preuve des Lemmes 6.3.3 et 6.3.4.

*Preuve du Lemme 6.3.3.*

1. D'après les formules (6.2.60), (6.2.3), (6.2.8) et en utilisant les changements de variables  $s = hy$  et  $t = hx$  il vient que :

$$\begin{aligned} \frac{X_k(h, R)}{\mathbb{E}[(M_1)^k]} &= \lambda_2 \int_0^h ds \left[ \int_0^s [f(t, s, R_t)]^k dt \right] \\ &= h^2 \lambda_2 \int_0^1 dy \left[ \int_0^y [f(hx, hy, \exp[\mu hx + \sigma B_{hx}])]^k dx \right]. \end{aligned} \quad (6.3.31)$$

Mais :

$$(B_{xh}, x \geq 0) \stackrel{(d)}{=} (\sqrt{h}B_x, x \geq 0). \quad (6.3.32)$$

En combinant (6.3.31) et (6.3.32) il suit que :

$$X_k(h, r) \stackrel{(d)}{=} \tilde{X}_k(h, R),$$

où  $\tilde{X}_k(h, R)$  est définie par (6.3.24). Ceci prouve la formule (6.3.23).

2. Le précédent changement de variable nous permet d'étudier le comportement asymptotique de  $\tilde{X}_k(h, R)$ . Nous souhaitons appliquer le théorème de convergence dominé. Avec (6.2.4) et (6.2.5) il suit que :

$$f(hx, hy, \exp[\sigma\sqrt{h}B_x + \mu hx]) = [\phi_{hx}(hy) - \gamma \exp\{\sigma\sqrt{h}B_x + \mu hx\}]_+.$$

Parce que  $x > 0$  et d'après (6.2.42) nous obtenons :

$$\begin{aligned} \lim_{h \rightarrow +\infty} f(hx, hy, \exp[\sigma\sqrt{h}B_x + \mu hx]) \\ = \alpha [\mathbb{1}_{\mu < 0} + \mathbb{1}_{\{\mu=0, B_x < 0\}}] + \mathbb{1}_{\{\mu=0, B_x=0\}}(\alpha - \gamma)_+, \text{ p.s..} \end{aligned}$$

D'autre part, l'inégalité (6.3.6) permet d'appliquer le théorème de convergence dominé. Nous avons si  $\mu < 0$  :

$$\begin{aligned} \lim_{h \rightarrow +\infty} \frac{\tilde{X}_k(h, R)}{h^2} &= \lambda_2 \mathbb{E}[(M_1)^k] \int_0^1 dy \left[ \int_0^y \alpha^k dx \right] \\ &= \lambda_2 \mathbb{E}[(M_1)^k] \frac{\alpha^k}{2} = C_k. \end{aligned}$$

Lorsque  $\mu = 0$  nous avons :

$$\begin{aligned} \lim_{h \rightarrow +\infty} \frac{\tilde{X}_k(h, R)}{h^2} &= \lambda_2 \mathbb{E}[(M_1)^k] \int_0^1 dy \int_0^y [\alpha \mathbb{1}_{\{B_x < 0\}} dx + (\alpha - \gamma)_+ \mathbb{1}_{\{B_x = 0\}}]^k dx \\ &= \lambda_2 \mathbb{E}[(M_1)^k] \int_0^1 (1-x) [\alpha^k \mathbb{1}_{\{B_x < 0\}} dx + [(\alpha - \gamma)_+]^k \mathbb{1}_{\{B_x = 0\}}] dx \end{aligned}$$

$$\begin{aligned}
 &= \lambda_2 \alpha^k \mathbb{E}[M_1] \int_0^1 (1-t) \mathbb{1}_{\{B_t < 0\}} dt \\
 &= 2C_k \int_0^1 (1-t) \mathbb{1}_{\{B_t < 0\}} dt, \quad p.s.,
 \end{aligned}$$

car  $\int_0^1 (1-t) \mathbb{1}_{\{B_t=0\}} dt$  est une variable aléatoire positive d'espérance nulle.

Ceci achève la preuve du Lemme 6.3.3.  $\square$

*Preuve du Lemme 6.3.4.*

1. D'après (6.2.60) et (6.3.27) il vient que :

$$\begin{aligned}
 \left[ \frac{X_1(h, R)}{h^2} - \tilde{C}_1(h) \right] h &= \lambda_2 \mathbb{E}[M_1] \left[ \frac{1}{h} \int_0^h ds \int_0^s [\phi_t(s) - \gamma R_t]_+ dt \right. \\
 &\quad \left. - \int_0^1 du \int_0^{hu} \phi_t(hu) dt \right] \\
 &\stackrel{s=hu}{=} \lambda_2 \mathbb{E}[M_1] \int_0^1 du \int_0^{hu} \left[ [\phi_t(hu) - \gamma R_t]_+ - \phi_t(hu) \right] dt \\
 &= -\lambda_2 \mathbb{E}[M_1] \int_0^1 du \int_0^{hu} \inf [\phi_t(hu), \gamma R_t] dt. \quad (6.3.33)
 \end{aligned}$$

Donc si  $\mu < 0$  il vient d'après (6.2.54) que :

$$\left[ \frac{X_1(h, R)}{h^2} - \tilde{C}_1(h) \right] h \xrightarrow[h \rightarrow +\infty]{p.s.} -\lambda_2 \mathbb{E}[M_1] \int_0^{+\infty} \inf (\alpha, \gamma R_t) dt,$$

car :

$$\int_0^{+\infty} \inf (\alpha, \gamma R_t) < +\infty.$$

Ce qui démontre la formule (6.3.28).

2. D'après (6.3.27) et (6.2.20) il vient que :

$$\begin{aligned}
 \left[ \tilde{C}_1(h) - C_1 \right] h &= \lambda_2 \mathbb{E}[M_1] \left[ \int_0^1 du \int_0^{hu} \phi_t(hu) dt - \frac{\alpha}{2} h \right] \\
 &= \lambda_2 \mathbb{E}[M_1] \left[ \int_0^1 du \int_0^{hu} [\phi_t(hu) - \alpha] dt \right]. \quad (6.3.34)
 \end{aligned}$$

En utilisant le résultat (6.2.89) il vient alors que :

$$\lim_{h \rightarrow +\infty} \left[ \tilde{C}_1(h) - C_1 \right] h = -2 \frac{e^{-\tau t^*}}{\tau} C_1. \quad (6.3.35)$$

Ce qui démontre la formule (6.3.29).  $\square$

### 6.3.4 Preuve de la Proposition 6.2.18

Soient  $\phi_1$  et  $\phi_2$  deux fonctions continues et bornées. Alors :

$$\begin{aligned}\Gamma_h &:= \mathbb{E} \left[ \phi_1 \left( \frac{P(h, R) - X_1(h, R)}{\sqrt{X_2(h, R)}} \right) \phi_2 \left( \left[ \frac{X_1(h, R)}{h^2} - C_1 \right] h \right) \right] \\ &= \mathbb{E} \left[ \tilde{\phi}_1(h, R) \phi_2 \left( \left[ \frac{X_1(h, R)}{h^2} - C_1 \right] h \right) \right],\end{aligned}\tag{6.3.36}$$

où :

$$\tilde{\phi}_1(h, r) := \mathbb{E} \left[ \phi_1 \left( \frac{P(h, r) - X_1(h, r)}{\sqrt{X_2(h, r)}} \right) \right].\tag{6.3.37}$$

En particulier d'après (6.2.71) il suit que :

$$\lim_{h \rightarrow +\infty} \tilde{\phi}_1(h, r) = \mathbb{E}[\phi_1(G)].\tag{6.3.38}$$

Puisque  $\phi_2$  est bornée et en utilisant (6.2.84), (6.3.36) et (6.3.37) il vient que :

$$\Gamma_h \xrightarrow{h \rightarrow +\infty} \mathbb{E} \left[ \phi_1(G) \phi_2 \left( -2C_1 \left[ \frac{e^{-\tau t^*}}{\tau} + \int_0^{+\infty} \inf \left( 1, \frac{\gamma}{\alpha} R_t \right) dt \right] \right) \right].$$

Ceci achève la preuve de la Proposition 6.2.18.

### 6.3.5 Preuve du Lemme 6.2.19

D'après (6.2.5) et pour  $h > t^*$  il vient que :

$$\begin{aligned}\int_0^1 dy \int_0^{hy} [\phi_t(hy) - \alpha] dt &= \int_{\frac{t^*}{h}}^1 dy \int_{t^*}^{hy} \alpha \left[ \frac{e^{\tau hy} - e^{\tau(t-t^*)}}{e^{\tau hy} - 1} - 1 \right] dt \\ &= \alpha \int_{\frac{t^*}{h}}^1 dy \int_{t^*}^{hy} \frac{1 - e^{\tau(t-t^*)}}{e^{\tau hy} - 1} dt \\ &= \alpha \int_{\frac{t^*}{h}}^1 \frac{dy}{e^{\tau hy} - 1} \left[ (hy - t^*) - \frac{e^{-\tau t^*}}{\tau} [e^{\tau hy} - e^{\tau t^*}] \right] dt \\ &= \alpha \left[ \int_{\frac{t^*}{h}}^1 \frac{hy - t^* + \frac{1}{\tau}(1 - e^{-\tau t^*})}{e^{\tau hy} - 1} dy - \frac{e^{-\tau t^*}}{\tau} \left( 1 - \frac{t^*}{h} \right) \right] \\ &= \alpha \left[ \frac{1}{h} \int_{t^*}^h \frac{s - t^* + \frac{1}{\tau}(1 - e^{-\tau t^*})}{e^{\tau s} - 1} ds - \frac{e^{-\tau t^*}}{\tau} \left( 1 - \frac{t^*}{h} \right) \right].\end{aligned}\tag{6.3.39}$$

Puisque :

$$\int_{t^*}^h \frac{s - t^* + \frac{1}{\tau}(1 - e^{-\tau t^*})}{e^{\tau s} - 1} ds < +\infty,\tag{6.3.40}$$

il suit que :

$$\lim_{h \rightarrow +\infty} \int_0^1 dy \int_0^{hy} [\phi_t(hy) - \alpha] dt = -\alpha \frac{e^{-\tau t^*}}{\tau}.\tag{6.3.41}$$

Ce qui achève la preuve du Lemme 6.2.19.





# Compléments numériques au calcul du besoin de solvabilité

---

Dans ce chapitre nous donnons quelques compléments numériques au chapitre précédent et en particulier aux Théorèmes 6.2.5 et 6.2.6.

Dans la Section 7.1 nous vérifierons que l'espérance asymptotique (Cf. la formule (6.2.24)) est proche de l'espérance théorique (Cf. la formule (5.4.2)) et de l'espérance empirique obtenue sur les simulations (Cf. la formule (5.4.7)). Nous analyserons l'influence des paramètres  $h$  et  $\mu$  sur l'espérance de la provision.

Dans la Section 7.2 nous nous intéresserons à la distribution empirique de la provision renormalisée :

$$\frac{P(h, R) - \mathbb{E}[P(h, R)]}{h\sqrt{C_2}}. \quad (7.0.1)$$

Pour ce faire nous utiliserons l'Algorithme 5.1 pour simuler la provision  $P(h, R)$ . Nous observerons l'influence de la valeur de  $\mu$  sur l'allure de la distribution empirique et montrerons que pour  $\mu$  "fortement" négatif l'allure de la distribution empirique de la provision renormalisée ressemble à celle d'une gaussienne. Nous montrerons que l'influence de l'intégrale d'une fonction du mouvement Brownien géométrique (qui s'ajoute à la v.a. gaussienne) est quasiment nulle lorsque  $\mu$  tend vers moins l'infini.

Enfin dans la Section 7.3 nous définirons le besoin de solvabilité empirique et nous donnerons une valeur numérique à cette quantité. Nous observerons finalement comment cette quantité fluctue selon le signe de  $\mu$ .

## 7.1 Espérances asymptotique et théorique

Dans le Chapitre 6 nous avons supposé que pour  $h = 276$  mois le comportement de la provision renormalisée (Cf. (7.0.1)) était proche de son comportement asymptotique lorsque  $h \rightarrow +\infty$ . Vérifions si c'est le cas en pratique. Considérons  $\mu = -0,07$  et traçons les courbes :

- Espérance asymptotique (courbe bleue) en fonction de  $h$  (Cf. (6.2.24)), i.e. :

$$\left[ h, h^2 \frac{\hat{\lambda}_2 \hat{\alpha} \bar{m}}{2} \right], \quad (7.1.1)$$

où  $\hat{\lambda}_2, \hat{\alpha}$  et  $\bar{m}$  sont définis dans la Section 5.4.2 (Cf. Figure 5.22).

- Espérance théorique (courbe noire) en fonction de  $h$ , i.e. :

$$\left[ h, \mathbb{E}^{theo}[P(h, R)] \right], \quad (7.1.2)$$

où  $\mathbb{E}^{theo}[P(h, R)]$  est la moyenne théorique de la provision calculée à la date 0 et concernant les prêts souscrits l'année 0, i.e. :

$$\begin{aligned} \mathbb{E}^{theo}[P(h, R)] = \hat{\lambda}_2 \bar{m} \times \int_{A_h} \left\{ \hat{\phi}_t(s) \Phi \left[ \frac{\ln \left[ \frac{\hat{\phi}_t(s)}{\hat{\gamma}} \right] - \hat{\mu}t}{\hat{\sigma} \sqrt{t}} \right] \right. \\ \left. - \hat{\gamma} e^{(\hat{\mu} + \frac{\hat{\sigma}^2}{2})t} \Phi \left[ \frac{\ln \left[ \frac{\hat{\phi}_t(s)}{\hat{\gamma}} \right] - (\hat{\mu} + \hat{\sigma}^2)t}{\hat{\sigma} \sqrt{t}} \right] \right\}, \end{aligned} \quad (7.1.3)$$

où  $\hat{\gamma}$ ,  $\hat{\mu}$ ,  $\hat{\sigma}$  et  $\hat{\phi}_t(s)$  sont définis dans la Section 5.4.2.

- La droite verticale jaune indiquera la valeur des espérances pour  $h = 276$  mois.

Nous ne traçons pas l'espérance empirique car nous avons vu dans la Section 5.4.2 que celle ci était très proche de l'espérance théorique. Nous donnerons également le ratio "espérance asymptotique sur espérance théorique" en rouge. Ce ratio sera indiqué sur le second axe des ordonnées, à droite.

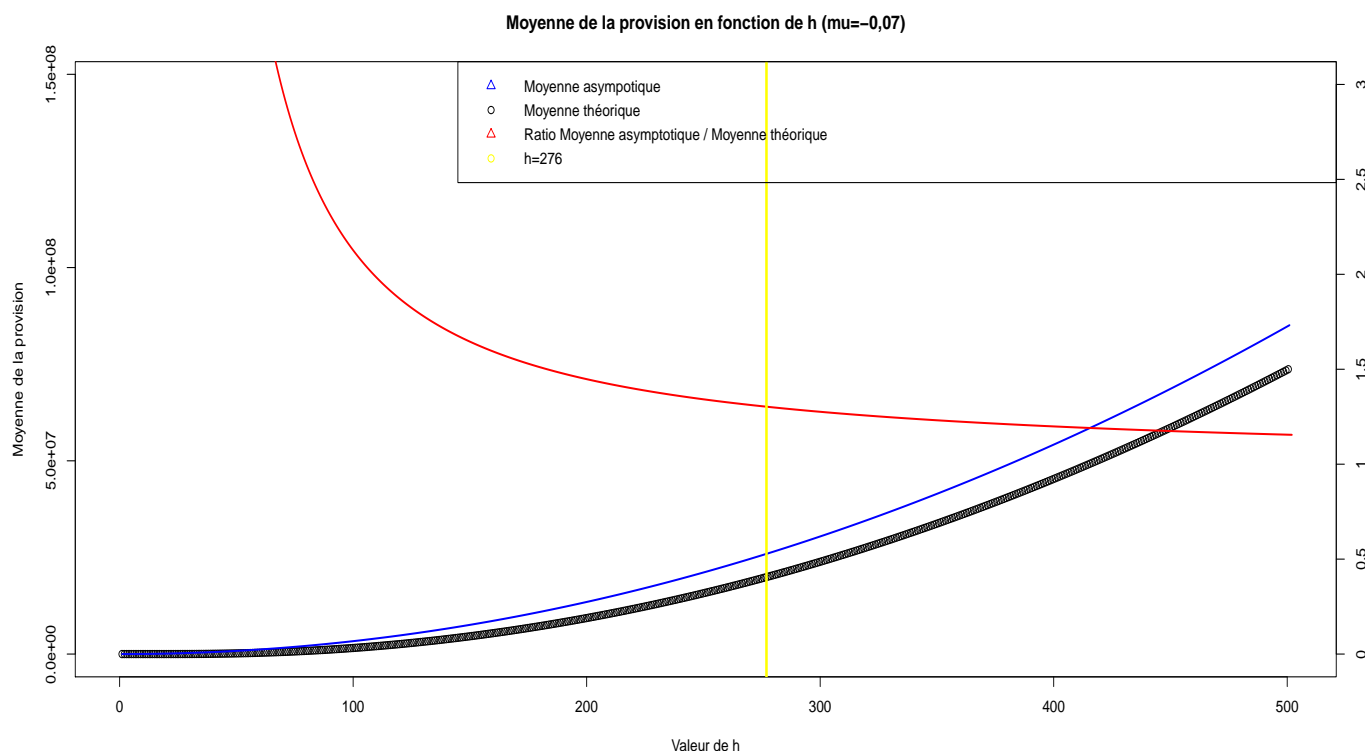


FIGURE 7.1 – Espérance de la provision en fonction de  $h$ .

## 7.1 Espérances asymptotique et théorique

Nous constatons que l'espérance asymptotique devient proche de l'espérance théorique plus la valeur de  $h$  est importante. Pour  $h = 276$  mois le ratio “espérance asymptotique sur espérance théorique” vaut environ 1,2.

Fixons  $h = 276$  mois. Nous souhaitons voir comment fluctue l'espérance asymptotique de la provision (Cf. (6.2.24)) et l'espérance théorique (Cf. (7.1.3)) en fonction de  $\mu$ . Traçons pour ce faire les courbes suivantes :

- Espérance asymptotique en fonction de  $\mu$  (Cf. (6.2.24)), i.e. :

$$\left[ \mu, \hat{\lambda}_2 \hat{\alpha} \bar{m} \hat{h}^2 \left( \frac{1}{2} \mathbb{1}_{\{\mu < 0\}} + \frac{1}{4} \mathbb{1}_{\{\mu = 0\}} \right) \right], \quad (7.1.4)$$

où  $\hat{\lambda}_2, \hat{\alpha}, \bar{m}$  et  $\hat{h}$  sont définis dans la Section 5.4.2.

- Espérance théorique (en noir) en fonction de  $\mu$ , i.e. :

$$\left[ \mu, \mathbb{E}^{theo}[P(h, R)] \right], \quad (7.1.5)$$

où  $\mathbb{E}^{theo}[P(h, R)]$  est définie par (7.1.3).

- La droite verticale jaune représentera la valeur des espérances pour la valeur de  $\mu$  estimée dans la Section 5.4.2, i.e. :  $\mu = -0,07$ .

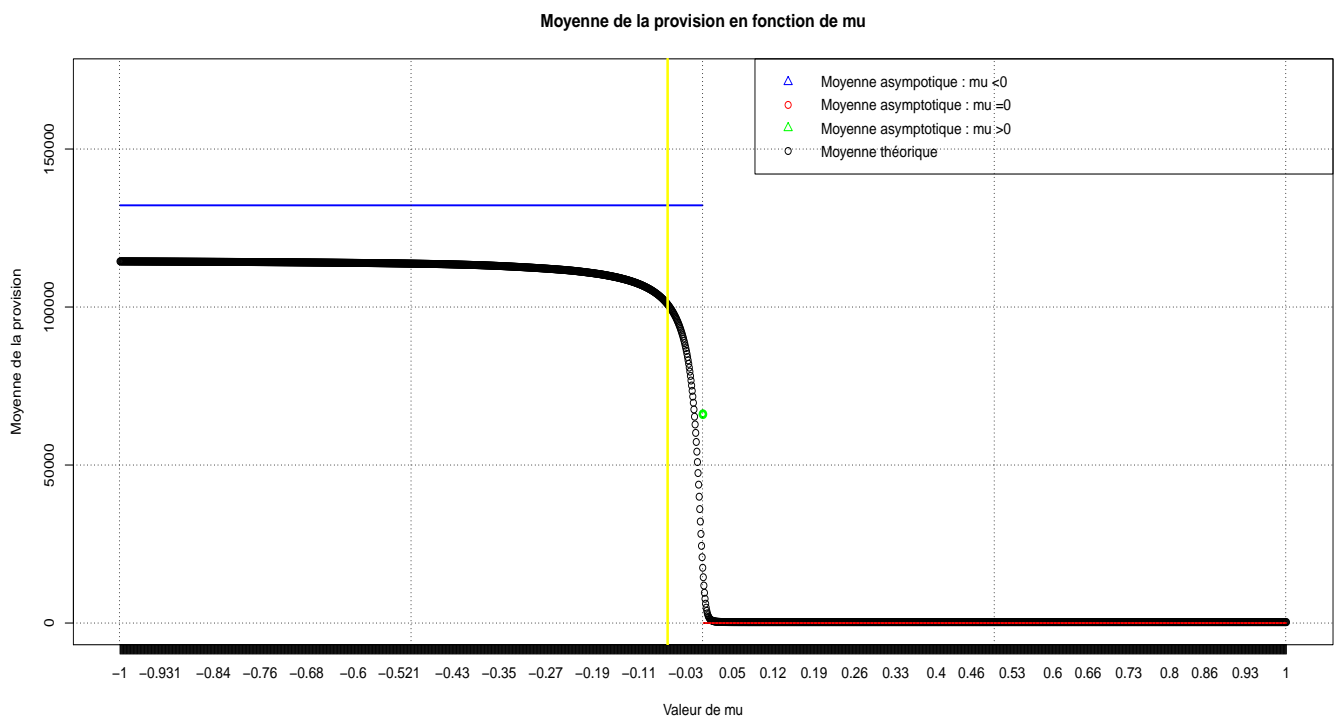


FIGURE 7.2 – Espérance de la provision en fonction de  $\mu$ .

Nous constatons que :

1. L'espérance asymptotique est proche de l'espérance théorique.
2. Les espérances théorique et asymptotique se comportent différemment selon le signe de  $\mu$ . Il y a trois régimes différents :
  - (a) Lorsque  $\mu < 0$  : la moyenne de la provision converge vers une constante strictement positive,
  - (b) Lorsque  $\mu$  est proche de 0 : la moyenne de la provision se rapproche de 0 ; il s'agit d'un état transitoire,
  - (c) Lorsque  $\mu > 0$  : la moyenne de la provision tend vers 0.

## 7.2 Distribution empirique de la provision

Dans le Théorème 6.2.5 nous avons montré que la provision, convenablement renormalisée, converge vers la somme de deux variables dont l'une est gaussienne, plus précisément :

$$\frac{P(h, R) - \mathbb{E}[P(h, R)]}{h\sqrt{C_2}} \xrightarrow{h \rightarrow +\infty} X^* + \kappa, \quad (7.2.1)$$

où les coefficients  $C_k$  et  $\kappa$  sont définis par (6.2.20) et (6.2.22) et :

$$X^* := G - \frac{2C_1}{\sqrt{C_2}} \left[ \frac{e^{-\tau t^*}}{\tau} + \int_0^{+\infty} \inf \left( 1, \frac{\gamma}{\alpha} R_t \right) dt \right]. \quad (7.2.2)$$

D'après (7.2.2) la variable  $X^*$  est la somme d'une constante et de deux variables :

1. Une variable gaussienne centrée et réduite,
2. Une intégrale d'une fonction d'un mouvement Brownien géométrique.

Dans cette section nous souhaitons montrer que l'influence de la seconde variable est faible lorsque  $\mu$  est négatif. Pour ce faire nous considérons la distribution empirique de :

$$P(h, R) - \mathbb{E}[P(h, R)]. \quad (7.2.3)$$

Nous souhaitons voir comment évolue cette distribution empirique en fonction de la valeur de  $\mu$ . A l'aide de l'Algorithme 5.1 il est possible d'obtenir un échantillon de taille  $l = 5\,000$  de la provision  $P(h, R)$ . Nous noterons  $(p_1^\mu, \dots, p_l^\mu)$  cet échantillon. Traçons l'histogramme empirique de :

$$\left( p_1^\mu - \frac{1}{l} \sum_{j=1}^l p_j^\mu, \dots, p_l^\mu - \frac{1}{l} \sum_{j=1}^l p_j^\mu \right), \quad (7.2.4)$$

pour différentes valeurs de  $\mu$  ( $\mu = -0, 1, -0, 07, -0, 04, -0, 01$ ) :

- La courbe bleue représente la densité empirique,
- La courbe rouge représente la densité d'une v.a. gaussienne centrée et d'écart type  $\hat{h}\sqrt{\hat{C}_2}$ .

## 7.2 Distribution empirique de la provision

Commençons par le cas  $\mu = -0,01$  :

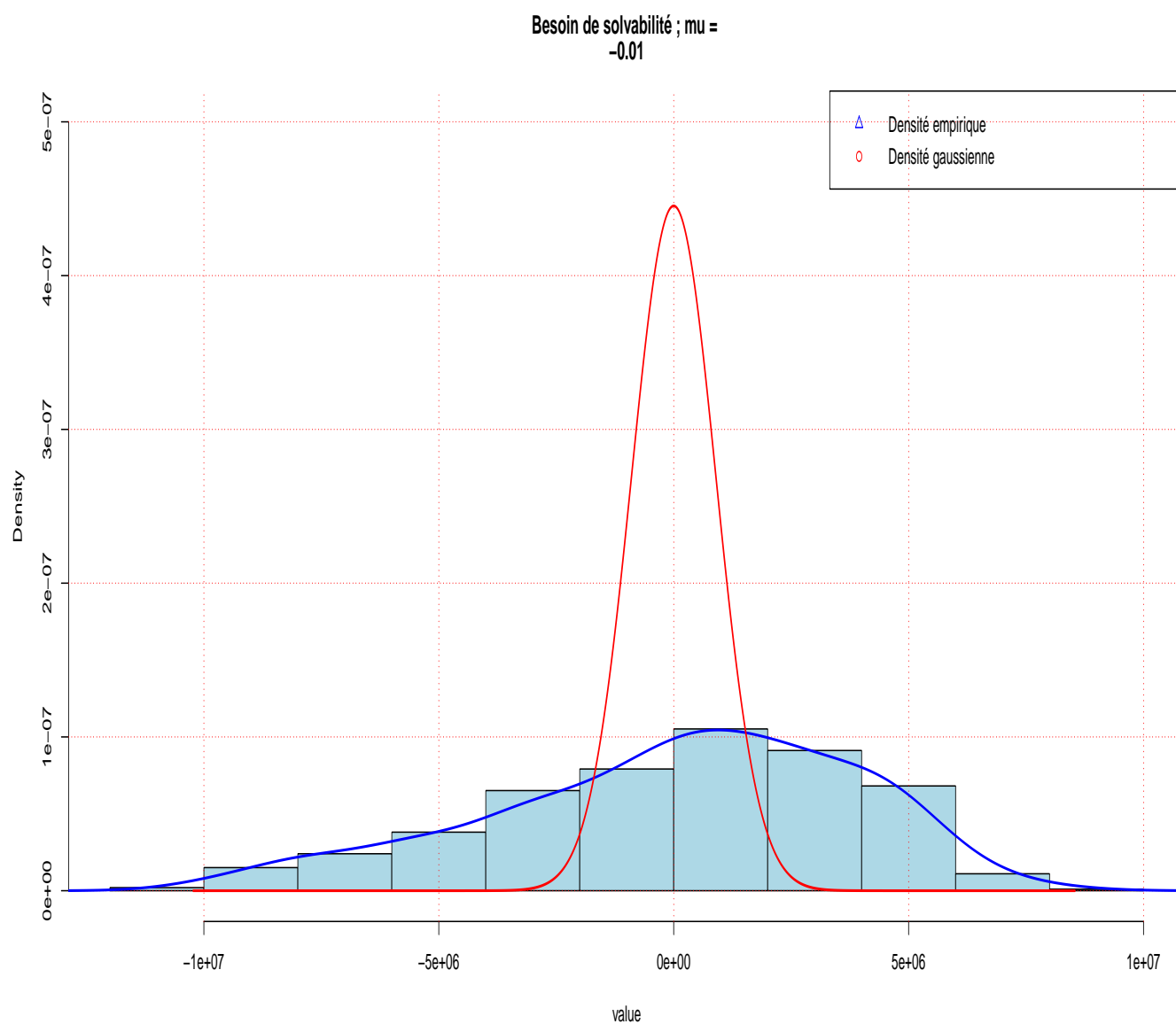


FIGURE 7.3 – Distribution empirique ;  $\mu = -0,01$ .

L'allure de la distribution empirique n'est clairement pas gaussienne. Les tests de normalité (Jarque- Bera, etc.) confirment cette observation tout comme le QQ-plot. Rappelons qu'un QQ-plot est un graphe représentant les quantiles observés en fonction des quantiles d'une v.a. gaussienne. Sur le prochain QQ-plot, nous trouvons en abscisse les quantiles empiriques de l'échantillon (7.2.4) et en ordonnée les quantiles d'une v.a. gaussienne centrée et d'écart type  $\hat{h}\sqrt{\hat{C}_2}$ .

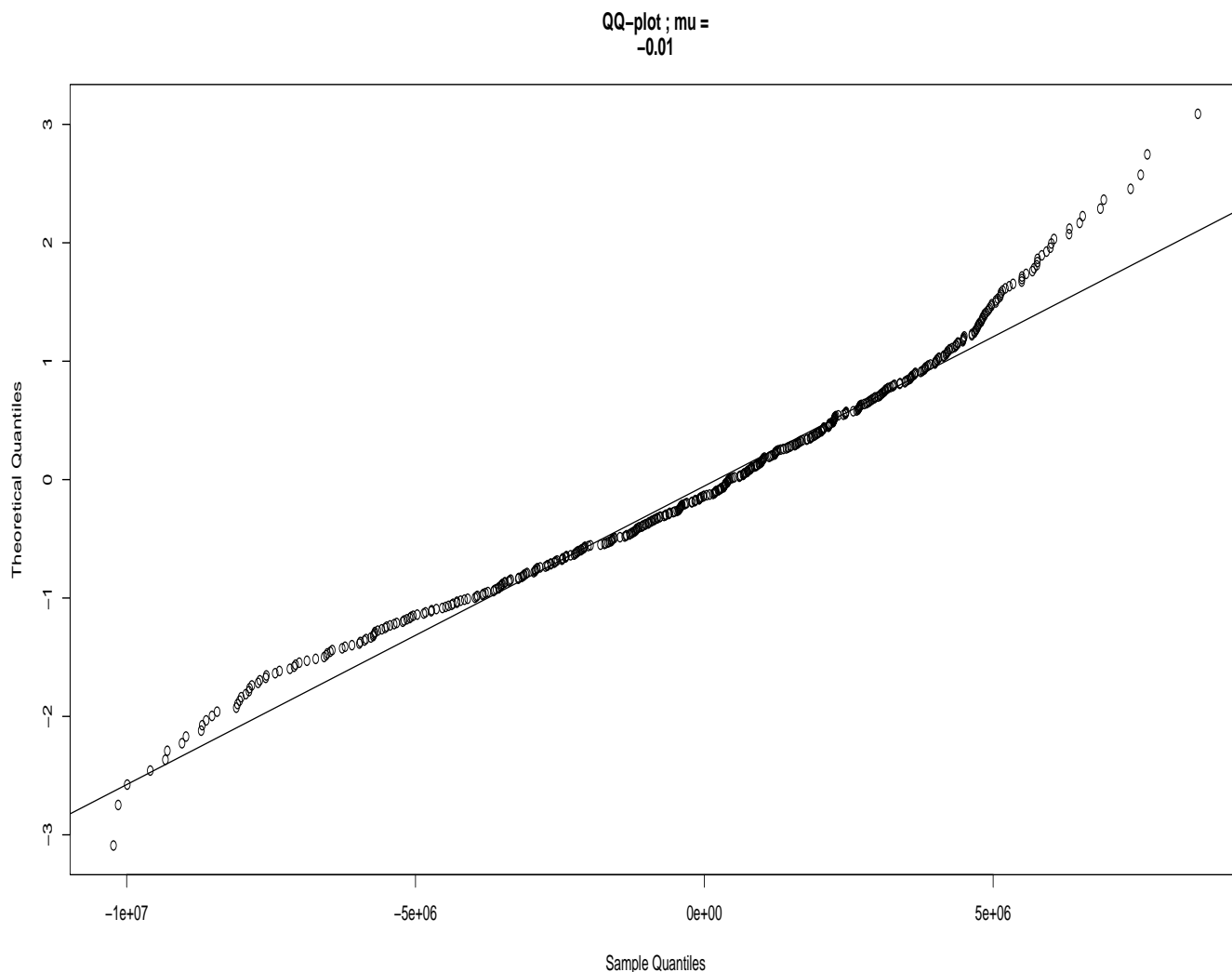


FIGURE 7.4 – QQ plot ;  $\mu = -0,01$ .

Comme le montre le QQ-plot la courbe formée par les quantiles empiriques et la droite des quantiles gaussiens ne coïncident pas, donc la distribution n'est pas gaussienne. Ainsi l'influence du terme additif dans  $X^*$ , i.e. l'intégrale d'une fonction du mouvement Brownien géométrique, n'est pas négligeable pour  $\mu = -0,01$ .

Les mêmes observations peuvent être réalisées pour le cas  $\mu = -0,04$ . En effet :

## 7.2 Distribution empirique de la provision

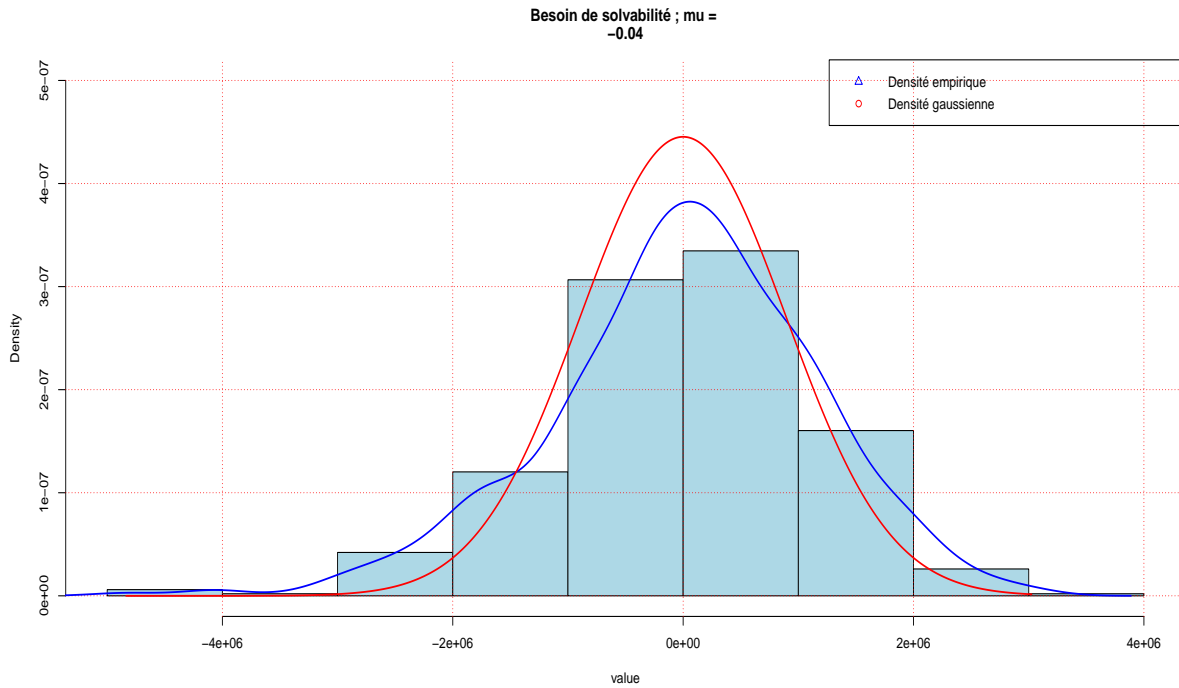


FIGURE 7.5 – Distribution empirique ;  $\mu = -0,04$ .

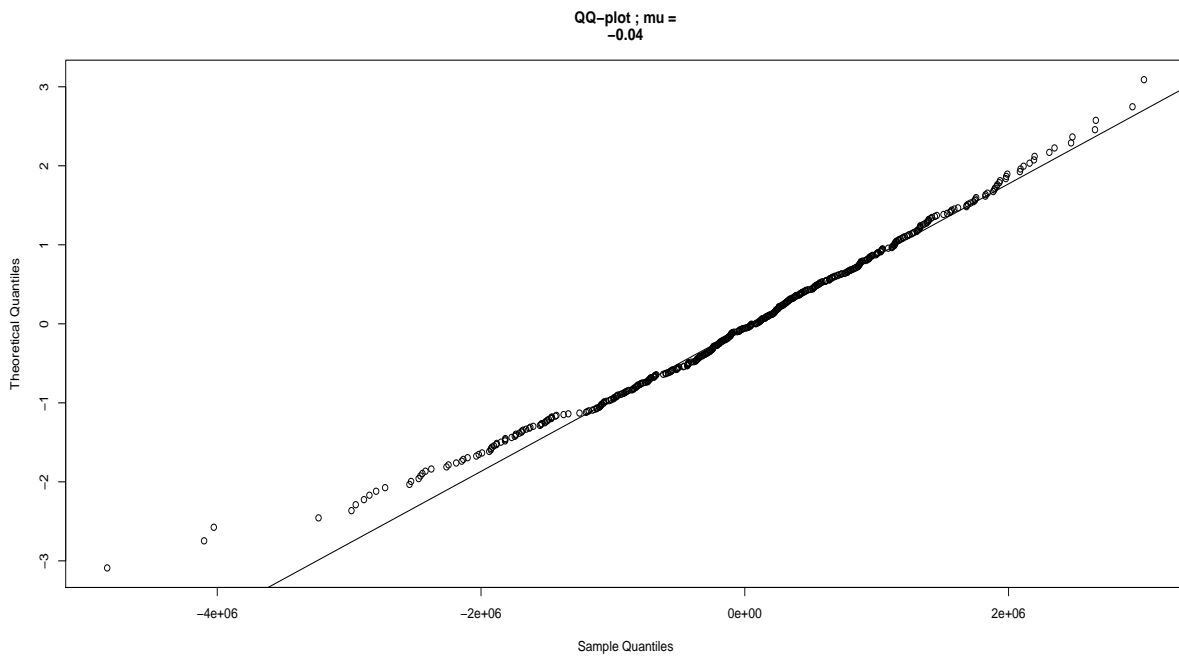


FIGURE 7.6 – QQ-plot ;  $\mu = -0,04$ .

En revanche pour  $\mu = -0,07$  et  $\mu = -0,1$  la densité de l'espérance centrée est proche de celle d'une v.a. gaussienne comme le montre les deux figures suivantes :

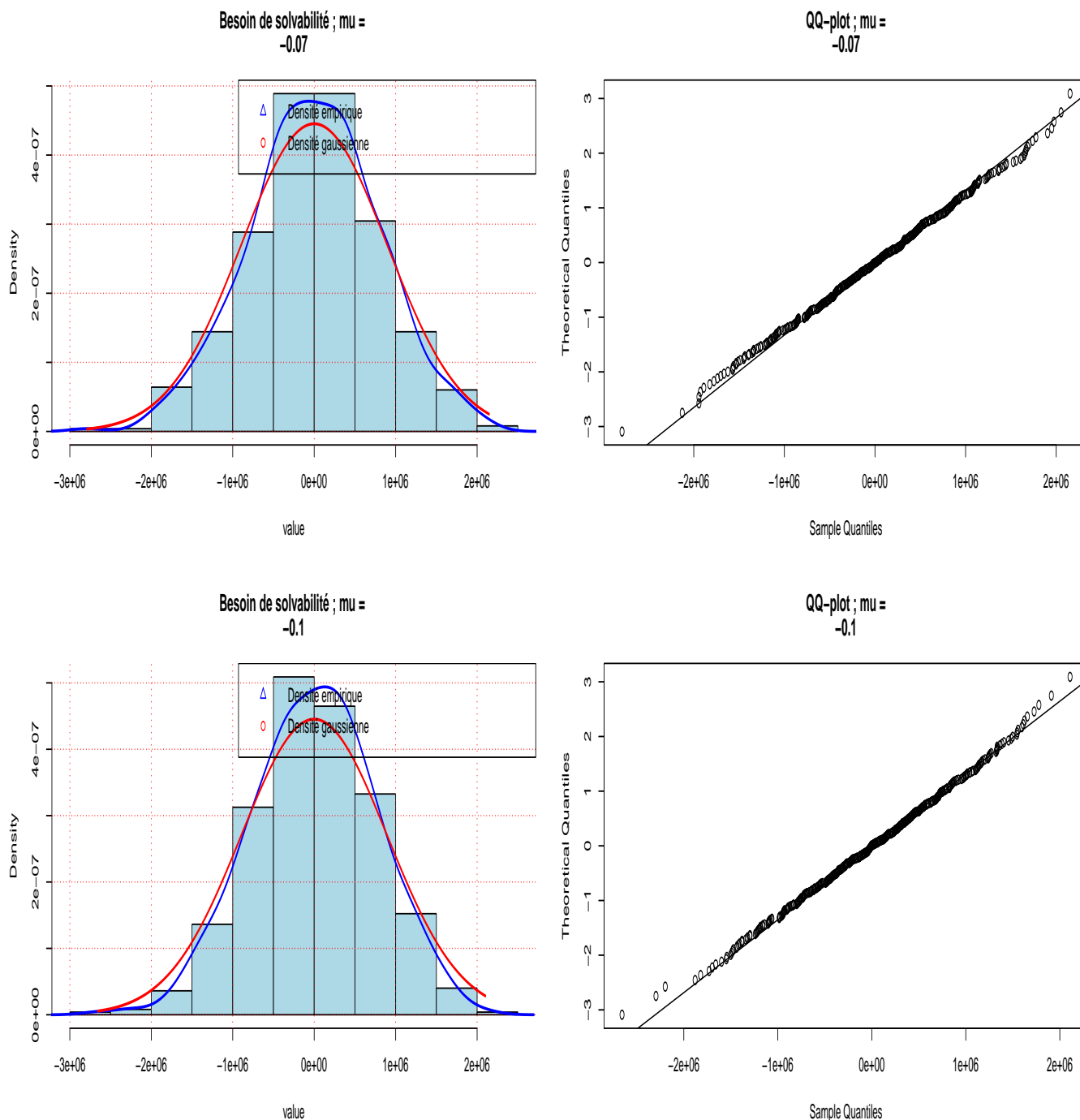


FIGURE 7.7 – Distribution empirique et QQ-plot ;  $\mu = -0,07$  et  $\mu = -0,1$ .



Lorsque  $\mu$  est négatif (loin de 0) la distribution empirique de  $P(h, R) - \mathbb{E}[P(h, R)]$  semble gaussienne.

Notons qu'il est attendu que l'influence de  $\int_0^{+\infty} \inf(1, \frac{\gamma}{\alpha} R_t) dt$  dans la loi limite  $P(h, R)$  est négligeable lorsque  $\mu \rightarrow -\infty$ . En effet :

$$\lim_{\mu \rightarrow -\infty} \int_0^{+\infty} \inf(1, \frac{\gamma}{\alpha} \exp[\mu t + \sigma B_t]) dt = 0. \quad (7.2.5)$$

Donc le terme additif dans la loi limite de la provision renormalisée ne joue pas de rôle lorsque  $\mu$  est très négatif. L'intérêt des simulations est de mettre en évidence que dès que  $\mu \leq -0,07$ , l'influence de  $X^*$  est négligeable.

### 7.3 Besoin de solvabilité empirique

Dans cette section nous calculons numériquement le besoin de solvabilité (BS). Comme énoncé dans le Chapitre 6, la méthode asymptotique ne permet pas de calculer numériquement le BS. En effet la formule (6.2.28) dit que :

$$BS^*(h) = h \left[ \mathcal{Q}_{X^*}(99, 5\%) + \kappa \right] \sqrt{\frac{\lambda_2 \alpha^2 \mathbb{E}[(M_1)^2]}{2}}, \quad (7.3.1)$$

où  $\mathbb{E}[(M_1)^2]$  est le moment d'ordre 2 du montant des prêts,  $\lambda_2, \alpha, \kappa$  sont définis par (6.2.8), (6.2.6), (6.2.22) et  $\mathcal{Q}_{X^*}(99, 5\%)$  (Cf. (6.2.13)) est le quantile d'ordre 99, 5% de la v.a.  $X^*$  définie par (7.2.2).

Puisque la loi de  $\int_0^{+\infty} \inf(1, \frac{\gamma}{\alpha} R_t) dt$  est inconnue, la formule (7.3.1) ne nous permet pas de calculer numériquement le BS. Puisque  $h$  est grand il est possible de réaliser des simulations de Monte Carlo. En effet nous avons vu dans les Chapitres 4 et 5 qu'il est possible de simuler la provision : l'Algorithme 5.1 nous fournit une méthode pour réaliser de telles simulations. Notons  $p^{\mu, l}$  l'échantillon de taille  $l = 5\,000$  obtenu via l'Algorithme 5.1, i.e. :

$$p^{\mu, l} = (p_1^\mu, \dots, p_l^\mu). \quad (7.3.2)$$

D'après la Définition 6.2.2 et la Remarque 6.2.3 (Cf. (6.2.14)), nous définissons le besoin de solvabilité empirique, noté  $BS_\mu^{emp}$ , de la manière suivante :

$$BS_\mu^{emp} := \inf \left\{ t : F_{p^{\mu, l}}(t) \geq 99, 5\% \right\} - \frac{1}{l} \sum_{j=1}^l p_j^\mu. \quad (7.3.3)$$

où  $F_{p^{\mu, l}}(t)$  est la fonction de répartition empirique du  $l$ -échantillon simulé.

Nous souhaitons voir comment évolue le besoin de solvabilité empirique en fonction de la valeur de  $\mu$ . Pour ce faire, nous traçons la courbe, noire,  $(\mu, BS_\mu^{emp})$  pour  $\mu \in [-1, 1]$  :

- $\mu$  sera représenté en abscisse,

- $BS_{\mu}^{emp}$  figurera en ordonnée.
- La droite verticale jaune représentera la valeur du besoin de solvabilité empirique pour la valeur du paramètre  $\mu$  estimé dans la Section 5.4.2, i.e.  $\mu = -0,07$ .

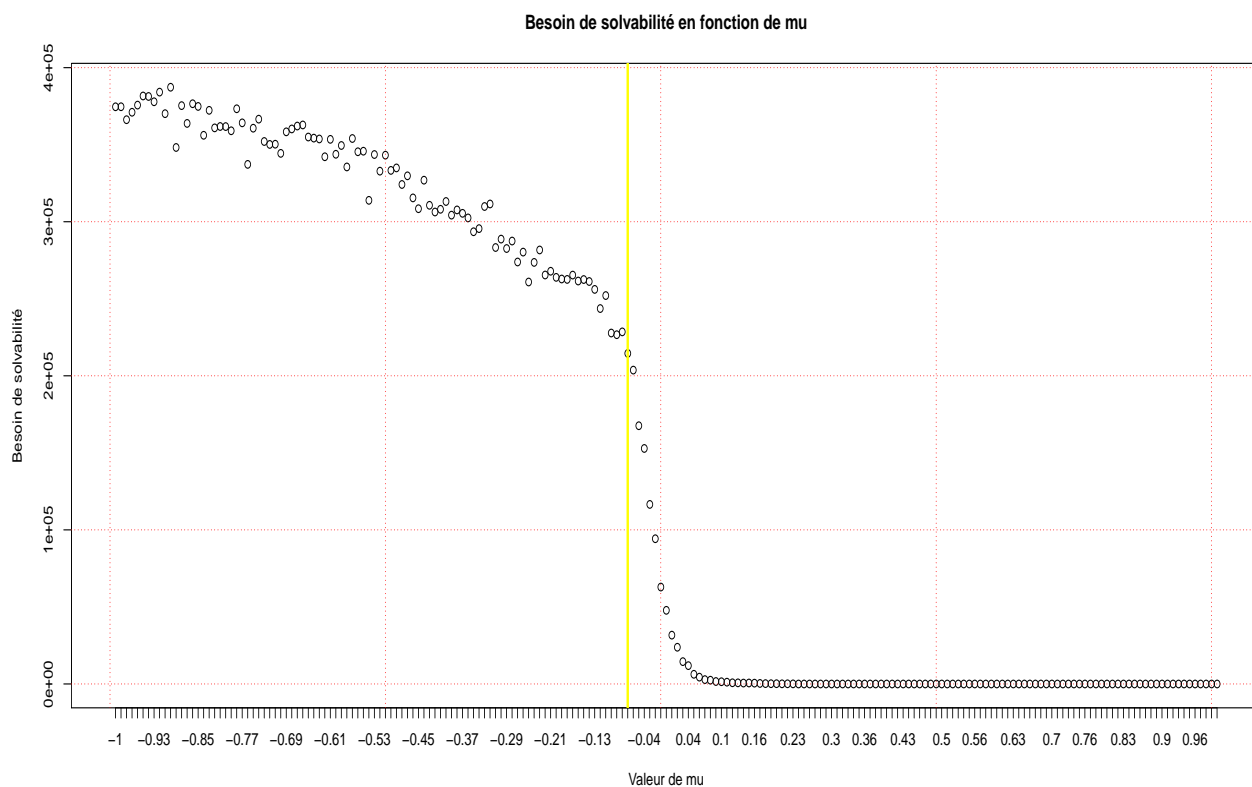


FIGURE 7.8 – Besoin de solvabilité en fonction de  $\mu$ .

Nous distinguons trois régimes différents dans le précédent graphique :

1. Lorsque  $\mu < 0$  : le besoin de solvabilité empirique se “rapproche” d’une constante positive lorsque  $\mu$  devient très négatif.
2. Lorsque  $\mu$  est proche de 0 : le besoin de solvabilité empirique se rapproche de 0. Il s’agit d’un état transitoire.
3. Lorsque  $\mu > 0$  : le besoin de solvabilité se rapproche de 0 lorsque  $\mu$  augmente.



# Bibliographie

- [1] *Technical Specification on the Long Term Guarantee Assessment*. EIOPA, 2013.
- [2] Hansjorg Albrecher and Onno J. Boxma. A ruin model with dependence between claim sizes and claim intervals. *Insurance : Mathematics and Economics*, 35(2) :245 – 254, 2004. Papers presented at the 7th {IME} Conference.
- [3] Hansjorg Albrecher, Corina Constantinescu, and Stephane Loisel. Explicit ruin formulas for models with dependence among risks. *Insurance : Mathematics and Economics*, 48(2) :265 – 270, 2011.
- [4] Aurélien Alfonsi. *Modélisation en risque de crédit. Calibration et discrétisation de modèles financiers*. PhD thesis, Ecole nationale des ponts et chaussées, 2006.
- [5] G. Allaire. *Analyse numérique et optimisation : une introduction à la modélisation mathématique et à la simulation numérique*. Mathématiques appliquées. École polytechnique, 2005.
- [6] Dassios Angelos and Jang Jiwook. A double shot noise process and its application in insurance. *Journal of Mathematics and System Science* 2, pages 82–93, 2012.
- [7] Elja Arjas. The claims reserving problem in non-life insurance : some structural ideas. *Astin Bulletin*, Vol. 19 :139 – 152, 1989.
- [8] Philippe Barbe, Anne-Laure Fougères, and Christian Genest. On the tail behaviour of sums of dependent risks. *ASTIN Bulletin*, 36(2) :361 – 373, 2006.
- [9] Glen Barnett and Benjamin Zehnwrith. Best estimates for reserves. *Proceedings of the Casualty Actuarial Society*, pages 245 – 303, 2000.
- [10] Gunnar Benktander. An approach to credibility in calculation ibnr for casualty excess reinsurance. *The Actuarial Review*, page 7, April 1976.
- [11] J. Bertoin and M. Yor. On the entire moments of self-similar markov process and exponential functionals of lévy processes. *Ann. Fac. Sci. Toulouse Math. (6)*, 11(1) :33–45, 2002.
- [12] Romain Biard. *Dépendance et événements extrêmes en théorie de la ruine : étude univariée et multivariée, problèmes d'allocation optimale*. Theses, Université Claude Bernard - Lyon I, October 2010.
- [13] Romain Biard, Claude Lefèvre, and Stéphane Loisel. Impact of correlation crises in risk theory : Asymptotics of finite-time ruin probabilities for heavy-tailed claim amounts when some independence and stationarity assumptions are relaxed. *Insurance : Mathematics and Economics*, 43(3) :412 – 421, 2008.

## BIBLIOGRAPHIE

---

- [14] Ronald L. Bornhuetter and Ronald E. Ferguson. The acutary and ibnr. *Proceedings of the Casualty Actuarial Society*, LIX :181–195, 1972.
- [15] Mathieu Boudreault. On a risk model with dependence between interclaim arrivals and claim sizes. *Scandinavian actuarial journal : Actuarial Society of Finland*, 2006.
- [16] Alexandre Boumezoued, Yoboua Angoua, Laurent Devineau, and Jean Philippe Boisseau. One-year reserve risk including a tail factor : closed formula and bootstrap approaches. Papers, arXiv.org, 2012.
- [17] Hans Buhlmann. Estimation of ibnr reserves by the methods chain ladder, cape cod and complimentary loss ratio. *International Summer School*, 1983.
- [18] Stavros Christofides. *Regression Models Based on Logincremental Payments*. Claims Reserving Manual. Institute and faculty of actuaries, 1990.
- [19] Erhan Çinlar. *Introduction to stochastic processes*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1975.
- [20] Anne Claude and Gérard Chaigneau. *Mathématiques financières*. Ellipses Marketing, 2010.
- [21] Christiane Coccozza-Thivent. *Processus stochastiques et fiabilité des systèmes*, volume 28 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin, 1997.
- [22] Gordon W. Crawford and Eric Rosenblatt. Efficient mortgage default option exercise : Evidence from loss severity. *Journal of Real Estate Research*, 10(5) :543–556, 1995.
- [23] Michel Denuit and Arthur Charpentier. *Mathématiques de l'assurance non-vie : Tarification et provisionnement*. Collection "Economie et statistiques avancées". Economica, 2005.
- [24] Jan Dhaene. *Solvency Capital, Risk Measures and Comonotonicity : A Review*. Katholieke Universiteit Leuven, [Faculty of Economics and Applied Economics], Department of Applied Economics, 2004.
- [25] Daniel Dufresne. The distribution of a perpetuity, with applications to risk theory and pension funding. *Scand. Actuar. J.*, 1990.
- [26] The european parliament and the council of the european union. *Directive 2009/138/EC of the european parliament and of the council of 25 November 2009 on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II)*. Official Journal of the European Union, 25 November 2009.

- 
- [27] The european parliament and the council of the european union. *Directive 2013/36/EU of the european parliament and of the council of 26 June 2013 on access to the activity of credit institutions and the prudential supervision of credit institutions and investment firms, amending Directive 2002/87/EC and repealing Directives 2006/48/EC and 2006/49/EC*. Official Journal of the European Union, 26 june 2013.
- [28] Spencer M. Gluck and Gary Venter. Stochastic trend models in casualty and life insurance. *Enterprise Risk Management Symposium*, 2009.
- [29] Joakim Hertig. *A Statistical Approach to IBNR-reserves in Marine Reinsurance*. Laboratory of Actuarial Mathematics, University of Copenhagen, Working Paper. Laboratory of Actuarial Mathematics, University of Copenhagen, 1983.
- [30] Douglas Hodes, Sholom Feldblum, and Gary Blumsohn. Workers compensation reserve uncertainty. *Proceedings of the Casualty Actuarial Society*, pages 263 – 392, 1999.
- [31] John Hull. *Options, Futures and Other Derivatives*. Prentice Hall finance series. Pearson/Prentice Hall, 2009.
- [32] Selma Jaziri. Méthodes de provisionnement non-vie et risque de réserve à un an. *ISFA*, 2011.
- [33] Christian Roholte Larsen. An individual claims reserving model. *Astin Bulletin*, Vol. 37 :95 – 115, 2007.
- [34] Thomas Mack. Distribution-free calculation of the standard error of chain ladder reserve estimates. *Astin Bulletin*, Vol. 23 :213 – 25, 1993.
- [35] Thomas Mack. The standard error of chain ladder reserve estimates : Recursive calculation and inclusion of a tail factor. *Astin Bulletin*, Vol. 29(2) :361 – 266, 1999.
- [36] William Margrabe. The value of an option to exchange one asset for another. *Journal of Finance*, 33(1) :177–86, March 1978.
- [37] Robert C Merton. On the Pricing of Corporate Debt : The Risk Structure of Interest Rates. *Journal of Finance*, 29(2) :449–70, May 1974.
- [38] Ragnar Norberg. Prediction of outstanding liabilities in non life insurance. *Astin Bulletin*, Vol. 23 :95 – 115, 1989.
- [39] Ivan Nourdin and Giovanni Peccati. *Normal approximations with Malliavin calculus : from Stein’s method to universality*. Cambridge tracts in mathematics. Cambridge University Press, Cambridge, New York, 2012. Informations complémentaires sur la publication à l’adresse <http://www.cambridge.org/9781107017771>.
- [40] E. Ohlsson and B. Johansson. *Non-Life Insurance Pricing with Generalized Linear Models*. EAA Series. Springer, 2010.

- [41] Esbjorn Ohlsson and Jan Lauzeningsks. The one-year non-life insurance risk. *Insurance : Mathematics and Economics*, 45(2) :203–208, 2009.
- [42] Robert Van Order. Modeling and evaluating the credit risk of mortgage loans : a primer. *The Journal of Risk Model Validation*, Vol. 2 :63 – 82, 2008.
- [43] Christian Partrat. *Provisionnement technique en Assurance non-vie : Perspectives actuarielles modernes*. Assurance, audit, actuariat. Economica, 2007.
- [44] Mickael Perrin. Calibration des undertaking specific parameters et leurs impacts sur les fonds propres. *EURIA*, 2012.
- [45] Mathieu Pigeon, Katrien Antonio, and Michel Denuit. Individual loss reserving with the multivariate skew normal distribution. Open access publications from katholieke universiteit leuven, Katholieke Universiteit Leuven, December 2012.
- [46] Frédéric Planchet, Quentin Guibert, and Marc Juillard. Un cadre de référence pour un modèle interne partiel en assurance de personnes. Post-print, HAL, 2010.
- [47] Frédéric Planchet, Quentin Guibert, and Marc Juillard. Measuring uncertainty of solvency coverage ratio in orsa for non-life insurance. *European Actuarial Journal*, 2(2) :205–226, 2012.
- [48] Gerhard Quarg and Thomas Mack. Munich chain ladder : A reserving method that reduces the gap between ibnr projections based on paid losses and ibnr projections based on incurred losses. *Variance*, Vol. 2(2) :266 – 299, 2008.
- [49] Filemon Ramirez-perez, Robert Serfling, and Edo Mex Cp. Asymptotic normality of shot noise on poisson cluster processes with cluster marks. Technical report, 2001.
- [50] Matthias Reitzner and Matthias Schulte. Central limit theorems for  $u$ -statistics of poisson point processes. *The Annals of Probability*, 41(6) :3879–3909, 11 2013.
- [51] A.E. Renshaw and R.J. Verrall. A stochastic model underlying the chain-ladder technique. *British Actuarial Journal*, 4 :903–923, 10 1998.
- [52] Tomasz Rolski, Hanspeter Schmidli, Volker Schmidt, and Jozef Teugels. *Stochastic processes for insurance and finance*. Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., Chichester, 1999.
- [53] Paavo Salminen and Marc Yor. Perpetual integral functionals as hitting times. 2004.
- [54] Thorsten Schmidt. Catastrophe insurance modeled by shot-noise processes. *Risks*, 2(1) :3–24, 2014.
- [55] Moez Souissi. An option pricing approach to stress-testing the canadian mortgage portfolio. *Paper presented at the Bank of Canada Annual Economic Conference on Developing a Framework to Assess Financial Stability, Ottawa, November 7-8, 2007*, 2007.

- 
- [56] Greg Taylor, Grainne McGuire, and James Sullivan. Centre for Actuarial Studies. *Individual Claim Loss Reserving Conditioned by Case Estimates*. Research paper. Centre for Actuarial Studies, Department of Economics, University of Melbourne, 2007.
- [57] Pierre Vallois. *Modélisations stochastiques et simulations*. Modélisations stochastiques et simulations. Ellipses, 2007.
- [58] Julien Vedani and Laurent Devineau. Solvency assessment within the ORSA framework : issues and quantitative methodologies. 2012.
- [59] Hans H.G. Verbeek. An approach to the analysis of claims experience in motor liability excess of loss reinsurance. *ASTIN Bulletin*, pages 195–202, 1972.
- [60] Bertrand Verdier and Artur Klinger. Jab chain : A model-based calculation of paid and incurred loss development factors. *ASTIN Colloquium*, 2005.
- [61] F. Etienne De Vylder. Estimation of ibnr claims by least squares. *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker*, pages 249–254, 1978.
- [62] Mario V. Wüthrich. Asymptotic value-at-risk estimates for sums of dependent random variables. *ASTIN Bulletin*, 33(1) :75–92, 2003.
- [63] Mario V. Wüthrich and Michael Merz. *Stochastic Claims Reserving Methods in Insurance*. The Wiley Finance Series. Wiley, 2008.
- [64] Mario V. Wüthrich and Michel Merz. *Stochastic Claims Reserving Methods in Insurance*. The Wiley Finance Series. Wiley, 2008.
- [65] Mario V. Wüthrich, Michel Merz, and Natalia Lysenko. Uncertainty of the claims development result in the chain ladder method. *Scandinavian Actuarial Journal*, 2009(1) :63–84, 2009.
- [66] Marc Yor. On some exponential functionals of Brownian motion. *Adv. in Appl. Probab.*, 24(3) :509–531, 1992.
- [67] Marc Jean Yor. *Exponential functionals of Brownian motion and related processes*. Springer finance. Springer, 2001.



*BIBLIOGRAPHIE*

---



## Part IV

# General introduction

---

|    |                                |     |
|----|--------------------------------|-----|
| 8  | Thesis overview                | 159 |
| 9  | Existing methods               | 167 |
| 10 | Conclusion and future projects | 181 |

---



# Thesis overview

---

In this chapter we will present the main problems which affect insurance companies and then look at the specific guarantee which will be considered in this document: the mortgage guarantee of real estate loans. Finally we will present the outline of this thesis.

## 8.1 Issues in insurance

Insurance companies (we will define this later) need to protect their policyholders (individual or corporate) from risks. Risks are random events and there are two main ways to categorize risks. The first option is:

1. Physical risk for individuals (accident, death, etc.),
2. Property risk (car, accommodation, etc.),
3. Liability risk (civil, professional, etc.).

And the second option is:

1. Life risk,
2. Non-life risk.

Protection of insurance company to the policyholder in the case of realization of the risk can be financial or material. In compensation of this protection, the policyholder needs to pay a premium to the insurance company in which support the risk. This premium is paid by the subscriber. However at the start of the policy the insurance company does not know how much it will need to pay when the risk materializes: this phenomenon is called “production cycle inversion”. This leads to two sources of uncertainty for the insurance company: the event considered as a risk may or may not occur, and the amount that the insurance company will need to pay if the event occurs is unknown. This in turn leads to the main problem that insurance companies have to solve:

### **Topic 1: “How to determine premiums”.**

There are several kinds of insurance company and we can distinguish them by their legal status. We can distinguish insurance companies regulated by the “Code des Assurances” (public limited company), mutual fund (non-profit making and restricted to personal insurance), and provident society regulated by the “Code de la Sécurité Sociale”. These companies can set their premiums by considering the personal characteristics of each individual policyholder. This gives us the second problem that insurance companies need to solve:

**Topic 2: “How to detect high-risk policyholders”.**

The duration of insurance contracts depends on the kind of insurance: life insurance contracts are usually taken over a longer period than non-life contracts. By law, insurance companies are required to retain sufficient funds, over the whole life of these contracts, to cover all of their commitments: this amount is called the “technical provision”. This introduces the third problem that insurance companies are confronted with:

**Topic 3: “How to calculate the provision”.**

Provisions are included in the accounts of the company and are then invested in stocks, bonds, real estate, etc.. The return on these investments is important for certain kinds of insurance (for example mortgage guarantee) as they are linked to the insurance company’s commitments to its policyholders, and most therefore meet strict conditions. The goal of these investments is twofold: to maximize profit and to minimize the risk of loss. This gives us the fourth problem which insurance companies face:

**Topic 4: “How to invest the premiums”.**

In addition to national regulation, insurance companies will soon have to comply with European reform “Solvency 2”. This reform (see Section 8.2.2 for more details) is designed to better protect and better inform policyholders by evaluating the solvency of insurance companies:

**Topic 5: “How to evaluate the solvency of insurance companies”.**

Let us now look at the specific guarantee considered in this thesis.

## 8.2 Mortgage guarantee of real estate loans

After presenting the insurance company CAMCA Assurance, we will introduce the specific guarantee proposed by this company and we will conclude by a presentation of the different provision that this insurance company needs to calculate.

### 8.2.1 CAMCA Assurance

In 1946, the “Caisses Régionales du Crédit Agricole” founded the mutual fund “CAMCA” (Caisse d’Assurances Mutuelles du Crédit Agricole) in order mutualize the risk of theft as it was difficult to find this protection on the insurance market at the end of the war.

At the end of the 1980s, CAMCA extended its activities to insurance “para-bancaire” (credit card insurance, etc.) and offered insurance and related benefits to the clients of

## 8.2 Mortgage guarantee of real estate loans

the “Caisses Régionales”. As CAMCA was a mutual fund, it could not offer insurance “para-bancaire” directly to these clients so in the 1990s it created a public limited company called CAMCA Assurance. CAMCA Assurance is based in Luxembourg for three main reasons:

1. At the time, Crédit Agricole wanted to expand its activities and provide similar services throughout Europe so used of “libre prestation de services”, which allows companies based in one EU country to work in another, to find a suitable location.
2. CAMCA already had a reinsurance company (CAMCA Réassurance) based in Luxembourg and so took advantage of the existing infrastructure and experience to create the new company, CAMCA Assurance.
3. The last reason is more pragmatic: in 1997 it only took eight months to create a company in Luxembourg but sixteen months in France.

Currently, the main activity of CAMCA Assurance is the mortgage indemnity of real estate loans for the “Caisses Régionales du Crédit Agricole” (called “banks” in the rest of this document). If a borrower fails to repay any or all of their monthly installments, they are in a default situation and the bank will submit an indemnity claim to the insurance company. This is an illustration of how the guarantee works:

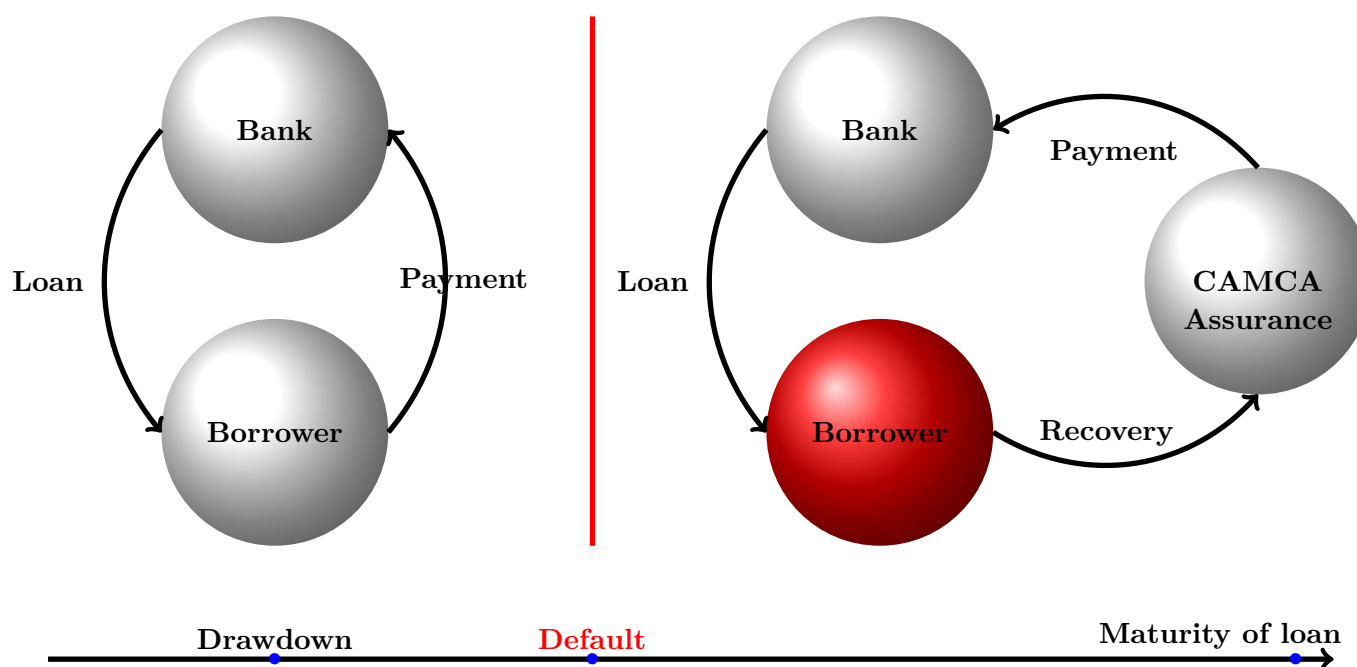


Illustration 8.1: Guarantee offered by CAMCA Assurance.

### **8.2.2 Mortgage guarantee**

A mortgage indemnity policy functions as follows: after the premium is paid by the borrower, CAMCA Assurance agrees to replace the borrower in the event of definitive payment defaults, that is to say when the borrower fails to repay any or all of their monthly installments. So CAMCA Assurance pays the bank when the borrower fails to repay definitively their monthly installments.

In this situation, CAMCA Assurance has the ability to use all legal means to recover debts: it can get the money directly from the borrower or it can try to repossess the property and sell it. Note that CAMCA Assurance delegates the recovery procedure to the bank: it is the bank that collects the amount of recovery from the borrower. If the recovery amount is lower than the borrower's debt then CAMCA Assurance pays the difference to the bank: CAMCA Assurance indemnifies the definitive loss net of recovery. Note that the loans considered are real estate loans.

Mortgage guarantees offered by CAMCA Assurance differ from more common borrower insurances which pay out in event of death or serious injury. It also differs from credit insurance where the policyholder is the bank.

In other words:

- CAMCA Assurance gives the borrower personal security: if the borrower is unable to pay his or her loan installments, CAMCA Assurance will replace them and reimburse the bank that financed the loan. The policy offered by CAMCA Assurance is a "simple mortgage guarantee" and the recovery procedure is delegated to the bank.
- The premium is paid by the borrower in a one-off installment at the beginning of the loan.
- The beneficiary of the guarantee is the bank that financed the loan. The bank is also the intermediary in the relationship between CAMCA Assurance and the borrower.
- The duration of the guarantee is equal to the loan duration: the guarantee expires when the loan is repaid by either the borrower or the insurance company.

Now we will give some detailed information about mortgage guarantees.

In some circumstances, banks can, in addition to the guarantee offered by CAMCA Assurance, add another guarantee to the loan. The two most common examples of this are:

1. When the bank is convinced the borrower will default in the near future or when the situation of the borrower changes (e.g. divorce) and so decides to take out an additional notarial mortgage over the property. As mentioned before, CAMCA



Assurance can repossess the property and sell it, but this is not an automatic process as is the case with notarial mortgages. The notarial mortgage, although more expensive, is safer (in terms of recovery) than the guarantee offered by CAMCA Assurance and these costs are borne by CAMCA Assurance.

2. When a borrower misses one or more installments the bank can commence legal action to recover its debt and this could mean the bank taking possession of the property (legal mortgage guarantee).

We will assume that this kind of mortgage security is systematic.

CAMCA Assurance has drawn up a profile covering loans and borrowers and if the loan and the borrower match this profile then cover is automatic. In other words, the bank does not need to contact CAMCA Assurance for approval. This is not the case when the loan or the borrower do not match the criteria and here the bank must seek approval from the insurance company. In the event of a claim where the loan or borrower criteria is not met and approval has not been given, the insurance company can refuse to settle any claim. However we will assume that there is no fraud or negligence by banks with regard to the relevant profiles.

It is also possible to compare the mortgage guarantee offered by CAMCA Assurance to a “credit default swap” (see Section 12.1) or with an exchange option (see Section 12.2).

### 8.2.3 Claims and provisioning

As mentioned before, a claim is submitted when a borrower defaults, so we will conclude our explanation of this guarantee by outlining the steps involved in a claim and the provisions that CAMCA Assurance needs to calculate in this regard.

In the event of a claim, CAMCA Assurance needs to reimburse the bank the amount outstanding after any recovery. This amount is called the “claim amount”, and there is no time limit on when this claim must be submitted. When a borrower defaults (not definitive a priori), the bank has two possibilities: make a claim immediately or wait for the borrower to pay the missing amount. But in this second situation this amount could fluctuate. However, for this thesis, we will assume that the bank submits the claim immediately. Note that a default could represent only a portion of the monthly payment but for simplicity we will assume the default amount is the entire installment.

During the indemnity process, CAMCA Assurance only pays the bank the outstanding amount owed by the borrower minus any property sale amount. However it is possible that the bank receives money directly from the borrower after the insurance company has paid and in this situation the bank will reimburse the insurance company:



Illustration 8.2: Recovery procedure.

This description allows us to distinguish two kinds of provision that CAMCA Assurance needs to calculate at date  $t_0$ :

1. A provision concerning claims submitted and indemnified before date  $t_0$ . This provision is an estimation of recoveries that the bank will receive after date  $t_0$  on claims already submitted. This provision will be negative. This is an illustration:

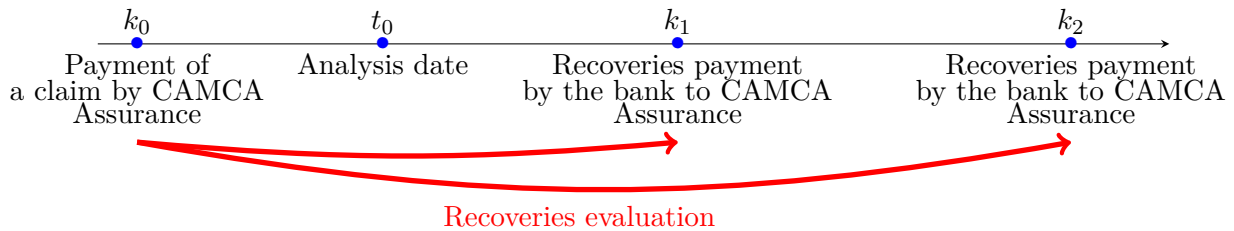


Illustration 8.3: Provision (negative) concerning recoveries.

2. A provision concerning claims that will be submitted after the date  $t_0$  but concerning loans drawn down before date  $t_0$ . CAMCA Assurance needs to calculate the claim amount, net of recoveries, that will be reported after date  $t_0$ . This is an illustration:

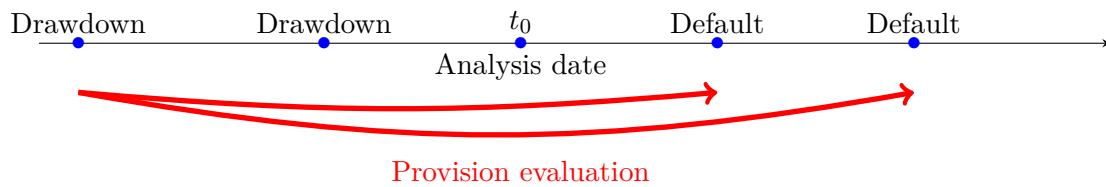


Illustration 8.4: Provision (positive).

Recovery received after CAMCA Assurance payment is extremely rare (less than 0.01% of the claim amount). So CAMCA Assurance only calculate the second type of provision.

### 8.3 Outline of Parts V and VI

The second part of this thesis is devoted to modeling the technical provisions that CAMCA Assurance needs to include in their accounts and to numerically evaluate these provisions. We will also calculate some risks indicators (mean, variance) and the law of the provision (see Chapter 11).

We will complete this part with remarks concerning the interpretation of the guarantee with the “credit default swap” (see Section 12.1), the link between exchange option and provisioning allows us to examine how the premiums it is invested (see Section 12.2), the justification of the assumptions of our model (see Section 12.3), and a numerical application of the theoretical results (see Section 12.4).

The third part will be devoted to the evaluation of CAMCA Assurance’s solvency under the Solvency 2 reform: in particular we will discuss the “solvency need” introduced in the second pillar of the reform. We will calculate the asymptotic law of the provision in Theorem 13.2.5 and the solvency need due to provisioning risk in Theorem 13.2.6. We will complete this chapter with a numerical application in Chapter 14.

In conclusion, this thesis will try to answer topics 3 and 5 which we introduced in Section 8, Topics 1, 2 and 4 are currently still a project.

Before this, however, we will present a quick overview of the existing methods used for provisioning and solvency evaluation.



# Existing methods

---

In this chapter we will present a quick overview of the existing methods used for provisioning and solvency evaluation under Solvency 2 reform.

## 9.1 Provisioning

This section draws on the works [23], [43], and [63].

Under article R-331-1 of the “code des assurances” insurers are obliged to determine the amount needed to cover their technical commitments; this amount is called the “technical provision” and in non-life insurance there are two classical methods used to calculate this provision: collective or individual.

### 9.1.1 Collective methods

Collective methods are based on the past development of claims and sometimes include exogenous information. Unfortunately these methods are unable to take into account the specific characteristics of the policyholders. In publications about provisioning problems there are two kinds of collective model: deterministic and stochastic.

Before presenting these methods, it is useful to introduce some notations and concepts.

#### Notation 9.1.1.

1. *Contracts are grouped by inception date (e.g. all contracts started in a specific month; all contracts one year old, etc.), by date of claim or by date of indemnity. This grouping are denoted by  $i$ :  $i \in \{1, \dots, n\}$  where  $n$  is the last grouping known. To simplify matters, we will only consider contract inception and a time unit of one year.*
2.  *$X_{i,j}$  is the claim amount concerning contracts drawdown grouping  $i$  and indemnify year  $i + j - 1$ ,  $j \in \{1, \dots, n\}$ . So  $X_{i,j}$  is a sum of individual claim amount.  $C_{i,j}$  is the claim amount concerning contracts drawdown grouping  $i$  and indemnify until year  $i + j - 1$ , i.e.:*

$$C_{i,j} := \sum_{k=1}^{k=j} X_{i,k}. \quad (9.1.1)$$

3. *At date  $n$ , the provision for grouping  $i$  is denoted by  $P_i$  and defined by:*

$$P_i := C_{i,n} - C_{i,n-i+1}. \quad (9.1.2)$$

### Deterministic methods

The most well-known method used by insurers is the Chain Ladder introduced around the 1930. This deterministic method (which is not based on any probability law) is based on the following development triangle and is used to evaluate the provision at date  $n$ :

$$\begin{pmatrix} C_{1,1} & C_{1,2} & \cdot & \cdot & C_{1,n} \\ C_{2,1} & C_{2,2} & \cdot & C_{2,n-1} & \\ \cdot & \cdot & \cdot & & \\ C_{n,1} & & & & \end{pmatrix} \quad (9.1.3)$$

The first assumption in the Chain Ladder model is that all contracts have the same maturity period and this period is  $n$ . The goal of this method, and more generally of the collective methods, is to evaluate the ultimate claim amount  $C_{i,n}$ , and so the provision  $P_i$ , by completing the development triangle defined above, i.e.:

$$\begin{pmatrix} C_{1,1} & C_{1,2} & \cdot & \cdot & C_{1,n} \\ C_{2,1} & C_{2,2} & \cdot & \cdot & C_{2,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ C_{n,1} & C_{n,2} & \cdot & \cdot & C_{n,n} \end{pmatrix} \Rightarrow \begin{pmatrix} P_1 = 0 \\ P_2 = C_{2,n} - C_{2,n-1} \\ \cdot \\ P_n = C_{n,n} - C_{n,1} \end{pmatrix} \quad (9.1.4)$$

Chain Ladder method was introduced to situations where the amount of data is considerable and where the development of claims does not depend on the grouping involved. This leads to the second assumption: that cumulative claim amount at date  $i + j$  for contracts drawn down in year  $i$  is proportionate to the cumulative claim amount at date  $i + j - 1$  for these contracts. The proportionality coefficient, denoted by  $\lambda_j$ , is deterministic and depends only on development year  $j$ , i.e.:

$$\exists \lambda_j \in \mathbb{R}_+ : C_{i,j+1} = \lambda_j C_{i,j}, \forall i, j : i + j \leq n. \quad (9.1.5)$$

When parameters  $(\lambda_j)_{1 \leq j \leq n-1}$  are estimated, it is possible to give a numerical value to the provision. Formula (9.1.5) is available for all  $i \leq n - j$ , parameter  $\lambda_j$  is then determined by:

$$\lambda_j = \frac{\sum_{i=1}^{n-j} C_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}}, \quad 1 \leq j \leq n-1, \quad (9.1.6)$$

Because  $C_{i,j+1} = \lambda_j C_{i,j}$ , the ultimate claim amount concerning grouping  $i$  is:

$$C_{i,n} = C_{i,n+1-i} \prod_{j=1}^{i-1} \lambda_{n-i+j}. \quad (9.1.7)$$

So the provision for grouping  $i$  is:

$$P_i = C_{i,n} - C_{i,n-i+1} = C_{i,n+1-i} \left[ \prod_{j=1}^{i-1} \lambda_{n-i+j} - 1 \right]. \quad (9.1.8)$$

Now we apply the Chain Ladder method on an example. Consider the following claims development triangle ( $n = 3$ ):

$$\begin{pmatrix} 1 & 4 & 8 \\ 7 & 9 & \\ 2 & & \end{pmatrix} \quad (9.1.9)$$

The objective is to calculate the ultimate claim amount for each grouping  $i$ , i.e. we want to calculate  $C_{1,3}$ ,  $C_{2,3}$  and  $C_{3,3}$ . The first grouping has matured so we know the ultimate claim amount  $C_{1,3}$ . From (9.1.6) we have:

$$\lambda_1 = \frac{4 + 9}{1 + 7} = \frac{13}{8}, \quad (9.1.10)$$

$$\lambda_2 = \frac{8}{4} = 2. \quad (9.1.11)$$

From (9.1.8) the provisions are:

$$P_1 = 0 \text{ (this grouping is matured),}$$

$$P_2 = 9 \times 2 - 9 = 9,$$

$$P_3 = 2 \times 2 \times \frac{13}{8} - 2 = 4,5.$$

The main disadvantage of the Chain Ladder method is its over-parametrization: at date  $n$  we need to estimate  $n - 1$  parameters,  $\lambda_1, \dots, \lambda_{n-1}$ , using  $\frac{n(n+1)}{2}$  data. Formula (9.1.6) means that old data is given the same weightings (or importance) as new data. In practice, new data is more comprehensive and representative than old data. The weighted Chain Ladder method allows us to take into account this phenomena by giving more weight to certain data (e.g. new data, gender-specific).

An other disadvantage is that the autoregressive assumption (9.1.5) of the Chain Ladder model is very restrictive although this can be relaxed as follows:

$$C_{i,j+1} = f(C_{i,j}), \quad \forall i, j : i + j \leq n, \quad (9.1.12)$$

where  $f$  is a function defined from  $\mathbb{R}_+$  to  $\mathbb{R}_+$ . A particular case of function  $f$  was proposed by S. Benjamin and L.M. Eagles in 1986 in the London Chain model:

$$\exists(\lambda_j, \alpha_j) \in (\mathbb{R}_+)^2 : f(C_{i,j}) = \lambda_j C_{i,j} + \alpha_j, \quad \forall i, j : i + j \leq n. \quad (9.1.13)$$

This method which allows us to take into account a situation where points  $(C_{i,j}, C_{i,j+1})_{i \geq 1}$  is a line which pass through 0, increases the over-parametrization problem of the Chain Ladder method. As a result, in 1988 E. Straub proposed the London Pivot model with this function  $f$ :

$$\exists(\lambda_j, \alpha) \in (\mathbb{R}_+)^2 : f(C_{i,j}) = \lambda_j C_{i,j} + \alpha(\lambda_j - 1), \quad \forall i, j : i + j \leq n. \quad (9.1.14)$$

A large number of Chain Ladder variants exist. All of these methods although easy to implement, have the same disadvantage: they do not take into account the draw-down year in the development of future claims. For certain types of insurance, it is

logical to consider the influence of the drawdown year by integrating exogenous information. This is the purpose of the Cape Code method (H. Bühlmann, 1983, see [17]), the Bornhuetter-Ferguson method (R.L. Bornhuetter and R.E. Ferguson, 1972, see [14]) and the Benktander method (G. Benktander, 1976, see [10]); exogenous information could be a premium amount or an ultimate claim ratio estimated by expert judgment. Other methods allow to be taken into account for the influence of grouping (factorial method of least squared, F. De Vylder, 1978, see [61]) and also for inflation (H.G. Verbeek and G. Taylor, 1972, see [59]).

All these methods are determinist so they only give punctual estimation without any fluctuation, i.e. they do not give a confidence interval.

Now we will present the stochastic methods used in non-life insurance provisioning.

### Stochastic methods

There are various stochastic methods used to calculate non-life insurance companies provisions and we provide a non-exhaustive classification in three parts: recursive models, models based on log-normal regression, and generalized linear model (GLM).

The most well-known recursive stochastic methods is the Mack model (T. Mack, 1993, see [34]) which is a stochastic version of the Chain Ladder model. The following assumptions are retained:

$$(C_{i,1}, \dots, C_{i,n}) \text{ and } (C_{k,1}, \dots, C_{k,n}) \text{ are two RVs independent } \forall i \neq k, \quad (9.1.15)$$

$$\exists \lambda_j \in \mathbb{R}_+ : \mathbb{E}[C_{i,j+1} | C_{i,1}, \dots, C_{i,j}] = \lambda_j C_{i,j}, \forall i, j : i + j \leq n, \quad (9.1.16)$$

$$\exists \sigma_j \in \mathbb{R}_+ : \mathbb{V}[C_{i,j+1} | C_{i,1}, \dots, C_{i,j}] = \sigma_j C_{i,j}, \forall i, j : i + j \leq n. \quad (9.1.17)$$

After estimating parameters  $(\lambda_j)_{1 \leq j \leq n-1}$  and  $(\sigma_j)_{1 \leq j \leq n-1}$ , these three assumptions allow us to give an estimation of the provision mean and also a confident interval associated with its mean. A great number of extensions of the Mack model exist such as the Schnauss model (1996), the Munich Chain Ladder model (G. Quarg and T. Mack, 2004, see [48]), and the JAB Chain model (B. Verdier and A. Klinger, 2005, see [60]).

The second type of stochastic model used is based on “log-normal” regression of development factors (Hertig model, 1985, see [29] and HFB model, 1985, see [30]) or on a “log-normal” regression of random variables  $X_{i,j}$  (S. Christophides model, 1990, see [18] and G. Barnett and B. Zehnirith model, 1999, see [9]).

An alternative is to consider GLM model. In these models, random variables  $X_{i,j}$  have an exponential density (normal, gamma, poisson, etc.), i.e.:

$$\exp \left[ \frac{x_{i,j} \theta_{i,j} - v(\theta_{i,j})}{u(\phi)} + w(x_{i,j}, \phi) \right], \quad (9.1.18)$$



where  $u$ ,  $v$ , and  $w$  are functions,  $\theta_{i,j}$ , and  $\phi$  are parameters called “natural parameter” of the exponential family and “dispersion parameter”. The main GLMs used in non-life insurance are those of A.E. Renshaw and R.J. Verrall (Poissonian model, 1998, see [51]) and those of S.M. Gluck (2000, see [28]).

All the collective methods (deterministic and stochastic) discussed have a number of common disadvantages. The main disadvantage is their over-parametrization. In addition they are unable to model claim amounts concerning specific policyholders or take into account the specific characteristics of individual policyholders. These methods do not take into account guarantees where contracts have different maturities (without the use of a tail factor). Another disadvantage is the provision amount is nil if the claim amount is also nil at the analysis date: this is the case for the specific guarantee which discussed previously. Finally, independence assumptions between groupings is not always confirmed in practice. So we will now look at an alternative to these methods: individual models.

### 9.1.2 Individual models

In individual models, the provision is defined as the sum of a random number of individual claim amounts concerning a specific policyholder. The most well-known model is the compound Poisson process (see [57]) in which the provision of grouping  $i$  at date  $n$  is denoted by  $P_i$  and defined by:

$$P_i := \sum_{j=N_n}^{N_{n+i}} S_j, \tag{9.1.19}$$

where:

- $(N_t)_{t \geq 0}$  is a Poisson process and  $N_t$  represents the number of claim submitted (or indemnified) in the period of time  $[0, t]$ ,
- $(S_j)_{j \geq 1}$  is a collection of random variables (RVs) independent and identically distributed (iid), where  $S_j$  represents the  $j$ -th claim concerning contracts drawn down in the grouping  $i$ ,
- The Poisson process  $(N_t)_{t \geq 0}$  is independent of claim amounts  $(S_j)_{j \geq 1}$ .

This kind of model is illustrated as follows:

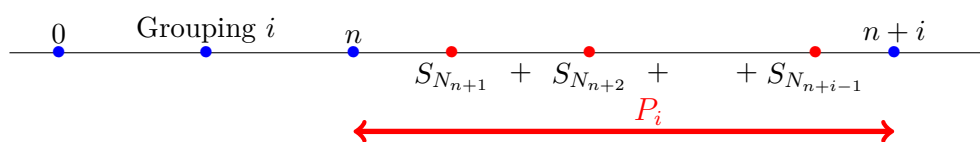


Illustration 9.1: Individual provision with the compound Poisson process.

There are numerous theoretical results concerning this process: moments (mean and

variance), asymptotic behavior, and law of the process have been determined. However this model does not take into account these dependencies:

- Because RVs  $(S_j)_{j \geq 1}$  are independent, it is not possible to consider the dependencies between “policyholders”.
- It is therefore impossible to take into account any external systemic risks (real estate, recession, etc.).
- Finally it is not possible to consider the link between the process representing the claim submission and the claim amount because the counting process  $(N_t)_{t \geq 0}$  is independent of RVs  $(S_j)_{j \geq 1}$ . In particular, the claim amount does not depend on the submission date (or indemnity date).

Certain models allow these dependencies to be taken into account. If we want to examine the dependency between the date of the claim and the claim amount we can use the “shot noise processes” (see [6], [49] and [54]) in which the provision is defined by:

$$P_i := \sum_{j=N_n}^{N_{n+i}} f(T_j, \xi_j), \quad (9.1.20)$$

where  $(N_t)_{t \geq 0}$  is a counting process where jump times are denoted by  $(T_j)_{j \geq 1}$ ,  $f$  is an explicit measurable function,  $(\xi_j)_{j \geq 1}$  is a collection of RVs iid and also independent of the counting process. So  $f(T_j, \xi_j)$  represents the amount of the  $j$  – *th* claim concerning contracts drawn down in the grouping  $i$ . With expression (9.1.20) it is clear that the amount of each claim depends on the submission date (or indemnify date).

Another method for dealing with this kind of dependency is to model the provision with  $U$ -statistics of a Poisson point process (see [19]), i.e.:

$$P_i := \sum_{(X_1, \dots, X_k) \in \eta_{\neq}^k} f(X_1, \dots, X_k), \quad (9.1.21)$$

where  $f$  is a symmetric function, random vector  $(X_j)_{j \geq 1}$  is a punctual Poisson process and  $\eta_{\neq}^k$  is the set of all  $k$ -tuples distinct points of the process.

The dependence between claim amount and claim intervals (applicable to natural catastrophes) is discussed in the paper by H. Albrecher and O.J. Boxma (see [2]), and in the thesis by R. Biard (see [12]).

With regard to the dependency between claim amounts, we can see copula, especially Archimedean copula explained in [3], [8] and [62]. It also interesting to consider marked Poisson process (see [33] and [38]).

There are also disadvantages with individual models:

1. The technicality of these methods implies that their use and explanation is more complicated than collectives models,

2. Computation times of these methods is longer than for collective methods,
3. The number of parameters can, sometimes, be significant,
4. Finally, these methods are not generic and can only be used for certain guarantees.

Comparison between collective methods and individual methods is as follows (R.J. Ver-  
ral):

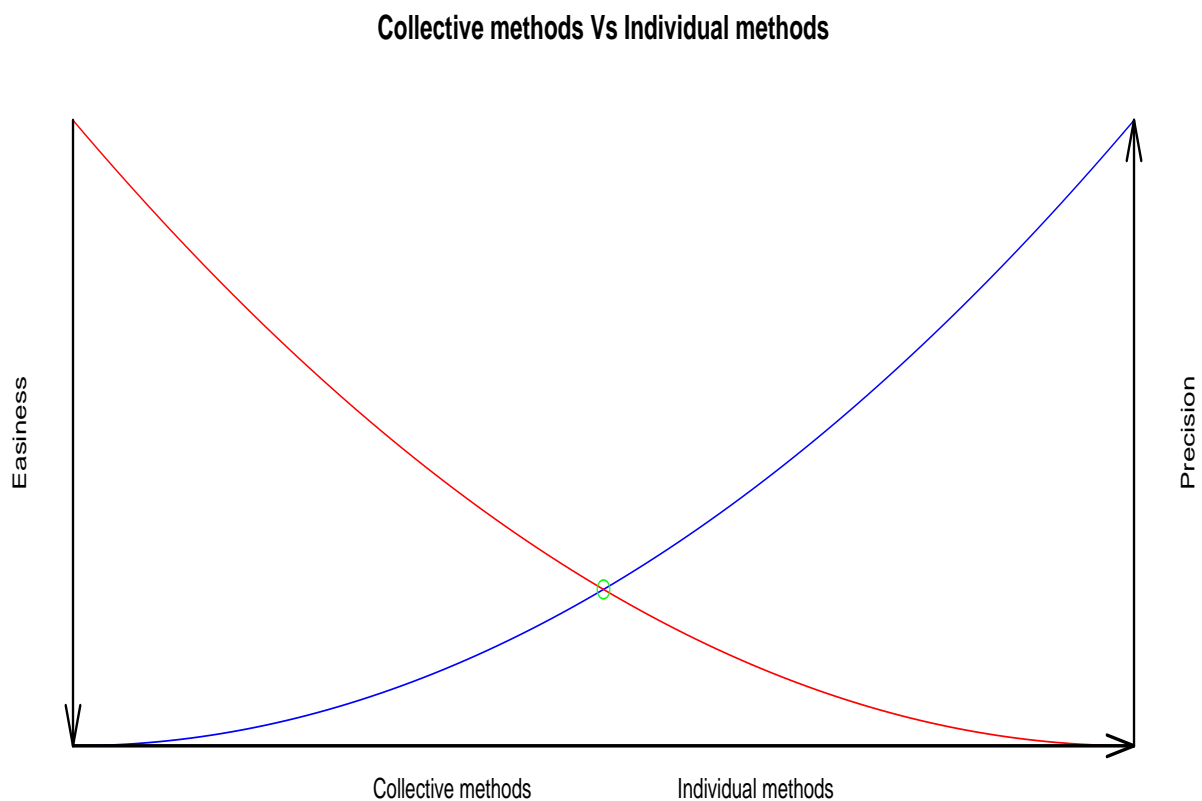


Illustration 9.2: Collective methods Vs individual methods.

Red line represents easiness of use and blue line represents precision. The ideal model is one which sits where the red and blue lines intersect, i.e. a model which maximizes precision and ease of use.

This concludes the overview concerning provisioning and we will now present the methods for evaluate solvency.

## 9.2 Solvency 2 reform

### 9.2.1 Regulatory overview

Solvency 2 reform will come into force on 1st January 2016 for life insurance companies, non-life insurance companies and mutual fund which trading in the European market. This reform has one main goal: to better protect and better inform insurance policyholders. In particular, standardizing the following regulation which insurance companies are subjected would allow this goal to be reached:

- Solvency calculation,
- Management and governance,
- Availability of information to the general public and to the authorities.

The main documents regarding this reform is Directive [26] in which the general structure of the reform is detailed. There are three pillars as in the banking sector reform Bale 2 (see [27]). Precise technicals specifications and guidelines are given by EIOPA (European Insurance and Occupational Pensions Authority) the European regulatory in charge of insurance market. Together with the “Banking Authority and the Financial European Market Authority” it forms the system of “European System of Financial Supervision”.

### 9.2.2 The three pillars

#### First pillar:

The first pillar of the Solvency 2 reform (see articles 100 to 127 of Directive [26]) deals with the quantitative evaluation of the “solvency capital requirement” (SCR) as a result of risks to which insurance companies are subjected, and detailed the amount that insurance companies must retain to cover these risks. Before see a graphical illustration we will give some explanations concerning the balance sheet of an insurance company:

- Insurance company assets are composed, essentially, of premiums received from policyholders which are subsequently invested in different products such as stocks, bonds, and real estate. These assets are owned by the insurance company and relatively quickly convert into cash (liquid asset).
- “Reinsurance” is the amount that insurance companies receive from the reinsurer in respect to cover all or part claims meet by policyholders. Although owned by insurance company, this money can be quickly obtained.
- Own fund is the capital invest by shareholders (the people who in theory own the company). This amount could, at any time, be repaid to the shareholders.
- “Provisions” are the estimated amounts that insurance companies may have to pay to honour their commitments. This amount is owned by the company.

Now we can see an graphical illustration of the SCR.

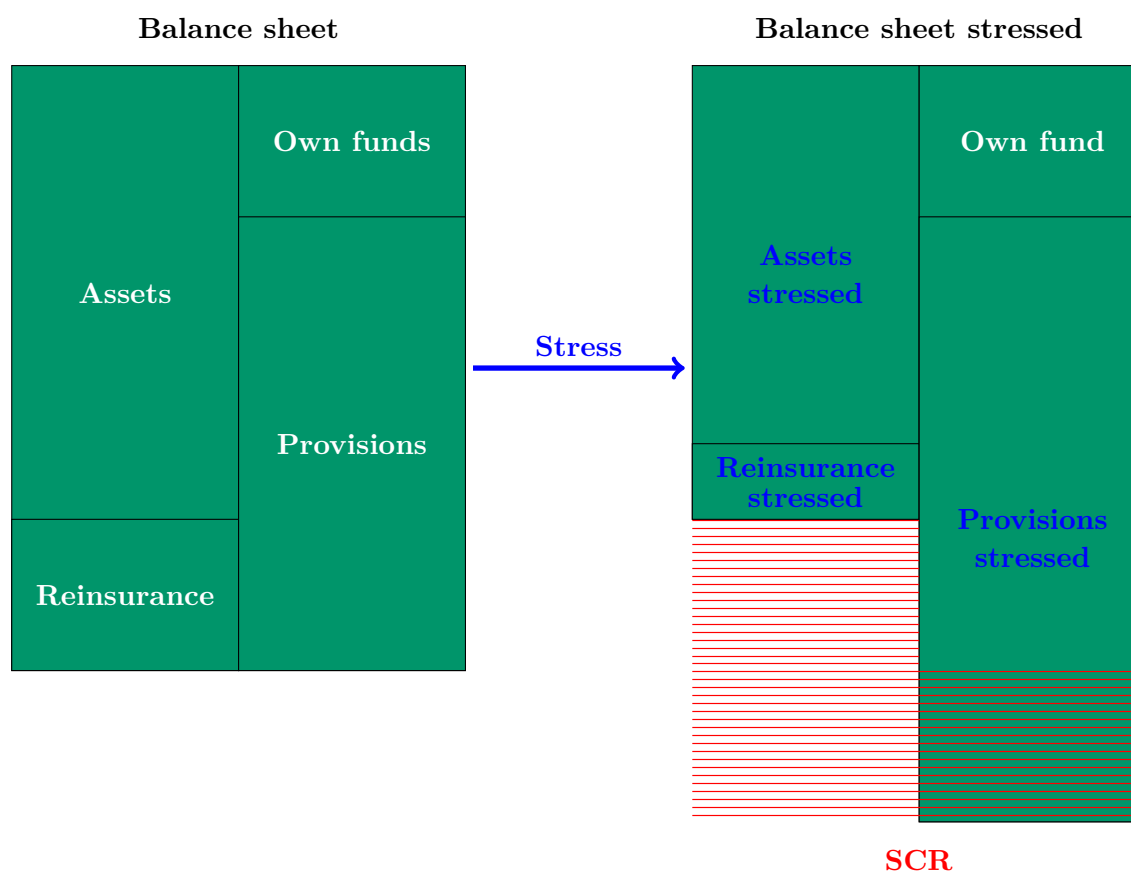


Illustration 9.3: Balance sheet stressed.

Risks evaluated in the SCR are imposed by the Directive and depend on the type of insurance company. The following risks relate to non-life insurance:

- Market risk: stocks, rate, spread, concentration (diversification), real estate, currency.
- Counterparty default risk: reinsurers, banks, etc..
- Underwriting risk: premium and provision, lapse, catastrophe.
- Operational risk: IT system, etc..

EIOPA proposes the use of a standard formula to evaluate the SCR associated to each risk. This SCR is “calibrated” as a value at risk over one year at 99.5%. Article 101 of Directive [26] specifies that:

*“It shall correspond to the Value-at-Risk of the basic own funds of an insurance or reinsurance undertaking subject to a confidence level of 99.5% over a one-year period.”*

For example the stocks SCR is calibrated as the amount lost on stock products over a one year period in 99.5% of all loss situations. EIOPA also gives a correlation matrix which

allows for the aggregation of SCRs and the calculation of a single SCR. The Directive allows insurance companies to use their own SCR formula if they prove the standard formula is unsuitable: this can be done for all risks (internal model) or only for certain risks (partial internal model).

### **Second pillar:**

The second pillar of Solvency 2 reform (see articles 41 to 50 of Directive [26]) deals with governance of insurance companies and also the evaluation of their own risks. It outlines the management structure that insurance companies must adhere to and also how this structure must operate. This information must be recorded and known and understood by all key personnel in the company. To do this, insurance companies need to determine and nominate four key functions: risk management, compliance, internal audit, and actuary. The relationship between these key functions, administrative board, and the company directors needs to be documented. The second pillar also provides guidelines as to how subcontractors are managed together with guidelines on the information which has to be made available to regulatory bodies, employees, and management board.

Also included in the second pillar is ORSA (Own Risk Solvency Assessment) where insurance companies evaluate their own risks and this process is outlined in article 45 of the Directive [26]:

*“As part of its risk-management system every insurance undertaking and reinsurance undertaking shall conduct its own risk and solvency assessment.”*

This evaluation is based on an exhaustive mapping of the insurance company’s internal processes and also on an exhaustive mapping of the risks to which insurance companies are subjected. For each quantifiable risk, insurance companies need to calculate the associated solvency need. This evaluation is similar to the one imposed in the first pillar but the analysis horizon and the risk indicator are not specified in the second pillar. So, in practice insurance companies can choose a risk indicator other than value at risk and an analysis horizon other than one year. Risks considered in the ORSA calculation need to be at least risks of the first pillar. The final step included in the ORSA is the stress test applied to the solvency calculation in the company’s business plan.

### **Third pillar:**

The third pillar gives detailed instructions as to what information is made available to regulatory bodies and to the general public, and how this information is provided.

## **9.2.3 Solvency capital and solvency need evaluation methods**

The second part of this thesis will be devoted to the calculation of the solvency need (ORSA) linked to the provision risk. Provision risk is the risk that provisions calculated

by insurance companies are not sufficient to meet their commitments. It can be measured by the difference between the provision and its mean (because this is the amount that insurance companies include in their accounts). This is illustrated in the following example:

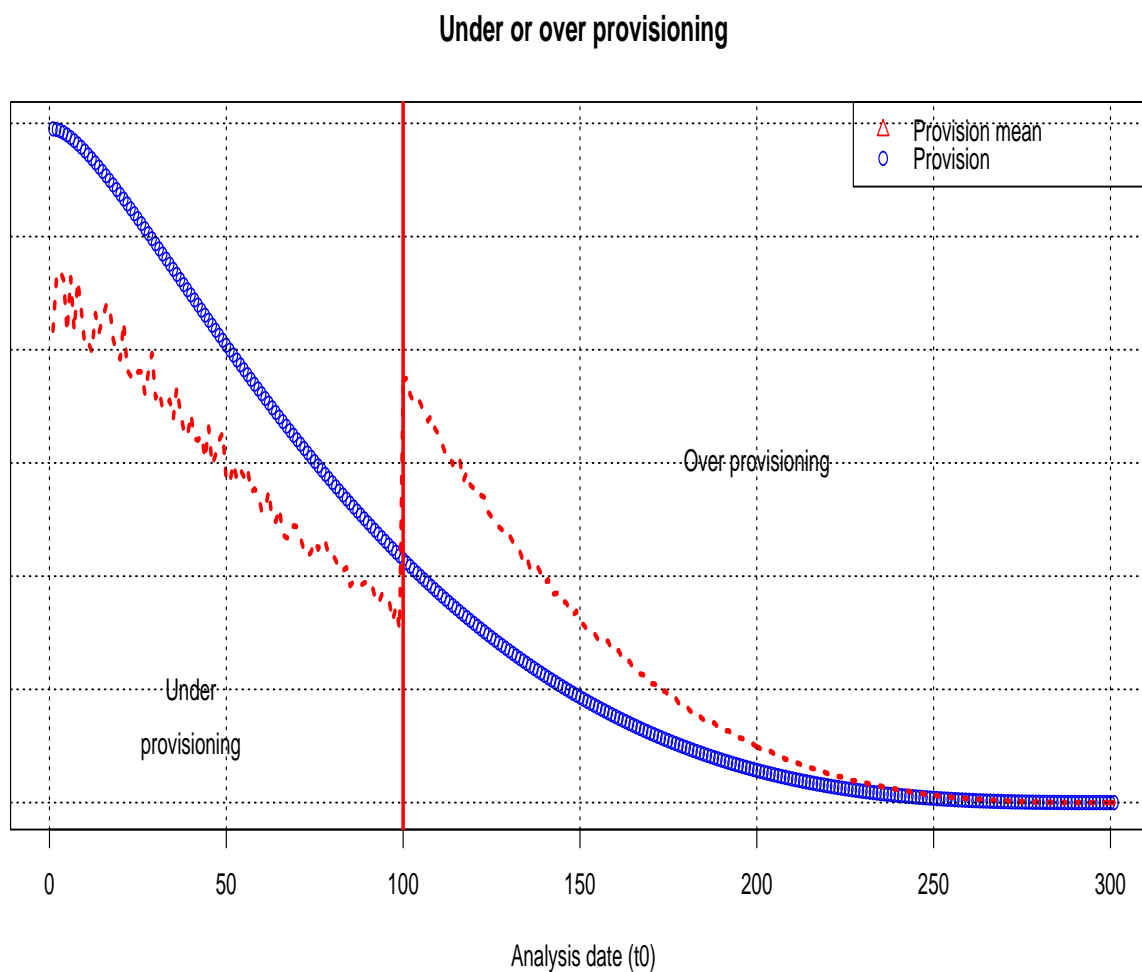


Illustration 9.4: Provisioning risk.

In the graph above:

- Analysis date,  $t_0$ , is on the  $x$  axis,
- The provision mean is the red line,
- The true provision is the blue line.

The difference between the true amount and the mean measures the risk of over or under provisioning: at the beginning of contracts there is an under-provisioning (mean is lower than the true value) and at the end of contracts there is a over-provisioning.

So the SCR linked to the provision risk is defined as the value at risk, at 99.5%, on the variable measuring risk over one year. But the standard formula from EIOPA does not differentiate between provisioning and premium risk calculation. In the standard formula, this SCR is approximated by:

$$SCR := 3\sigma V, \quad (9.2.1)$$

where  $V$  is an activity volume including the provisions amount and the premium amount, and  $\sigma$  represents the global volatility associated with premiums and provisions. We will not provide any details concerning the calculations for  $V$  and  $\sigma$  because these can be found in the technical specifications from EIOPA (see [1]).

The Directive allows insurance companies to calculate parameter  $\sigma$  using their own data (Undertaking Specific Parameters, USP). EIOPA gives three methods for calculating USP, and in turn for calculating volatility. These methods are detailed in “consultation paper 75” of EIOPA and in reports [32] and [44]. The second and third methods are based on a methodology outlined by Wuthrich in [65]. Note that the use of USP by insurance companies is inside the scope of the standard formula: it is not an internal model.

As mentioned before, the Directive allows insurance companies to calculate the SCR linked to the provision risk using their own formula. Examples are given in papers [41] and [46] which are devoted to this subject.

Contrary to SCR calculations, there are no specified methods concerning the solvency need calculation in the second pillar. The calculation of solvency need can be interpreted as an internal model for the SCR where the risk measure and the horizon analysis are not specified. See papers [16], [47], and [58] for more details on this subject.







# Conclusion and future projects

---

## 10.1 Conclusion

In this thesis we have assumed that an insurance company is required to indemnify a bank against losses related to a borrower defaulting on payments.

In the first part of this thesis we modeled the amount of losses resulting from the borrowers defaulting on a fixed period of time ; this amount is the provision. The claim amount is the key to our model. For one borrower  $j$  and for a fixed default date  $T_j$ , this amount is:

$$M_j \times \max [\phi_{T_j}(D_j) - R_{T_j}; 0 ], \quad (10.1.1)$$

where:

- $M_j$  is the borrowed amount and  $D_j$  is the term of the loan,
- $M_j \phi_{T_j}(D_j)$  is the outstanding amount owed by the borrower and depends on the specific characteristics of borrowers,
- $M_j R_{T_j}$  is the property sale amount. The proportionality coefficient  $R_{T_j}$  is modeled by a geometric Brownian motion and represents the market price fluctuation. We will assume that this coefficient is the same for each borrower, so there is a dependency between each claim amount.

We also modeled the law of the couples (Maturity of the loan  $T_j$ , Term of the loan  $D_j$ ) by a Poisson point process which allows us to define the provision  $P_h$ , where  $h$  is the maximum duration of loans, as the sum of a random number of individual claims:

$$P_h := \sum_{j \geq 1} \mathbb{1}_{\{(T_j, D_j) \in A_h\}} M_j \max [\phi_{T_j}(D_j) - R_{T_j}; 0 ], \quad (10.1.2)$$

where  $A_h$  is the set where defaults appears possible. Both theoretical and applied, formula (10.1.2) gives an explicit formula for the mean and for the variance, and we also can simulate the provision. The latter enables us to calculate different risk indicators including confidence interval and quantile. The amount of data available on each borrower allows us to estimate the parameters of our model.

In the second part we studied the solvency need due to provisioning risk (in relation to the proposed European reform Solvency 2). The solvency need, denoted by  $BS_h$ , is defined as the quantile of the centered provision, i.e.:

$$BS(h) := \mathcal{Q}_{P_h - \mathbb{E}(P_h)}(99.5\%), \quad (10.1.3)$$

To do this we studied the asymptotic behaviour of  $P_h$  when  $h \rightarrow +\infty$  and found this result:

$$\frac{P_h - \mathbb{E}(P_h)}{h\theta_1} \xrightarrow[h \rightarrow +\infty]{\mathcal{L}} G - \int_0^{+\infty} \inf(\theta_2, \theta_3 R_t) dt + \theta_4, \quad (10.1.4)$$

where  $\theta_1, \theta_2, \theta_3$ , and  $\theta_4$  are coefficients which can be written with the parameters of our model and  $G$  is a standard Gaussian random variable, and independent of  $(R_t)_{t \geq 0}$ . So we approximated  $BS(h)$  by the following quantity  $BS^*(h)$ :

$$BS^*(h) := h\theta_1 \times q_{99.5\%}, \quad (10.1.5)$$

where  $q_{99.5\%}$  is the 99.5% quantile of:

$$G - \int_0^{+\infty} \inf(\theta_2, \theta_3 R_t) dt + \theta_4. \quad (10.1.6)$$


This formula is only theoretical as the law of  $\int_0^{+\infty} \inf(\theta_2, \theta_3 R_t) dt$  is unknown. However as we mentioned before it is possible to simulate the provision and the results allows us to give a numerical value to the solvency need.

Note that if there is no correlation between claim amounts, then the limit law of the provision will converge to a standard Gaussian random variable. The additional term which is written with the integral of a functional of Brownian motion takes these correlations into account, so there is a “separation of risks sources”, one coming from the personal characteristics of the borrowers and the other coming from the real estate price fluctuation which is the same for each borrower.

## 10.2 Future projects

### Implementation of our model

The first project will be to implement our model in CAMCA Assurance. To do this we need:

1. To test (and “backtest”) our model on all the available data.
2. To explain our model and the test results together with the assumptions related to it.
3. To implement our model in a practical way by developing a software package  in conjunction with CAMCA Assurance’s existing IT system.

### Premium investment

In Section 12.2 we will show that the guarantee offered by CAMCA Assurance can be interpreted as an exchange option which leads us to the question of premium investment. To determine an investment strategy, we need to define a market with exogenous assets.

When exchange option maturity is fixed there is strategy to cover this option. Here the difficulty will be to take into account random number of defaults and the random time of these defaults, i.e. the random maturity of exchange options.

### Risk profile

An other interesting area to study will be the risk profile of borrowers. We could determine the main characteristics of borrowers who default and also the characteristics which predict future defaults. The tool used would be analysis of data.

### Solvency 2

Finally we could examine Solvency 2 reform in more detail as this thesis only studied the solvency need for a specific grouping. However the asymptotic method that we proposed is only theoretical and by looking at formula (3.1.4) we see that the law limit of the centered provision depends on the following variable:

$$\int_0^{+\infty} \inf(\theta_2, \theta_3 R_t) dt. \quad (10.2.1)$$

To date the law of this variable is unknown so further study into the law of this integral is a possibility.

Further study into the solvency need associated with other risks such as market risk, counterparty risk, etc. could also be considered.



## Part V

# Provisioning against borrowers default risk

---

|  |            |
|--|------------|
| <b>11 Provisioning against borrowers default risk</b>              | <b>187</b> |
| 11.1 Borrowers' default  | 187        |
| 11.2 Loss provision modeling                                       | 189        |
| 11.3 Results   | 196        |
| 11.4 Parameters estimation   | 199        |
| 11.5 Proof   | 203        |
| <b>12 Additional information concerning the provisioning model</b> | <b>217</b> |
| 12.1 Guarantee caution with the "credit default swap"              | 217        |
| 12.2 Provision with exchange option                                | 219        |
| 12.3 Justification of the assumptions                              | 221        |
| 12.4 Numerical application   | 228        |

---





# Provisioning against borrowers default risk

---

This chapter correspond to a paper submitted in february 2014 in “Insurance: Mathematics and Economics”.

This chapter focuses on the risk of loan default from the point of view of an insurer required to indemnify a bank for losses resulting from a borrower defaulting.

The main objective of this chapter is to model the provision (or claim reserve) against borrowers’ default. Unlike traditionally models used, our model depends on the specific characteristics of the borrowers (amount borrowed and term of loan). This chapter will also take into account three kinds of dependence: the dependence between each claim amount, the dependence between the date of default and the claim amount, and the dependence between the number of defaults and the claim amount. We will precise the link between our model and those existing.

One of the main result detailed in this chapter is the calculation of the mean and the variance provision. We will also give the law of the provision given the number of defaults. Another characteristic of our model is that is practical. When the parameters are estimated, we can numerically evaluate the mean, the variance and we can formulate an algorithm to simulate the provision which allows for the calculation of quantiles. To conclude we will give a methodology to estimate the parameters of our model.

**Key words:** *Borrower default risk; Individual stochastic provisioning; Poisson point process; Geometric Brownian motion; Time of default; Quantile.*

## 11.1 Borrowers’ default

### 1) *Overview:*

In this chapter, it is assumed that an insurance company is required to indemnify a bank against losses related to a borrower defaulting on payments. The loans under consideration are real estate loans. A borrower is considered to be in a state of “claim” at a date  $T$  when he or she can no longer pay their monthly loan installments.

The main issue for the insurance company is to evaluate, at time  $t_0$ , potential future claim amounts from borrowers whose loans are drawn down prior to  $t_0$  and who have not defaulted prior to  $t_0$ . This amount is the “provision”. It is important to note that in the event of default, the bank can get the money directly from the borrower or try to repossess the property and sell it. Obviously the provision is unknown at time  $t_0$  as it depends on future possible events. The main aim of this paper is to propose a stochastic model which is not only simple but also focuses on the key factors which affect it.

The claim amount is the key to our model. For one borrower and for a fixed default date, this amount is equal to  $\max(S_T^1 - S_T^2; 0)$ , where  $S_T^1$  is the outstanding amount owed by the borrower at date  $T$  and  $S_T^2$  is the property sale amount (see Section 11.2.2). The amount  $S_T^1$  is proportionate to the borrowed amount; the proportionality coefficient depends on the term of the loan and the date of default (i.e. the maturity). The quantity  $S_T^2$  is proportionate to the borrowed amount; the proportionality coefficient represents the market price fluctuation, the borrower’s deposit and the selling costs of property. This is modeled by a geometric Brownian motion. We assume that the process representing the market price fluctuation is the same for each borrower, so there is a dependence between each claim amount. We will model the law of the couples (Maturity of the loan, Term of the loan) by a point Poisson process (see Section 11.2.3), which will allow us to define the provision as the sum of individual claims (see Definition 11.2.5). For more details about individual models see [7], [33], [38], [45], and [56].

In Section 11.2.4 we also examine the links between our model and those of Cramer - Lundberg and Mack. Unlike those ours considers three different dependences (which is one of the main focus of this chapter):

1. The dependence between individual claim amounts (because the market price fluctuation is common for each borrower),
2. The dependence between claim amounts and default dates,
3. The dependence between the term of loan and the date of default.

In Section 11.3, by making additional assumptions, we will calculate the mean (see Theorem 11.3.2) and the variance (see Theorem 11.3.4) of the provision. In addition, our model allows us to go even further than the calculation of these two parameters. Indeed we provide an algorithm which permits to simulate the provision (see Theorem 11.3.6). The amount of data available on each borrower allows us to give a numerical value to the parameters of our model (see Section 11.4). When the parameters are estimated, we will obtain a numerical value of the mean, the variance, and the quantile of the provision.

All proofs will be given in Section 11.5.

## 2) Existing methods:

Models normally used by insurance companies to calculate such provisions are collective

models (Chain Ladder, Mack; see [34], [35], and [64]) because they aggregate data and do not allow for the calculation of a single claim amount. The disadvantage of this models is double: firstly the aggregation of data imply an over-parameterization and secondly there is no reason to assume that the accident years (or drawdown years in our case) are independent. There is a second type of model, seldom used by insurance companies, but which provides this type of guarantee: these are models derived from the ruin theory, the best known of which is the Cramer - Lundberg model (see [52]). In this model the provision is a random sum of individual claims concerning a given borrower. The drawback of such models is that they take into account neither the dependance between each claims amount, nor the dependance between a claim amount and the date of the claim, and nor dependance between the number of claims and the amount of claim. You can find more information in Section 9.

### 3) Conclusion:

This chapter describes an alternative method for provisioning adapted to the specific insurance contracts considered. It defines an individual stochastic model which takes into account the borrowers' characteristics together with a number of dependences. Both theoretical and applied, this model gives explicit formula for the mean and the variance, and we also can simulate the provision. The latter enables us to calculate different risk indicators including confidence interval and quantile. However, we have made a couple of assumptions which in everyday circumstances, would not hold true. Firstly, that the interest rate charged for each borrower is identical and secondly, that any default is immediate and irreversible. However we had try to justify these assumptions in Section 12.3.

## 11.2 Loss provision modeling

In Section 11.2.1 we will discuss notations and assumptions. Then, in Section 11.2.2, we will define a claim amount. In Section 11.2.3 we will model the couples (Maturity of the loan, Term of the loan) by a Poisson point process (see assumption  $(A_7)$ ) then we will define the fluctuation set of this couples (see Definition 11.2.4). This allows us to model the provision by totaling the random sum of individual claim amounts reported on random dates (see Definition 11.2.5). This is the key concept and will allow us to make a quantitative analysis (mean, variance, and simulations). Finally, in Section 11.2.4, we will examine the links between our model and those of Cramer - Lundberg and Mack.

### 11.2.1 Notations and assumptions

In this paper we will assume a continuous time. The time  $t$  is defined as a fraction of year and  $t^*$  is one month. We denote:

- $B$  ('B' for "borrower") the borrower set,
- The borrower by  $j$ :  $j \in B$ ,

- $I_j$  ('I' for "inception") the drawdown date of the borrower  $j$ :  $I_j \in [0, t_0]$ , where  $t_0$  is the last date under consideration.  $I_j$  is the date on which the funds are released.
- $D_j$  (resp.  $M_j$ ), the positive r.v. (random variable) representing the loan term (resp. loan amount) of borrower  $j$ ,

To each borrower  $j$ , is associated the couple  $(D_j, M_j)$ .

Borrower  $j$  draws down his loan in month  $i$  if  $i = \lfloor \frac{I_j}{t^*} \rfloor$ , where  $\lfloor x \rfloor$  represents the floor of  $x$ . In the following we will assume that a loan drawn down in month  $i$  is drawn down at the beginning of this month, that is to say:

$(A_1)$  : If borrower  $j$  draws down his loan at date  $I_j$  then this corresponds to month  $i$ , we will assume that:  $I_j = t^*i$ .

See Section 12.3.1 for a justification of this assumption.

We suppose that the duration of loans is capped by an unknown quantity  $h > 0$  such as:

$$(A_2) : \begin{cases} D_j \leq h, j \in B, & (11.2.1) \\ t_0 < I_j + h, j \in B. & (11.2.2) \end{cases}$$

See Section 12.3.2 for a justification of this assumption. This can be illustrated as follows:



Illustration 11.1: Duration capped.

Recall that 0 is the first date of drawdown of the loans.

In the rest of the paper we will consider real estate loans with fixed monthly payments and fixed interest rate for the term of the loan and for every borrower. We will use  $\tau$  to denote the exponential rate for all borrowers. Finally we will assume that early payments are not possible. However we will assume that it is possible to repay the loan in its entirety before the maturity date  $I_j + D_j$ . Therefore we will assume that:

$$(A_3) : \begin{cases} \text{Monthly repayments are fixed,} & (11.2.3) \\ \tau \text{ is the exponential and deterministic rate,} & (11.2.4) \\ \text{Early partial payments are prohibited.} & (11.2.5) \end{cases}$$

When a borrower defaults (not definitive a priori), the bank has two possibilities: make a claim or wait for the borrower to pay the missing monthly installment. In the second situation, the number of missing monthly installments can increase. However, we will

assume that the bank, in the event of default, submits the claim immediately. Note that a default could represent only a portion of the monthly installment but for simplicity we will assume a default concerns the entire monthly installment. Therefore we will assume that:

$$(A_4) : \begin{cases} \text{Default is submitted immediately as a claim,} & (11.2.6) \\ \text{Default concerns the entire monthly installment.} & (11.2.7) \end{cases}$$

See Section 12.3.4 for a justification of this assumption.

However, the bank can recover the borrower's debt in the event of default. This amount is called the recoveries amount. There are two types of recovery:

1. The bank may resell the property financed by the loan. Resale will be automatic if the bank took a mortgage guarantee but less certain if it didn't.
2. The bank may also be repaid by the borrower: inheritance, return to credit, etc..

Recoveries from borrowers are not so frequent. A lack of funds is the most common cause of default and recoveries usually result from the sale of the property. Once again for simplicity, we will assume that borrower recoveries are always zero and that any property sale is made immediately at the time of default. In reality this is seldom, if ever, the case. We assume that the property's value at time  $T$  is proportionate to the borrowed amount, the market price variation, the borrower's deposit, and the resale costs (including notary fees, legal procedures fees, etc.). We will assume that market price variation depends only on the time interval between drawdown date  $I_j$  and default date  $T$ . We can summarize all the assumptions put forward as follows:

$$(A_5) : \begin{cases} \text{Borrower recoveries are zero,} \\ \text{Property recoveries received at date } T \text{ are:} \\ \gamma M_j R_{T-I_j}, j \in B, I_j \leq T \leq I_j + D_j, \end{cases} \quad (11.2.8)$$

where  $\gamma \in \mathbb{R}_+$  is a coefficient independent of borrowers (see (12.3.11)), and  $(R_t)_{t \geq 0}$  represents the market price fluctuation with  $R_0 = 1$ . Through (11.2.8) we see that  $R_t$  does not depend on the borrowers, so there is a dependence between the sale amount of each borrower. Details on this model are given in Section 12.3.5.

For numerical applications we will assume that  $(R_t)_{t \geq 0}$  is the geometric Brownian motion with drift  $\mu$  and volatility  $\sigma$  so that  $R_0 = 1$ , that is to say:

$$R_t = \exp \left\{ \sigma B_t + \mu t \right\}, \quad (11.2.9)$$

where  $(B_t)_{t \geq 0}$  is a standard Brownian motion.

### 11.2.2 Claim amount

We define the claim amount concerning a borrower who has a default at time  $T$ . This is the key notion of our model.

**Definition 11.2.1.** Let us consider borrower  $j$  who has default at date  $T$ . We denote:

1.  $S_T^{1,j}$  the sum, at date  $T$ , of unpaid capital until the maturity, unpaid interest until the date of default, and additional interest for late payment,
2.  $S_T^{2,j}$  the recoveries.

The claim amount is defined by:

$$(S_{T-I_j}^{1,j} - S_{T-I_j}^{2,j})_+, \quad I_j \leq T \leq I_j + D_j. \quad (11.2.10)$$

Note that even if the claim occurs at date  $T$  our notations take into account the fact that loan was drawdown at date  $I_j$  and not at date 0, it is why we have a shift in our notation :  $T - I_j$ .

It is clear that (11.2.10) implies that the claim amount cannot be negative. Indeed, if the sum of unpaid capital, unpaid interest, and additional interest is lower than the recoveries then the insurance company will not pay anything to the bank. The claim amount does not represent the actual amount of loss but only the amount that the insurance company will repay to the bank in the event of default.

Under assumptions  $(A_1)$  to  $(A_5)$ , the quantities  $S^{1,j}$  and  $S^{2,j}$  can be expressed in terms of individual characteristics of borrower  $j$  ( $D_j$  and  $M_j$ ),  $\gamma$  (see (12.3.11)), and the market price fluctuation  $((R_t)_{t \geq 0})$ .

**Proposition 11.2.2.** Let assumptions  $(A_1)$  to  $(A_5)$  hold.

1. We have:

$$S_t^{1,j} := M_j \phi_t(D_j), \quad (11.2.11)$$

$$S_t^{2,j} := \gamma M_j R_t, \quad (11.2.12)$$

where  $\gamma$  and  $(R_t)_{t \geq 0}$  are introduced in Assumption  $(A_5)$  and:

$$\phi_t(d) := \begin{cases} \alpha, & \text{if } s \leq t^*, \\ \alpha, & \text{if } t \leq t^* \leq s, \\ \alpha \frac{e^{\tau s} - e^{\tau(t-t^*)}}{e^{\tau s} - 1}, & \text{if } 0 < t^* < t \leq s, \end{cases} \quad (11.2.13)$$

$$\alpha := e^{\tau t^*} (1 + t^* p) - t^* p, \quad (11.2.14)$$

with  $p \in \mathbb{R}_+$  represents coefficient due to late payment interest (see (11.5.7) for more details).

2. So by combining (11.2.10), (11.2.11), and (11.2.12) we have:

$$(S_{T-I_j}^{1,j} - S_{T-I_j}^{2,j})_+ := M_j f(T - I_j, D_j, R_{T-I_j})_+, \quad (11.2.15)$$

where the function  $f$  is defined by:

$$f(t, d, u) = (\phi_t(d) - \gamma u)_+. \quad (11.2.16)$$

*Proof.* See Section 11.5.1. □

### 11.2.3 Characteristic of claims, a model of provision

Let  $T_j^i$  be the r.v. representing the maturity of the loan drawn down during month  $i$  by borrower  $j$ . We deduce from notations introduced in Section 11.2.1:

$$t^*i \leq T_j^i \leq t^*i + D_j. \quad (11.2.17)$$

**Remark 11.2.3.**

1. *To date, in practice, there have been no defaults submitted at the end of a loan. On the other hand, if a default appears at the end of a loan then the claim amount will be insignificant. So we do not take into account claims submitted at date  $T_j^i = t^*i + D_j$ . See Section 12.3.9 for a justification of this assumption. It is important to distinguish two different situations in the previous inequality:*
  - $T_j^i = t^*i + D_j$ , if the borrower repays his loan normally until maturity,
  - $T_j^i < t^*i + D_j$ , if the borrower defaults before maturity.
2. *It is clear that relation (11.2.17) implies that when a claim occurs, there is a dependence between the loan duration and the date of default.*

The fact that a borrower may or may not default leads to a partition of the “fluctuation set” of the couples (Maturity of the loan, Term of the loan).

**Definition 11.2.4.** We denote  $A_{t_0-}^i$  (respectively  $A_{t_0+}^i$ ) as the “fluctuation set” of the couples (Maturity of the loan, Term of the loan) concerning loans drawn down during month  $i$  and which have defaulted in the period  $[t^*i, t_0]$  (respectively  $]t_0, t^*i + h[$ ):

$$A_{t_0-}^i := \{(d, t) \in \mathbb{R}_+^2 : t^*i \leq t \leq t_0; t < t^*i + d; d \leq h\}, \quad (11.2.18)$$

$$A_{t_0+}^i := \{(d, t) \in \mathbb{R}_+^2 : t_0 < t < t^*i + d; d \leq h\}, \quad (11.2.19)$$

where  $h$  is introduced in assumption  $(A_2)$ . This is an illustration:

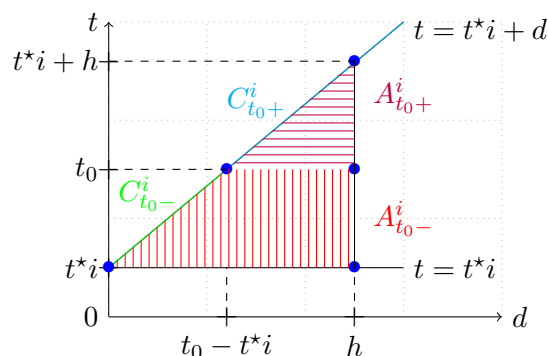


Illustration 11.2: Sets  $A_{t_0-}^i$  and  $A_{t_0+}^i$ .

Now we would like to consider loans drawn down during month  $i$ . We assume:

$$(A_6) : \text{The cloud of points } (T_j^i, D_j)_{j \in B} \text{ is a Poisson point process (PPP for short) with intensity } \Lambda_i. \quad (11.2.20)$$

See [19] and [21] for more details on these processes and Section 12.3.6 for a justification of this assumption.

If we assume that the drawdown rules do not change from one month to another then a default can only be explained by a change in the borrower's situation personal, the month of drawdown does not affect on the default. So we will assume that:

$$(A_7) : \text{The PPP } \left( (T_j^i, D_j)_{j \in B} \right)_{i \geq 0} \text{ are independent,} \quad (11.2.21)$$

and:

$$(A_8) : \begin{cases} (T_j^i, D_j)_{j \in B}, (M_j)_{j \in B}, \text{ and } (R_t)_{t \geq 0} \text{ are independent,} & (11.2.22) \\ \text{The variables } (M_j)_{j \in B} \text{ are i.i.d..} & (11.2.23) \end{cases}$$

See Sections 12.3.7 and 12.3.8 for a justifications of theses assumptions.

Now we define the provision on  $]t_0, t^*i + h[$  in a stochastic way as a sum a individual claims amounts.

**Definition 11.2.5.** *Provision on  $]t_0, t^*i + h[$ .*

1. *The claims amount for all loans drawn down during month  $i$ ,  $t^*i \leq t_0$ , and submitted as claims in  $]t_0, t^*i + h[$  is called “provision for month  $i$ ”, denoted by  $P_{t_0}^i$  (‘ $P$ ’ for “provision”) and defined as the sum of each individual claim amount:*

$$P_{t_0}^i := \sum_{j \geq 1} \mathbb{1}_{\{(T_j^i, D_j) \in A_{t_0+}^i\}} M_j f(T_j^i - t^*i, D_j, R_{T_j^i - t^*i}), \quad (11.2.24)$$

where  $A_{t_0+}^i$  is defined by (11.2.19) and  $M_j f(T_j^i - t^*i, D_j, R_{T_j^i - t^*i})$  is given by (13.2.4) and represents the claim amount at time  $T_j^i$  concerning borrower  $j$  who drew down his loan during month  $i$ .

2. *The claims amount for all loan drawdowns prior to  $t_0$  and occurring as claims in  $]t_0, t_0 + h[$  is called the “total provision”, denoted by  $P_{t_0}$ , and defined as the sum of monthly provisions:*

$$P_{t_0} := \sum_{i=0}^{\lfloor \frac{t_0}{t^*} \rfloor + 1} P_{t_0}^i. \quad (11.2.25)$$

**Remark 11.2.6.**

1. *Determining the total provision as the sum of monthly provisions is a constraint imposed by the insurance company.*



2. As we take into account the date of defaults in their amounts we have not included discounting in the modeling of the claim amount. However if we do this by adding an exponential factor in Definition (11.2.24), our results would be true.
3. It is possible to interpret the provision with exchange options: see Section 12.2 for more details. This interpretation will allow us to speak about the investment of premiums.

#### 11.2.4 Relation with other models

In Cramer - Lundberg's model (see [52]) the provision on  $]t_0, t_1]$ ,  $0 \leq t_0 < t_1 < +\infty$ , is modeled by:

$$\sum_{j \geq 1} Y^j \mathbb{1}_{\{t_0 < T_j \leq t_1\}},$$

where  $T_j$  is the default time of  $j$ -th claim and  $(Y^j)_{j \geq 1}$  is a sequence of i.i.d. r.v. representing the claim amounts. The r.v.'s  $(Y^j)_{j \geq 1}$  are assumed to be independent of  $(T_j)_{j \geq 1}$ . So this model does not allow us to consider the dependence between the date of default and the claim amount, the dependence between the borrower's situation and the claim amount or the dependence between all the claim amounts.

In the Mack model (see [35]),  $C_{i,k+1}$  is the claim amount, for the accident year  $i$  (corresponding to the drawn down date in our case) and the development year  $k$ , submitted as claims in  $]i; i+k+1]$ . Therefore, according to Definition 11.2.5 this quantity can be expressed as:

$$C_{i,k+1} = C_{i,k} + \sum_{j \geq 1} \mathbb{1}_{\{(T_j^i, D_j) \in A_{(i+k+1)^+}^i \setminus A_{(i+k)^+}^i\}} M_j f(T_j^i - t^*i, D_j, R_{T_j^i - t^*i}), \quad (11.2.26)$$

Let  $\mathcal{F}_{i,k}$  be the information available at date  $i+k$ , i.e.:

$$\mathcal{F}_{i,k} := \sigma \left[ (T_j^i, T_j^i \leq it^* + k)_{j \in B}; (R_t)_{t \leq k}; (D_j)_{j \in B: I_j = t^*i}; (M_j)_{j \in B: I_j = t^*i} \right]. \quad (11.2.27)$$

Assume that (11.2.9) holds, since  $(T_j^i, D_j)_{j \in B}$  is a PPP and conditionally to  $\mathcal{F}_{i,k}$ :

$$(R_{s+it^*+k}; 0 \leq s \leq 1) \stackrel{(d)}{=} (R_{it^*+k} R'_s; 0 \leq s \leq 1), \quad (11.2.28)$$

where  $R'$  is an independent copy of  $R$ , it follows that:

$$E \left[ C_{i,k+1} | \mathcal{F}_{i,k} \right] = C_{i,k} + \beta_{i,k}(R_k), \quad (11.2.29)$$

where:

$$\beta_{i,k}(x) := \mathbb{E} \left[ \sum_{j \geq 1} \mathbb{1}_{\{(T_j^i, D_j) \in A_{(i+k+1)^+}^i \setminus A_{(i+k)^+}^i\}} M_j f(T_j^i - t^*i, D_j, x R_{T_j^i - t^*i}) \right]. \quad (11.2.30)$$

So our model is drastically different from the Mack model, where the main assumption is:

$$E\left[C_{i,k+1}|\mathcal{F}_{i,k}\right] = f_k C_{i,k}, \quad (11.2.31)$$

where  $f_k$  is a deterministic coefficient independent of  $i$ . Our model, from a collective view point, is an additive model where the coefficient  $\beta_{i,k}$  (see (11.2.30)) takes into account the month of draw down  $i$  and the development date  $k$ .

## 11.3 Results

We will start by calculating the mean and the variance of the provision then we will deal with the cumulative distribution function (c.d.f.) of this r.v..

### 11.3.1 Mean and variance of the provision

**Notation 11.3.1.** *It is convenient to adopt the following notations:*

$$F_1(t, x, \mu, \sigma) := x\Phi\left[\frac{\ln(x) - \mu t}{\sigma\sqrt{t}}\right] - e^{(\mu + \frac{\sigma^2}{2})t}\Phi\left[\frac{\ln(x) - (\mu + \sigma^2)t}{\sigma\sqrt{t}}\right], \quad (11.3.1)$$

where  $x \in \mathbb{R}_+$ ,  $\Phi$  is the c.d.f. of the standard normal distribution, and  $\mu, \sigma$  are the parameters of the geometric Brownian motion defined by (11.2.9).

**Theorem 11.3.2.** *Let assumptions  $(A_1)$  to  $(A_8)$  hold. We also assume that  $(R_t)_{t \geq 0}$  is the geometric Brownian motion with drift  $\mu$  and volatility  $\sigma$  defined by (11.2.9). So we have:*

$$\mathbb{E}\left[P_{t_0}^i\right] = \gamma \mathbb{E}[M_1] \int_{A_{t_0+}^i} F_1\left[t - t^*i, \frac{\phi_{t-t^*i}(d)}{\gamma}, \mu, \sigma\right] d\Lambda_i(t, d), \quad (11.3.2)$$

$$\mathbb{E}\left[P_{t_0}\right] = \gamma \mathbb{E}[M_1] \sum_{i=0}^{\lfloor \frac{t_0}{t^*} \rfloor} \int_{A_{t_0+}^i} F_1\left[t - t^*i, \frac{\phi_{t-t^*i}(d)}{\gamma}, \mu, \sigma\right] d\Lambda_i(t, d), \quad (11.3.3)$$

where  $F_1$  and  $\phi_t(d)$  are respectively defined by (11.3.1) and (11.2.13), and  $\mathbb{E}[M_1]$  is the mean of borrowed amount.

The proof is given in Section 11.5.2. Formula (11.3.2) and (11.3.3) are not only theoretical. When the measure  $\Lambda_i$  is the Lebesgue measure and when the parameters are given, it is also possible to calculate a numerical value of the provision mean, see Section 12.4.2.

Now we will focus on the calculation of variance of the provision.

**Notation 11.3.3.** *Let  $(\alpha, \beta) \in \mathbb{R}^2$ . We denote:*

$$\tilde{d}(x, y, z, t) := \frac{\ln(x) - yt}{z\sqrt{t}}, \quad (11.3.4)$$

$$F_2(t, d) := 2 \frac{\phi_t(d)}{\gamma} F_1 \left[ t, \frac{\phi_t(d)}{\gamma}, \mu, \sigma \right] - F_1 \left[ t, \left( \frac{\phi_t(d)}{\gamma} \right)^2, 2\mu, 2\sigma \right], \quad (11.3.5)$$

$$F_3(t, d, t', d', x) := \frac{e^{\mu t + \sigma \sqrt{t} x - \frac{x^2}{2}}}{\sqrt{2\pi}} \left( \frac{\phi_t(d)}{\gamma} - e^{\mu t + \sigma \sqrt{t} x} \right) F_1 \left[ t' - t, \frac{\phi_{t'}(d')}{\gamma e^{\mu t + \sigma \sqrt{t} x}}, \mu, \sigma \right], \quad (11.3.6)$$

where  $\phi_t(d)$  and  $F_1$  are respectively defined by (11.2.13) and (11.3.1).

**Theorem 11.3.4.** *Let assumptions  $(A_1)$  to  $(A_8)$  hold. We also assume that  $(R_t)_{t \geq 0}$  is the geometric Brownian motion with drift  $\mu$  and volatility  $\sigma$  defined by (11.2.9). The variance of provision of month  $i$  is:*

$$\begin{aligned} \mathbb{V}[P_{t_0}^i] &= \gamma^2 \mathbb{E}[(M_1)^2] \int_{A_{t_0+}^i} F_2(t - t^*i, d) d\Lambda_i(t, d) + \gamma^2 [\mathbb{E}(M_1)]^2 \\ &\times \left[ \int_{A_{t_0+}^i \times A_{t_0+}^i} \left\{ \int_{-\infty}^{\tilde{d} \left( \frac{\phi_{t-t^*i}(d)}{\gamma}, \mu, \sigma, t-t^*i \right)} F_3 \left[ t - t^*i, d, t' - t^*i, d', x \right] dx \right\} \right. \\ &\left. d\Lambda_i(t, d) d\Lambda_i(t', d') - \left\{ \int_{A_{t_0+}^i} F_1 \left[ t - t^*i, \frac{\phi_{t-t^*i}(d)}{\gamma}, \mu, \sigma \right] d\Lambda_i(t, d) \right\}^2 \right], \quad (11.3.7) \end{aligned}$$

where  $\mathbb{E}[(M_1)^2]$  is the second moment of the borrowed amount.

The proof of this result is given in Section 11.5.3. The variance of the total provision can also be determined, see Section 11.5.6.

Formula (11.3.7) is not trivial. This is due to the fact that the claim amounts are not independent due to the process  $(R_t)_{t \geq 0}$ . Therefore, we have to take into account the correlation between two claim amounts.

**Remark 11.3.5.** *Without assuming that  $(R_t)_{t \geq 0}$  is the geometric Brownian motion defined by (11.2.9) we get:*

$$\begin{aligned} \mathbb{V}[P_{t_0}^i] &= \mathbb{V}[N_i(A_{t_0+}^i)] \mathbb{E} \left[ \left( M_1 f(T_1^i - t^*i, D_1, R_{T_1^i - t^*i}) \right)^2 \right] \\ &+ \left[ \mathbb{E}[N_i(A_{t_0+}^i)] \right]^2 \left\{ C^{i,i} - \left[ \mathbb{E}[M_1 f(T_1^i - t^*i, D_1, R_{T_1^i - t^*i})] \right]^2 \right\}, \quad (11.3.8) \end{aligned}$$

where the couple  $(T_1^i, D_1)$  represents a generic point of the PPP  $(T_j^i, D_j)_{j \in B}$  and:

$$N_i(A) := \sum_{j \geq 1} \mathbb{1} \left\{ (T_j^i, D_j) \in A \right\}, \quad (11.3.9)$$

$$C^{i,i'} := \left[ \mathbb{E}[M_1] \right]^2 \mathbb{E} \left[ g \left( T_1^i - t^*i, D_1, T_2^i - t^*i, D_2 \right) \right], \quad (11.3.10)$$

$$g(t, d, t', d') := E \left[ \left( \phi_t(d) - \gamma R_t \right)_+ \left( \phi_{t'}(d') - \gamma R_{t'} \right)_+ \right], \quad (11.3.11)$$

where the couple  $(T_2^i, D_2)$  represents an other generic point of the PPP  $(T_j^i, D_j)_{j \in B}$  and independent of  $(T_1^i, D_1)$ . Result (11.3.8) (proved in Section 11.5.3) allows us to identify sources of risk. The first term of this expression represents variability of the number of claims around the mean of the number of claims. The second term of (11.3.8) is, in the classical case of the compound Poisson process (independence of claim amount), equal to the variance of the claim amount. So, the second term represents variability of the claim amount around its mean.

When  $(R_t)_{t \geq 0}$  is the geometric Brownian motion defined by (11.2.9) we can explicitly evaluate the function  $g$  and so deduce an expression of the variance. We could replace the geometric Brownian motion with another process. However to get explicit formulas it need that  $\mathbb{E}[(\phi_t(s) - \gamma R_t)_+]$  and  $g(t, d, t', d')$  to be calculable.

Formula (11.3.7), as the one related to the mean, is not just theoretical. When the measure  $\Lambda_i$  is the Lebesgue measure and when the parameters are fixed, a numerical value of the variance of the provision can also be given.

### 11.3.2 The law of $P_{t_0}^i$

It is useful to order the loans maturities as follows:

$$T_{(1)}^i < T_{(2)}^i < \dots < T_{(n)}^i < \dots \quad (11.3.12)$$

We denote by  $D_{(j)}$  the loan duration associated to the date  $T_{(j)}^i$ .

**Theorem 11.3.6.** *Let assumptions  $(A_1)$  to  $(A_8)$  hold.*

1. Recall that:

$$\mathcal{L}[N_i(A_{t_0+}^i)] = \mathcal{P}[\Lambda_i(A_{t_0+}^i)].$$

2. The conditional law of  $P_{t_0}^i$  given  $N_i(A_{t_0+}^i)$  is:

$$\mathcal{L}[P_{t_0}^i | N_i(A_{t_0+}^i) = n] = \mathcal{L}[u_n((T_{(j)}^i)_{1 \leq j \leq n}, (D_{(j)})_{1 \leq j \leq n}, (M_j)_{1 \leq j \leq n}, (G_j)_{1 \leq j \leq n})], \quad (11.3.13)$$

where:

$$\begin{aligned} & u_n((t_{(j)}^i)_{1 \leq j \leq n}, (d_{(j)})_{1 \leq j \leq n}, (m_j)_{1 \leq j \leq n}, (g_j)_{1 \leq j \leq n}) \\ & := \sum_{j=1}^n m_j \left\{ \phi_{t_{(j)}^i - t^* i}(d_{(j)}) - \gamma \prod_{k=1}^j r_k \right\}_+, \end{aligned} \quad (11.3.14)$$

$$r_k = \exp \left[ \sigma \sqrt{t_{(k)}^i - t_{(k-1)}^i} g^k + \mu(t_{(k)}^i - t_{(k-1)}^i) \right], \text{ with this agreement: } t_{(0)}^i = 0, \quad (11.3.15)$$

$(G_k)_{k \geq 1}$  is a collection of Gaussian r.v. i.i.d. and independent of

$$((T_{(j)}^i, D_{(j)})_{1 \leq j \leq n}, (M_j)_{1 \leq j \leq n}). \quad (11.3.16)$$

*Proof.* See Section 11.5.4. □

Clearly the law of the provision is not explicit. However we formulated the law of the provision as a complicated function but dependent on a sequence of r.v.'s which can be easily simulated. Theorem 11.3.6 allows us to realize Monte Carlo simulation. However, we need to know the law of the amounts  $M_j$ . We have, for loans drawdown month  $i = 0$ ,  $b^i$  observations of the borrowed amount ;  $b^i$  is around 300, so it is reasonable to replace the c.d.f. of  $M_1$  by the empirical c.d.f.. We denote  $(m_j)_{1 \leq j \leq b^i}$  the borrowed amount observed and let  $F_M$  the empirical c.d.f., i.e.:

$$F_M(m) := \frac{1}{b^i} \sum_{j=1}^{b^i} \mathbb{1}_{\{m_j \leq m\}}.$$

We simulate the borrowed amount by reversing the empirical c.d.f.  $F_M$ , i.e. we simulate a random variable uniformly distributed on  $[0, 1]$  and we denote  $\zeta$  its realization. And then we calculate:  $m_{\lfloor b^i \zeta \rfloor + 1}$ . This allows us to get the following algorithm to simulate the provision (we assume that  $\Lambda_i$  is known):

---

**Algorithm 11.1** Simulation of a sample of length  $l$  for the r.v.  $P_{t_0}^i$ .

---

- 1: **for**  $k = 1$  to  $l$  **do**
- 2: Simulate a r.v. with distribution  $\mathcal{P}(\Lambda_i(A_{t_0+}^i))$ . Its realization is denoted by  $n$ .
- 3: Simulate  $n$  r.v. i.i.d. uniformly distributed on  $A_{t_0+}^i$ . The realizations are denoted by  $(t_j, d_j)_{1 \leq j \leq n}$ . Order dates  $(t_j)_{1 \leq j \leq n} : t_{(1)} \leq \dots \leq t_{(n)}$ . Duration associated at time  $t_{(j)}$  is denoted by  $d_{(j)}$ .
- 4: The variables  $(M_j)_{1 \leq j \leq n}$  are simulated by reversing the empirical c.d.f.. The realizations are denoted by  $(\tilde{m}_j)_{1 \leq j \leq n}$ .
- 5: Simulate  $n$  r.v.'s i.i.d. with standard normal distribution. The realizations are denoted by  $(g_j)_{1 \leq j \leq n}$ .
- 6: Calculate:

$$p_k^i = \sum_{j=1}^n \tilde{m}_j \left\{ \phi_{t_{(j)} - t^* i}(d_{(j)}) - \gamma \prod_{k=1}^j r_k \right\}_+, \quad (11.3.17)$$

where  $\phi$  and  $r_k$  are respectively defined by (11.2.13) and (11.3.15).

- 7: **end for**
- 

This algorithm will allow us, when the intensity of the PPP will be specified and when the parameters will take fixed values, to simulate the provision and then to deduce some risks indicators.

## 11.4 Parameters estimation

The theoretical model that we just brought forward has been created to calculate explicitly the provision with data from CAMCA Assurance. The theoretical analysis presented

in Section 11.3, with Theorems 11.3.2, 11.3.4, and 11.3.6, allows us to calculate numerically the mean, the variance, and the quantile of the provision when:

1. The measure  $\Lambda_i$  is specified,
2. The parameters are estimated.

The amount of data available on each borrower (amount borrowed, term, and default date) allows us to estimate the parameters of our model (see Section 11.4.2). But firstly we start by specifying the measure  $\Lambda_i$ .

### 11.4.1 Choice of the intensity $\Lambda_i$

Remembering the partition,  $A_{t_0-}^i$  and  $A_{t_0+}^i$  (see (11.2.18) and (11.2.19)), of the fluctuation set of the couples (Maturity of the loan, Term of the loan) concerns loans which have defaulted.

**Definition 11.4.1.** We denote  $C_{t_0-}^i$  (resp.  $C_{t_0+}^i$ ) as the “fluctuation set” of the couples (Maturity of the loan, Term of the loan) concerning loans drawn down during month  $i$  and which have matured in the period  $[t^*i, t_0]$  (resp.  $]t_0, t^*i + h]$ ):

$$C_{t_0-}^i := \{(d, t) \in \mathbb{R}_+^2 : t = t^*i + d; t \leq t_0; d \leq h\}, \quad (11.4.1)$$

$$C_{t_0+}^i := \{(d, t) \in \mathbb{R}_+^2 : t = t^*i + d; t > t_0; d \leq h\}. \quad (11.4.2)$$

This is an illustration:

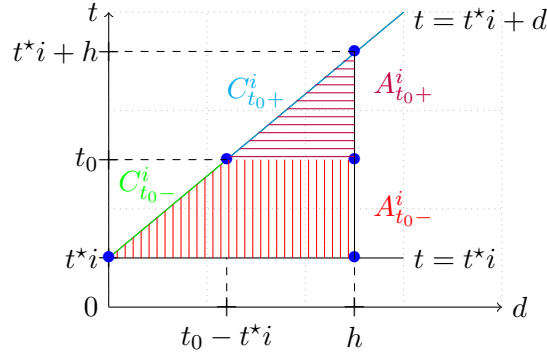


Illustration 11.3: Sets  $C_{t_0-}^i$  and  $C_{t_0+}^i$ .

Having disposed of this preliminary step, we can now choose the intensity  $\Lambda_i$ . We will assume that:

$$\Lambda_i := \lambda_1 \text{Leb}_1|_{C_{t_0-}^i \cup C_{t_0+}^i} + \lambda_2 \text{Leb}_2|_{A_{t_0-}^i \cup A_{t_0+}^i}, \quad (11.4.3)$$

where:

- $\lambda_1$  and  $\lambda_2$  are two real and strictly positive parameters,

- $Leb_{1|C_{t_0-}^i \cup C_{t_0+}^i}$  (resp.  $Leb_{2|A_{t_0-}^i \cup A_{t_0+}^i}$ ) is the Lebesgue measure on  $\mathbb{R}$  (resp. on  $\mathbb{R}^2$ ) restricted on  $C_{t_0-}^i \cup C_{t_0+}^i$  (resp.  $A_{t_0-}^i \cup A_{t_0+}^i$ ).

See Section 12.3.7 for a justification of this assumption.

### 11.4.2 Estimation of parameters

#### List of parameters:

We will start by listing the parameters:

1. The mean of borrowed amount:  $\mathbb{E}(M_1)$ ,
2. The monthly rate:  $\tau$  defined by (11.2.4),
3. The coefficient  $\gamma$  associated to the borrower's deposit and the selling costs of property, see (12.3.11),
4. The penalty coefficient  $p$ , in relation with late payment, see (11.5.7),
5. Coefficients  $\mu$  and  $\sigma$  of the geometric Brownian motion defined by (11.2.9),
6. Parameters  $\lambda_1$  and  $\lambda_2$  linked to the intensity of the PPP  $\Lambda_i$ , see (11.4.3),
7. Parameter  $h$  representing the maximal duration, see (11.2.1).

#### Data available:

At date  $t_0$ , we dispose of three different sources of data: data shared by all borrowers, data linked to claims, and data exogenous to the insurance company. Note that the first two sources are private because they are supplied by the insurance company while third is public. For each borrower  $j$  having drawn down his loan during month  $i$  so that  $it^* \leq t_0$ , we have the term  $d_j$ , the amount  $m_j$  and the maturity  $t_j^i$  of the loan. We also dispose of:

- The number  $b^i$  of loans drawn down during month  $i$ . The total number of loans drawn down prior to  $t_0$  is denoted by:

$$b := \sum_{i=0}^{i_0} b^i, \quad (11.4.4)$$

where:

$$i_0 := \lfloor \frac{t_0}{t^*} \rfloor \in \mathbb{N}^*. \quad (11.4.5)$$

- The number  $n^i$  of defaults submitted until date  $t_0$ ,
- The number  $v^i$  of loans which have matured prior to  $t_0$ .

So the number of loan “in force” at date  $t_0$ , that is to say loans which have not defaulted or matured, is given by  $b^i - n^i - v^i$ .

Finally, we have exogenous data provided by INSEE (“Institut National de la Statistique et des Etudes Economiques”) and by the broker “Empruntis”. INSEE gives an index of house prices in France which is denoted by  $u_t$  at date  $t$ . The broker “Empruntis” provides annual statistics about real estate loans, in particular the mean of the borrower’s deposit and the mean of actuarial loan rate. See Section 12.4.1 for descriptive statistics.

### Estimation of parameters

So the mean of borrowed amount will be estimated by empirical mean. The rate of each loan will be estimated with the actuarial rate given by “Empruntis” (actuarial rate estimate to 5%). Coefficient  $\gamma$  (see 12.3.11) is determined by the statistics from “Empruntis” (coefficient  $a$ , see (12.3.9), representing the borrower’s deposit is determined at 30%) and by “expert judgement” (coefficient  $c$ , see (12.3.10), representing selling costs of property is determined at 10%). Coefficient  $p$ , due to late payment interest, will be estimated by “expert judgement” at 2 (see (11.5.7)). Finally coefficients  $\mu$  and  $\sigma$  will be estimated by INSEE’s data.

We turn now to the estimation of  $\lambda_1$ ,  $\lambda_2$ , and the maximum term  $h$ . The maximum-likelihood estimators are denoted by  $\hat{\lambda}_1$ ,  $\hat{\lambda}_2$ , and  $\hat{h}$ .

**Proposition 11.4.2.** *Let assumptions  $(A_1)$  to  $(A_8)$  hold. Then  $x = \hat{\lambda}_1$ ,  $y = \hat{\lambda}_2$ , and  $z = \hat{h}$  are solutions of the following system:*

$$\sum_{i=0}^{i_0} \left[ \frac{v^i}{x} + \frac{(b^i - n^i - v^i)\sqrt{2}(z - t_0 + t^*i)}{x\sqrt{2}(z - t_0 + t^*i) + \frac{y(z-t_0+t^*i)^2}{2}} \right] = \sqrt{2}(i_0 + 1)z, \quad (11.4.6)$$

$$(S) : \sum_{i=0}^{i_0} \left[ \frac{n^i}{y} + \frac{(b^i - n^i - v^i)(z - t_0 + t^*i)^2}{2x\sqrt{2}(z - t_0 + t^*i) + y(z - t_0 + t^*i)^2} \right] = \frac{(i_0 + 1)z^2}{2}, \quad (11.4.7)$$

$$\begin{aligned} & \sum_{i=0}^{i_0} \left[ \frac{n^i(t_0 - t^*i)}{(t_0 - t^*i)\left(z - \frac{1}{2}(t_0 - t^*i)\right)} + \frac{(b^i - n^i - v^i)\left(\sqrt{2}x + y(z - t_0 + t^*i)\right)}{x\sqrt{2}(z - t_0 + t^*i) + \frac{y(z-t_0+t^*i)^2}{2}} \right] \\ & = (x\sqrt{2} + yz)(i_0 + 1). \end{aligned} \quad (11.4.8)$$

The proof of the Proposition 11.4.2 is given in Section 11.5.5. Note that the above system can be resolved numerically using  $\mathbb{R}$ .

Now we summarize the parameters estimation:



| Parameter                 | Notation                                    | Estimation   | Source                         |
|---------------------------|---|--|--------------------------------|
| $\mathbb{E}(M_1)$         | $\bar{m}$                                   | $\frac{1}{b} \sum_{i=1}^b m_j$                             | Insurance company              |
| $\tau$                    | $\hat{\tau}$                                | $\ln\left(1 + \frac{5}{100} \frac{1}{12}\right)$           | Empruntis                      |
| $\gamma$                  | $\hat{\gamma}$                              | 1, 17  | Expert judgement and Empruntis |
| $p$                       | $\hat{p}$                                   | 2  | Expert judgement               |
| $\mu$                     | $\hat{\mu}$                                 | $\frac{\sum_{i=1}^{i_0} x_i}{i_0 t^*}$                     | INSEE                          |
| $\sigma$                  | $\hat{\sigma}$                              | $\frac{\sum_{i=1}^{i_0} (x_i - \hat{\mu} t^*)^2}{i_0 t^*}$ | INSEE                          |
| $\lambda_1, \lambda_2, h$ | $\hat{\lambda}_1, \hat{\lambda}_2, \hat{h}$ | Maximum likelihood   | Insurance company              |

Illustration 11.4: Estimations.

where:

$$x_t = \ln(u_t) - \ln(u_{t-1}). \quad (11.4.9)$$

We will see in Section 12.4.2 of Chapter 12 how use these estimations to get numerical value of the mean, the variance and the quantiles of the provision.

## 11.5 Proof

### 11.5.1 Proof of Proposition 11.2.2

Let us consider borrower  $j$  having drawn down his loan at date  $I_j$  for an amount  $M_j$  and a term  $D_j$ . The constant rate of return is denoted by  $\tilde{\tau}$ . So:

$$(1 + \tilde{\tau})^{t^*} = e^{\tau t^*}. \quad (11.5.1)$$

From [20], pages 172 to 178, we obtain:

1. Capital repaid by borrower  $j$  at date  $t$  is:

$$C_t^j = \begin{cases} (e^{\tau t^*} - 1)M_j \left( \frac{e^{\tau(t-I_j-t^*)}}{e^{\tau D_j} - 1} \right), & \text{if } I_j < t \leq I_j + D_j, \\ 0, & \text{if } t = I_j. \end{cases} \quad (11.5.2)$$

2. Interest repaid by borrower  $j$  at date  $t$  is:

$$U_t^j = \begin{cases} (e^{\tau t^*} - 1)M_j \left( \frac{e^{\tau D_j} - e^{\tau(t-I_j-t^*)}}{e^{\tau D_j} - 1} \right), & \text{if } I_j < t \leq I_j + D_j, \\ 0, & \text{if } t = I_j. \end{cases} \quad (11.5.3)$$

3. Monthly installment paid by borrower  $j$  at date  $t$  is:

$$ME_t^j = \begin{cases} (e^{\tau t^*} - 1)M_j \left( \frac{e^{\tau D_j}}{e^{\tau D_j} - 1} \right), & \text{if } I_j < t \leq I_j + D_j, \\ 0, & \text{if } t = I_j. \end{cases} \quad (11.5.4)$$

Now we assume that this borrower defaults at the deterministic date  $T$ . The assumption  $(A_4)$  implies that the bank reports the claim immediately. Definition 11.2.1, formula (11.5.1), (11.5.2), (11.5.3), and (11.5.4) allow us to determine the different factors making up the claim amount:

1. The unpaid capital from date  $T$  until the loan maturity,  $I_j + D_j$ , is:

$$\sum_{t=T}^{I_j+D_j} C_t^j = \begin{cases} M_j \left( \frac{e^{\tau D_j} - e^{\tau(T-I_j-t^*)}}{e^{\tau D_j} - 1} \right), & \text{if } I_j < T \leq I_j + D_j, \\ M_j, & \text{if } T = I_j. \end{cases} \quad (11.5.5)$$

2. The unpaid interest for the monthly installment at date  $T$  is:

$$U_T^j = \begin{cases} (e^{\tau t^*} - 1)M_j \left( \frac{e^{\tau D_j} - e^{\tau(T-I_j-t^*)}}{e^{\tau D_j} - 1} \right), & \text{if } I_j < T \leq I_j + D_j, \\ 0, & \text{if } T = I_j. \end{cases} \quad (11.5.6)$$

3. Any missed monthly installment will incur penalty interest charges proportional to the delay. This penalty interest is calculated from the date of the last payment until the claim is submitted. The penalty interest is equal to the monthly interest multiplied by a coefficient. The multiplicative coefficient due to late payment interest is fixed for every borrowers and this coefficient will not change. It will be denoted by  $p$  (“ $p$ ” for “penalty”). The additional interest on the unpaid monthly installments only is:

$$t^*pU_T^j = \begin{cases} t^*p(e^{\tau t^*} - 1)M_j \left( \frac{e^{\tau D_j} - e^{\tau(T-I_j-t^*)}}{e^{\tau D_j} - 1} \right), & \text{if } I_j < T \leq I_j + D_j, \\ 0, & \text{if } T = I_j, \end{cases} \quad (11.5.7)$$

The additional interest is multiplied by the number of months that the monthly payment remains unpaid: in this case we multiply by  $t^*$  because the duration of unpaid interest until the reporting date is one month (from date  $T - t^*$  to date  $T$ ) according on the assumption  $(A_4)$ .

4. The recoveries received at date  $T$  is:

$$\gamma M_j R_{T-I_j}, \quad (11.5.8)$$

where  $\gamma$  is defined by (12.3.11). See assumption  $(A_5)$  for more details.

By combining (11.5.5), (11.5.6), and (11.5.7) we prove (11.2.11). Formula (11.2.12) is given by (11.5.8).

### 11.5.2 Proof of Theorem 11.3.2

The proof of Theorem 11.3.2 is based on the following lemma:

**Lemma 11.5.1.** *Let  $K \in \mathbb{R}_+$ . We consider  $(\alpha, \beta) \in \mathbb{R}^2$  and  $(B_t)_{t \geq 0}$  a standard Brownian motion. Then:*

$$\mathbb{E}\left[\left(K - \exp[\alpha B_t + \beta t]\right)_+\right] = F_1(t, K, \beta, \alpha), \quad (11.5.9)$$

where the function  $F_1$  is defined by (11.3.1).

*Proof of Lemma 11.5.1.* Analysis similar to that in [31] page 303 shows that (11.5.9) is proved.  $\square$

Having disposed of this preliminary step, we can now prove the Theorem 11.3.2. Let assumptions  $(A_1)$  to  $(A_8)$  hold. We assume that  $(R_t)_{t \geq 0}$  is the geometric Brownian motion defined by (11.2.9).  $(T_j^i, D_j)_{j \in B}$  is a PPP, so given  $N_i(A_{t_0+}^i) = k$  the couples  $(T_j^i, D_j)_{1 \leq j \leq k}$  are i.i.d., with law  $\frac{1}{\Lambda_i(A_{t_0+}^i)} \Lambda_i|_{A_{t_0+}^i}$ , where  $N_i$  is defined by (11.3.9). So:

$$\begin{aligned} \mathbb{E}[P_{t_0}^i] &= \sum_{k \geq 0} P[N_i(A_{t_0+}^i) = k] \mathbb{E}\left[\sum_{j=1}^k S(T_j^i - t^*i, D_j, M_j, R_{T_j^i - t^*i})\right] \\ &= \mathbb{E}[N_i(A_{t_0+}^i)] \mathbb{E}[M_1 f(T_1^i - t^*i, D_1, R_{T_1^i - t^*i})] \end{aligned} \quad (11.5.10)$$

$$= \mathbb{E}[N_i(A_{t_0+}^i)] \mathbb{E}[M_1] \int_{A_{t_0+}^i} \mathbb{E}\left[\left\{\phi_{t-t^*i}(d) - \gamma R_{t-t^*i}\right\}_+\right] \frac{d\Lambda_i(t, d)}{\Lambda_i(A_{t_0+}^i)}, \quad (11.5.11)$$

where  $\phi_t(d)$  is defined by (11.2.13). By combining (11.5.11) and (11.5.9) we have:

$$\mathbb{E}[P_{t_0}^i] = \gamma \mathbb{E}[M_1] \int_{A_{t_0+}^i} F_1\left(t - t^*i, \frac{\phi_{t-t^*i}(d)}{\gamma}, \mu, \sigma\right) d\Lambda_i(t, d).$$

This established (11.3.2). It follows easily that formulas (11.3.3) are proved.

### 11.5.3 Proof of Theorem 11.3.4

Let assumptions  $(A_1)$  to  $(A_8)$  hold. Our proof starts with a technical lemma:

**Lemma 11.5.2.** *Let assumptions  $(A_1)$  to  $(A_8)$  hold.*

1. *For all  $j \neq l$  we have:*

$$\mathbb{E}\left[M_j f(T_j^i - t^*i, D_j, R_{T_j^i - t^*i}) M_l f(T_l^i - t^*i, D_l, R_{T_l^i - t^*i})\right] = C^{i,i}, \quad (11.5.12)$$

where  $C^{i,i}$  is defined by (11.3.10). For all  $i \neq i'$  we have:

$$\mathbb{E}\left[M_j f(T_j^i - t^*i, D_j, R_{T_j^i - t^*i}) M_l f(T_l^{i'} - t^*i', D_l, R_{T_l^{i'} - t^*i'})\right] = C^{i,i'}, \quad (11.5.13)$$

2. Let  $K \in \mathbb{R}_+$ . Consider  $(\alpha, \beta) \in \mathbb{R}^2$  and  $(B_t)_{t \geq 0}$  a standard Brownian motion. Then:

$$\mathbb{E}\left[\left\{\left(K - \exp[\alpha B_t + \beta t]\right)_+\right\}^2\right] = 2KF_1(t, K, \beta, \alpha) - F_1(t, K^2, 2\beta, 2\alpha), \quad (11.5.14)$$

where  $F_1$  is defined by (11.3.1).

3. We have:

$$\begin{aligned} \mathbb{V}[P_{t_0}^i] &= \mathbb{V}[N_i(A_{t_0+}^i)] \mathbb{E}\left[\left(M_1 f(T_1^i - t^*i, D_1, R_{T_1^i - t^*i})\right)^2\right] \\ &\quad + \left[\mathbb{E}[N_i(A_{t_0+}^i)]\right]^2 \left\{C^{i,i} - \left[\mathbb{E}[M_1 f(T_1^i - t^*i, D_1, R_{T_1^i - t^*i})]\right]^2\right\}. \end{aligned} \quad (11.5.15)$$

4. If  $(R_i)_{t \geq 0}$  is the geometric Brownian motion defined by (11.2.9) then:

$$\mathbb{E}\left[\left(M_1 f(T_1^i - t^*i, D_1, R_{T_1^i - t^*i})\right)^2\right] = \gamma^2 \mathbb{E}[(M_1)^2] \int_{A_{t_0+}^i} F_2(t - t^*i, d) \frac{d\Lambda_i(t, d)}{\Lambda_i(A_{t_0+}^i)}, \quad (11.5.16)$$

$$\begin{aligned} C^{i,i'} &= \gamma^2 \left[\mathbb{E}[M_1]\right]^2 \int_{A_{t_0+}^i \times A_{t_0+}^{i'}} \left\{ \int_{-\infty}^{\tilde{d}\left(\frac{\phi_{t-t^*i}(d)}{\gamma}, \mu, \sigma, t-t^*i\right)} F_3\left(t - t^*i, d, t' - t^*i', d', x\right) dx \right\} \\ &\quad \frac{d\Lambda_i(t, d) d\Lambda_{i'}(t', d')}{\Lambda_i(A_{t_0+}^i) \Lambda_{i'}(A_{t_0+}^{i'})}, \end{aligned} \quad (11.5.17)$$

where functions  $\tilde{d}$ ,  $F_2(t, d)$ , and  $F_3$  are defined by (11.3.4), (11.3.5), and (11.3.6).

The proof of this lemma will be made at the end of this section. Having disposed of this preliminary step, we can now prove the Theorem 11.3.4. We want to calculate the variance of the monthly provision. By combining (11.5.15), (11.3.2), (11.5.16), and (11.5.17) we obtain:

$$\begin{aligned} \mathbb{V}[P_{t_0}^i] &= \gamma^2 \mathbb{E}[(M_1)^2] \int_{A_{t_0+}^i} F_2(t - t^*i, d) d\Lambda_i(t, d) + \gamma^2 \left[\mathbb{E}[M_1]\right]^2 \\ &\quad \times \left[ \int_{A_{t_0+}^i \times A_{t_0+}^{i'}} \left\{ \int_{-\infty}^{\tilde{d}\left(\frac{\phi_{t-t^*i}(d)}{\gamma}, \mu, \sigma, t-t^*i\right)} F_3\left(t - t^*i, d, t' - t^*i', d', x\right) dx \right\} \right. \\ &\quad \left. d\Lambda_i(t, d) d\Lambda_{i'}(t', d') - \left\{ \int_{A_{t_0+}^i} F_1\left(t - t^*i, \frac{\phi_{t-t^*i}(d)}{\gamma}, \mu, \sigma\right) d\Lambda_i(t, d) \right\}^2 \right]. \end{aligned}$$

This established (11.3.7). Let us come back to the proof of Lemma 11.5.2.

*Proof of Lemma 11.5.2.*

1. By combining (11.2.16) and (11.3.11) we obtain that  $\mathbb{E}\left[M_j f(T_j^i - t^*i, D_j, R_{T_j^i - t^*i}) M_l f(T_l^i - t^*i, D_l, R_{T_l^i - t^*i})\right]$  is equal to:

$$\begin{aligned} & \mathbb{E}\left[M_j M_l \left\{ \phi_{T_j^i - t^*i}(D_j) - \gamma R_{T_j^i - t^*i} \right\}_+ \left\{ \phi_{T_l^i - t^*i}(D_l) - \gamma R_{T_l^i - t^*i} \right\}_+\right] \\ &= \left[\mathbb{E}[M_1]\right]^2 \mathbb{E}\left[\left\{ \phi_{T_j^i - t^*i}(D_j) - \gamma R_{T_j^i - t^*i} \right\}_+ \left\{ \phi_{T_l^i - t^*i}(D_l) - \gamma R_{T_l^i - t^*i} \right\}_+\right] \\ &= \left[\mathbb{E}[M_1]\right]^2 \mathbb{E}\left[g\left(T_1^i - t^*i, D_1, T_2^i - t^*i, D_2\right)\right] \\ &= C^{i,i}, \end{aligned}$$

as  $(T_j^i, D_j)_{j \in B}$  is a PPP then the r.v.'s  $(T_j^i, D_j)_{j \geq 1}$  are i.i.d.. This established (11.5.12).

In the same manner  $\mathbb{E}\left[M_j f(T_j^i - t^*i, D_j, R_{T_j^i - t^*i}) M_l f(T_l^{i'} - t^*i', D_l, R_{T_l^{i'} - t^*i'})\right]$  is equal to:

$$\begin{aligned} & \mathbb{E}\left[M_j M_l \left\{ \phi_{T_j^i - t^*i}(D_j) - \gamma R_{T_j^i - t^*i} \right\}_+ \left\{ \phi_{T_l^{i'} - t^*i'}(D_l) - \gamma R_{T_l^{i'} - t^*i'} \right\}_+\right] \\ &= \left[\mathbb{E}[M_1]\right]^2 \mathbb{E}\left[\left\{ \phi_{T_j^i - t^*i}(D_j) - \gamma R_{T_j^i - t^*i} \right\}_+ \left\{ \phi_{T_l^{i'} - t^*i'}(D_l) - \gamma R_{T_l^{i'} - t^*i'} \right\}_+\right] \\ &= \left[\mathbb{E}[M_1]\right]^2 \mathbb{E}\left[g\left(T_j^i - t^*i, D_j, T_l^{i'} - t^*i', D_l\right)\right] \\ &= \left[\mathbb{E}[M_1]\right]^2 \mathbb{E}\left[g\left(T_1^i - t^*i, D_1, T_{i',2} - t^*i', D_2\right)\right] \\ &= C^{i,i'}. \end{aligned}$$

This established the formula (11.5.13).

2. If  $x \geq 0$  then:

$$\left((K - x)_+\right)^2 = 2K(K - x)_+ - (K^2 - x^2)_+,$$

Applying Lemma 11.5.1 and the last observation we can conclude:

$$\begin{aligned} & \mathbb{E}\left[\left\{ \left(K - \exp[\alpha B_t + \beta t]\right)_+ \right\}^2\right] \\ &= 2K \mathbb{E}\left[\left(K - \exp[\alpha B_t + \beta t]\right)_+\right] - E\left[\left(K^2 - \exp[2\alpha B_t + 2\beta t]\right)_+\right], \\ &= 2KF_1(t, K, \beta, \alpha) - F_1(t, K^2, 2\beta, 2\alpha). \end{aligned}$$

This established the formula (11.5.14).

3. First we calculate the second moment of the monthly provision. Then we can deduce the variance of the monthly provision. We have:

$$\mathbb{E}\left[(P_{t_0}^i)^2\right] = \sum_{k \geq 1} \mathbb{P}\left[N_i(A_{t_0+}^i) = k\right] \mathbb{E}\left[\left\{ \sum_{j=1}^k M_j f(T_j^i - t^*i, D_j, R_{T_j^i - t^*i}) \right\}^2\right]$$

$$= A + B, \quad (11.5.18)$$

where:

$$A = \sum_{k \geq 1} \mathbb{P} \left[ N_i(A_{t_0+}^i) = k \right] \mathbb{E} \left[ \sum_{j=1}^k \left( M_j f(T_j^i - t^* i, D_j, R_{T_j^i - t^* i}) \right)^2 \right],$$

and

$$B = \sum_{k \geq 1} \mathbb{P} \left[ N_i(A_{t_0+}^i) = k \right] \mathbb{E} \left[ 2 \sum_{1 \leq j < l \leq k} M_j f(T_j^i - t^* i, D_j, R_{T_j^i - t^* i}) \right. \\ \left. \times M_l f(T_l^i - t^* i, D_l, R_{T_l^i - t^* i}) \right].$$

We begin by calculating  $A$ . Analysis similar to that in the proof of Theorem 11.3.2, shows that:

$$A = \sum_{k \geq 1} \mathbb{P} \left[ N_i(A_{t_0+}^i) = k \right] \sum_{j=1}^k \mathbb{E} \left[ \left( M_1 f(T_1^i - t^* i, D_1, R_{T_1^i - t^* i}) \right)^2 \right] \\ = \sum_{k \geq 1} \mathbb{P} \left[ N_i(A_{t_0+}^i) = k \right] k \mathbb{E} \left[ \left( M_1 f(T_1^i - t^* i, D_1, R_{T_1^i - t^* i}) \right)^2 \right] \\ = \mathbb{E} \left[ N_i(A_{t_0+}^i) \right] \mathbb{E} \left[ \left( M_1 f(T_1^i - t^* i, D_1, R_{T_1^i - t^* i}) \right)^2 \right]. \quad (11.5.19)$$

Now we calculate  $B$ . (11.5.12) implies:

$$B = \sum_{k \geq 1} \mathbb{P} \left[ N_i(A_{t_0+}^i) = k \right] \times 2 \sum_{1 \leq j < l \leq k} C^{i,i} \\ = \sum_{k \geq 1} \mathbb{P} \left[ N_i(A_{t_0+}^i) = k \right] k(k-1) C^{i,i} \\ = \mathbb{E} \left[ N_j^2(A_{t_0+}^i) \right] C^{i,i} - \mathbb{E} \left[ N_i(A_{t_0+}^i) \right] C^{i,i} \\ = \left[ \mathbb{E} \left[ N_i(A_{t_0+}^i) \right] \right]^2 C^{i,i}, \quad (11.5.20)$$

where  $C^{i,i}$  is defined by (11.3.10). By combining (11.5.18), (11.5.19), and (11.5.20) we obtain:

$$\mathbb{E} \left[ (P_{t_0}^i)^2 \right] = \mathbb{E} \left[ N_i(A_{t_0+}^i) \right] \mathbb{E} \left[ \left( M_1 f(T_1^i - t^* i, D_1, R_{T_1^i - t^* i}) \right)^2 \right] + \left[ \mathbb{E} \left[ N_i(A_{t_0+}^i) \right] \right]^2 C^{i,i}. \quad (11.5.21)$$

Let us come back to calculation of the monthly provision variance. Combining (11.5.10) and (11.5.21) we conclude that:

$$\mathbb{V} \left[ P_{t_0}^i \right] = \mathbb{E} \left[ (P_{t_0}^i)^2 \right] - \left[ \mathbb{E} \left[ P_{t_0}^i \right] \right]^2$$

$$\begin{aligned}
&= \mathbb{V} \left[ N_i(A_{t_0+}^i) \right] \mathbb{E} \left[ \left( M_1 f(T_1^i - t^* i, D_1, R_{T_1^i - t^* i}) \right)^2 \right] \\
&+ \left[ \mathbb{E} \left[ N_i(A_{t_0+}^i) \right] \right]^2 \left\{ C^{i,i} - \mathbb{E}^2 \left[ M_1 f(T_1^i - t^* i, D_1, R_{T_1^i - t^* i}) \right] \right\}.
\end{aligned}$$

This established the formula (11.5.15).

4. Assumption  $(A_8)$ , (13.2.4) and (11.2.9) imply:

$$\mathbb{E} \left[ \left( M_1 f(T_1^i - t^* i, D_1, R_{T_1^i - t^* i}) \right)^2 \right] = \mathbb{E} \left[ (M_1)^2 \right] \int_{A_{t_0+}^i} \mathbb{E} \left[ \left\{ (\phi_{t-t^* i}(d) - \gamma R_{t-t^* i})_+ \right\}^2 \right] \frac{d\Lambda_i(t, d)}{\Lambda_i(A_{t_0+}^i)}.$$

Applying (11.5.14) we obtain:

$$\mathbb{E} \left[ \left( M_1 f(T_1^i - t^* i, D_1, R_{T_1^i - t^* i}) \right)^2 \right] = \gamma^2 \mathbb{E} \left[ (M_1)^2 \right] \int_{A_{t_0+}^i} F_2(t - t^* i, d) \frac{d\Lambda_i(t, d)}{\Lambda_i(A_{t_0+}^i)},$$

where  $F_2$  is defined by (11.3.5). So formula (11.5.16) is proved.

Now we want to calculate  $C^{i,i'}$  which is defined by (11.3.10). We start by evaluating  $g(t, d, t', d')$ , for  $t \neq t'$ , where  $g$  is defined by (11.3.11). Assume for example  $t' > t$ . Let  $0 < t_1 < \dots < t_n$ . Definition of  $R$ . shows that given  $R_{t_1} = x_1, \dots, R_{t_{n-1}} = x_{n-1}$ ,  $R_{t_n}$  has the same law as  $x_{n-1} R_{t_n - t_{n-1}}$ , i.e.:

$$\mathcal{L} \left[ R_{t_n} | R_{t_1} = x_1, \dots, R_{t_{n-1}} = x_{n-1} \right] = \mathcal{L} \left[ x_{n-1} R_{t_n - t_{n-1}} \right]. \quad (11.5.22)$$

In particular,

$$\mathcal{L} \left( R_t, R_{t'} \right) = \mathcal{L} \left( R_t, R'_{t'-t} \right), \quad (11.5.23)$$

where:

$$R'_{t'-t} := \exp \left[ \sigma (B_{t'} - B_t) + \mu (t' - t) \right]. \quad (11.5.24)$$

So for  $t' \geq t$ :

$$\begin{aligned}
g(t, d, t', d') &= \gamma^2 \mathbb{E} \left[ \left\{ \frac{\phi_t(d)}{\gamma} - R_t \right\}_+ \left\{ \frac{\phi_{t'}(d')}{\gamma} - R_{t'} \right\}_+ \right] \\
&= \gamma^2 \mathbb{E} \left[ \left\{ \frac{\phi_t(d)}{\gamma} - R_t \right\}_+ \left\{ \frac{\phi_{t'}(d')}{\gamma} - R_t R'_{t'-t} \right\}_+ \right] \\
&= \gamma^2 \mathbb{E} \left[ \left\{ \frac{\phi_t(d)}{\gamma} - R_t \right\}_+ \tilde{g}(R_t) \right], \quad (11.5.25)
\end{aligned}$$

where:

$$\tilde{g}(z) := z \mathbb{E} \left[ \left\{ \frac{\phi_{t'}(d')}{z\gamma} - R_{t'-t} \right\}_+ \right].$$

Formula (11.5.9) implies:

$$\tilde{g}(z) = z F_1 \left( t' - t, \frac{\phi_{t'}(d')}{z\gamma}, \mu, \sigma \right). \quad (11.5.26)$$

Let us come back to the calculation of  $g$ . Using (11.3.4), (11.5.25), and (11.5.26) we obtain:

$$\begin{aligned}
 g(t, d, t', d') &= \gamma^2 \mathbb{E} \left[ \left\{ \frac{\phi_t(d)}{\gamma} - R_t \right\}_+ R_t F_1 \left( t' - t, \frac{\phi_{t'}(d')}{R_t \gamma}, \mu, \sigma \right) \right] \\
 &= \gamma^2 \int_{-\infty}^{\tilde{d}(\frac{\phi_t(d)}{\gamma}, \mu, \sigma, t)} \frac{e^{\mu t + \sigma \sqrt{t} x - \frac{x^2}{2}}}{\sqrt{2\pi}} \left( \frac{\phi_t(d)}{\gamma} - e^{\mu t + \sigma \sqrt{t} x} \right) \\
 &\quad \times F_1 \left( t' - t, \frac{\phi_{t'}(d')}{\gamma e^{\mu t + \sigma \sqrt{t} x}}, \mu, \sigma \right) dx \\
 &= \int_{-\infty}^{\tilde{d}(\frac{\phi_t(d)}{\gamma}, \mu, \sigma, t)} \gamma^2 F_3(t, d, t', d', x) dx, \tag{11.5.27}
 \end{aligned}$$

where the function  $F_3$  is defined by (11.3.6). Formulas (11.3.11) and (11.5.27) imply:

$$\begin{aligned}
 C^{i, i'} &= \gamma^2 \left[ \mathbb{E}[M_1] \right]^2 \int_{A_{t_0+}^i \times A_{t_0+}^{i'}} \left\{ \int_{-\infty}^{\tilde{d}(\frac{\phi_{t-t^*i}(d)}{\gamma}, \mu, \sigma, t-t^*i)} F_3(t-t^*i, d, t'-t^*i', d', x) dx \right\} \\
 &\quad \frac{d\Lambda_i(t, d) d\Lambda_{i'}(t', d')}{\Lambda_i(A_{t_0+}^i) \Lambda_{i'}(A_{t_0+}^{i'})}.
 \end{aligned}$$

This established the formula (11.5.17) and achieve the proof of Lemma 11.5.2.  $\square$

#### 11.5.4 Proof of Theorem 11.3.6

To simplify the proof we will denote:

$$T_j^i = T_j.$$

First we order the dates of default  $(T_j)_{1 \leq j \leq n}$  as follows:

$$T_{(1)} = T_{j_1} < \dots < T_{(n)} = T_{j_n}, \tag{11.5.28}$$

where  $(j_1, \dots, j_n) \in B^n : j_k \neq j_l \forall k \neq l \in \{1, \dots, n\}$  and with agreement  $T_{(0)} = 0$ .

The term and the amount associated to the loan which defaulted at date  $T_{(j)}$  are denoted by  $D_{(j)}$  and  $M_{(j)}$  respectively. Then, for example,  $D_{(1)} = D_{j_1}$ .

With Definition 11.2.5 and Proposition 11.2.2 we can rewrite the provision for month  $i$ :

$$P_{t_0}^i = \sum_{j \geq 1} \mathbb{1}_{\{(T_{(j)}, D_{(j)}) \in A_{t_0+}^i\}} M_{(j)} \left\{ \phi_{T_{(j)}-t^*i}(D_{(j)}) - \gamma R_{T_{(j)}-t^*i} \right\}_+.$$

The identity in law (11.5.22) implies:

$$\mathcal{L} \left[ \left( R_{T_{(1)}}, \dots, R_{T_{(n)}} \right) \right] = \mathcal{L} \left[ \left( \exp \left( \sigma \sqrt{T_{(1)}} G_1 + \mu T_{(1)} \right), \dots, \right. \right.$$



$$\prod_{k=1}^n \exp \left( \sigma \sqrt{T_{(k)} - T_{(k-1)}} G_k + \mu(T_{(k)} - T_{(k-1)}) \right), \quad (11.5.29)$$

where  $(G_j)_{1 \leq j \leq n}$  is a collection of r.v.'s i.i.d. with standard normal distribution and independent of  $\left( (T_j, D_j)_{1 \leq j \leq n}, (M_j)_{1 \leq j \leq n} \right)$ .

So the law of  $P_{t_0}^i$  given  $N_i(A_{t_0+}^i) = n$  is:

$$\mathcal{L} \left[ P_{t_0}^i | N_i(A_{t_0+}^i) = n \right] = \mathcal{L} \left[ u_n \left( (T_{(j)})_{1 \leq j \leq n}, (D_{(j)})_{1 \leq j \leq n}, (M_{(j)})_{1 \leq j \leq n}, (G_j)_{1 \leq j \leq n} \right) \right],$$

where:

$$u_n \left( (t_{(j)})_{1 \leq j \leq n}, (d_{(j)})_{1 \leq j \leq n}, (m_{(j)})_{1 \leq j \leq n}, (g_j)_{1 \leq j \leq n} \right) := \sum_{i=1}^n m_{(j)} \left\{ \phi_{t_{(j)} - t_{*i}}(d_{(j)}) - \gamma \prod_{k=1}^j r_k \right\}_+,$$

$$r_k = \exp \left[ \sigma \sqrt{t_{(k)} - t_{(k-1)}} g_k + \mu(t_{(k)} - t_{(k-1)}) \right], \text{ with this agreement: } t_{(0)} = 0.$$

In other way, random subscripts  $j_1, \dots, j_n$  defined by (11.5.28) do not depend on  $T_1, \dots, T_n$ . Yet  $(M_j)_{1 \leq j \leq n}$  are independent of  $(T_j)_{1 \leq j \leq n}$ , so:

$$\mathcal{L} \left[ (M_{j_1}, \dots, M_{j_n}) \right] = \mathcal{L} \left[ (M_1, \dots, M_n) \right]. \quad (11.5.30)$$

Therefore:

$$\mathcal{L} \left[ P_{t_0}^i | N_i(A_{t_0+}^i) = n \right] = \mathcal{L} \left[ u_n \left( (T_{(j)})_{1 \leq j \leq n}, (D_{(j)})_{1 \leq j \leq n}, (M_j)_{1 \leq j \leq n}, (G_j)_{1 \leq j \leq n} \right) \right].$$

### 11.5.5 Proof of Proposition 11.4.2

Our proof starts with a technical lemma:

#### Lemma 11.5.3.

1. Consider the collection  $(Z_i, Z'_i, Z''_i)_{0 \leq i \leq i_0}$  of r.v. satisfying:

- (a) The r.v.'s  $\{Z_i, Z'_i, Z''_i, 0 \leq i \leq i_0\}$  are independent,
- (b) The marginal distributions are:

$$\mathcal{L}(Z_i) = \mathcal{P}(f_i(\theta)), \quad \mathcal{L}(Z'_i) = \mathcal{P}(f'_i(\theta)), \quad \mathcal{L}(Z''_i) = \mathcal{P}(f''_i(\theta)), \quad (11.5.31)$$

where  $f_i, f'_i,$  and  $f''_i$  are  $C^1(\mathbb{R}^3)$  functions and  $\theta = (\theta_1, \theta_2, \theta_3) \in \mathbb{R}^3$ .

Let  $(z_i, z'_i, z''_i)_{0 \leq i \leq i_0}$  be a realisation of the collection  $(Z_i, Z'_i, Z''_i)_{0 \leq i \leq i_0}$ . The three likelihood equations are:

$$\sum_{i=0}^{i_0} \left[ z_i \frac{\partial f_i(\theta)}{\partial \theta_k} + z'_i \frac{\partial f'_i(\theta)}{\partial \theta_k} + z''_i \frac{\partial f''_i(\theta)}{\partial \theta_k} \right] - \sum_{i=0}^{i_0} \frac{\partial}{\partial \theta_k} \left[ f_i(\theta) + f'_i(\theta) + f''_i(\theta) \right] = 0, \quad k = 1, 2, 3. \quad (11.5.32)$$

2. Let  $Leb_1$  (respectively  $Leb_2$ ) the Lebesgue measure on  $\mathbb{R}$  (respectively on  $\mathbb{R}^2$ ). So:

$$Leb_1(C_{t_0-}^i) = \sqrt{2}(t_0 - t^*i), \quad (11.5.33)$$

$$Leb_2(A_{t_0-}^i) = (t_0 - t^*i) \left( h - \frac{1}{2}(t_0 - t^*i) \right), \quad (11.5.34)$$

$$Leb_2(A_{t_0+}^i) = \frac{1}{2}(h - t_0 + t^*i)^2, \quad (11.5.35)$$

$$Leb_1(C_{t_0+}^i) = \sqrt{2}(h - t_0 + t^*i), \quad (11.5.36)$$

where sets  $C_{t_0-}^i$ ,  $A_{t_0-}^i$ ,  $A_{t_0+}^i$ , and  $C_{t_0+}^i$  are defined by (11.4.1), (11.2.18), (11.2.19), and (11.4.2).

The proof of this lemma will be given at the end of this section. Having disposed of this preliminary step, we can now prove the Proposition 11.4.2. We denote:

$$\begin{aligned} Z_i &= N_i(C_{t_0-}^i) \text{ and } z_i = v^i, \\ Z'_i &= N_i(A_{t_0-}^i) \text{ and } z'_i = n^i, \\ Z''_i &= N_i(A_{t_0+}^i \cup C_{t_0+}^i) \text{ and } z''_i = b^i - n^i - v^i, \\ \theta &= (\theta_1, \theta_2, \theta_3) = (\lambda_1, \lambda_2, h). \end{aligned}$$

Using (11.3.9), (11.4.3), and Lemma 11.5.3 we can deduce that (11.5.31) is true and we have:

$$\begin{aligned} f_i(\theta) &= \lambda_1 Leb_1(C_{t_0-}^i) = \lambda_1 \sqrt{2}(t_0 - t^*i), \\ f'_i(\theta) &= \lambda_2 Leb_2(A_{t_0-}^i) = \lambda_2 (t_0 - t^*i) \left( h - \frac{1}{2}(t_0 - t^*i) \right), \\ f''_i(\theta) &= \lambda_1 Leb_1(C_{t_0+}^i) + \lambda_2 Leb_2(A_{t_0+}^i) = \lambda_1 \sqrt{2}(h - t_0 + t^*i) + \frac{\lambda_2}{2}(h - t_0 + t^*i)^2, \\ f_i(\theta) + f'_i(\theta) + f''_i(\theta) &= \lambda_1 \sqrt{2}h + \frac{\lambda_2}{2}(h)^2. \end{aligned}$$

Variables  $Z_i = N_i(C_{t_0-}^i)$ ,  $Z'_i = N_i(A_{t_0-}^i)$ , and  $Z''_i = N_i(A_{t_0+}^i \cup C_{t_0+}^i)$  are independent so we can use Lemma 11.5.3. The likelihood equations, where  $z_i, z'_i$ , and  $z''_i$  are exchanged by  $v^i, n^i$ , and  $b^i - n^i - v^i$  (see Section 11.4.2), lead to the following system:

$$\begin{aligned} \sum_{i=0}^{i_0} \left[ v^i \frac{\frac{\partial \lambda_1 Leb_1(C_{t_0-}^i)}{\partial \theta_k}}{\lambda_1 Leb_1(C_{t_0-}^i)} + n^i \frac{\frac{\partial \lambda_2 Leb_2(A_{t_0-}^i)}{\partial \theta_k}}{\lambda_2 Leb_2(A_{t_0-}^i)} + (b^i - n^i - v^i) \frac{\frac{\partial (\lambda_1 Leb_1(C_{t_0+}^i) + \lambda_2 Leb_2(A_{t_0+}^i))}{\partial \theta_k}}{\lambda_1 Leb_1(C_{t_0+}^i) + \lambda_2 Leb_2(A_{t_0+}^i)} \right] \\ - \sum_{i=0}^{i_0} \frac{\partial}{\partial \theta_k} \left[ \lambda_1 Leb_1(C_{t_0-}^i) + \lambda_2 Leb_2(A_{t_0-}^i) + \lambda_1 Leb_1(C_{t_0+}^i) + \lambda_2 Leb_2(A_{t_0+}^i) \right] = 0, \end{aligned}$$

where  $k = 1, 2, 3$ . Using Lemma 11.5.3 the likelihood equations become:

$$\sum_{i=0}^{i_0} \left[ \frac{v^i}{\lambda_1} + \frac{(b^i - n^i - v^i) \sqrt{2}(h - t_0 + t^*i)}{\lambda_1 \sqrt{2}(h - t_0 + t^*i) + \frac{\lambda_2(h - t_0 + t^*i)^2}{2}} \right] = \sqrt{2}(i_0 + 1)h,$$

$$\begin{aligned}
 (S) : \sum_{i=0}^{i_0} \left[ \frac{n^i}{\lambda_2} + \frac{(b^i - n^i - v^i)(h - t_0 + t^*i)^2}{2\lambda_1\sqrt{2}(h - t_0 + t^*i) + \lambda_2(h - t_0 + t^*i)^2} \right] &= \frac{(i_0 + 1)(h)^2}{2}, \\
 \sum_{i=0}^{i_0} \left[ \frac{n^i(t_0 - t^*i)}{(t_0 - t^*i)\left(h - \frac{1}{2}(t_0 - t^*i)\right)} + \frac{(b^i - n^i - v^i)\left(\sqrt{2}\lambda_1 + \lambda_2(h - t_0 + t^*i)\right)}{\lambda_1\sqrt{2}(h - t_0 + t^*i) + \frac{\lambda_2(h - t_0 + t^*i)^2}{2}} \right] \\
 &= (\lambda_1\sqrt{2} + \lambda_2h)(i_0 + 1).
 \end{aligned}$$

We denote  $(\hat{\lambda}_1, \hat{\lambda}_2, \hat{h})$  the solution of this system which corresponds to the likelihood maximum estimators of parameters  $\lambda_1, \lambda_2$ , and  $h$ . This achieves the proof of Proposition 11.4.2.

*Proof of Lemma 11.5.3.*

1. The likelihood of  $(Z_i, Z'_i, Z''_i)_{0 \leq i \leq i_0}$  is denoted by  $L_{i_0}$ . Using the independence of variables  $\{Z_i, Z'_i, Z''_i, 0 \leq i \leq i_0\}$  we obtain:

$$L_{i_0} = \frac{\prod_{i=0}^{i_0} f_i^{z_i}(\theta) [f'_i(\theta)]^{z'_i} [f''_i(\theta)]^{z''_i}}{\prod_{i=0}^{i_0} z_i! z'_i! z''_i!} \exp \left[ - \sum_{i=0}^{i_0} (f_i(\theta) + f'_i(\theta) + f''_i(\theta)) \right],$$

where  $z_i, z'_i$ , and  $z''_i$  are contained in  $\mathbb{N}$ . So:

$$\begin{aligned}
 \ln(L_{i_0}) &= - \ln \left( \prod_{i=0}^{i_0} z_i! z'_i! z''_i! \right) + \sum_{i=0}^{i_0} \left[ z_i \ln [f_i(\theta)] + z'_i \ln [f'_i(\theta)] + z''_i \ln [f''_i(\theta)] \right] \\
 &\quad - \sum_{i=0}^{i_0} \left[ f_i(\theta) + f'_i(\theta) + f''_i(\theta) \right]. \tag{11.5.37}
 \end{aligned}$$

The likelihood equations are:

$$\frac{\partial}{\partial \theta_k} (\ln(L_{i_0})) = 0, \quad k = 1, 2, 3. \tag{11.5.38}$$

Combining (11.5.37) and (11.5.38) the likelihood equations lead to this system:

$$\sum_{i=0}^{i_0} \left[ z_i \frac{\partial f_i(\theta)}{\partial \theta_k} + z'_i \frac{\partial f'_i(\theta)}{\partial \theta_k} + z''_i \frac{\partial f''_i(\theta)}{\partial \theta_k} \right] - \sum_{i=0}^{i_0} \frac{\partial}{\partial \theta_k} \left[ f_i(\theta) + f'_i(\theta) + f''_i(\theta) \right] = 0, \quad k = 1, 2, 3.$$

2. The proof is straightforward with definitions of  $C_{t_0-}^i, A_{t_0-}^i, A_{t_0+}^i$ , and  $C_{t_0+}^i$ . □

### 11.5.6 Variance of the total provision

Now we want to calculate the total provision variance. We have:

$$\mathbb{V}[P_{t_0}] = \sum_{i=0}^{\lfloor \frac{t_0}{t^*} \rfloor + 1} \mathbb{V}[P_{t_0}^i] + 2 \sum_{0 \leq i < i' \leq \lfloor \frac{t_0}{t^*} \rfloor} \left\{ \mathbb{E}[P_{t_0}^i P_{t_0}^{i'}] - \mathbb{E}[P_{t_0}^i] \mathbb{E}[P_{t_0}^{i'}] \right\}. \quad (11.5.39)$$

We calculate  $\mathbb{E}[P_{t_0}^i P_{t_0}^{i'}]$ .

$$\begin{aligned} \mathbb{E}[P_{t_0}^i P_{t_0}^{i'}] &= \mathbb{E} \left[ \left\{ \sum_{j \geq 1} \mathbb{1}_{\{(T_j^i, D_j) \in A_{t_0+}^i\}} M_j f(T_j^i - t^*i, D_j, R_{T_j^i - t^*i}) \right\} \right. \\ &\quad \times \left. \left\{ \sum_{l \geq 1} \mathbb{1}_{\{(T_l^{i'}, D_l) \in A_{t_0+}^{i'}\}} M_l f(T_l^{i'} - t^*i', D_l, R_{T_l^{i'} - t^*i'}) \right\} \right] \\ &= \sum_{k \geq 0, h \geq 0} \left\{ \mathbb{P}[N_i(A_{t_0+}^i) = k] \mathbb{P}[N_{j'}(A_{t_0+}^{i'}) = h] \right\} \\ &\quad \times \sum_{1 \leq j \leq k, 1 \leq l \leq h} \mathbb{E} \left[ M_j f(T_j^i - t^*i, D_j, R_{T_j^i - t^*i}) M_l f(T_l^{i'} - t^*i', D_l, R_{T_l^{i'} - t^*i'}) \right] \\ &= \sum_{k \geq 0, h \geq 0} \left\{ \mathbb{P}[N_i(A_{t_0+}^i) = k] \mathbb{P}[N_{j'}(A_{t_0+}^{i'}) = h] \right\} C^{i,i'} k h \\ &= C^{i,i'} \mathbb{E}[N_i(A_{t_0+}^i)] \mathbb{E}[N_{j'}(A_{t_0+}^{i'})]. \end{aligned} \quad (11.5.40)$$

Combining (11.5.39), (11.5.40), (11.5.17), and (11.3.2) we have:

$$\begin{aligned} \mathbb{V}[P_{t_0}] &= \sum_{i=0}^{\lfloor \frac{t_0}{t^*} \rfloor + 1} \mathbb{V}[P_{t_0}^i] + 2 \sum_{0 \leq i < i' \leq \lfloor \frac{t_0}{t^*} \rfloor} \gamma^2 [\mathbb{E}[M_1]]^2 \left\{ \int_{A_{t_0+}^i \times A_{t_0+}^{i'}} \left\{ \int_{-\infty}^{\tilde{d}(\frac{\phi_{t-t^*i}(d)}{\gamma}, \mu, \sigma, t-t^*i)} \right. \right. \\ &\quad F_3 \left( t - t^*i, d, t' - t^*i', d', x \right) dx \Big\} d\Lambda_i(t, d) d\Lambda_{i'}(t', d') \\ &\quad - \int_{A_{t_0+}^i} F_1 \left( t - t^*i, \frac{\phi_{t-t^*i}(d)}{\gamma}, \mu, \sigma \right) d\Lambda_i(t, d) \\ &\quad \times \left. \int_{A_{t_0+}^{i'}} F_1 \left( t - t^*i', \frac{\phi_{t-t^*i'}(d)}{\gamma}, \mu, \sigma \right) d\Lambda_{i'}(t, d) \right\}, \end{aligned}$$

where  $\mathbb{V}[P_{t_0}^i]$  is obtained by (11.3.7).





# Additional information concerning the provisioning model

---

In this chapter we will provide some additional information on results obtained in Chapter 11. We will start by explaining the link between the guarantee offered by CAMCA Assurance and the “credit default swap”. Then we will interpret the provision as a random sum of exchange option payoff with random maturity; this allows us to speak about the premium investment. Then we will try to justify, as much as possible, the assumptions linked to our provisioning model. We will also propose a numerical application in which we will give some descriptive statistics of the loans involved, a quantitative analysis of the results obtained in the previous chapter (theoretical mean, mean on simulations, etc.) and an analysis of the sensitivity within certain parameters. We will conclude this chapter by presenting the numerical method used to solve the non-linear system presented in Theorem 11.4.2.

## 12.1 Guarantee caution with the “credit default swap”

In this section we will show the link between a “credit default swap” (CDS) and the guarantee offered by CAMCA Assurance.

A CDS is a financial protection contract in which:

1. The buyer of the protection is an entity (person, company, etc.) denoted by A,
2. The seller is denoted by B,
3. And the entity sensitive to the risk is denoted by C.

Buyer A pays a premium (in general quarterly) to seller B to get the protection against an situation affecting entity C. In the event of this situation occurring (the event is defined on the contract), then B indemnifies A with the amount defined in the contract.

Situations affecting C could be: a default payment, bankruptcy, a reduction in credit rating, etc..

The aim of the CDS is to transfer the risk covered by A to B if A has a direct link to C. However A does not necessarily have any link with C: in this case the CDS is a market speculation tool. This is an illustration:

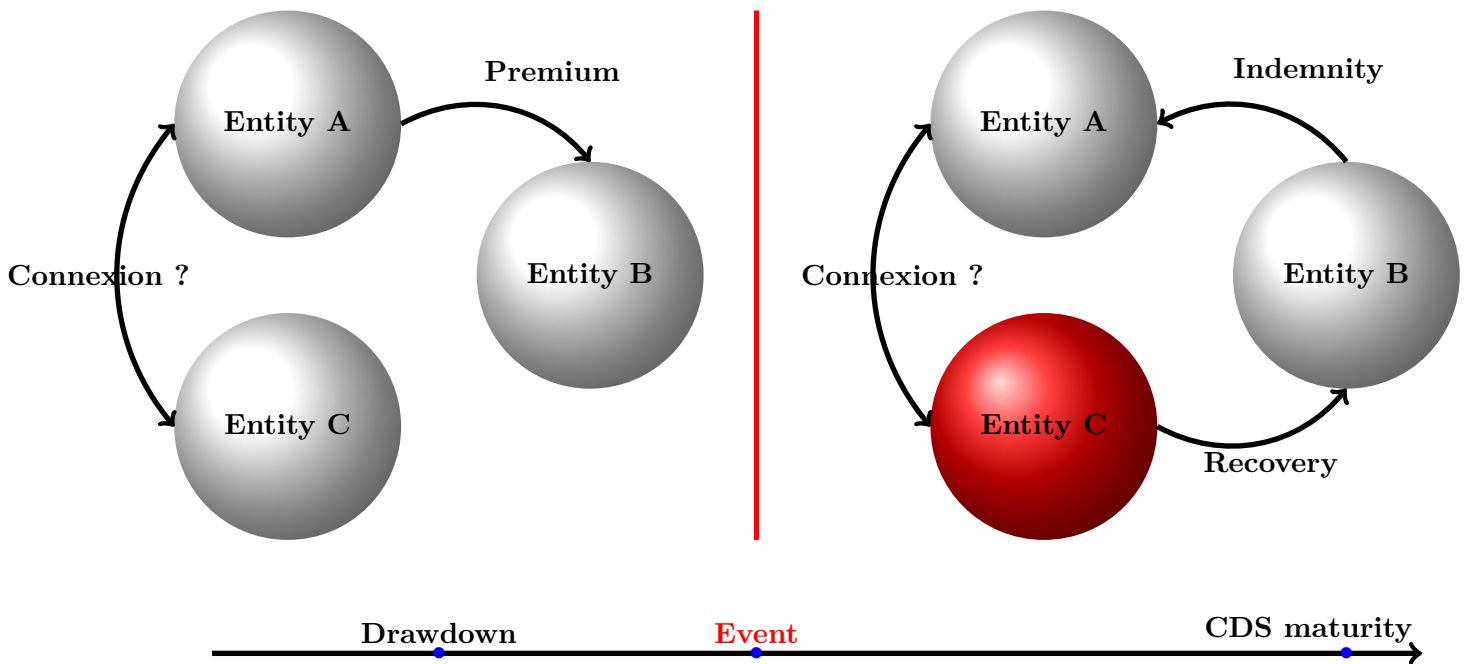


Illustration 12.1: The “credit defaults swap”.

The illustration above is similar to the Illustration 8.1 presented for the description of the guarantee offered by CAMCA Assurance. We will compare this guarantee with the CDS and explain the differences. In the guarantee offered by the insurance company:

- Entity A is the bank,
- Entity B is the insurance company,
- Entity C is the borrower.

The main differences to a CDS are as follows:

1. Entity C pays the CDS premium.
2. In a typical CDS, the subscriber and the beneficiary of the contract are the same, i.e. entity A. In the guarantee offered by CAMCA Assurance, the subscriber is entity C (the borrower) and the beneficiary is entity A (the bank).
3. The premium is paid at the beginning of the contract in one payment.
4. The trigger for the CDS is the submission of a claim by entity A: the submission of claim is decided by the entity A.

Now we will look at how CDS premiums are calculated. We will consider a CDS with nominal  $N$ , maturity  $D$  and let  $t_0 = 0 < t_1 < \dots < t_n = D$  the payment dates of premium  $s$ . Let  $T$  be the random time of default and  $R$  the random rate of recovery.



This is illustrated as follows:



Illustration 12.2: Illustration of payment in a “credit default swap”.

So to calculate the CDS premium we need to:

1. Calculate the mean of the recovery rate:  $\mathbb{E}[R]$ ,
2. Evaluate the c.d.f. of the default date,
3. Estimate the risk free interest rate.

This kind of model is different to ours because the claim number is not random: a provision can be calculated on each contract.

## 12.2 Provision with exchange option

Let us interpret the claim amount, submitted at random time  $T_j^i$  and concerning a borrower with situation  $(D_j, M_j)$ , as an exchange option payoff. More precisely we want to interpret (11.2.10) as an exchange option payoff with random maturity  $T_j^i$  (see [31] and [36]), i.e.:

$$(S_{T-I_j}^{1,j} - S_{T-I_j}^{2,j})_+, \quad I_j \leq T \leq I_j + D_j. \quad (12.2.1)$$

Therefore from (11.2.24) the provision can be viewed as a random sum of exchange option payoffs with random maturity.

As a reminder we will now look at the mechanism of the exchange option. A european exchange option is an option which allows the buyer (of this option) to exchange an asset with a price of  $S_T^2$  for an asset with a price of  $S_T^1$  at a time  $T$  called maturity. The price of this option is called the premium. The payoff of this option is denoted by  $f_T$  and defined by:

$$f_T = (S_T^1 - S_T^2)_+.$$

When the option matures, the buyer has two choices:

1. If  $S_T^1 > S_T^2$  then the buyer buys at time  $T$  an asset with a price of  $S_T^2$  and exchanges this for an asset with a price of  $S_T^1$  that he or she resells immediately. So the profit for the buyer is:

$$S_T^1 - S_T^2 = (S_T^1 - S_T^2)_+. \quad (12.2.2)$$

2. If  $S_T^1 \leq S_T^2$  then the buyer does nothing. The profit for the buyer is:

$$0 = (S_T^1 - S_T^2)_+. \quad (12.2.3)$$

Now we would like to explain the similarity between the policy offered by the insurance company and an exchange option.

**Example 12.2.1.** Policy offered by the insurance company.

We will assume that borrower wishes to borrow a certain amount from the bank in month  $i = 0$ . To protect against the risk default the borrower will take out an insurance policy, which will effectively replace borrower should this situation arise. We consider an exchange option with a random maturity  $T$  which correspond to the date of borrower default and we consider the following assets:

- Asset with price  $(S_t^{1,j})_{t \geq 0}$  which represent the amount own by the borrower  $j$  at date  $t$ ,
- Asset with price  $(S_t^{2,j})_{t \geq 0}$  which represent the property sale amount cash at date  $t$ .

The insurance company is seen as the exchange option seller on the asset with a price of  $S_t^{2,j}$ , representing the property sale amount, for an asset with a price of  $S_t^{1,j}$ , representing the outstanding amount owed by the borrower. The option buyer is the borrower. The option random maturity is the time of default  $T_j^0$ . So at time 0:

- The borrower pays a premium to hold the exchange option. Thank to borrowed amount  $M_j$  and to the borrower deposit, the borrower has a property with price of  $(1 + a)M_j$ .
- The insurance company receives the option premium.

When the option matures, which corresponds to the time of default, the borrower has two choices:

1. If the recoveries amount is higher than or equal to the outstanding amount owed by the borrower, then the borrower does not exercise its exchange option. The property is sold at price  $S_{T_j^0}^{2,j}$  and the borrower covers the debt and may even show a profit. The claim amount, i.e. the option payoff, is zero for the insurance company.
2. If the recoveries amount is lower than the outstanding amount owed by the borrower, then the borrower exercises its exchange option. It exchanges its asset  $S_{T_j^0}^{2,j}$  for asset  $S_{T_j^0}^{1,j}$  and the claim amount for the insurance company is  $S_{T_j^0}^{1,j} - S_{T_j^0}^{2,j}$ . So, the insurance company buys, at time  $T_j^0$ , an asset with a price of  $S_{T_j^0}^{1,j}$  and exchanges with the borrower for an asset with a price of  $S_{T_j^0}^{2,j}$  that it resells immediately.

Remember that the option payoff does not represent the “profit” for the borrower in the event of default, but only the amount that it receives from the insurance company.

The previous example shows that it is possible to use an exchange option payoff between two assets with maturity  $T_j^i$  to model the claim amount submitted at time  $T_j^i$ . So provision for the loan drawn down during month  $i$  can be viewed as a random sum of exchange option payoffs of an asset with a price of  $S_t^{2,j}$  for an asset with a price of  $S_t^{1,j}$

and with random maturity equal to the time of default  $T_j^i$ . The number of this option is Poissonian. For more details about the default risk modelling with exchange option we refer the reader to [22], [42], and [55].

This interpretation of the exchange option leads us to the question of the how the premiums received by the insurance company are invested. For borrower  $j \in B$  the premium amount received by the company is denoted by  $p_j$ . We had shown that the guarantee offered by the company can be interpreted as an exchange option between two assets and the premium amount of this exchange option is  $p_j$ . The maturity  $T_j^i$  of the option is however random. In addition the number of contracts involved, and also the number of exchange options, is also random and is equal to:

$$\sum_{i=0}^{i_0} N_i (C_{t_0+}^i \cup C_{t_0-}^i \cup A_{t_0+}^i \cup A_{t_0-}^i), \quad (12.2.4)$$

where  $N_i$ ,  $C_{t_0-}^i$ ,  $A_{t_0-}^i$ ,  $A_{t_0+}^i$  et  $C_{t_0+}^i$  are defined by (11.3.9), (11.4.1), (11.2.18), (11.2.19), and (11.4.2).

So to determine a hedging strategy for premium received we have to define a market which depends on exogenous assets (other than  $S_t^{1,j}$  and  $S_t^{2,j}$ ) and also takes into account two sources of uncertainty:

1. Uncertainty over exchange option maturity dates,
2. Uncertainty over the number of exchange options.

## 12.3 Justification of the assumptions

In this section we will try to justify all the assumptions realized in Chapter 11.

### 12.3.1 Assumption 1

Recall the first assumption of our model:

(A<sub>1</sub>) : If borrower  $j$  draws down his loan at date  $I_j$  then this corresponds to month  $i$ , we will assume that:  $I_j = t^*i$ .

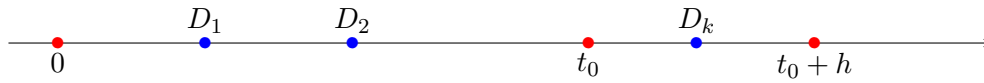
We assume that loans are drawn down at the beginning of each month although this assumption has no real justification. The loans are drawn down in a continually way but one month is a short period of time in comparison to the loan duration. In addition CAMCA Assurance only knows the month and year of drawdown not the actual date.

### 12.3.2 Assumption 2

RVs representing the loan duration are upper bounded by an unknown quantity  $h > 0$  so that:

$$(A_2) : \begin{cases} D_j \leq h, & j \in B, & (12.3.1) \\ t_0 < I_j + h, & j \in B. & (12.3.2) \end{cases}$$

This is an illustration:



As mentioned in 8.2.2, CAMCA Assurance has drawn up a profile covering loans and borrowers and if the loan and the borrower match this profile then cover is automatic. In this profile the insurance company has included a maximum loan duration concerning loan so it is logical to assume that RVs representing this are capped by a deterministic quantity.

However if the borrower does not match the profile then the bank must seek approval from the insurance company. The loan duration can in this case be higher than the maximum duration included in the profile. In addition loans can renegotiated during their life and so the maturity can change: we do not take this situation into account because the number of loans renegotiated is insignificant (0.43% of all loans). This is why we assume that RVs representing loan duration are upper bounded by an unknown deterministic quantity.

### 12.3.3 Assumption 3

Recall the third assumption:

$$(A_3) : \begin{cases} \text{Monthly repayments are fixed,} & (12.3.3) \\ \tau \text{ is the exponential and deterministic rate,} & (12.3.4) \\ \text{Early partial payments are prohibited.} & (12.3.5) \end{cases}$$

Assumption (12.3.3) is confirmed in practice.

Assumption (12.3.4) is not confirmed in practice: there is no common rate for every borrower. The impact on the provision will be discussed in 12.4.3.

Assumption (12.3.5) is not confirmed in practice. However if we assume (12.3.5) is true the provision is higher than if we assume (12.3.5) is not true, so this leads to the insurance company being more cautious.

### 12.3.4 Assumption 4

Recall the fourth assumption of our model:

$$(A_4) : \begin{cases} \text{Default is submitted immediately as a claim,} & (12.3.6) \\ \text{Default concerns the entire monthly installment.} & (12.3.7) \end{cases}$$

Assumption (12.3.7) is confirmed in practice.

However assumption (12.3.6) is not confirmed in practice: the delay between the first default and the declaration of the claim is around 24 months. Indeed bank can submit the claim when it want. Look at a typical example (we denoted ‘P’ a payment, ‘D’ for default payment and ‘DS’ for claim submission):

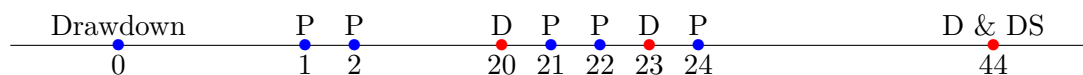


Illustration 12.3: Submission of a claim.

In the example above, the bank waited until the borrower had defaulted on three payments before submitting a claim. This is a typical example of how banks deal with default situation.

Recall that the outstanding amount owed by the borrower is the sum of unpaid capital and penalty interest. In the example above, penalty interest is higher than in our model: interest for monthly installments 20, 23 and 44 is unpaid while in our model the claim would be submitted directly after the first default payment and only the interest for monthly installment 20 would be unpaid. So, in our model, we underestimate the claim amount concerning penalty interest. However unpaid interest is usually insignificant in comparison to unpaid capital. This is an illustration:

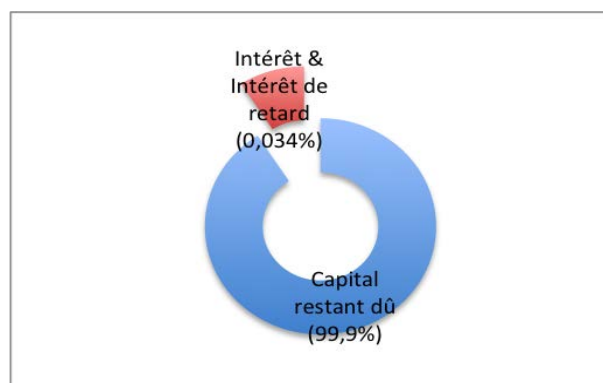


Illustration 12.4: Unpaid capital Vs Unpaid interest.

In our model the claim is submitted at date 20 so monthly payments 21, 22, ..., 43, and

44 are unpaid. In the example above the borrower has made monthly payments 21, 22 and 24 after the first default at date 20. This means that we overestimate the claim amount concerning unpaid capital.

It is also possible that between the first default payment and the submission of the claim, the borrower repays part of the unpaid capital and so the unpaid capital would decrease. Using our previous example we assume the borrower repays one of the missing installment and this is denoted by ‘ $R_k$ ’:

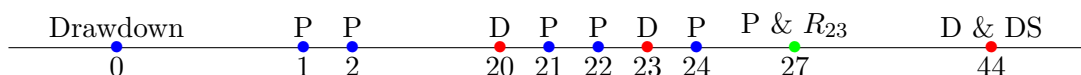


Illustration 12.5: Reimbursement of a missing monthly installment.

At date 27 the borrower pays the normal monthly installment plus the missing monthly installment 23. In this case, our model overestimate the claim amount concerning capital and unpaid interest.

### 12.3.5 Assumption 5

The fifth assumption is:

$$(A_5) : \begin{cases} \text{Borrower recoveries are zero,} \\ \text{Property recoveries received at date } T \text{ are:} \\ \gamma M_j R_{T-I_j}, \quad j \in B, \quad I_j \leq T \leq I_j + D_j, \end{cases} \quad (12.3.8)$$

Recoveries from borrowers are not so frequent. A lack of funds is the most common cause of default and recoveries usually result from the sale of the property. This is why we assume that borrower recoveries are nil.

We assume that any property sale is made immediately at the time of default. However the bank only repossesses the property in 70% of all defaults cases: this is a disadvantage of our model.

On the drawdown date, we assume that the property’s value is equal to the amount borrowed plus the borrower’s deposit. We can reasonably assume that the borrower’s deposit is proportionate to the borrowed amount since this is usually a condition of the loan. We assume that the proportionality coefficient is fixed and that it is the same for all borrowers. Then we denote:

$$a \in \mathbb{R}_+ \text{ the coefficient representing the borrower’s deposit.} \quad (12.3.9)$$

We have not taken into account the fact that the deposit fluctuates depending on the borrowers or that the coefficient  $a$  is random.

We can further assume that the property's value at time  $t$  depends on market price; improvements or damage to the house are not taken into account. We assume in particular that the property's value at time  $t$  is equal to the property's value at the drawdown date multiplied by a r.v.  $R_{T,I_j}$  (' $R$ ' for "Real estate"). This represents the market price variation between times  $I_j$  and  $T$ . This variable is assumed to be identical for all borrowers. Again to simplify, we assume that coefficient  $R_{T,I_j}$  depends only on the time interval between drawdown date  $I_j$  and default date  $T$ , that is to say:

$$R_{T,I_j} = R_{T-I_j},$$

with  $R_0 = 1$ .

Property sale in the event of default incurs costs, including notary fees, legal procedure fees, etc.. These costs are assumed to be proportional to the property's value at default time. The proportionality coefficient is assumed to be identical for all borrowers. Then we denote:

$$c \in [0; 1] \text{ the coefficient representing selling costs of property.} \quad (12.3.10)$$

So the amount realized from the potential sale of the property in the event of default at date  $T$  is:

$$(1 + a)M_j R_{T-I_j} - c(1 + a)M_j R_{T-I_j} = \gamma M_j R_{T-I_j},$$

where:

$$\gamma = (1 - c)(1 + a). \quad (12.3.11)$$

### 12.3.6 Assumption 6

We assume that:

( $A_6$ ) : The cloud of points  $(T_j^i, D_j)_{j \in B}$  is a PPP with intensity  $\Lambda_i$  :

$$\forall i : \Lambda_i := \lambda_1 \text{Leb}_{|C_{t_0-}^i \cup C_{t_0+}^i} + \lambda_2 \text{Leb}_{|A_{t_0-}^i \cup A_{t_0+}^i}, \quad (12.3.12)$$

where:

- $\lambda_1$  et  $\lambda_2$  are two positives parameters,
- $\text{Leb}_{|C_{t_0-}^i \cup C_{t_0+}^i}$  (resp.  $\text{Leb}_{|A_{t_0-}^i \cup A_{t_0+}^i}$ ) is the Lebesgue measure restricted on  $C_{t_0-}^i \cup C_{t_0+}^i$  (resp.  $A_{t_0-}^i \cup A_{t_0+}^i$ ) defined by (11.4.2) and (11.4.1) (resp. (11.2.19) and (11.2.18)).

It is classical assumption without justification.

### 12.3.7 Assumption 7

Recall this assumption:

$$(A_7) : \text{The PPP } \left( (T_j^i, D_j)_{j \in B} \right)_{i \geq 0} \text{ are independent.} \quad (12.3.13)$$

A-priori one borrower has no link between another borrower. It is why we make this assumption.

### 12.3.8 Assumption 8

This assumption is as follows:

$$(A_8) : \begin{cases} (T_j^i, D_j)_{j \in B}, (M_j)_{j \in B} \text{ and } (R_t)_{t \geq 0} \text{ are independents,} & (12.3.14) \\ \text{The RVs } (M_j)_{j \in B} \text{ are i.i.d..} & (12.3.15) \end{cases}$$

$R_t$  represents the fluctuation of the property value due to a fluctuation in real estate price. We could assume, for example, that a significant fluctuation in real estate price leads to an increase in claims: a real estate market crash can imply a more general crisis (systemic crisis). We will assume such a link, impossible to measure, does not exist, so the PPP is independent of real estate price fluctuation, i.e.:

$$(T_j^i, D_j)_{j \in B} \text{ and } (R_t)_{t \geq 0} \text{ are independent.} \quad (12.3.16)$$

We could also think that real estate price fluctuates in relation to the property value. For example:

- Big fluctuation for cheap houses,
- Small fluctuation for expensive houses.

This fluctuation difference is significant for a large range of property values. However for CAMCA Assurance the loan amounts are relatively concentrated (see Section 12.4.1), this is why we assume that:

$$(M_j)_{j \in B} \text{ and } (R_t)_{t \geq 0} \text{ are independent.} \quad (12.3.17)$$

In addition the loan amount does not impact on a claim: a claim is due to a change in the personal situation of the borrower not in the characteristics of the loan, this is why we assume:

$$(T_j^i, D_j)_{j \in B} \text{ and } (M_j)_{j \in B} \text{ are independent.} \quad (12.3.18)$$

Assumption (12.3.15) is confirmed because loans are relatively homogeneous. This is due to the profile introduced by CAMCA Assurance. In addition the independence assumption between each claim amount is justified because a link between two borrowers is extremely rare.

### 12.3.9 Additional information to Remark 11.2.3

In this section we assume  $i = 0$ .

In Remark 11.2.3 we explained that defaults which occurred at the end of the loan are not taken into account in our model. If a default payment occurs at this date, then the default amount will be small.



We have seen that function  $f$  (see (11.2.16)) depends on the time unit  $t^*$ . It follows that the limit of a claim amount submitted at random date  $T = D$  concerning a loan with duration  $D$  and amount  $M$  is nil when the time unit  $t^*$  becomes small:

$$\begin{aligned}\lim_{t^* \rightarrow 0} Mf(D, D, R_D) &= \lim_{t^* \rightarrow 0} M \left[ \alpha \frac{e^{\tau D} - e^{\tau(D-t^*)}}{e^{\tau D} - 1} - \gamma R_D \right]_+ \\ &= 0,\end{aligned}\tag{12.3.19}$$

where function  $f$  is defined by (11.2.16).

Let us move on to numerical application of the theoretical results obtained in Chapter 11.

## 12.4 Numerical application

In Section 12.4.1 we will provide some descriptive statistics on the loans considered and in Section 12.4.2 we will compare the simulation results using Algorithm 11.1 of those obtained with analytical formulae presented in Theorems 11.3.2 and 11.3.4. We will also study the fluctuations of provisions based on the date of analysis and under a run-off assumption. In Section 12.4.3 we analyze the sensitivity of the provision mean when some parameters fluctuate and in Section 12.4.4 we will study the probability of a claim amount being nil. Finally, we will explain in Section 12.4.5, the numerical method used to solve the system of equations presented in Theorem 11.4.2.

In order to make the presentation easier, we will use the year as our time unit to measure groupings and not the month as in the previous chapter. Throughout this section data used is that collected by the insurance company from January 2000 to December 2013 for a particular bank. We will analyze 45,126 contracts distributed by grouping as follows:

| Grouping | 00  | 01  | 02  | 03  | 04  | 05    | 06    |
|----------|-----|-----|-----|-----|-----|-------|-------|
| Number   | 361 | 409 | 429 | 330 | 394 | 1,734 | 3,041 |
| In %     | 0.8 | 0.9 | 0.9 | 0.7 | 0.9 | 3.8   | 6.7   |

| Grouping | 07    | 08    | 09    | 10    | 11    | 12    | 13    |
|----------|-------|-------|-------|-------|-------|-------|-------|
| Number   | 6,722 | 6,048 | 3,545 | 5,763 | 6,356 | 4,581 | 5,413 |
| In %     | 14.9  | 13.4  | 7.8   | 12.8  | 14.1  | 10.1  | 12    |

Illustration 12.6: Number of loans per grouping.

00 means grouping 2000, 01 means grouping 2001, etc..

### 12.4.1 Descriptive statistics

Here we present some statistics describing the following data:

- Amount and duration of loans,
- Outstanding amount owed and interest,
- Number of claims,
- Loans matured.

#### Amount and duration of loans

|                      | Mean   | Standard deviation | Minimum | Maximum | Median |
|----------------------|--------|--------------------|---------|---------|--------|
| Amount of loan (EUR) | 65,378 | 60,093             | 150     | 700,000 | 50,000 |
| Duration (month)     | 201    | 69.9               | 12      | 360     | 180    |

Illustration 12.7: Descriptive statistics on contracts involved.

## 12.4 Numerical application

---

Considering all grouping in 12/31/2013, the average amount of loans involved is EUR 65,378 for an average duration of 201 months. The extent of the amounts and durations is quite large (EUR 699,850 for the amounts and 348 months for the duration). This is due to:

- atypical applications for loans that do not match the standard profile concerning the property financed, see Section 8,
- refinancing operation for certain loans (low amount and duration),
- and finally the bridge loans.

However, standard deviations (EUR 60 093 for the amounts and 70 months for durations) show that the non respect of the standard profile is more commonly associated with an amount higher the maximum duration than an atypical duration.

Now observe the distribution of the loans:

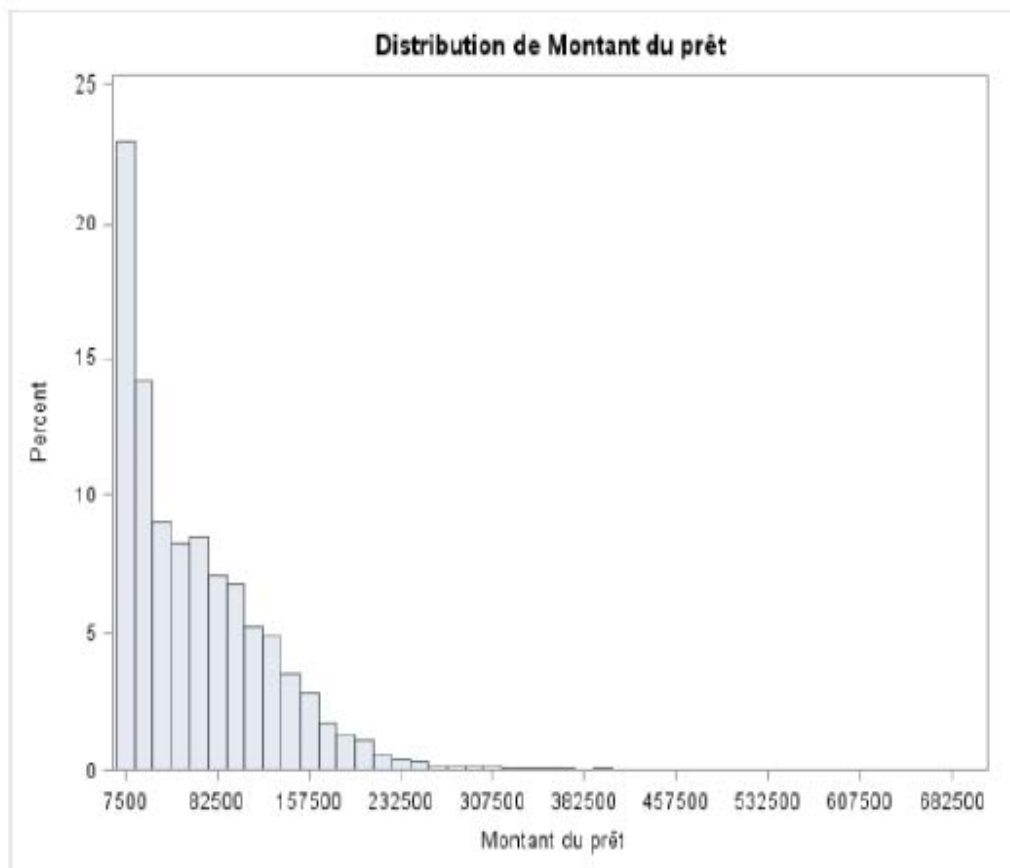


Illustration 12.8: Histogram of the loan amount.

The previous histogram is determined by class of loan amounts. We chose a division

into  $k$  classes, where  $k$  is (Sturges's method):

$$k := 1 + 3,3 \times \log(n), \tag{12.4.1}$$

with  $n$  is the number of data considered; about 36 classes in our case. We can notice that the distribution is unimodal.

We observe now the histogram the loan durations:

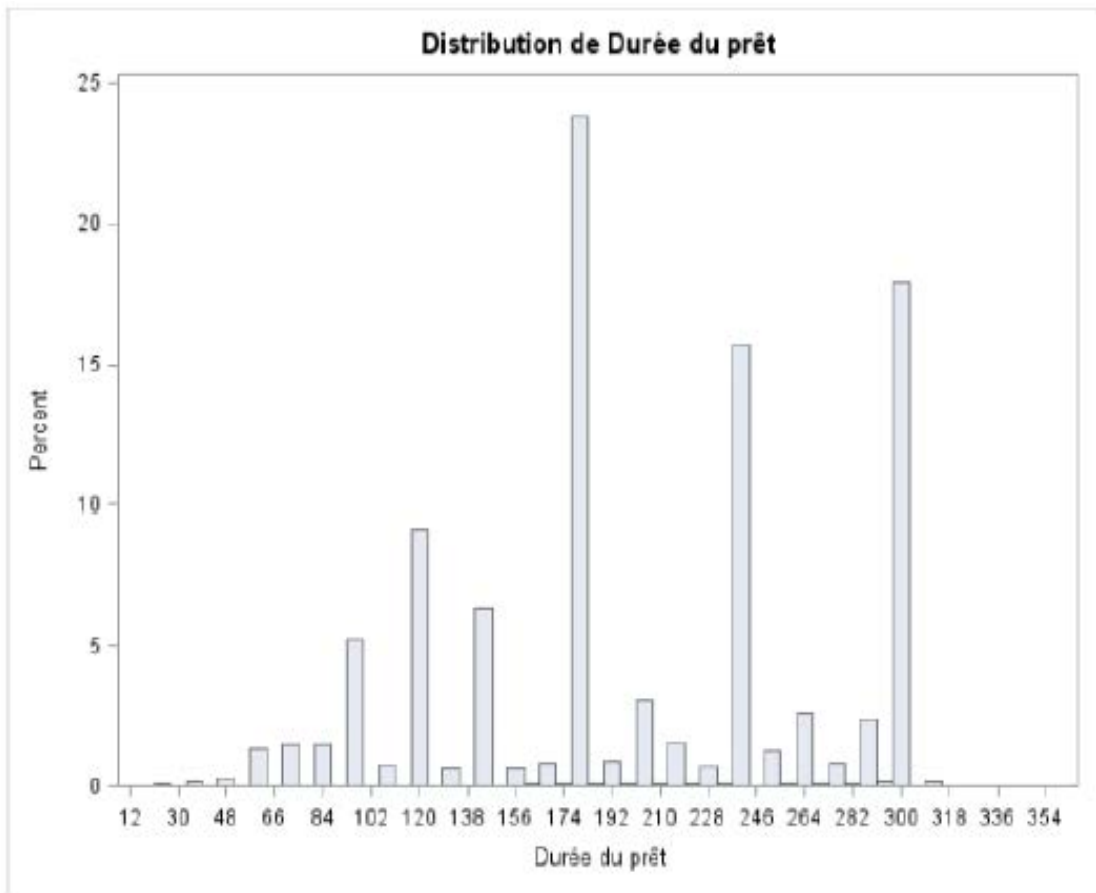
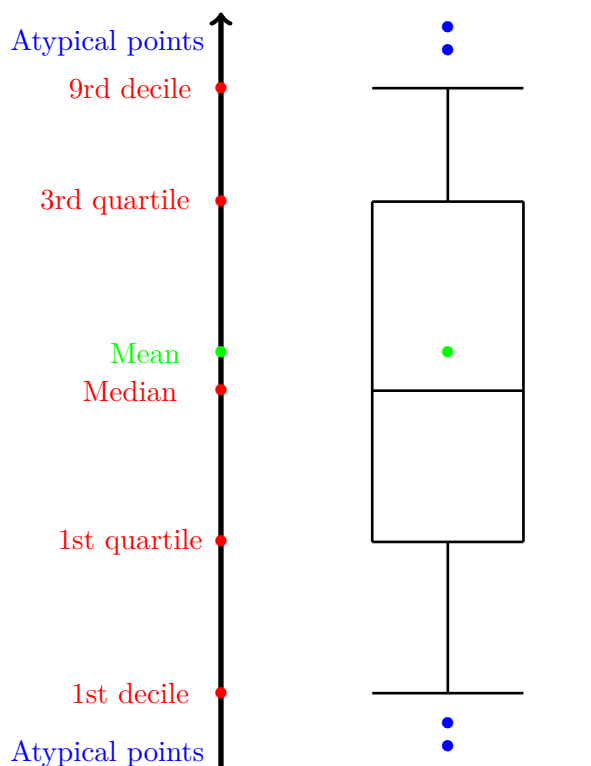


Illustration 12.9: Histogram of the duration of the loans.

We find that the empirical distribution of loan duration is multi-modal and that the durations are fairly scattered around the median (180 months).

We will now analyze box-plots of which we recall the interpretation:



Now we can observe the box-plots of duration and amounts of loans:

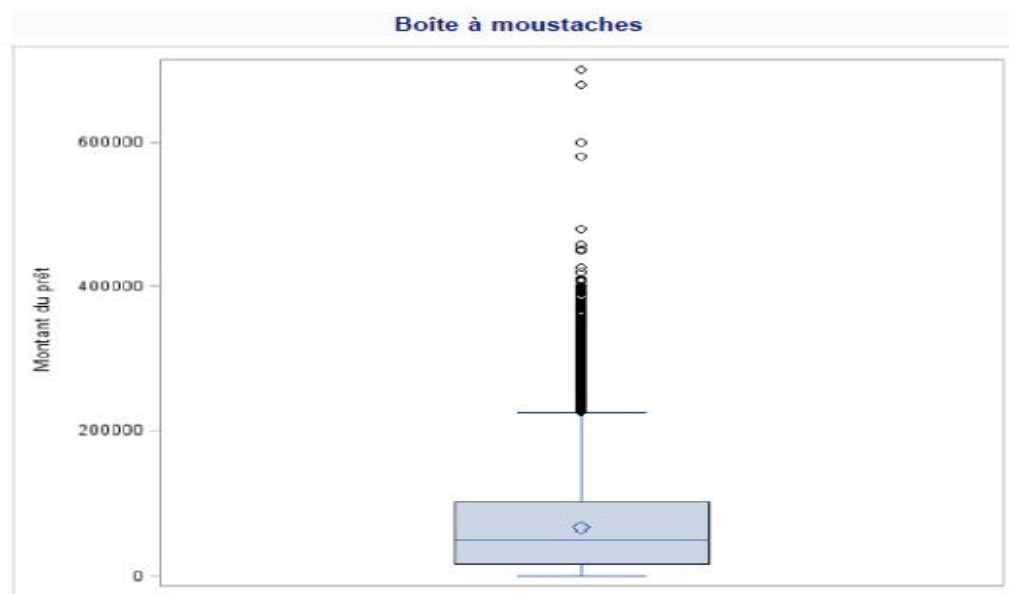


Illustration 12.10: Boxplot of loan amount.

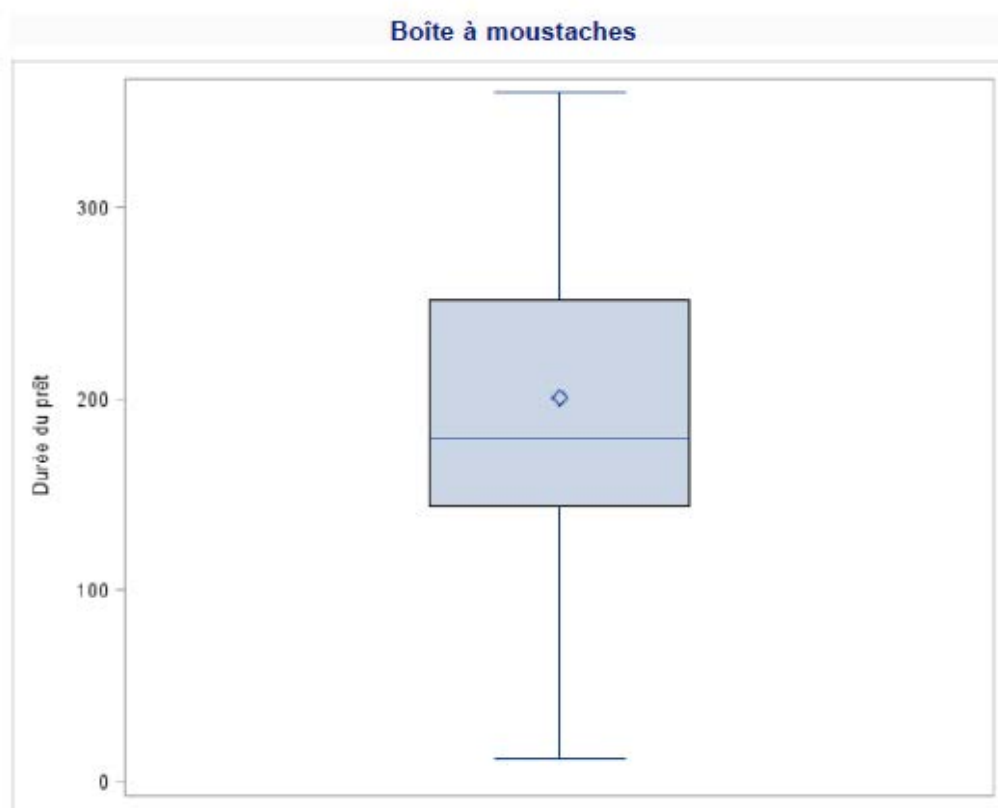


Illustration 12.11: Boxplot of loan duration.

Let's analyze now the data for each grouping:

| Grouping                     | 00  | 01  | 02  | 03 | 04  | 05  | 06  |
|------------------------------|-----|-----|-----|----|-----|-----|-----|
| Amt grouping / Amt total (%) | 0.1 | 0.1 | 0.1 | 0  | 0.5 | 3.6 | 6.7 |

| Grouping                     | 07   | 08 | 09  | 10   | 11   | 12   | 13   |
|------------------------------|------|----|-----|------|------|------|------|
| Amt grouping / Amt total (%) | 16.1 | 12 | 6.2 | 14.2 | 14.9 | 11.2 | 14.3 |

Illustration 12.12: Weight (in terms of loan amount) by grouping.

where “Amt grouping” means “Loan amounts by grouping” and “Amt total” means “Loan amount for all grouping”. This table indicates that the grouping from 2006 to 2013 alone represent nearly 95% of the total amount of guaranteed loans.

Let us conclude with box-plots of grouping:

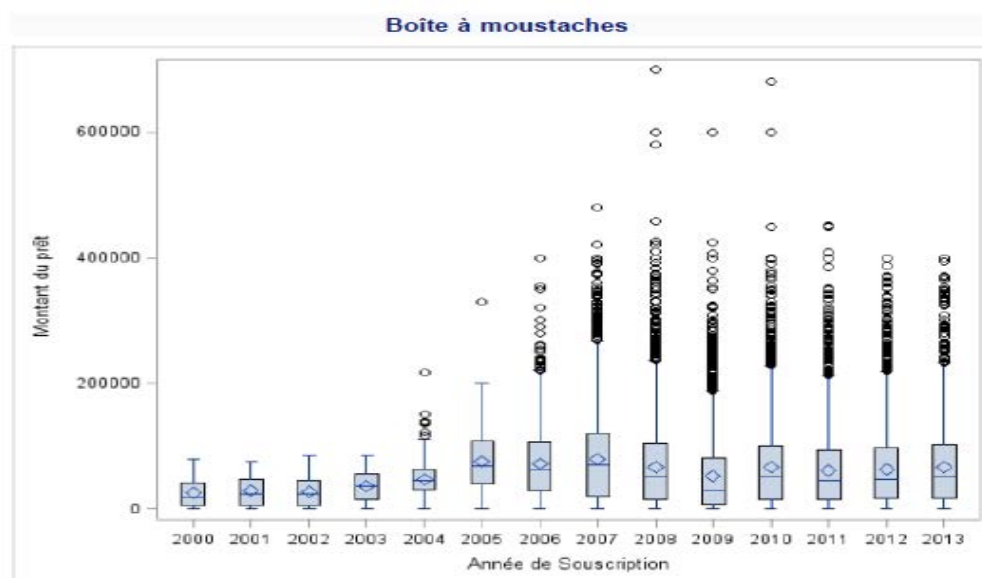


Illustration 12.13: Box-plot of loan amount by grouping.

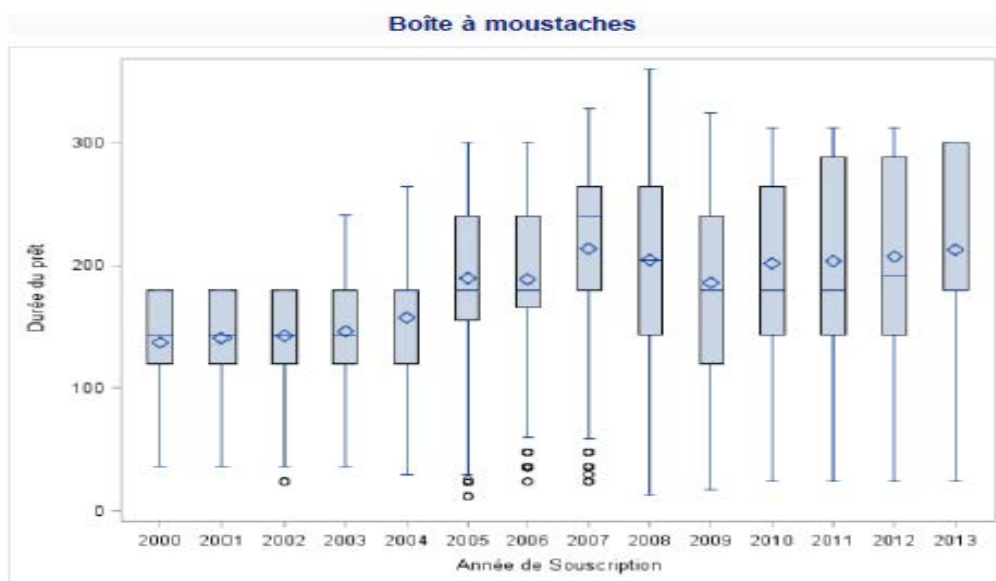


Illustration 12.14: Box-plot of loan duration by grouping.

The mean of loan amount tends to remain constant over the last groupings. Nevertheless, we can note a slight increase in the average duration. We can notice that the number of atypical loan amount (synonym of exemption) does not increase for groupings from 2006 to 2013. This reflects that deviation requests by the bank to obtain a amount borrowed superior to the limit remains stable: an increase could lead to higher risks. The box-plot

of the loans duration confirms the fact that durations are less concentrated around the mean than the amounts.

### Outstanding amount owed and penalty interest

The amount of penalty interest is 0.034% of the total outstanding amount owed and penalty interest. Considering this, we will focus on the analysis of the outstanding amount owed. The average outstanding amount owed is EUR 53,945 (for an average amount loan of EUR 65,378). By grouping, the evolution of box-plots is as follows:

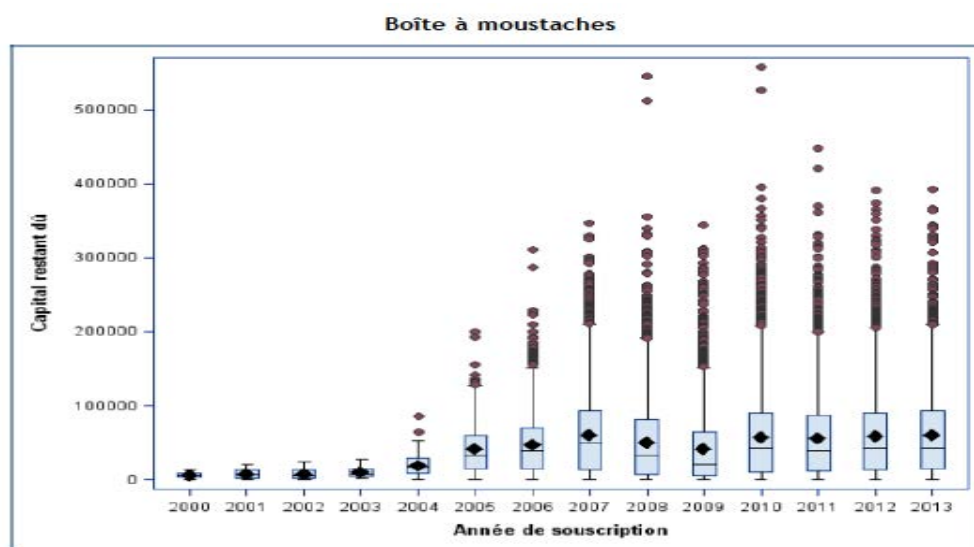


Illustration 12.15: Box-plot of outstanding amount owed.

Several observations can be made:

- The outstanding amount owed represents the capital under risk (penalty interest omitted): the maximum amount of provisions can not exceed the outstanding amount owed.
- The average of outstanding amount is more important for new grouping than for old grouping.

Now compare the outstanding amount owed with the loans amount:

| Grouping                        | 00   | 01 | 02   | 03   | 04   | 05 | 06   |
|---------------------------------|------|----|------|------|------|----|------|
| OAo grouping / Amt grouping (%) | 15.6 | 24 | 23.4 | 25.2 | 38.6 | 53 | 63.1 |

| Grouping                        | 07   | 08   | 09   | 10   | 11   | 12   | 13   |
|---------------------------------|------|------|------|------|------|------|------|
| OAo grouping / Amt grouping (%) | 74.1 | 77.2 | 80.2 | 86.7 | 90.5 | 93.8 | 90.2 |

Illustration 12.16: Ratio OAo over loan amount.

where “OAo grouping” means “outstanding amount owed by grouping” and recall “Amt



grouping” means “Loan amount by grouping”. As before, the outstanding amount ratio increase in relation to grouping. On the other hand the holder groupings (before 2006) have a ratio lower than 50%: in case of claim, the resale of the property could cover the outstanding amount owed by the borrower. To conclude observe the ratio “OAO grouping  $i$  / OAO all grouping”:

| Grouping                     | 00 | 01 | 02 | 03 | 04  | 05  | 06  |
|------------------------------|----|----|----|----|-----|-----|-----|
| OAO grouping / OAO total (%) | 0  | 0  | 0  | 0  | 0.2 | 2.3 | 5.2 |

| Grouping                     | 07   | 08   | 09  | 10   | 11   | 12   | 13   |
|------------------------------|------|------|-----|------|------|------|------|
| OAO grouping / OAO total (%) | 14.6 | 11.3 | 6.1 | 15.1 | 16.5 | 12.8 | 15.7 |

Illustration 12.17: Weight (in OAO) by grouping.

where “OAO total” means “outstanding amount owed for all grouping”. Here is a graphical illustration of this table:

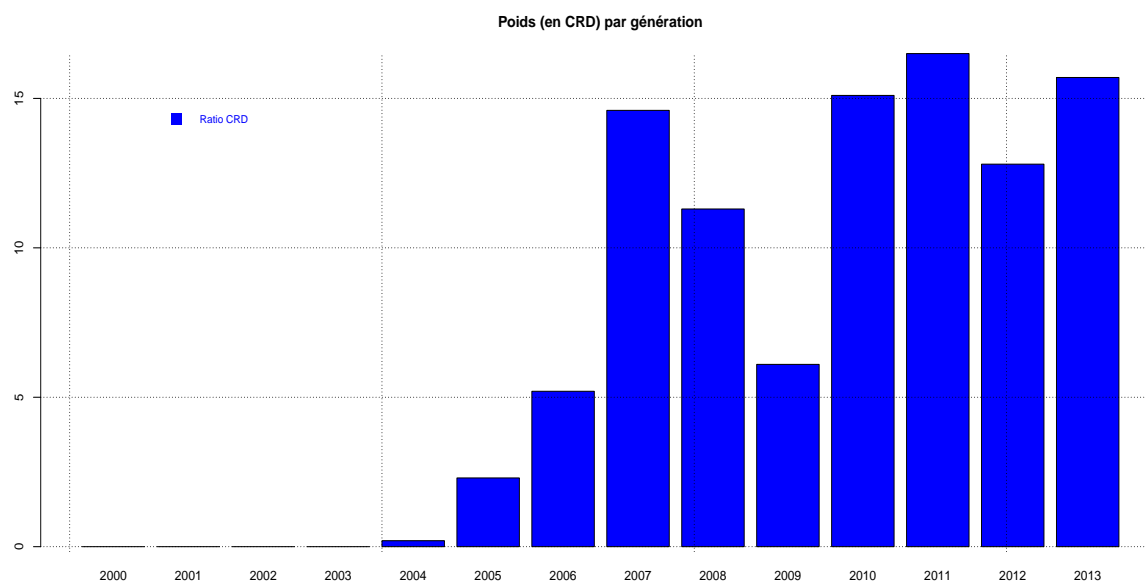


Illustration 12.18: Weight (in OAO) by grouping - Graphic.

Note that grouping 2006 to 2014 represents over than 95% of the total outstanding amount owed: these grouping represent the majority of risks.

### Claims

For your privacy we will not provide the data in terms of amount but only in terms of number. This is the distribution by grouping of 330 claims submitted between the first january 2000 and the 31 december 2013:

*Additional information concerning the provisioning model*

| Grouping                                   | 00   | 01  | 02  | 03   | 04  | 05   | 06   |
|--|------|-----|-----|------|-----|------|------|
| Nb claims grouping                         | 41   | 30  | 32  | 36   | 15  | 64   | 37   |
| Nb claims grouping / Nb claims total (%)   | 12.4 | 9.1 | 9.7 | 11   | 4.5 | 19.4 | 11.2 |
| Nb claims grouping / Nb loans grouping (%) | 11.3 | 7.3 | 7.4 | 10.9 | 3.8 | 3.7  | 1.2  |

| Grouping                                   | 07   | 08  | 09  | 10  | 11  | 12  | 13 |
|--|------|-----|-----|-----|-----|-----|----|
| Nb claims grouping                         | 35   | 13  | 21  | 2   | 2   | 2   | 0  |
| Nb claims grouping / Nb claims total (%)   | 10.6 | 3.9 | 6.4 | 0.6 | 0.6 | 0.6 | 0  |
| Nb claims grouping / Nb loans grouping (%) | 0.5  | 0.2 | 0.6 | 0   | 0   | 0   | 0  |

Illustration 12.19: Number of claims.

where “Nb claims grouping” means “number of claims by grouping”, “Nb claims total” means “number of claims for all grouping”, and “Nb loans grouping” means “number of loans by grouping”.

As mentioned in Section 12.3.9, a claim made in the early years of the life of the loan does not lead, in general, reimbursement from the insurance company. This is supported by previous data: grouping 2013 has not need compensation from the insurance company and the company intervened twice for each generation between 2010 and 2012.

The third row of the above table indicates that groupings 2000, 2003 , 2005, 2006, and 2007 are the generations that have the largest number of claims: they represent nearly 65% of total claims paid but for a weight around 27% of total loans.

The last line of the above table gives us the ratio, by grouping, of the number of claims divided the number of initial loans; this ratio is called the loss ratio. We logically find that the older grouping have rates loss higher than recent grouping.

### Loans matured

The number of loans matured in a natural way, i.e. without payment default, is:

| Grouping                                    | 00  | 01  | 02  | 03  | 04  | 05  | 06  |
|---|-----|-----|-----|-----|-----|-----|-----|
| Nb matured grouping                         | 279 | 292 | 269 | 191 | 147 | 483 | 770 |
| Nb matured grouping / Nb loans grouping (%) | 77  | 71  | 63  | 58  | 37  | 28  | 25  |

| Grouping                                    | 07   | 08   | 09  | 10  | 11  | 12  | 13  |
|---|------|------|-----|-----|-----|-----|-----|
| Nb matured grouping                         | 1695 | 1422 | 523 | 439 | 299 | 133 | 15  |
| Nb matured grouping / Nb loans grouping (%) | 25   | 23   | 15  | 8   | 5   | 3   | 0,3 |

Illustration 12.20: Loans matured.

where “Nb matured grouping” means “number of loans matured normally by grouping”. As expected the number of loans matured increase in function of grouping.

### 12.4.2 Mean, confident interval, and claim number estimation

As we explained in Chapter 11:

1. We can obtain a numerical value of the mean and variance of the provision with formulas (11.3.2) and (11.3.7) of Theorems 11.3.2 and 11.3.4.
2. We can simulate the provision with the Monte Carlo method proposed in Algorithm 11.1 of Section 11.3.2 and obtain a numerical value of quantile (mean, variance, etc.) on the simulated sample.

These two methods, one theoretical and the other based on simulations, should provide quite close numerical results for mean and variance. First, we need to obtain a numerical value of the parameters of our model. To do so, we will use the estimation method proposed in Section 11.4 applied to the portfolio of contracts presented in the previous section. That leads to following estimates:

| Parameter         | Notation          | Value  |
|-------------------|-------------------|--------|
| $\mathbb{E}(M_1)$ | $\bar{m}$         | 65,378 |
| $\tau$            | $\hat{\tau}$      | 0.07   |
| $\gamma$          | $\hat{\gamma}$    | 1.17   |
| $p$               | $\hat{p}$         | 2      |
| $\mu$             | $\hat{\mu}$       | -0.07  |
| $\sigma$          | $\hat{\sigma}$    | 0.1958 |
| $\lambda_1$       | $\hat{\lambda}_1$ | 94.89  |
| $\lambda_2$       | $\hat{\lambda}_2$ | 0.46   |
| $h$ (année)       | $\hat{h}$         | 23, 26 |

Illustration 12.21: Synthesis estimation parameters.

Note that according to what we have explained in Section 11.4, parameters related to the geometric Brownian motion were fixed by expert judgement. These parameters represent the evolution of real estate up to 30 years: it seems more appropriate to set these parameters rather than using data to estimate them. The choice made corresponds to a 30% drop in prices of real estate on a period of 10 years. This scenario partially matches the benchmark used by the European Central bank for stress tests in the banking sector ; This exercise is called “Asset Quality Review” or AQR.

We will present several results concerning the provision:

1. Comparison between theoretical mean and mean obtained by simulation,
2. Evolution of mean based on the date of analysis (grouping fixed),
3. Evolution of mean based on grouping (at a fixed analysis date),
4. Evolution of mean based on the analysis date and in run-off assumption,
5. Comparison with the provisioning methodology used by the insurance company.

### Theoretical mean Vs simulated mean

Here we make the comparison for the grouping 2000 (i.e.  $i = 0$ ) and by changing the analysis date  $t_0$ :

1. The mean of the provision obtained by replacing the value of parameters by their estimation in the formula (11.3.2) of Theorem 11.3.2, which will be denoted by:

$$\begin{aligned} \mathbb{E}^{theo}[P_{t_0}^0] := & \hat{\lambda}_2 \bar{m} \times \int_{\hat{A}_{t_0+}^0} \left\{ \hat{\phi}_t(s) \Phi \left[ \frac{\ln \left[ \frac{\hat{\phi}_t(s)}{\hat{\gamma}} \right] - \hat{\mu}t}{\hat{\sigma}\sqrt{t}} \right] \right. \\ & \left. - \hat{\gamma} e^{(\hat{\mu} + \frac{\hat{\sigma}^2}{2})t} \Phi \left[ \frac{\ln \left[ \frac{\hat{\phi}_t(s)}{\hat{\gamma}} \right] - (\hat{\mu} + \hat{\sigma}^2)t}{\hat{\sigma}\sqrt{t}} \right] \right\} ds dt, \end{aligned} \quad (12.4.2)$$

where:

$$\hat{A}_{t_0+}^0 := \left\{ (s, t) \in (\mathbb{R}_+)^2 : t_0 < t < s; s \leq \hat{h} \right\}, \quad (12.4.3)$$

$$\hat{\phi}_t(s) := \begin{cases} \left( e^{\hat{\tau}t^*} (1 + t^* \hat{p}) - t^* \hat{p} \right) \times \frac{e^{\hat{\tau}s} - e^{\hat{\tau}(t-t^*)}}{e^{\hat{\tau}s} - 1}, & \text{if } t^* < t \leq s, \\ \left( e^{\hat{\tau}t^*} (1 + t^* \hat{p}) - t^* \hat{p} \right), & \text{if } t \leq t^* \text{ or } s \leq t^*. \end{cases} \quad (12.4.4)$$

2. And mean of the provision that we obtained by simulating the provision with the following algorithm:

---

**Algorithm 12.1** Simulation of a sample of size  $l$  of variable  $P_{t_0}^0$ .

---

- 1: **for**  $k = 1$  to  $l$  **do**
- 2: Simulate a RV with law  $\mathcal{P}(\hat{\lambda}_2 \text{Leb}(\hat{A}_{t_0+}^0))$ . We denoted by  $n$  its realization.
- 3: Simulate  $n$  RVs iid with uniform law on  $\hat{A}_{t_0+}^0$ . We denoted by  $(t_j, d_j)_{1 \leq j \leq n}$  its realizations. Order dates  $(t_j)_{1 \leq j \leq n} : t_{(1)} \leq \dots \leq t_{(n)}$ . We denoted by  $d_{(j)}$  the duration associated to  $t_{(j)}$ .
- 4: RVs  $(M_j)_{1 \leq j \leq n}$  are simulated by reversing the empirical cdf. We denoted by  $(\tilde{m}_j)_{1 \leq j \leq n}$  its realizations.
- 5: Simulate  $n$  RVs iid with normal law. We denoted by  $(g_j)_{1 \leq j \leq n}$  its realizations.
- 6: Calculate:

$$p_k^0 = \sum_{j=1}^n \tilde{m}_j \left\{ \hat{\phi}_{t_{(j)}}(d_{(j)}) - \hat{\gamma} \prod_{k=1}^j \hat{r}_k \right\}_+, \quad (12.4.5)$$

where  $\hat{\phi}$  is defined by (12.4.4) and:

$$\hat{r}_k := \exp \left[ \hat{\sigma} \sqrt{t_{(k)} - t_{(k-1)}} g_k + \hat{\mu}(t_{(k)} - t_{(k-1)}) \right]. \quad (12.4.6)$$

- 7: **end for**
- 

Note that the sample size used is equal to  $l = 5\,000$ . The mean obtained on the

simulated sample is denoted by  $\mathbb{E}^{emp}(P_{t_0}^0)$  and defined by:

$$\mathbb{E}^{emp}(P_{t_0}^0) := \frac{1}{l} \sum_{k=1}^l p_k^0. \quad (12.4.7)$$

Now plot the curves  $[t_0, \mathbb{E}^{theo}(P_{t_0}^0)]$  (see (12.4.2)) and  $[t_0, \mathbb{E}^{emp}(P_{t_0}^0)]$  (see (12.4.7)) for  $t_0 \in \{0, \dots, [\hat{h}]\}$  (see Illustration 12.21):

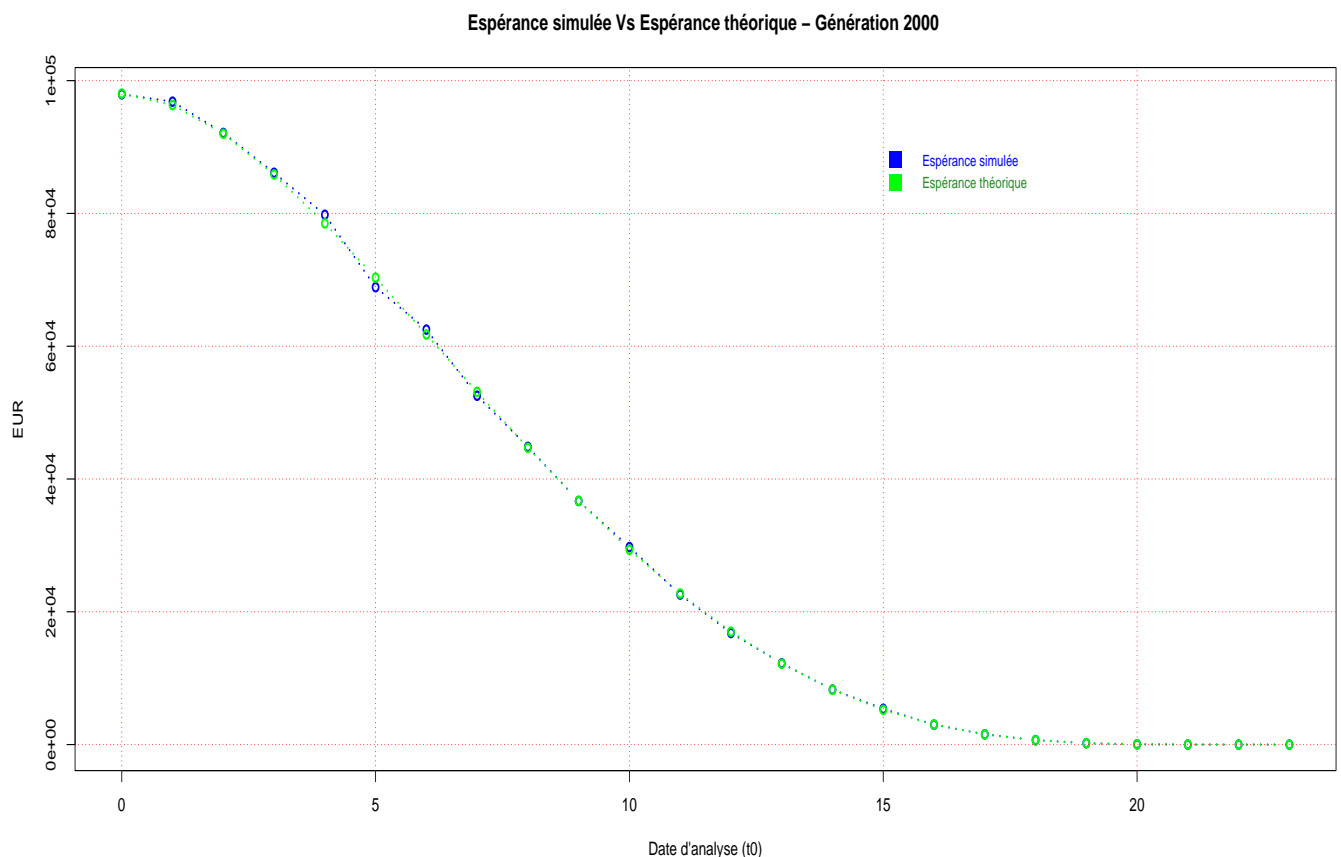


Illustration 12.22: Theoretical mean Vs simulated mean.

A few observations:

1. The theoretical mean and mean obtained by simulation are extremely close whatever the analysis date.
2. As expected, the provision mean is decreasing with the date analysis but this decrease is not linear.
3. Finally, from 2018 the mean of the provision is almost zero. This would allow us to determine numerically the time  $t_0^*$  (in this example  $t_0^* = 18$  years) from which the mean of the provision becomes insignificant, i.e.:

$$t_0^* := \inf\{t_0 \geq 0 : \mathbb{E}[P_{t_0}^0] \leq \beta \}, \quad (12.4.8)$$

where  $\beta$  is set very small. This would provide the insurance company the maximum provisioning duration for each grouping, i.e. the period during which the company should record a provision in its accounts.

We can notice that with the Monte Carlo simulations it is possible to obtain a confident interval of mean (at the 99.5% in the example below):

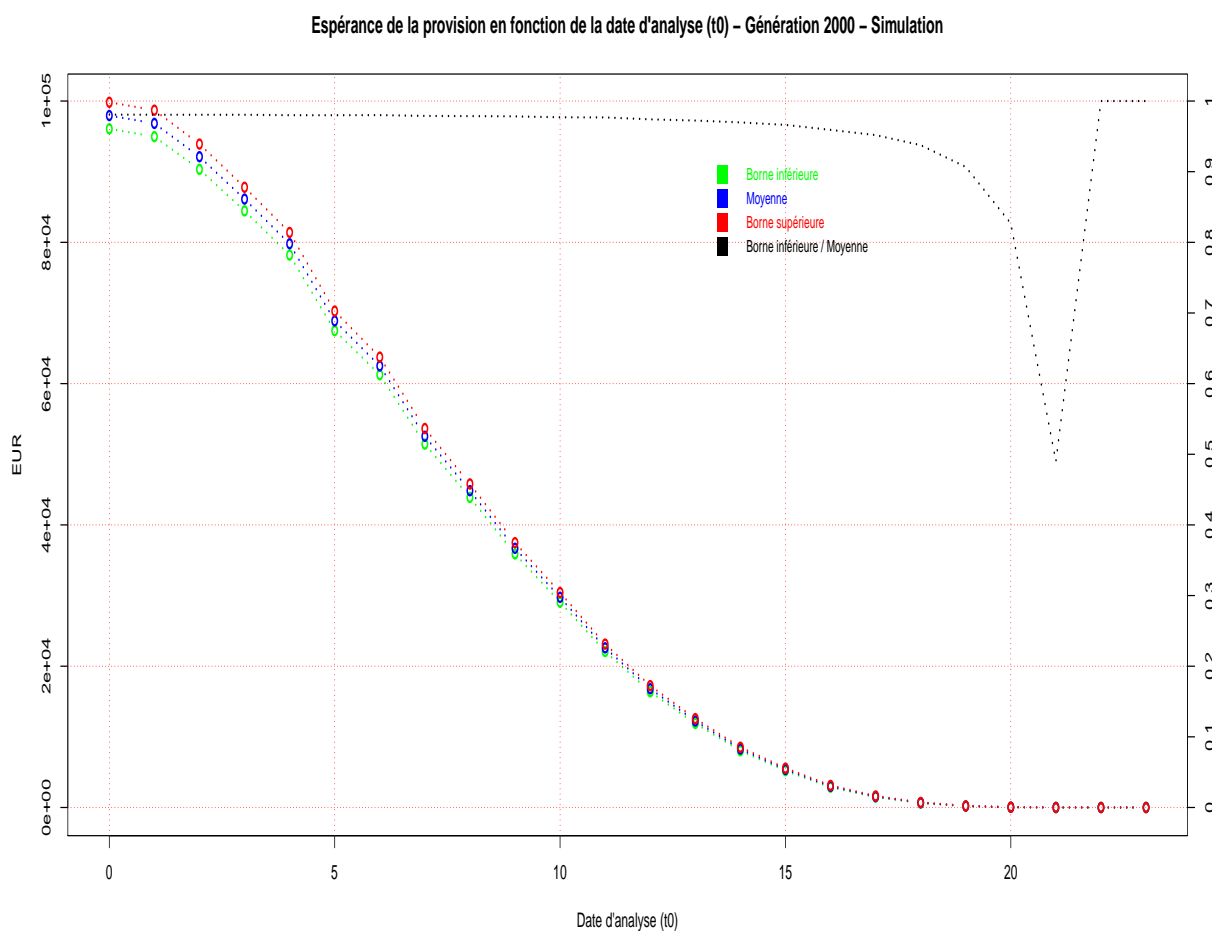


Illustration 12.23: Mean and confident interval based on the analysis date.

With regard to the simplicity of the algorithm used and the number of simulations run, confident intervals are very accurate, as we can see on the black curve representing the relative gap between the lower bound of the confident interval and mean of the provision.

### Evolution of mean in function of groupings

Here we set the analysis date to December 2013 the 31th and we make groupings of contracts vary. We only present the results obtained via simulations, i.e. we plot the curve:  $[i, \mathbb{E}^{emp}(P_{t_0}^i)]$  (see (12.4.7)) for  $i \in \{0, \dots, 13\}$ .

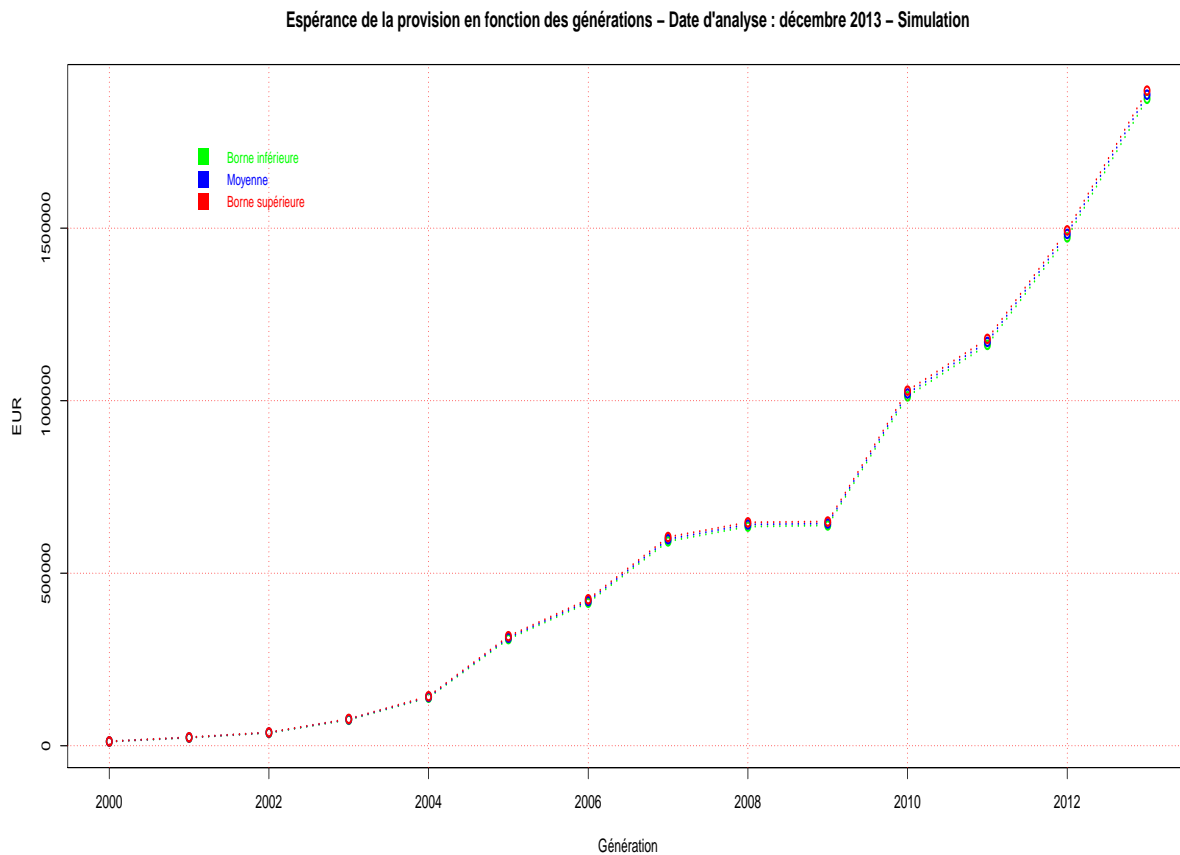


Illustration 12.24: Provision mean in relation to groupings.

As expected, the most recent the grouping is, the higher the provision mean increases.

It is interesting to compare the ratio  $\mathbb{E}^{emp}(P_{t_0}^i) / \mathbb{E}^{emp}(P_{t_0})$  to the ratio previously introduced “OAO grouping  $i$  over OAO all grouping”, where  $\mathbb{E}^{emp}(P_{t_0})$  is defined by:

$$\mathbb{E}^{emp}(P_{t_0}) := \sum_{i=0}^{\lfloor \frac{t_0}{t^*} \rfloor + 1} \mathbb{E}^{emp}(P_{t_0}^i). \quad (12.4.9)$$

Indeed the percentage of provision represented by each grouping must be substantially the same as the outstanding amount owed which is the capital under risk. We try to verify that the provision is “correctly distributed” between each grouping. We can thus

notice that because of the possibility to sell the financed real estate in case of claim, it would not be illogical that the provisioning ratio would be inferior to outstanding amount owed for the oldest groupings. There are the ratios:

| Grouping                                    | 00  | 01  | 02  | 03  | 04  | 05  | 06  |
|---|-----|-----|-----|-----|-----|-----|-----|
| <b>OAo grouping / OAo total (%)</b>         | 0   | 0   | 0   | 0   | 0.2 | 2.3 | 5.2 |
| $E^{emp}(P_{t_0}^i) / E^{emp}(P_{t_0})$ (%) | 0.4 | 0.7 | 1.1 | 1.6 | 2.3 | 3.2 | 4.3 |

| Grouping                                    | 07   | 08   | 09  | 10   | 11   | 12   | 13   |
|---|------|------|-----|------|------|------|------|
| <b>OAo grouping / OAo total (%)</b>         | 14.6 | 11.3 | 6.1 | 15.1 | 16.5 | 12.8 | 15.7 |
| $E^{emp}(P_{t_0}^i) / E^{emp}(P_{t_0})$ (%) | 5.6  | 7.1  | 9   | 11.7 | 14   | 17.5 | 21.4 |

Illustration 12.25: Ratio OAo Vs Ratio mean (%).

Remember that “OAo grouping” means “outstanding amount owed for each grouping of contracts” and “total OAo” means “outstanding amount owed for all groupings of contracts”. Here is an graphical illustration of this chart:

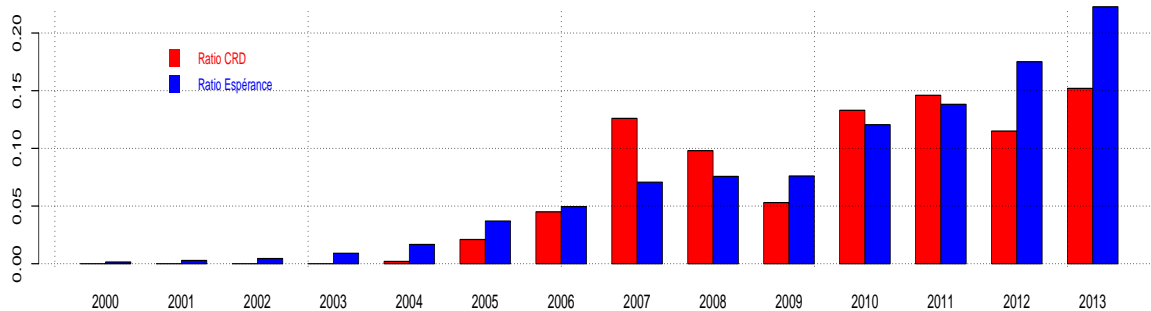
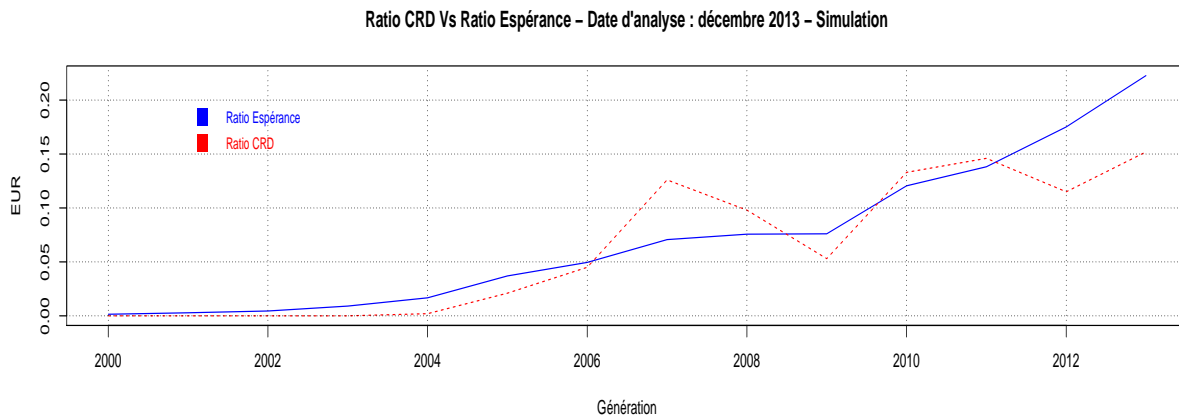


Illustration 12.26: Ratio OAo Vs Ratio mean (graphic).



We have to admit that the ratio of provisions “roughly” follows the same distribution by grouping as the outstanding amount owed ratio. However, we can notice that for groupings 2007, 2008, 2010 and 2011 this ratio (OAO Ratio) exceeds the ratio of provisions: it illustrates a weakness of the model.

The first thought of practitioners is that groupings previous 2008 are the riskiest. The crisis of 2008 would have changed the subscription policy of the “Caisses Régionales du Crédit Agricole”: contracts insured by CAMCA Assurance since this date would probably be less risky. Nevertheless, it is too early to judge whether this supposition is confirmed: upper groupings to 2008 have not yet matured enough to be compared to previous groupings. In any case, our PPP intensity modeling does not enable to take into account this phenomenon because we assumed that the parameters are independent of groupings.

We also find that the ratio of provisions is superior to the ratio of outstanding amount owed for groupings 2012 and 2013: this fact is really interesting for the insurance company to the extent that it prefers to give itself some leeway during the first years of a grouping.

To conclude, we can observe the number of claims predicted by the model as well as the ratio “Total number of claims / number of initial loans”. We neglect the groupings from 2000 to 2005 in the following analysis, which are quite insignificant.

| Grouping                                      | 06   | 07   | 08 | 09   | 10   | 11 | 12   | 13   |
|---|------|------|----|------|------|----|------|------|
| Nb claims estimated by grouping               | 39,1 | 43,9 | 49 | 54,4 | 60,1 | 66 | 72,2 | 78,7 |
| Nb claims total grouping / Nb loans total (%) | 2,5  | 1,2  | 1  | 2,1  | 1    | 1  | 1,6  | 1,4  |

Illustration 12.27: Number of predicted claims.

where “Nb claims estimated by grouping” means “mean of the number of claims by grouping” and “Nb claims total grouping” means “number of claims indemnified plus mean of the number of claims by grouping”.

We notice several factors:

1. The rate concerning grouping 2013 is entirely due to the model since there has not yet been a claims paid by the company at the date of calculation,
2. Previous groupings to 2008 do not appear riskier than the the other one contrary to practitioners supposition,
3. PPP intensity does not take into account groupings, thus the number of predicted claims ignores the number of loans taken out by grouping: it illustrates a weakness of the model.

**Evolution of provision mean in function of analysis date and under run-off assumption**

We focus on the evolution of the provision mean (for all grouping) in function of the analysis date  $t_0$  and under a run-off assumption from 31 december 2013, that is to say without assuming new production after this date. To do this we plot the following curve  $[t_0, \mathbb{E}^{emp}(P_{t_0}^{Runoff})]$  for  $t_0 \in \{0, \dots, \lfloor \hat{h} + 13 \rfloor\}$  where  $P_{t_0}^{Runoff}$  is here defined by:

$$P_{t_0}^{Runoff} := \sum_{i=0}^{\min(13, \lfloor \frac{t_0}{\tau^*} \rfloor)} P_{t_0}^i, \tag{12.4.10}$$

with  $P_{t_0}^i$  defined by (11.2.24). This is an illustration:

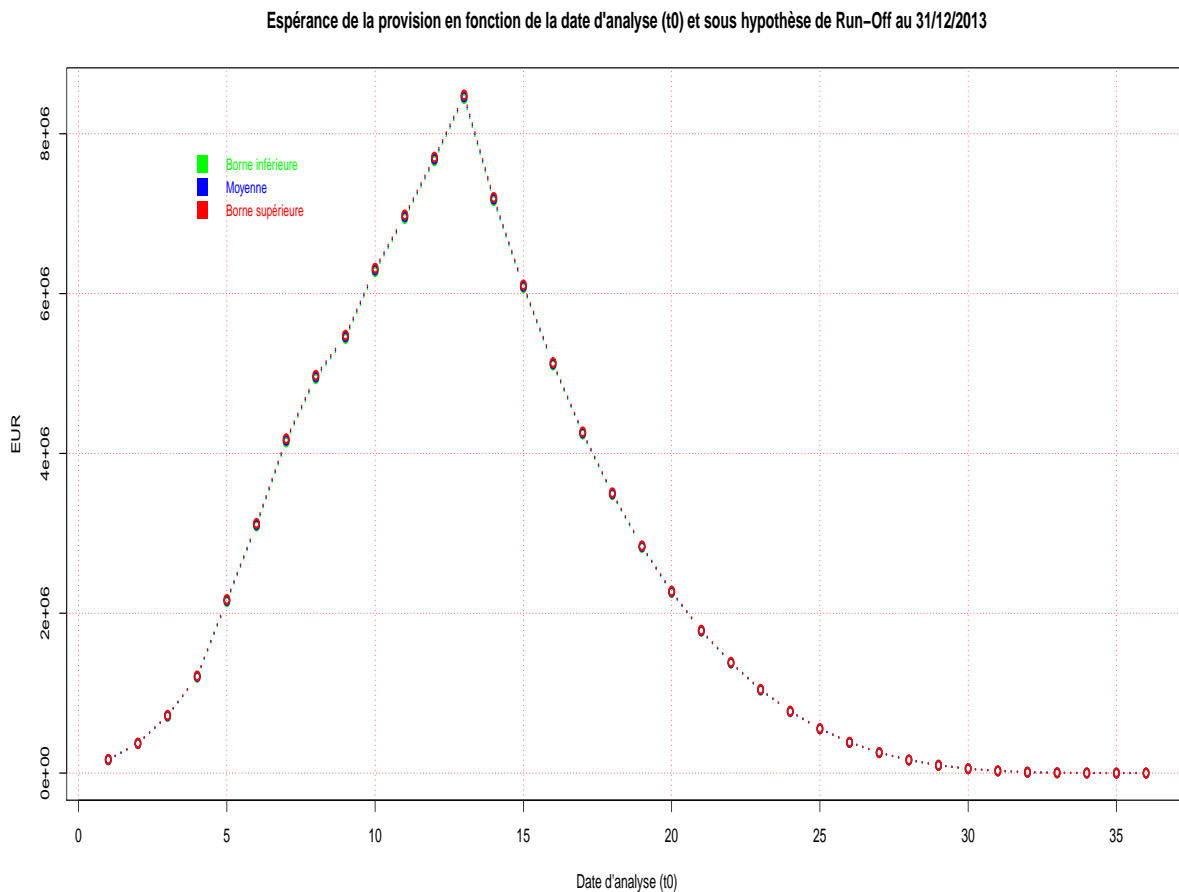


Illustration 12.28: Provision mean under run-off assumption.

We see that the provision mean is nil after 2031 when we assume a run-off in 2013.

### Comparison with CAMCA Assurance provisioning

The insurance company defines two types of provisions concerning mortgage guarantee of real estate loans: an individual provision (for each contract) for risky contracts and a collective provision for the others ; insurance company is able to isolate the most contracts risky thanks to different information.

The individual provision determined for each contract is calculated as “the expected loss bilan” or ELB. The methodology used is that of banking reform Basel 2, i.e. the ELB is equal to the probability of default of the contract multiplied by the loss at default (“loss given default”) multiplied by the exposure at time of default (“exposure at default”). These provisions are provided by the bank. We will do not enter into any real detail on the calculation of this provision because we do not have all the necessary information.

The second type of provision is a collective provision “Mean of claim amount \* Number of Claims ”, where:

- The mean of claim amount is determined for each grouping as the outstanding amount owed (including penalty interest),
- The number of claims is determined with a pseudo Chain ladder method.

CAMCA Assurance included these provisions in its accounts; it is statutory provisions.

Future Solvency 2 reform (see Section 9.1.2) requires insurance companies, to calculate technical provisions with “Best Estimate” principle (BE for short) in beginning of 2016. The BE of technical provisions is defined in the technical specifications provided by EIOPA as follows (see [1]):

*“The best estimate should correspond to the probability weighted average of future cash-flows taking account of the time value of money”.*

Schematically, the BE of provisions corresponds to a discounting of technical statutory provisions without taking a larger prudential safety margin: thus the BE provisions are generally lower than the statutory ones. CAMCA insurance proceeds, since 2010, with a BE calculation of these provisions. This calculation is a rehabilitation of the method used to calculate statutory provisions that we will not detail here.

On 31 December 2013, we compare the amount of statutory provisions calculated by the insurance company, the amount of BE of these provisions and the amount calculated using our model (but this time discounted): the first two provisions are provided by the company. We present the results by grouping of contracts. In the graph below, the red bars represent the statutory provisions calculated by CAMCA Assurance, the blue bars represent the BE of these provisions, the green ones represent provisions calculated with our method and the x axis represents the groupings of contracts. We ignore the

groupings from 2000 to 2005.

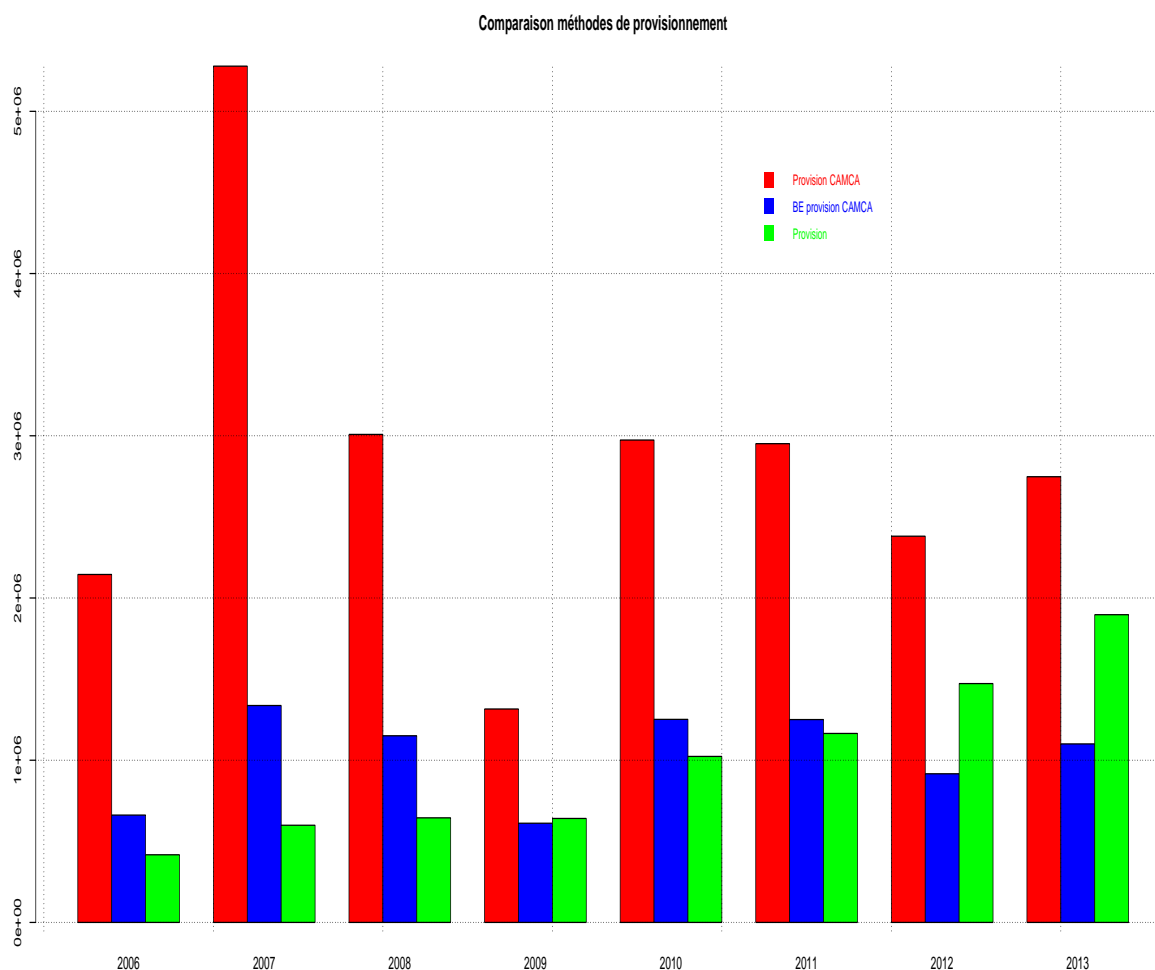


Illustration 12.29: Comparison of provisioning method.

We can draw from Illustration 12.29 the following consequences:

1. As stated above, statutory provisions are higher to the BE ones.
2. We also notice that the statutory provisions and the BE provisions take into account the character, supposedly riskier of groupings before 2008.
3. Provisions of our model and the BE provisions are relatively close for groupings from 2009 to 2011, but diverge somewhat for others.

Now we compare the ratios “Provision grouping  $i$  over Provisions all groupings” and “OAO grouping  $i$  OAO all groupings”. This comparison will tell us to what extent the provisions are “correctly” distributed by grouping.

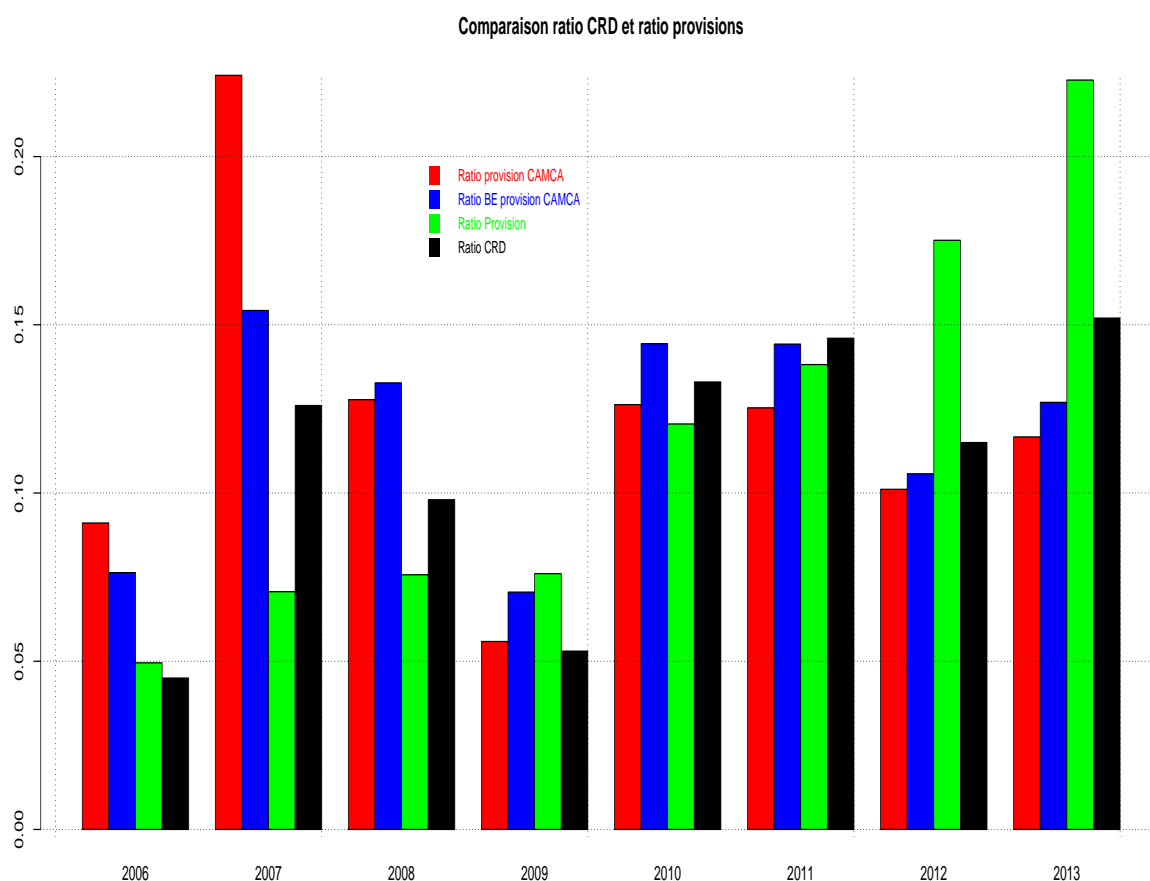


Illustration 12.30: Ratio OAO Vs Ratio mean in function of provisioning method.

A few observations:

1. As mentioned above we find that only our model accrues provisioning for the last two groupings (i.e. 2012 and 2013): this fact is interesting for the insurance company.
2. In contrast, the statutory provisions and the BE of these provisions take account, unlike our model, character supposedly riskier of groupings 2008 and the previous ones.
3. For groupings from 2009 to 2011 ratios of provisions and OAO are similar.

To conclude we can compare the claims ratios of these different methods of provisioning i.e. we compare the ratio “Provisions + amount of paid claims divided by insurance premiums collected”.

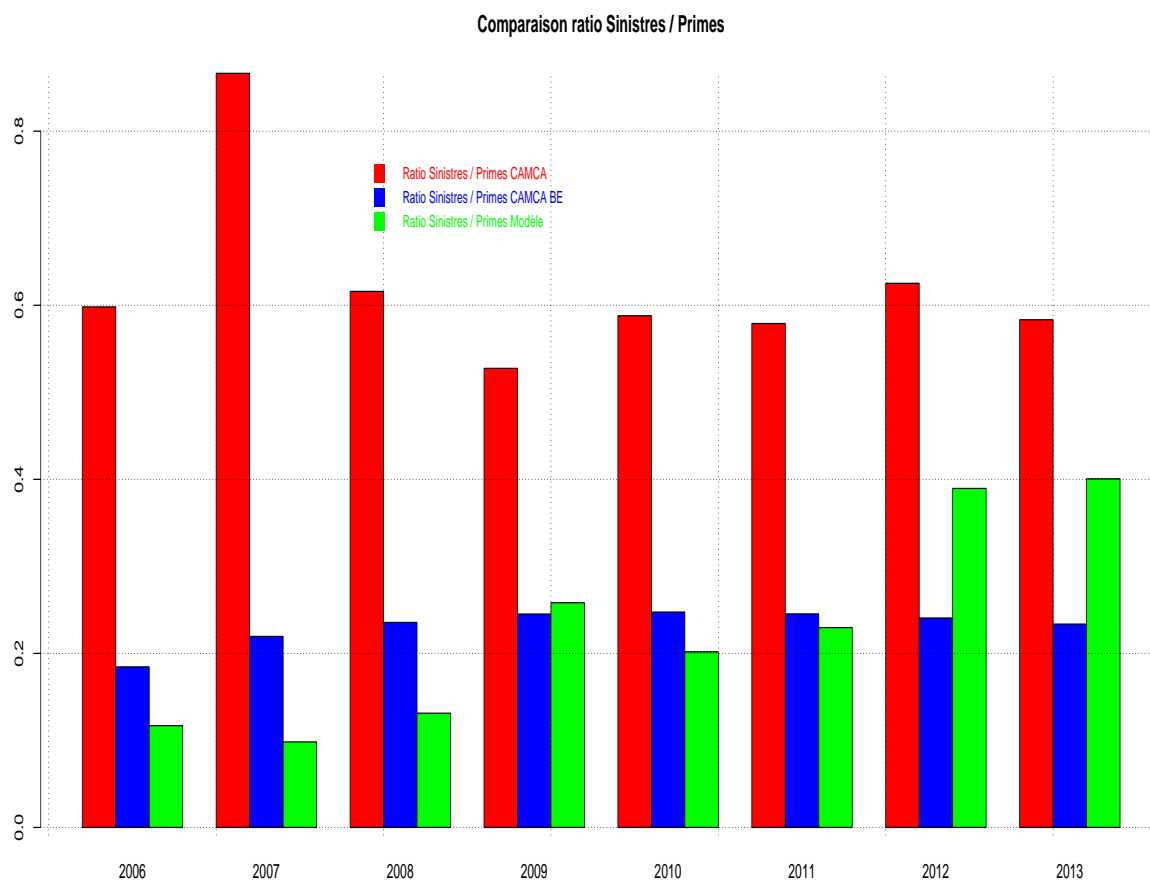


Illustration 12.31: Ratio Claim / Premium in function of provisioning method.

A few observations:

1. Claims ratio determined with the statutory provisions are largely superior to the other claims ratio.
2. Claims ratio of BE provisions are almost constant for each grouping.
3. Claims ratio determined using our model are more important for groupings 2012 and 2013 and are quite similar with those calculated using the BE method of provisions for other groupings.

Our model suggests that the 2008 and earlier groupings would not be more risky than others; it reverses the a-priori of practitioners. The claims ratio known at the date of calculation reinforces this conclusion.

### 12.4.3 Sensitivity with regard to parameters

In this section we consider grouping  $i = 0$ . From Theorem 11.3.2 we have:

$$\mathbb{E}[P_{t_0}^0] = \lambda_2 \gamma \mathbb{E}[M_1] \int_{A_{t_0+}^0} F_1 \left[ t, \frac{\phi_t(s, \tau)}{\gamma}, \mu, \sigma \right] dt ds. \quad (12.4.11)$$

where  $F_1$  is defined by (11.3.1),  $\mathbb{E}[M_1]$  is the borrowed amount mean, and:

$$\phi_t(s, \tau) := \begin{cases} \left[ e^{\tau t^*} (1 + t^* p) - t^* p \right] \frac{e^{\tau s} - e^{\tau(t-t^*)}}{e^{\tau s} - 1}, & \text{if } t^* < t \leq s, \\ \left[ e^{\tau t^*} (1 + t^* p) - t^* p \right], & \text{if } t \leq t^* \text{ or } s \leq t^*, \end{cases} \quad (12.4.12)$$

We denote:

$$g(\lambda_2, \tau, \mu, \sigma) := \mathbb{E}[P_{t_0}^0], \quad 0 \leq t_0 \leq h.$$

We want to measure the sensitivity of the provision mean subsequently a fluctuation of some parameters, i.e. measure the fluctuation of  $\mathbb{E}(P_{t_0}^0)$  following the fluctuation:

1. of the PPP intensity:  $\lambda_2$ ,
2. of the loan rate:  $\tau$ ,
3. of the parameters of the geometric Brownian motion:  $\mu$  and  $\sigma$ .

We will compare the evolution of these sensitivities when the analysis date  $t_0$  fluctuates.

#### Sensitivity with regard to PPP intensity

Here we want to quantify the impact on the provision mean of a fluctuation of the PPP intensity: that is to say the impact on  $\mathbb{E}(P_{t_0}^0)$  of a fluctuation of the mean of the number of claim by time unit when we assume that the other parameter do not fluctuate.

**Definition 12.4.1.** *Sensitivity of the provision mean subsequently a fluctuation of  $\lambda_2$  is denoted by  $S_{\lambda_2}$  ( $S$  for sensitivity) and defined by:*

$$S_{\lambda_2} := \frac{\partial g(\lambda_2, \tau, \mu, \sigma)}{\partial \lambda_2}. \quad (12.4.13)$$

It is clear that  $\mathbb{E}[P_{t_0}^0]$  is proportionate to  $\lambda_2$ , so:

$$S_{\lambda_2} = \gamma \mathbb{E}[M_1] \int_{A_{t_0+}^0} F_1 \left[ t, \frac{\phi_t(s, \tau)}{\gamma}, \mu, \sigma \right] dt ds. \quad (12.4.14)$$

We see that  $S_{\lambda_2}$  is not function of  $\lambda_2$ , that is to say the fluctuation of  $\mathbb{E}(P_{t_0}^0)$  subsequently the fluctuation of  $\lambda_2$  is the same for all value of  $\lambda_2$ .

### Sensitivity with regard to the loan rate

Here we want to determine how the provision mean fluctuates when the rate  $\tau$  fluctuates.

**Definition 12.4.2.** *The sensitivity of the provision mean subsequently a fluctuation of the rate  $\tau$  is denoted by  $S_\tau$  ( $S$  for sensitivity) and defined by:*

$$S_\tau := \frac{\partial g(\lambda_2, \tau, \mu, \sigma)}{\partial \tau}. \quad (12.4.15)$$

Let  $\Psi$  the density of a Gaussian RV centered with variance equal to one. From (12.4.11) it follows:

$$\begin{aligned} S_\tau = \lambda_2 \mathbb{E}[M_1] \int_0^h \left\{ \int_0^s \left[ \frac{\partial}{\partial \tau} \phi_t(s, \tau) \Phi \left[ \tilde{d} \left( \frac{\phi_t(s, \tau)}{\gamma}, \mu, \sigma, t \right) \right] \right. \right. \\ \left. \left. + \frac{\frac{\partial}{\partial \tau} \phi_t(s, \tau)}{\sigma \sqrt{t}} \left( \Psi \left[ \tilde{d} \left( \frac{\phi_t(s, \tau)}{\gamma}, \mu, \sigma, t \right) \right] - \frac{\gamma e^{(\mu + \frac{\sigma^2}{2})t}}{\phi_t(s, \tau)} \Psi \left[ \tilde{d} \left( \frac{\phi_t(s, \tau)}{\gamma}, \mu + \sigma^2, \sigma, t \right) \right] \right) \right] dt \right\} ds, \end{aligned} \quad (12.4.16)$$

where:

$$\frac{\partial}{\partial \tau} \phi_t(s, \tau) = \begin{cases} t^*(1 + t^*p) \frac{e^{\tau s} - e^{\tau(t-t^*)}}{e^{\tau s} - 1} + [e^{\tau t^*} (1 + t^*p) - t^*p] \\ \quad \times \frac{[se^{\tau s} - (t - t^*)e^{\tau(t-t^*)}](e^{\tau s} - 1) - se^{\tau s}[e^{\tau s} - e^{\tau(t-t^*)}]}{(e^{\tau s} - 1)^2}, & \text{if } t^* < t \leq s, \\ t^* e^{\tau t^*} (1 + t^*p), & \text{if } t \leq t^* \text{ or } s \leq t^*, \end{cases} \quad (12.4.17)$$

where function  $\phi$  is defined by (12.4.12).

We plot the curve  $(\tau, S_\tau)$  for  $\tau \in [\frac{\hat{\tau}}{2}; 2\hat{\tau}]$  and for several value of  $t_0$ :

- Green curve represent the evolution of sensitivity in function of  $\tau$  and for analysis date  $t_0 = 0$ ,
- Blue curve represent the evolution of sensitivity in function of  $\tau$  and for analysis date  $t_0 = 5$ ,
- Red curve represent the evolution of sensitivity in function of  $\tau$  and for analysis date  $t_0 = 10$ ,
- Black curve represent the evolution of sensitivity in function of  $\tau$  and for analysis date  $t_0 = 15$ ,

Variable  $\tau$  is on the x axis and the sensitivity value on the y axis. The yellow vertical curve represents value of sensitivity for the value of  $\tau$  given in Section 12.4.2, i.e.  $\hat{\tau} = 0.07$ .



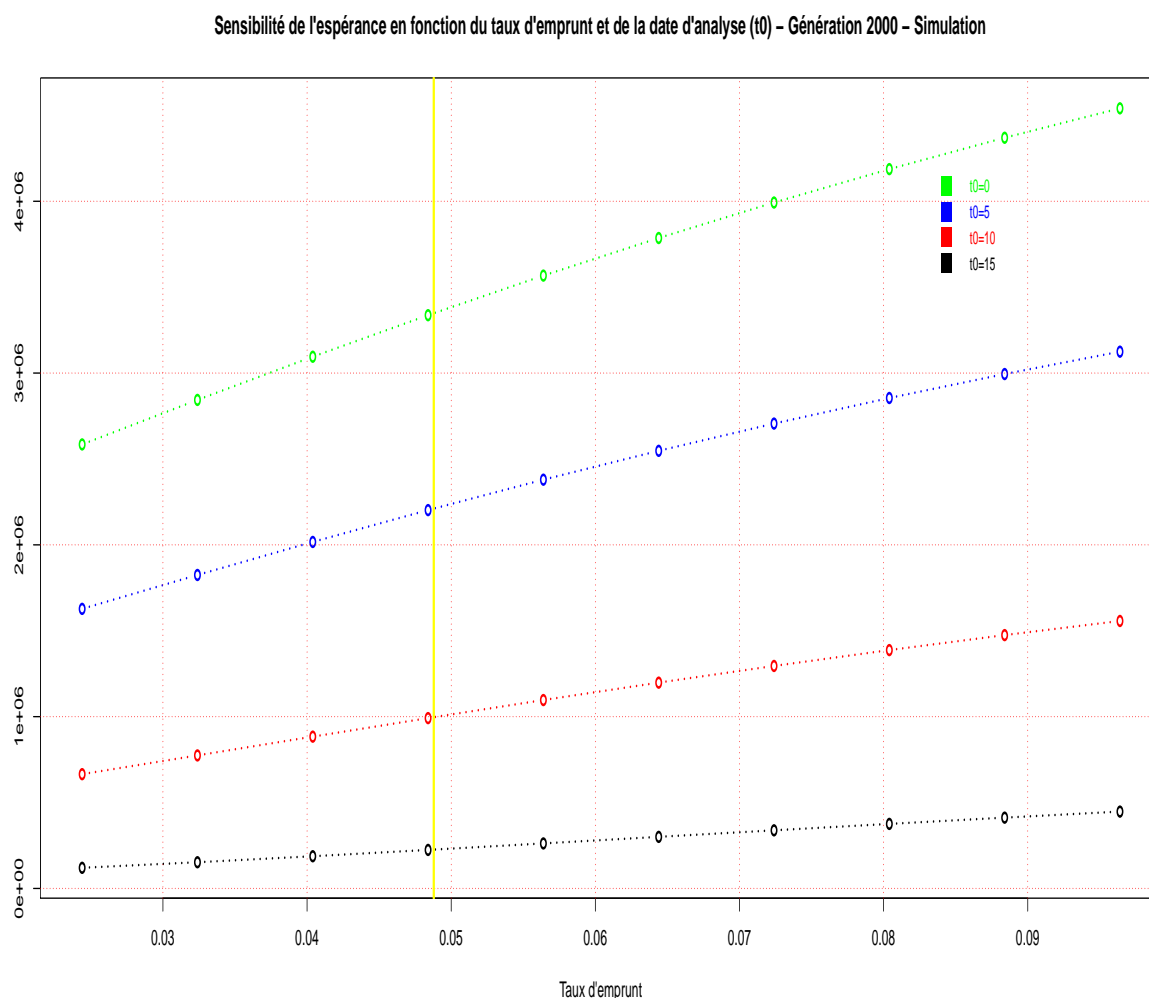


Illustration 12.32: Sensitivity of the provision mean with regard to  $\tau$ .

Some observations:

1. As expected, the sensitivity is positive, that is to say the provision mean increase when  $\tau$  increase.
2. in addition influence of parameter  $\tau$  decrease in function of analysis date  $t_0$ .
3. We observe also that that the sensitivity seems to be convex in function of  $\tau$ .

### Sensitivity with regard to the parameters of real estate

Here we want to measure the impact of a fluctuation of parameters  $\mu$  and  $\sigma$  on the provision mean by assuming that all the other parameters are constant. These two sensitivities gives us an indication of how  $\mathbb{E}(P_{t_0}^0)$  fluctuates when the real estate price fluctuate. We define these two sensitivities:

**Definition 12.4.3.** *Sensitivity of provision mean subsequently the fluctuation of  $\mu$  (respectively  $\sigma$ ) is denoted by  $S_\mu$  (respectively  $S_\sigma$ ) and defined by:*

$$S_\mu := \frac{\partial g(\lambda_2, \tau, \mu, \sigma)}{\partial \mu}, \quad (12.4.18)$$

$$S_\sigma := \frac{\partial g(\lambda_2, \tau, \mu, \sigma)}{\partial \sigma}. \quad (12.4.19)$$

From formula (12.4.11) it comes that:

$$\begin{aligned} S_\mu = & -\lambda_2 \mathbb{E}[M_1] \int_0^h \left\{ \int_0^s \frac{\sqrt{t} \phi_t(s, \tau)}{\sigma} \Psi \left[ \tilde{d} \left( \frac{\phi_t(s, \tau)}{\gamma}, \mu, \sigma, t \right) \right] \right. \\ & \left. + \gamma e^{(\mu + \frac{\sigma^2}{2})t} \left[ t \Phi \left[ \tilde{d} \left( \frac{\phi_t(s, \tau)}{\gamma}, \mu + \sigma^2, \sigma, t \right) \right] - \frac{\sqrt{t}}{\sigma} \Psi \left[ \tilde{d} \left( \frac{\phi_t(s, \tau)}{\gamma}, \mu + \sigma^2, \sigma, t \right) \right] \right] dt \right\} ds \end{aligned} \quad (12.4.20)$$

$$\begin{aligned} S_\sigma = & -\lambda_2 \mathbb{E}[M_1] \int_0^h \left\{ \int_0^s \frac{\phi_t(s, \tau)}{\sigma} \tilde{d} \left( \frac{\phi_t(s, \tau)}{\gamma}, \mu, \sigma, t \right) \Psi \left[ \tilde{d} \left( \frac{\phi_t(s, \tau)}{\gamma}, \mu, \sigma, t \right) \right] \right. \\ & \left. + \gamma e^{(\mu + \frac{\sigma^2}{2})t} \left[ \sigma t \Phi \left[ \tilde{d} \left( \frac{\phi_t(s, \tau)}{\gamma}, \mu + \sigma^2, \sigma, t \right) \right] - \left( 2\sqrt{t} + \frac{\tilde{d} \left( \frac{\phi_t(s, \tau)}{\gamma}, \mu + \sigma^2, \sigma, t \right)}{\sigma} \right) \right. \right. \\ & \left. \left. \times \Psi \left[ \tilde{d} \left( \frac{\phi_t(s, \tau)}{\gamma}, \mu + \sigma^2, \sigma, t \right) \right] \right] dt \right\} ds, \end{aligned} \quad (12.4.21)$$

where  $\Psi$  is the density of a Gaussian RV centered with variance equal to one.

Here we plot curves  $(\mu, S_\mu)$  and  $(\sigma, S_\sigma)$  for  $\mu \in [-1; 1]$ ,  $\sigma \in [0; 1]$  and for several value of  $t_0$ :

- Green curves represent evolution of sensitivities with regard to  $\mu$  and  $\sigma$  and for the analysis date  $t_0 = 0$ ,
- Blue curves represent evolution of sensitivities with regard to  $\mu$  and  $\sigma$  and for the analysis date  $t_0 = 5$ ,
- Red curves represent evolution of sensitivities with regard to  $\mu$  and  $\sigma$  and for the analysis date  $t_0 = 10$ ,
- Black curves represent evolution of sensitivities with regard to  $\mu$  and  $\sigma$  and for the analysis date  $t_0 = 15$ ,

Parameter  $\mu$  (resp.  $\sigma$ ) is on the x axis and the sensitivity on the y axis. The vertical yellow curve represent the sensitivities values for values of  $\mu$  and  $\sigma$  given in Section 12.4.2, i.e.  $\mu = -0.07$  and  $\sigma = 0.1958$ .

12.4 Numerical application

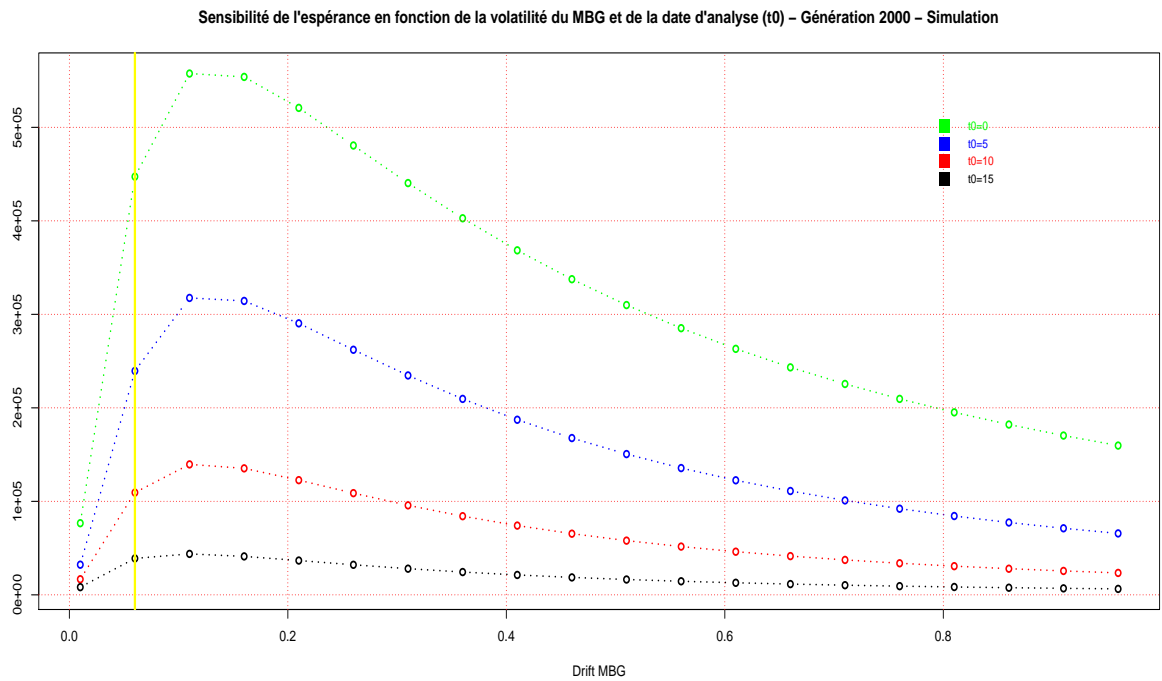
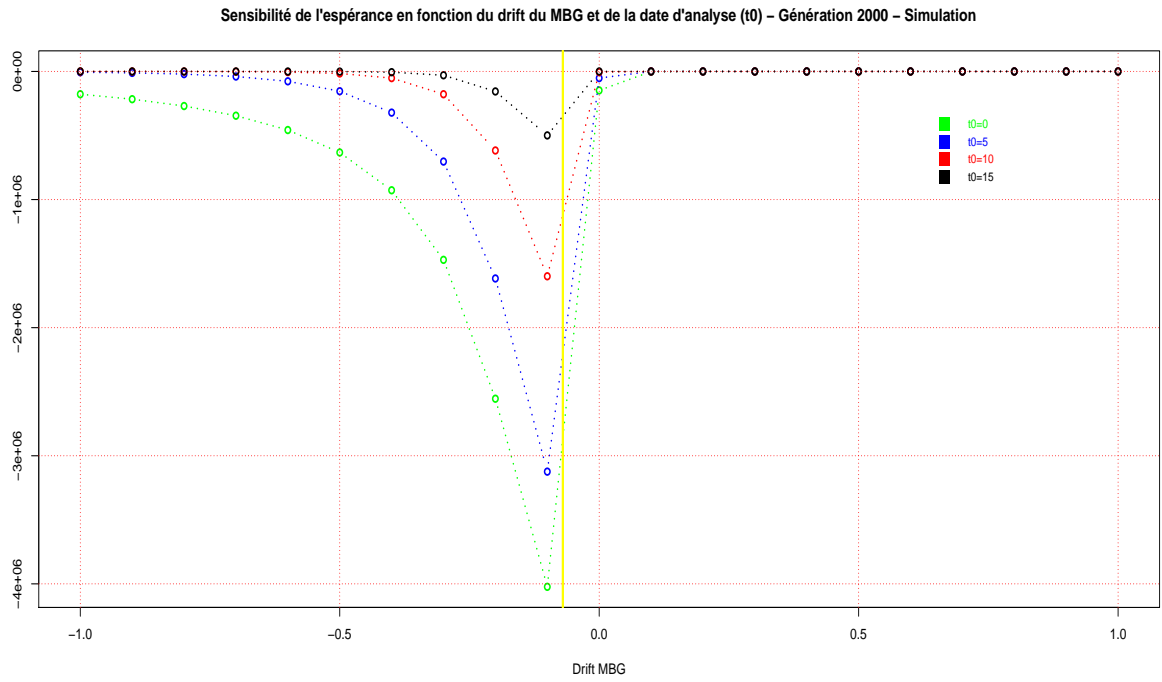


Illustration 12.33: Sensitivity of the provision mean with regard to  $\mu$  and  $\sigma$ .

From Illustration 12.33 we can conclude:

1. As expected, sensitivity of the provision mean according to  $\mu$  is negative. This mean that the provision mean decreases when  $\mu$  increases. Indeed parameter  $\mu$  represents the trend of real estate price: when the prices increases, the recovery on property increases also so the claim amount decreases. We see also that provision mean is concave with regard to  $\mu$ . In addition influence of parameter  $\mu$  decreases when  $t_0$  increases.
2. Sensitivity of the provision mean subsequently of  $\sigma$  is positive. This mean that the provision mean increases when  $\sigma$  increases. This sensitivity decreases when  $t_0$  increases. We can also note that the provision mean is convex function of  $\sigma$ .
3. To date we are unable to explain why there is a failure for  $\mu \sim -0,1$  in Illustration 12.33.

### Comparison between sensitivities

Here we want to compare impact of parameters on the provision mean in function of the analysis date  $t_0$ . We plot the following curves  $(t_0, S_k)$  where:  $k \in \{\hat{\lambda}_2, \hat{\tau}, \hat{\mu}, \hat{\sigma}\}$ . These curves represent the evolution in function of  $t_0$  of the sensitivities.

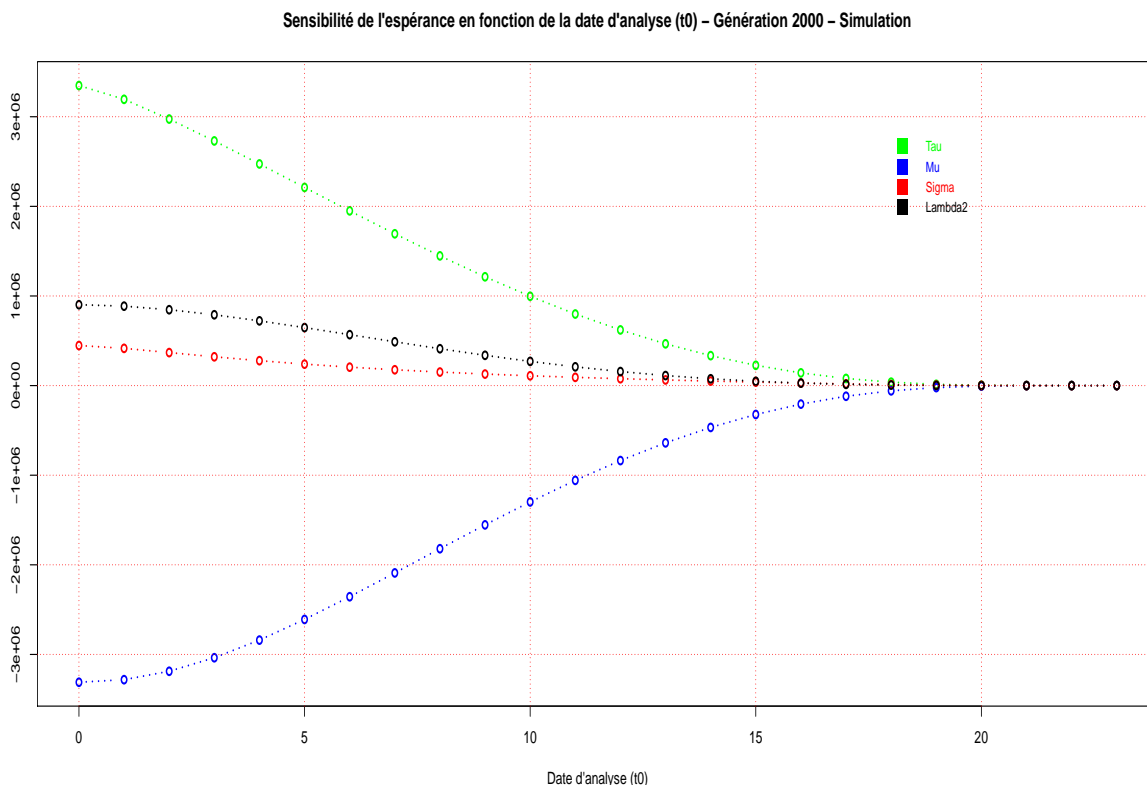


Illustration 12.34: Comparison of sensitivities in function of  $t_0$ .

We see that:

1. Only  $\mu$  allows to reduce the provision mean. Provision mean increases when  $\tau$ ,  $\sigma$  and  $\lambda_2$  increase.
2. We also see that the rate  $\tau$  has much impact than the other parameters.

#### 12.4.4 Probability of nullity of claim amount

It is interesting to look at the probability, given  $[(T, D) = (t, d)]$ , that the claim amount is nil. We denoted by  $p(t, d, t^*)$  this probability, i.e.:

$$p(t, d, t^*) := \mathbb{P}\left[Mf(T, D, R_T) = 0 \mid (D, T) = (d, t)\right], \quad (12.4.22)$$

where function  $f$  is defined by (11.2.16). Note that the probability  $p$  depends on time  $t^*$  because the function  $f$  implicitly depends on  $t^*$ . We analyze further the evolution of this probability in function of the values of the ratio  $\frac{t}{d}$ .

By definition of  $f$  it follows:

$$\begin{aligned} Mf(t, d, R_t) = 0 &\Leftrightarrow M[\phi_t(d) - R_t\gamma]_+ = 0 \\ &\Leftrightarrow \phi_t(d) - R_t\gamma \leq 0 \\ &\Leftrightarrow R_t \geq \frac{\phi_t(d)}{\gamma} \\ &\Leftrightarrow B_t \geq \frac{\ln\left[\frac{\phi_t(d)}{\gamma}\right] - \mu t}{\sigma t}. \end{aligned} \quad (12.4.23)$$

By definition of a Brownian motion it comes:

$$p(t, d, t^*) = 1 - \Phi\left[\frac{\ln\left[\frac{\phi_t(d)}{\gamma}\right] - \mu t}{\sigma\sqrt{t}}\right], \quad (12.4.24)$$

where  $\phi$  is defined by (11.2.13) and  $\Phi$  is the cdf of a standard Gaussian RV.

We want to see how the probability  $p(t, d, t^*)$  fluctuates in function of:

$$\frac{t}{d} =: \beta, \quad (12.4.25)$$

i.e. in function of the step of a loan.

To do this we plot the curve  $\left(\beta, \min_d(p(\beta d, d, t^*))\right)$  for parameter given in Section 12.4.2:

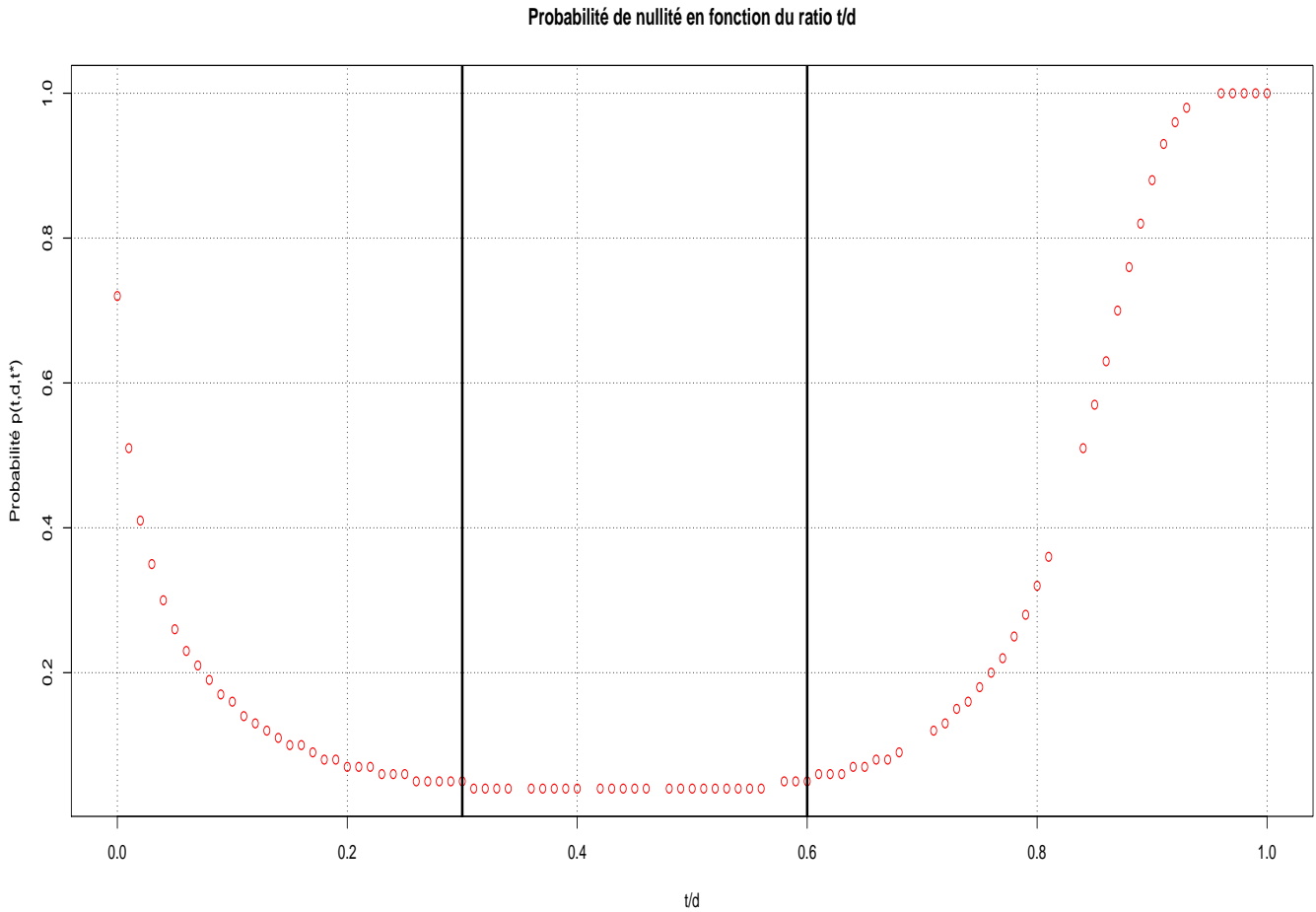


Illustration 12.35: Probability of nullity of a claim amount.

The previous chart points out two important things.

1. If a claim is declared at the beginning of the life of a loan then the amount of such claim is likely to be zero. This observation confirms the a-priori of practitioners. According to practitioners, the explanation is that if a claim occurs at the beginning of the life of the loan then the amount owed by the borrower is lower than the property resale amount. We explain in more detail this point. Consider a reported claim at time 0, then:

$$\lim_{t^* \rightarrow 0} Mf(0, D, R_0) = M[\alpha - (1 + a)(1 - c)]_+, \quad (12.4.26)$$

where recall that  $a$  represents the borrower deposit and  $c$  represents cost due the resale of property. So if  $\frac{1}{c} - \frac{1}{a} \geq \alpha$  then the claim amount submitted at date 0 is nil. Given our estimates it follows:

$$(1 + a)(1 - c) = 1,17 \text{ and } \alpha = 1,15, \quad (12.4.27)$$

so:

$$\lim_{t^* \rightarrow 0} Mf(0, D, R_0) = 0. \quad (12.4.28)$$

2. On the other hand we find that the probability  $p(t, d, t^*)$ :

- Decreases in the early life of the loans,
- Stay level to 0 between ratios 0.3 and 0.6,
- Increases from ratio 0.6 to achieve a probability close to 1 for ratio 1.

This observation is also confirmed by practitioners. One reason advanced is that early in the life of the loan, borrower deposit is sufficient to cancel the amount of the claim (that is why it is high at the beginning of life loans). On the other hand more a loan is advanced in time, less the outstanding amount owed is important, that is the reason why  $p(t, d, t^*)$  increases from the ratio 0.6.

#### 12.4.5 Function “nleqslv” of software

No translation, see page [106](#).





## Part VI

# Solvency need resulting from provisioning risk in a ORSA context

---

|  |            |
|--|------------|
| <b>13 Solvency need resulting from provisioning risk in a ORSA context</b> | <b>261</b> |
| 13.1 Solvency 2 . . . . .  | 261        |
| 13.2 Solvency Need resulting from provisioning risk . . . . .              | 264        |
| 13.3 Technicals proofs . . . . .   | 279        |
| <b>14 Additionnal information concerning the solvency need calculation</b> | <b>289</b> |
| 14.1 Asymptotic and theoretical means . . . . .                            | 289        |
| 14.2 Empirical distribution of provision . . . . .                         | 292        |
| 14.3 Empirical solvency need . . . . .                                     | 297        |

---



# Solvency need resulting from provisioning risk in a ORSA context

---

This chapter focuses on the risk of loan default from the point of view of an insurer required to indemnify a bank for losses resulting from a borrower defaulting.

The main purpose of this chapter is to calculate the Solvency Need resulting from the provisioning risk in an ORSA (Own Risk Solvency Assessment) context imposed by Solvency 2 reform (see Section 9.1.2). For this, we will use an individual model of provision introduced in Chapter 11. The main difference between our method and existing methods is that ours takes into account, in the Solvency Need evaluation, the specific characteristics of the borrowers and the dependence between default amounts with a systemic risk factor (real estate). From a mathematics point of view we will study the convergence in law of a sum of a random number of random variables not independent and not identically distributed.

**Key words:** *Solvency 2; ORSA; Solvency Need; Provisioning risk; Quantile; Convergence in law; Perpetual integral functional of Brownian motion.*

## 13.1 Solvency 2

1) *Overview:*

As mentioned in Section 9.1.2, the purpose of Solvency 2 reform (see Directive [26]) is to better protect and better inform insurance policyholders. This reform, like in the reform process in the banking sector Bâle 2/3 (see [27]), will determine the loss associated from the different risks to which insurance companies are subjected. Indeed, article 45 of the Directive [26] specifies that:

*“As part of its risk-management system every insurance undertaking and reinsurance undertaking shall conduct its own risk and solvency assessment. That assessment shall include at least the following:*

1. *the overall solvency needs taking into account the specific risk profile, approved risk tolerance limits and the business strategy of the undertaking,*
2. *...”*

This evaluation is commonly known as the “ORSA system” (Own Risk Solvency Assessment) and is found in the second pillar of the Solvency 2 reform.

In this chapter we will focus on the evaluation of the provisioning risk, that we will call, in accordance with article 45, the *Solvency Need* (SN) resulting from the provisioning risk.

To do this, we will consider an insurance company that is required to indemnify a bank against losses related to a borrower defaulting on payments and we will calculate the SN at the drawdown date of the contracts involved.

2) *Existing methods and outline of chapter:*

The calculation of the SN requires to define a risk indicator, i.e. a risk measure, and an analysis horizon: these two elements are not defined in the Directive. The SN calculation differs from the Solvency Capital Requirement (SCR) calculation because in the SCR evaluation the risk indicator and the analysis horizon are defined. Remember that for the SCR, article 101 of Directive [26] states that:

*“It shall correspond to the Value-at-Risk of the basic own funds of an insurance or reinsurance undertaking subject to a confidence level of 99.5% over a one-year period.”*

The SN has to mitigate against the failings of the calculation methods of the SCR. As mentioned in [41], using a period of one year in the calculation of the SCR appears unsuitable for insurance companies with long maturity contracts and commitments. Unlike [58], where the SN is defined as the maximum of future losses, we will define the SN to the maximum duration of contracts (see Definition 13.2.2 and Remark 13.2.3). The risk measure considered in this chapter will be the 99.5% Value At Risk (see [24] for more details). As in [16], and to simplify calculations, we do not take into account discount effect or risk margin in this chapter (see [47] for more information on these elements).

To calculate the SN we will use an individual model of the provision which is introduced in Chapter 11 and described in Section 13.2.1. Unlike in [46], this individual model allows us to take into account the dependence between individual borrowers using a systemic risk factor (real estate), and also the dependance between default amounts and the counting process. We refer the reader to [8], [13], [15], [62] and references therein for more details concerning the consideration of some dependences.

We will see, in Section 13.2.2, that to calculate the SN, the key step is to evaluate the quantile of the provision, i.e. invert the cumulative distribution function (c.d.f.) of the provision. As calculating the provision c.d.f. is impossible, we will approximate the quantile with the asymptotic study of the provision when the maximum duration becomes high (see Theorem 13.2.5) and we will give the proof of this convergence with some interesting intermediate results in Section 13.2.3. This result of convergence will

allow us to give an approximation of the SN, see Theorem 13.2.6.

To simplify the presentation of results, notations of Chapter 11 will be lightly amended. All technical proofs will be given in Section 13.3.

3) *Principal result:*

The random variable (RV) measuring provisioning risk is defined as the difference between the provision and its mean (since it is this mean that insurance companies includes in their accounts). So the SN at date 0, concerning loans drawn down in month 0, is defined by:

$$SN(h) := \mathcal{Q}_{P(h,R) - \mathbb{E}[P(h,R)]}(99.5\%), \quad (13.1.1)$$

where  $h$  is the maximum duration of loans,  $\mathcal{Q}_{P(h,R) - \mathbb{E}[P(h,R)]}$  is the quantile of  $P(h, R) - \mathbb{E}[P(h, R)]$ , and  $\mathbb{E}[P(h, R)]$  is the mean of the RV  $P(h, R)$  which represents the provision calculated at time 0 and concerning loans drawn down in month 0. Recall that (see (13.2.1)):

$$P(h, R) := \sum_{j \geq 1} \mathbb{1}_{\{(T_j, D_j) \in A_h\}} M_j f(T_j, D_j, R_{T_j}), \quad (13.1.2)$$

with  $(T_j, D_j)_{j \geq 1}$  a Poisson point process representing the cloud of points (Term of loan, Duration of loan),  $A_h$  the set where defaults appears possible (see (13.2.3)),  $(M_j)_{j \geq 1}$  a collection of independent and identically distributed (i.i.d.) RVs representing amount borrowed, the function  $f$  defined by (13.2.4), and  $(R_t)_{t \geq 0}$  the geometric Brownian motion representing market price fluctuation and defined by (13.2.7). We will also assume that RVs  $(T_j, D_j)_{j \geq 1}$ ,  $(M_j)_{j \geq 1}$  and the process  $(R_t)_{t \geq 0}$  are independent.

We see from formula (13.1.1) that the calculation of  $SN(h)$  requires to invert the c.d.f. of  $P(h, R) - \mathbb{E}[P(h, R)]$ . In practice  $h = 276$  months, which is considered as a long period. From a mathematical point of view we are in an asymptotic behaviour, i.e.  $h \rightarrow +\infty$ . In Theorem 13.2.5 we will show that the provision correctly normalized converges in law to a mix of Gaussian and the integral functional of a geometric Brownian motion, i.e.:

$$\frac{P(h, R) - \mathbb{E}[P(h, R)]}{h\theta_1} \xrightarrow[h \rightarrow +\infty]{\mathcal{L}} G - \int_0^{+\infty} \inf(\theta_2, \theta_3 R_t) dt + \theta_4, \quad (13.1.3)$$

where  $\theta_1, \theta_2, \theta_3$  and  $\theta_4$  are constants which are defined with parameters of our model and  $G$  is a Gaussian centered with variance equal to one, and independent of  $(R_t)_{t \geq 0}$ . Consequently we will approximate  $SN(h)$  by the quantity  $SN^*(h)$  (see Theorem 13.2.6 for more details) defined by:

$$SN^*(h) := q_{99.5\%} h \theta_1, \quad (13.1.4)$$

where  $q_{99.5\%}$  is the 99.5% quantile of:

$$G - \int_0^{+\infty} \inf(\theta_2, \theta_3 R_t) dt + \theta_4. \quad (13.1.5)$$

Formula (13.1.4) is only theoretical. Indeed the law of  $\int_0^{+\infty} \inf(\theta_2, \theta_3 R_t) dt$  is unknown. However we will discuss in Section 14.3 of the possibility to simulate the provision and then deduce a numerical value of the SN.

Note that if there is no correlation between the claim amount in each claim, the left term in (13.1.3) would converge to a Gaussian RV. The additional term which is written with the integral function of a geometric Brownian motion takes any correlations into account. So there is a “separation of risk sources”, one coming from the personal characteristics of borrowers who are independent and an other coming from the cost of real estate which affects all borrowers.

## 13.2 Solvency Need resulting from provisioning risk

In this chapter we will assume a continuous time. The time  $t$  is defined as a fraction of year and  $t^*$  is one month.

In Section 13.2.1 we will start by explaining the context of provisioning related to a borrower defaulting on payments presented in Chapter 11. We will also explain the main assumptions and notations linked to this model. Section 13.2.2 will be devoted to the calculation of the SN, i.e. obtaining formula (13.1.4). Section 13.2.3 will be devoted to the proof of result (13.1.3).

### 13.2.1 Provisioning against borrowers default risk

An insurance company is required to indemnify a bank against losses related to a borrower defaulting on payments. The loans under consideration are real estate loans. The main issue for the insurance company is to calculate, at date  $t_0 = 0$ , potential future claim amounts from borrowers whose loans are drawn down in month 0. This amount is the “provision”. It is essential to note that in the event of default, the bank can try to repossess the property and sell it.

In Chapter 11 the provision calculated at time 0 and concerning loans drawn down in month 0 is denoted by  $P(h, R)$  and defined by (we will amended the notations):

$$P(h, R) := \sum_{j \geq 1} \mathbf{1}_{\{(T_j, D_j) \in A_h\}} M_j f(T_j, D_j, R_{T_j}), \quad (13.2.1)$$

where:

- $D_j$  (resp.  $M_j$ ) is the duration (resp. the amount) of the loan drawn down in month 0 by the borrower  $j$ .  $h$  is the maximum duration (unknown) of loans:

$$D_j \leq h, \quad j \geq 1. \quad (13.2.2)$$

- $T_j$  represents the maturity of the loan of borrower  $j$ , i.e. the real duration. It is important to distinguish two different possibilities:

1.  $T_j = D_j$  means that the borrower repays his loan normally until maturity,
2.  $T_j < D_j$  means the borrower defaults before maturity.

$R := (R_t, t \geq 0)$  is the market price fluctuation between date 0 and  $t$ . The set  $A_h$  is defined by:

$$A_h := \{(d, t) \in (\mathbb{R}_+)^2 : 0 \leq t < d; d \leq h\}. \quad (13.2.3)$$

This is an illustration:

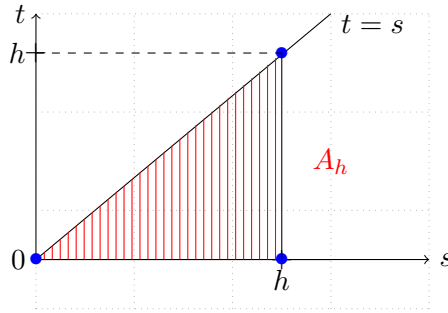


Illustration 13.1: Set  $A_h$ .

- $M_j f(T_j, D_j, R_{T_j})$  is the amount that the insurance company will repay to the bank in the event of default concerning borrower  $j$  at time  $T_j$ . Function  $f$  is defined by:

$$f(t, d, u) := [\phi_t(d) - \gamma u]_+, \quad (13.2.4)$$

where  $t, d, u > 0$ ,  $\gamma \in ]0, 1[$  and:

$$\phi_t(s) := \begin{cases} \alpha, & \text{if } t \leq s \leq t^*, \\ \alpha, & \text{if } t \leq t^* \leq s, \\ \alpha \frac{e^{\tau s} - e^{\tau(t-t^*)}}{e^{\tau s} - 1}, & \text{if } 0 < t^* < t \leq s, \end{cases} \quad (13.2.5)$$

$$\alpha := e^{\tau t^*} (1 + t^* p) - t^* p, \quad (13.2.6)$$

with  $p \in \mathbb{R}_+$  represents coefficient due to late payment interest, and  $\tau \in \mathbb{R}_+$  is the constant rate of return.

Note that from (13.2.1) and (13.2.4),  $M_j \phi_{T_j}(D_j)$  represents the outstanding amount owed by the borrower  $j$  at date  $T_j$  and  $\gamma M_j R_{T_j}$  is the property sale amount at date  $T_j$ . The claim amount cannot be negative. When the recovery amount is higher than the outstanding amount owed by the borrower, then the insurance company will not pay anything to the bank (because in reality it is the bank that collects the recovery amount). The claim amount does not represent the actual amount of loss but only the amount that the insurance company will repay to the bank in the event of default.

Recall the main assumptions presented in Chapter 11:

**Assumption 13.2.1.**

(A<sub>5</sub>):  $(R_t)_{t \geq 0}$  is the geometric Brownian motion with drift  $\mu$  and volatility  $\sigma$  so that  $R_0 = 1$ , i.e.:

$$R_t := \exp [\sigma B_t + \mu t], \quad (13.2.7)$$

where  $(B_t)_{t \geq 0}$  is a standard Brownian motion starting from 0.

(A<sub>6</sub>): The clouds of points  $(T_j, D_j)_{j \geq 1}$  is a Poisson point process with intensity  $\Lambda$ :

$$\Lambda := \lambda_1 \text{Leb}_{|C_h} + \lambda_2 \text{Leb}_{|A_h}, \quad (13.2.8)$$

where  $(\lambda_1, \lambda_2) \in (\mathbb{R}_+)^2$ ,  $C_h := \{(t, t) \in (\mathbb{R}_+)^2\}$ , and  $\text{Leb}_{|C_h}$  (resp.  $\text{Leb}_{|A_h}$ ) is the Lebesgue measure restricted on  $C_h$  (resp.  $A_h$ ).

(A<sub>8</sub>): The RVs  $(T_j, D_j)_{j \geq 1}, (M_j)_{j \geq 1}$  and the process  $(R_t)_{t \geq 0}$  are independent, and the RVs  $(M_j)_{j \geq 1}$  are i.i.d..

We will also assume that:

$$(A_9) : \quad \mathbb{E}[(M_1)^4] < +\infty \text{ et } M_1 > 0 \text{ a.s.} \quad (13.2.9)$$

Let us now move on the calculation of the SN.

### 13.2.2 Calculation of the Solvency Need

As mentioned in Section 13.1, the SN corresponds to the amount that an insurance company needs to cover the risk of under-provisioning over the maximum duration of the loans.

In this section, we firstly define the RV measuring the provisioning risk (see Definition 13.2.2). In Remark 13.2.3 we will speak about the model proposed by Wüthrich et al. in [65]. In this paper they are interested in two risk indicators: the variance and the conditional mean square error. These two indicators have been also studied in [16]. However we will use, as in [41] and [47], the Value At Risk as the risk indicator to define the SN. Then we will give the main theoretical result of the chapter (see Theorem 13.2.5), i.e. the convergence of:

$$\frac{P(h, R) - \mathbb{E}[P(h, R)]}{h}, \quad (13.2.10)$$

when  $h \rightarrow +\infty$ . Finally this theoretical result will allow us to estimate the SN, see Theorem 13.2.6. We conclude this section with some remarks concerning the link with other model or the influence of parameter  $\mu$ , see Remarks 13.2.7, 13.2.8, 13.2.9, 13.2.10, and 13.2.11.



**Definition 13.2.2.**

1. The RV measuring the provisioning risk for month 0 is defined by:

$$P(h, R) - \mathbb{E}[P(h, R)], \quad (13.2.11)$$

where  $P(h, R)$  is defined by (13.2.1).

2. The SN resulting from provisioning risk for month 0, denoted by  $SN(h)$ , is the 99.5% quantile (or Value At Risk) of  $P(h, R) - \mathbb{E}[P(h, R)]$ , i.e.:

$$SN(h) := \mathcal{Q}_{P(h, R) - \mathbb{E}[P(h, R)]}(99.5\%), \quad (13.2.12)$$

where  $\mathcal{Q}_X$  is the quantile of the RV  $X$  at 99.5%, i.e.:

$$\mathbb{P}[X \leq \mathcal{Q}_X(99.5\%)] = 99.5\%. \quad (13.2.13)$$

**Remark 13.2.3.**

1. Note that if  $c \in \mathbb{R}$  and  $Y$  is a RV then:

$$\mathcal{Q}_{Y+c}(99.5\%) = \mathcal{Q}_Y(99.5\%) + c. \quad (13.2.14)$$

2. In Section 2.3 of [65], the RV measuring the provisioning risk over one year and calculated at date 0 is denoted by CDR (Claim Development Result) and defined by (we use our notations):

$$\begin{aligned} CDR_1 := & \mathbb{E}[P(h, R)|\mathcal{F}_0] - \left\{ \sum_{j \geq 1} \mathbb{1}_{\{(T_j, D_j) \in A_h, T_j \leq 1\}} M_j f(T_j, D_j, R_{T_j}) \right. \\ & \left. + \mathbb{E} \left[ \sum_{j \geq 1} \mathbb{1}_{\{(T_j, D_j) \in A_h, T_j > 1\}} M_j f(T_j, D_j, R_{T_j}) \middle| \mathcal{F}_1 \right] \right\}, \end{aligned} \quad (13.2.15)$$

Sums  $\sum_{j \geq 1} \mathbb{1}_{\{(T_j, D_j) \in A_h, T_j \leq 1\}} M_j f(T_j, D_j, R_{T_j})$  and  $\sum_{j \geq 1} \mathbb{1}_{\{(T_j, D_j) \in A_h, T_j > 1\}} M_j f(T_j, D_j, R_{T_j})$  are provisions for the period of time  $[0, 1]$  and  $]1, h]$ ,  $\mathcal{F}_0$  and  $\mathcal{F}_1$  are the “information available” at dates 0 and 1, i.e.:

$$\mathcal{F}_t := \sigma \left[ (D_j)_{j \geq 1}, (M_j)_{j \geq 1}, (R_s)_{s \leq t}, (T_j \wedge t)_{j \geq 1} \right]. \quad (13.2.16)$$

As mentioned in their paper, Wuthrich et al. are interested in a “short term” risk analysis as imposed in the first pillar of the Solvency 2 reform. However we are interested, in this paper, in a “long-term” risk analysis. If we take the Definition (13.2.15) proposed by Wuthrich et al. not to one year but to maximum duration of loans,  $h$ , we obtain:

$$CDR_h := - \left[ P(h, R) - \mathbb{E}[P(h, R)|\mathcal{F}_0] \right]. \quad (13.2.17)$$

So our definition of RV measuring provisioning risk is similar to the one proposed by Wuthrich et al. with some adjustments:

- (a) We do not take into account information available at date 0,
- (b) We use an individual model of provision. However Wutrich et al. use a collective model of provision (Chain Ladder).
- (c) The analysis horizon is the maximum duration of loans.

$\mathbb{E}[P(h, R)]$  is known (see Theorem 11.3.2 of Chapter 11). But the law of  $P(h, R)$  is unknown but it can be simulated. In Theorem (13.2.5) we will prove that the quantity defined by (13.2.10) converges in law when  $\mu < 0$ . Consequently we will approximate  $SN(h)$  by the quantity  $SN^*(h)$  defined by:

$$SN^*(h) := hQ_X(99.5\%), \quad (13.2.18)$$

where  $X$  is the limit of  $\frac{P(h, R) - \mathbb{E}[P(h, R)]}{h}$  when  $h \rightarrow +\infty$  and  $Q_X$  is defined by (13.2.13).

To calculate  $SN^*(h)$  it is convenient to adopt the following notations:

**Notation 13.2.4.** We denote:

$$G \text{ a Gaussian RV centered with unit variance, and independent of } (R_t)_{t \geq 0}, \quad (13.2.19)$$

$$C_k := \lambda_2 \frac{\alpha^k}{2} \mathbb{E}[(M_1)^k], k \geq 1, \quad (13.2.20)$$

$$X^* := G - \frac{2C_1}{\sqrt{C_2}} \left[ \frac{e^{-\tau t^*}}{\tau} + \int_0^{+\infty} \inf(1, \frac{\gamma}{\alpha} R_t) dt \right], \quad (13.2.21)$$

$$\kappa := \frac{\lambda_2 \mathbb{E}[M_1]}{\sqrt{C_2}} \int_0^{+\infty} \left( \alpha \Phi \left[ \frac{\mu t - \ln \left[ \frac{\alpha}{\gamma} \right]}{\sigma \sqrt{t}} \right] + \gamma e^{(\mu + \frac{\sigma^2}{2})t} \Phi \left[ \frac{\ln \left[ \frac{\alpha}{\gamma} \right] - (\mu + \sigma^2)t}{\sigma \sqrt{t}} \right] \right) dt, \quad (13.2.22)$$

$$\eta := \lambda_2 \mathbb{E}[M_1] \int_0^{+\infty} \left\{ \alpha \Phi \left[ \frac{\ln \left[ \frac{\alpha}{\gamma} \right] - \mu t}{\sigma \sqrt{t}} \right] - \gamma e^{(\mu + \frac{\sigma^2}{2})t} \Phi \left[ \frac{\ln \left[ \frac{\alpha}{\gamma} \right] - (\mu + \sigma^2)t}{\sigma \sqrt{t}} \right] \right\} dt, \quad (13.2.23)$$

where  $\Phi$  is the c.d.f. of  $G$ ,  $\mathbb{E}[(M_1)^k]$  is the  $k$  moment of  $M_1$ ,  $\lambda_2$ ,  $\alpha$ ,  $\tau$ ,  $\gamma$ ,  $\mu$  and  $\sigma$  are defined by (13.2.8), (13.2.6), (11.2.13), (13.2.4), (13.2.7), and (13.2.7).

**Theorem 13.2.5.** Let assumptions (A<sub>1</sub>) to (A<sub>9</sub>) hold.

1. The asymptotic behaviour of the provision mean is the following:

$$\lim_{h \rightarrow +\infty} \frac{\mathbb{E}[P(h, R)]}{h^2} = \lambda_2 \alpha \mathbb{E}[M_1] \left( \frac{1}{2} \mathbf{1}_{\{\mu < 0\}} + \frac{1}{4} \mathbf{1}_{\{\mu = 0\}} \right). \quad (13.2.24)$$

In addition, when  $\mu > 0$  we have:

$$\lim_{h \rightarrow +\infty} \frac{\mathbb{E}[P(h, R)]}{h} = \eta, \quad (13.2.25)$$

where  $\eta$  is defined by (13.2.23).

2. If  $\mu < 0$  then the provision normalized converges in law:

$$\frac{P(h, R) - h^2 C_1}{h\sqrt{C_2}} \xrightarrow[h \rightarrow +\infty]{\mathcal{L}} X^*, \quad (13.2.26)$$

$$\frac{P(h, R) - \mathbb{E}[P(h, R)]}{h\sqrt{C_2}} \xrightarrow[h \rightarrow +\infty]{\mathcal{L}} X^* + \kappa, \quad (13.2.27)$$

where  $C_k$ ,  $X^*$ , and  $\kappa$  are defined by (13.2.20), (13.2.21), and (13.2.22).

Proof of Theorem 13.2.5 is given in Section 13.2.3: this allows us to show some interesting intermediate results. We will also give comments related to Theorem 13.2.5, see Remarks 13.2.8, 13.2.9, 13.2.10, and 13.2.11, but now we want to start by stating the main consequence of Theorem 13.2.5, i.e. the calculation of  $SN^*(h)$ .

**Theorem 13.2.6.** *Let assumptions  $(A_1)$  to  $(A_9)$  hold. If  $\mu < 0$  then:*

$$BS^*(h) = h \left[ \kappa + \mathcal{Q}_{X^*}(99.5\%) \right] \sqrt{\frac{\lambda_2 \alpha^2 \mathbb{E}[(M_1)^2]}{2}}, \quad (13.2.28)$$

where  $\mathcal{Q}_{X^*}(99.5\%)$  is defined by (13.2.13) and the constant  $\kappa$  is defined by (13.2.22).

*Proof.* It is a direct consequence of the definition of  $SN^*(h)$ , see (13.2.18), and Theorem 13.2.5, see (13.2.27), and (13.2.14).  $\square$

**Remark 13.2.7.**

1. In Chapter 11 we had proposed a methodology to estimate parameters with real data. This methodology gives to  $\mu$  a negative value it is why we had only considered the case  $\mu < 0$  in the previous theorem.
2. Formula (13.2.28) is only theoretical. Indeed, the law of:

$$\int_0^{+\infty} \inf \left( 1, \frac{\gamma}{\alpha} R_t \right) dt, \quad (13.2.29)$$

is unknown. However we had seen in Section 12.4.2 that it is possible to simulate the provision: we will discuss in Section 14.3 of the possibility to evaluate the solvency need with these simulations.

Now we do some remarks concerning Theorem 13.2.5.

**Remark 13.2.8.**

It is possible to give another expression of the RV  $X^*$  defined by (13.2.21). If we remark that:

$$\inf(1, x) = x - (x - 1)_+, \quad (13.2.30)$$

it comes that:

$$\int_0^{+\infty} \inf \left( 1, \frac{\gamma}{\alpha} R_t \right) dt = \int_0^{+\infty} \frac{\gamma}{\alpha} R_t dt - \int_0^{+\infty} \left[ \frac{\gamma}{\alpha} R_t - 1 \right]_+ dt. \quad (13.2.31)$$

In [25] it is shown that:

$$\int_0^{+\infty} R_t dt \stackrel{(d)}{=} \frac{2}{\sigma^2 \Gamma}, \quad (13.2.32)$$

where  $\Gamma$  is a Gamma RV with parameters  $-\frac{2\mu}{\sigma^2}$ . So by combining (13.2.32), (13.2.31), and (13.2.21) it comes:

$$X^* \stackrel{(d)}{=} G - \frac{2C_1}{\sqrt{C_2}} \left[ \frac{e^{-\tau t^*}}{\tau} + \frac{2\gamma}{\alpha \sigma^2 \Gamma} - \int_0^{+\infty} \left[ \frac{\gamma}{\alpha} R_t - 1 \right]_+ dt \right]. \quad (13.2.33)$$

However the RV  $\Gamma$  in (13.2.33) is not independent of  $\int_0^{+\infty} \left[ \frac{\gamma}{\alpha} R_t - 1 \right]_+ dt$ . Integral as (13.2.32) are called “perpetual integral functionals of Brownian motion” (see [11], [25], [53], [66], and [67] for more details).

**Remark 13.2.9.**

The sign of  $\mu$  plays an important role in the asymptotic behaviour of the provision.

1. If  $\mu < 0$  then  $R_h \xrightarrow{h \rightarrow +\infty} 0$ . It means that we consider no property resale and leads to the insurance company being more cautious. This is the most interesting case for the insurance company.
2. If  $\mu > 0$  then  $R_h \xrightarrow{h \rightarrow +\infty} +\infty$ . So  $\mu > 0$  means that we consider an “exponentially” increase of the real estate market. This case is not interesting in a practical point of view.
3.  $\mu = 0$  means that there is only fluctuations on the property value. In this case we are not able to calculate the asymptotic law of  $P(h, R) - \mathbb{E}[P(h, R)]$ .

**Remark 13.2.10.**

1. The renormalized factor  $h\sqrt{C_2}$  (see (13.2.26) and (13.2.27)) is not equal to the variance of the provision as in the usual central limit theorem.
2. Note that if there is no correlation between claim amounts, RV  $X^*$  (see (13.2.21)) will be equal to a Gaussian centered with variance equal to one. The additional term in the RV  $X^*$  which is written with an integral of a Brownian motion:

$$\frac{2C_1}{\sqrt{C_2}} \int_0^{+\infty} \inf \left( 1, \frac{\gamma}{\alpha} R_t \right) dt, \quad (13.2.34)$$

takes into account the correlation between the claim amounts. The term:

$$\frac{2C_1}{\sqrt{C_2}} \frac{e^{-\tau t^*}}{\tau}, \quad (13.2.35)$$

represents the error of renormalization. As mentioned in the introduction there is a separation of “risk sources”, one coming from the personal characteristics of the borrowers and the other coming from the real estate price fluctuations which are the same for each borrower.

**Remark 13.2.11.** *Comparison with models which take into account some dependencies. Recall two PPP model with dependences:*

1. *U-statistics of PPP are defined by:*

$$\sum_{(X_1, \dots, X_k) \in \eta_{\neq}^k} f(X_1, \dots, X_k), \quad (13.2.36)$$

where  $f$  is a symmetric function, the set of  $(X_i)_{i \geq 1}$  is a PPP, and  $\eta_{\neq}^k$  is the set of all  $k$ -tuples of distinct points of the process.

2. *The “shot noise” processes are defined by:*

$$\sum_{j=1}^{N_h} f(T_j, \xi_j), \quad 0 < t < +\infty,$$

where  $f$  is an explicit measurable function,  $(T_j)_{j \geq 1}$  are jump times of the counting process  $(N_s)_{s \geq 0}$ ,  $(\xi_j)_{j \geq 1}$  is a collection of RVs i.i.d., and independent of  $(N_s)_{s \geq 0}$ .

For this kind of model it is possible to show a central limit theorem (see [49] and [50]). However these model are not able to take into account the process  $(R_t)_{t \geq 0}$ .

### 13.2.3 Proof of Theorem 13.2.5

#### Proof of item 1 of Theorem 13.2.5:

Firstly we prove that the mean of the provision converges to a constant which depends on  $\mu$  as  $h \rightarrow +\infty$ . Recall that  $i = 0$  and:

$$\Lambda(dt ds) = \lambda_2 dt ds, \quad (13.2.37)$$

where  $\Lambda$  and  $\lambda_2$  are defined by (13.2.8). Using Theorem 11.3.2 of Chapter 11 and the changes of variables  $s = hy$  and  $t = hx$  we get:

$$\mathbb{E}[P(h, R)] = \lambda_2 \mathbb{E}[M_1] \int_0^h ds \left[ \int_0^s \{g_1(t, s) - g_2(t, s)\} dt \right] \quad (13.2.38)$$

$$= h^2 \lambda_2 \mathbb{E}[M_1] \int_0^1 dy \left[ \int_0^y \{g_1(hx, hy) - g_2(hx, hy)\} dx \right], \quad (13.2.39)$$

where:

$$g_1(u, v) := \phi_u(v) \Phi \left[ \frac{\ln \left[ \frac{\phi_u(v)}{\gamma} \right] - \mu u}{\sigma \sqrt{u}} \right], \quad u < v, \quad (13.2.40)$$

$$g_2(u, v) := \gamma e^{(\mu + \frac{\sigma^2}{2})u} \Phi \left[ \frac{\ln \left[ \frac{\phi_u(v)}{\gamma} \right] - (\mu + \sigma^2)u}{\sigma \sqrt{u}} \right], \quad u < v, \quad (13.2.41)$$

with  $\Phi$  the c.d.f. of a standard Gaussian RV and  $\phi$  defined by (11.2.13). Note the condition  $0 < x < y$  implies:

$$\lim_{h \rightarrow +\infty} \phi_{hx}(hy) = \alpha, \quad (13.2.42)$$

with  $\alpha$  defined by (13.2.6). So we have:

$$\lim_{h \rightarrow +\infty} g_1(hx, hy) = \alpha \left( \mathbb{1}_{\{\mu < 0\}} + \frac{1}{2} \mathbb{1}_{\{\mu = 0\}} \right). \quad (13.2.43)$$

To study the asymptotical behaviour of  $g_2(hx, hy)$  we will distinguish two cases:

1. If  $\mu + \sigma^2 \leq 0$ , then  $\mu + \frac{\sigma^2}{2} < 0$ . Because  $\Phi(z) \leq 1$  it follows:

$$\lim_{h \rightarrow +\infty} g_2(hx, hy) = 0. \quad (13.2.44)$$

2. If  $\mu + \sigma^2 > 0$ , recall that:

$$\Phi(z) \sim \frac{1}{\sqrt{2\pi}|z|} e^{-\frac{z^2}{2}}, \quad z \rightarrow -\infty. \quad (13.2.45)$$

So when  $h \rightarrow +\infty$  we have:

$$\begin{aligned} g_2(hx, hy) &\sim e^{(\mu + \frac{\sigma^2}{2})hx} \frac{\gamma}{\sqrt{2\pi}} \frac{\sigma\sqrt{hx}}{|\ln[\frac{\phi_{hx}(hy)}{\gamma}] - (\mu + \sigma^2)hx|} \\ &\times \exp \left[ -\frac{1}{2\sigma^2 hx} \left\{ \ln[\frac{\phi_{hx}(hy)}{\gamma}] - (\mu + \sigma^2)hx \right\}^2 \right] \\ &\sim \frac{\sigma\gamma}{\sqrt{2\pi}(\mu + \sigma^2)\sqrt{hx}} \exp \left[ \frac{(\mu + \sigma^2) \ln[\frac{\phi_{hx}(hy)}{\gamma}]}{\sigma^2} - \frac{\left( \ln[\frac{\phi_{hx}(hy)}{\gamma}] \right)^2}{2\sigma^2 hx} \right. \\ &\quad \left. + \left[ \left( \mu + \frac{\sigma^2}{2} \right) - \frac{(\mu + \sigma^2)^2}{2\sigma^2} \right] hx \right] \\ &\sim \frac{\sigma\gamma}{\sqrt{2\pi}(\mu + \sigma^2)\sqrt{hx}} \exp \left[ \frac{(\mu + \sigma^2) \ln(\frac{\alpha}{\gamma})}{\sigma^2} \right] \exp \left[ -\frac{\mu^2}{2\sigma^2} hx \right]. \end{aligned}$$

So:

$$\lim_{h \rightarrow +\infty} g_2(hx, hy) = 0. \quad (13.2.46)$$

So combining (13.2.44) and (13.2.46) we have:

$$\lim_{h \rightarrow +\infty} g_2(hx, hy) = 0. \quad (13.2.47)$$

An application of Lemma 11.5.1 gives:

$$\mathbb{E}[f(hx, hy, R_{hx})] = g_1(hx, hy) - g_2(hx, hy), \quad (13.2.48)$$

so it follows from formula (13.3.6) of Lemma 13.3.2:

$$0 \leq g_1(hx, hy) - g_2(hx, hy) \leq \alpha, \forall h > 0, \forall 0 < x < y. \quad (13.2.49)$$

Consequently using (13.2.39), (13.2.43), (13.2.47), (13.2.49) and the dominated convergence theorem we have:

$$\lim_{h \rightarrow +\infty} \frac{\mathbb{E}[P(h, R)]}{h^2} = \lambda_2 \alpha \mathbb{E}[M_1] \left( \frac{1}{2} \mathbb{1}_{\{\mu < 0\}} + \frac{1}{4} \mathbb{1}_{\{\mu = 0\}} \right). \quad (13.2.50)$$

This established formula (13.2.24).

Now we want to study the limit of  $\frac{\mathbb{E}[P(h, R)]}{h}$  when  $\mu > 0$  and  $h \rightarrow +\infty$ . We start with a lemma useful for the next:

**Lemme 13.2.12.** *Let  $a \in \mathbb{R}$ ,  $b \in \mathbb{R}$ , and  $c > 0$ . Then:*

$$\int_0^{+\infty} \Phi \left[ \frac{a - bt}{\sqrt{t}} \right] dt < +\infty, \quad b > 0, \quad (13.2.51)$$

$$\int_0^{+\infty} e^{(b + \frac{c^2}{2})t} \Phi \left[ \frac{a - (b + c^2)t}{c\sqrt{t}} \right] dt < +\infty, \quad (13.2.52)$$

where  $\Phi$  is the c.d.f. of a Gaussian centered with variance equal to one.

The proof of this lemma will be made in Section 13.3.1. Let us come back to study of the limit of  $\frac{\mathbb{E}[P(h, R)]}{h}$ . Using (13.2.38) and the change of variable  $s = hy$  we obtain:

$$\begin{aligned} \Gamma_h &:= \frac{\mathbb{E}[P(h, R)]}{h \lambda_2 \mathbb{E}[M_1]} = \frac{1}{h} \int_0^h ds \left[ \int_0^s \{g_1(t, s) - g_2(t, s)\} dt \right] \\ &\stackrel{s=hy}{=} \int_0^1 dy \left[ \int_0^{hy} \{g_1(t, hy) - g_2(t, hy)\} dt \right], \end{aligned} \quad (13.2.53)$$

where functions  $g_1$  and  $g_2$  are defined by (13.2.40) and (13.2.41) respectively. Because  $t \leq hy$  it follows:

$$\lim_{h \rightarrow +\infty} \phi_t(hx) = \alpha, \quad t > 0, x > 0. \quad (13.2.54)$$

So:

$$\lim_{h \rightarrow +\infty} g_1(t, hy) = \alpha \Phi \left[ \frac{\ln \left[ \frac{\alpha}{\gamma} \right] - \mu t}{\sigma \sqrt{t}} \right], \quad (13.2.55)$$

$$\lim_{h \rightarrow +\infty} g_2(t, hy) = \gamma e^{(\mu + \frac{\sigma^2}{2})t} \Phi \left[ \frac{\ln \left[ \frac{\alpha}{\gamma} \right] - (\mu + \sigma^2)t}{\sigma \sqrt{t}} \right]. \quad (13.2.56)$$

From (13.3.6) we have:

$$\phi_t(hy) \leq \alpha. \quad (13.2.57)$$

So:

$$0 \leq g_1(t, hy) \leq \alpha \Phi \left[ \frac{\ln \left( \frac{\alpha}{\gamma} \right) - \mu t}{\sigma \sqrt{t}} \right], \quad (13.2.58)$$

$$0 \leq g_2(t, hy) \leq \gamma e^{(\mu + \frac{\sigma^2}{2})t} \Phi \left[ \frac{\ln(\frac{\alpha}{\gamma}) - (\mu + \sigma^2)t}{\sigma\sqrt{t}} \right]. \quad (13.2.59)$$

Combining (13.2.53), (13.2.55), (13.2.56), (13.2.58), (13.2.59), (13.2.51), and (13.2.52) it follows:

$$\lim_{h \rightarrow +\infty} \Gamma_h = \int_0^{+\infty} \left\{ \alpha \Phi \left[ \frac{\ln \left[ \frac{\alpha}{\gamma} \right] - \mu t}{\sigma\sqrt{t}} \right] - \gamma e^{(\mu + \frac{\sigma^2}{2})t} \Phi \left[ \frac{\ln \left[ \frac{\alpha}{\gamma} \right] - (\mu + \sigma^2)t}{\sigma\sqrt{t}} \right] \right\} dt.$$

This established (13.2.25).

**Proof of formula (13.2.26) of Theorem 13.2.5:**

We introduce some notations.

**Notation 13.2.13.**

$$X_k(h, r) := \lambda_2 \mathbb{E} \left[ (M_1)^k \right] \int_{A_h} [f(t, s, r_t)]^k dt ds, k \geq 1, \quad (13.2.60)$$

$$P(h, r) := \sum_{j \geq 1} \mathbb{1}_{\{(T_j, D_j) \in A_h\}} M_j f(T_j, D_j, r_{T_j}), \quad (13.2.61)$$

$$r : \mathbb{R}_+ \longrightarrow \mathbb{R}_+ \\ t \longmapsto r_t, \quad (13.2.62)$$

$$r := (r_t, t \geq 0), \quad (13.2.63)$$

where  $A_h$  and  $f$  are defined by (13.2.3) and (13.2.4) respectively. Note that  $X_k(h, R)$  is a RV which is  $\sigma[R_t, t \leq h]$  measurable.

We will study the convergence in law of the provision given that  $(R_t = r_t, t \geq 0)$ . This mean that we will consider the asymptotic behaviour of  $P(h, r)$  defined by (13.2.61). We will prove (see Proposition 13.2.14) that this variable, correctly renormalized, converges in law to a Gaussian RV when  $h \rightarrow +\infty$ . In a second step we will study the asymptotic behaviour of the renormalization factors, i.e.  $\frac{X_k(h, R)}{h^2}$ , according to the value of  $\mu$  and we will prove that  $\frac{X_k(h, R)}{h^2}$  converges in probability to a constant when  $\mu < 0$ , see (13.2.82). Finally we will prove that  $\left( \frac{X_1(h, R)}{h^2} - C_1 \right) h$  converges almost surely to the integral of a function of a geometric Brownian motion when  $\mu < 0$ , see (13.2.86).

These results will allow us to prove that the provision, correctly renormalized, converges in law to a mix of Gaussian RV and the integral of a functional of geometric Brownian motion.

We start by the asymptotical study of the provision  $P(h, r)$ . To do this we need to



introduce the Wasserstein distance, denoted by  $d_W$ , and defined for two RVs  $X$  and  $Y$  by:

$$d_W(X, Y) := \sup_{\xi \in Lip(1)} \left| \mathbb{E}[\xi(X)] - \mathbb{E}[\xi(Y)] \right|, \quad (13.2.64)$$

where  $Lip(1)$  is the set of all function  $\xi : \mathbb{R} \rightarrow \mathbb{R}$  which are 1-lipschitziennes, i.e. verifying:

$$|\xi(x) - \xi(y)| \leq |x - y|, \quad \forall x, y \in \mathbb{R}. \quad (13.2.65)$$

**Proposition 13.2.14.** *Let assumptions  $(A_1)$  to  $(A_9)$  hold.*

1. We have:

$$d_W \left[ \frac{P(h, r) - \mathbb{E}[P(h, r)]}{\sqrt{\mathbb{V}[P(h, r)]}}, G \right] \leq \frac{\sqrt{X_4(h, r)}}{X_2(h, r)}, \quad (13.2.66)$$

where  $X_k(h, r)$  is defined by (13.2.60).

2. In addition:

$$\mathbb{E}[P(h, r)] = X_1(h, r) \quad (13.2.67)$$

$$\mathbb{V}[P(h, r)] = X_2(h, r). \quad (13.2.68)$$

3. If the function  $t \rightarrow r_t$  verify:

$$\lim_{t \rightarrow +\infty} r_t = 0, \quad (13.2.69)$$

then:

$$\lim_{h \rightarrow +\infty} \frac{\sqrt{X_4(h, r)}}{X_2(h, r)} = 0. \quad (13.2.70)$$

4. So:

$$\frac{P(h, r) - \mathbb{E}[P(h, r)]}{\sqrt{\mathbb{V}[P(h, r)]}} \xrightarrow[h \rightarrow +\infty]{\mathcal{L}} G. \quad (13.2.71)$$

The proof of this proposition will be made in Section 13.3.2.

**Remark 13.2.15.** *The Kolmogorow distance between two RVs  $X$  and  $Y$  is denoted by  $d_{Kol}$  and defined by:*

$$d_{Kol}(X, Y) := \sup_{z \in \mathbb{R}} \left| \mathbb{P}[X \leq z] - \mathbb{P}[Y \leq z] \right|. \quad (13.2.72)$$

In [39] it is recalled that if  $G$  is a Gaussian RV centered with variance equal to one then:

$$d_{Kol}(X, G) \leq 2\sqrt{d_W(X, G)}. \quad (13.2.73)$$

By combining (13.2.66) and (13.2.73) it follows:

$$\left| \mathbb{P} \left[ \frac{P(h, r) - \mathbb{E}[P(h, r)]}{\sqrt{\mathbb{V}[P(h, r)]}} \leq x \right] - \mathbb{P}[G \leq x] \right| \leq 2\sqrt{\frac{\sqrt{X_4(h, r)}}{X_2(h, r)}}, \quad x > 0. \quad (13.2.74)$$

And from (13.3.13) it follows:

$$\left| \mathbb{P} \left[ \frac{P(h, r) - \mathbb{E}[P(h, r)]}{\sqrt{\mathbb{V}[P(h, r)]}} \leq x \right] - \mathbb{P}[G \leq x] \right| \leq 2 \left[ \frac{\alpha^2 \mathbb{E}[(M_1)^4]}{\lambda_2 [\mathbb{E}[(M_1)^2]]^2 \int_{A_h} [f(t, s, r_t)]^2 dt ds} \right]^{\frac{1}{4}}, \quad (13.2.75)$$

From (13.3.22) we have:

$$\begin{aligned} \int_{A_h} [f(t, s, r_t)]^2 dt ds &\geq h \int_0^1 dy \left[ \int_0^{+\infty} \mathbb{1}_{t \leq hy} \left( [\alpha(1 - e^{-\tau t^*}) - r_t]_+ \right)^2 dt \right] \\ &\geq h \int_0^{+\infty} \mathbb{1}_{\frac{t}{h} \leq 1} \left( [\alpha(1 - e^{-\tau t^*}) - r_t]_+ \right)^2 \left(1 - \frac{t}{h}\right) dt. \end{aligned} \quad (13.2.76)$$

From (13.2.69) :  $\exists \tilde{t} : \forall t \geq \tilde{t} : r_t < \alpha(1 - e^{-\tau t^*})$ . So:

$$\int_{A_h} [f(t, s, r_t)]^2 dt ds \geq h \int_{\tilde{t}}^{+\infty} \mathbb{1}_{\frac{t}{h} \leq 1} \left( [\alpha(1 - e^{-\tau t^*}) - r_t]_+ \right)^2 \left(1 - \frac{t}{h}\right) dt. \quad (13.2.77)$$

Let  $\tilde{t}'$  such as:  $\forall t > \tilde{t}' : r_t < \frac{\alpha}{2}(1 - e^{-\tau t^*})$  and:

$$\tilde{C} := \frac{\alpha}{2}(1 - e^{-\tau t^*}). \quad (13.2.78)$$

For  $h > 2\tilde{t}'$  it comes:

$$\begin{aligned} \int_{A_h} [f(t, s, r_t)]^2 dt ds &\geq h\tilde{C} \int_{\tilde{t}'}^h \left(1 - \frac{t}{h}\right) dt \\ &\geq h^2 \tilde{C} \int_{\frac{\tilde{t}'}{h}}^1 (1 - u) du \\ &\geq h^2 \frac{\tilde{C}}{2}. \end{aligned} \quad (13.2.79)$$

By combining (13.2.80) and (13.2.79) it follows:

$$\left| \mathbb{P} \left[ \frac{P(h, r) - \mathbb{E}[P(h, r)]}{\sqrt{\mathbb{V}[P(h, r)]}} \leq x \right] - \mathbb{P}[G \leq x] \right| \leq \frac{C}{\sqrt{h}}, x > 0, \quad (13.2.80)$$

where:

$$C = 2 \sqrt{\frac{\alpha}{\mathbb{E}[(M_1)^2]}} \left( \frac{2\mathbb{E}[(M_1)^4]}{\lambda_2 \tilde{C}} \right)^{\frac{1}{4}}. \quad (13.2.81)$$

Result (13.2.80) gives the rate of convergence of the provision renormalized to a Gaussian RV when  $(R_t)_{t \geq 0}$  is given.

Now we study the asymptotic behaviour of RVs  $X_k(h, R)$  according to the value of  $\mu$ :

**Proposition 13.2.16.** *Let assumptions  $(A_1)$  to  $(A_8)$  hold and  $k \geq 1$ .*

1. If  $\mu < 0$  then:

$$\frac{X_k(h, R)}{h^2} \xrightarrow[h \rightarrow +\infty]{\mathcal{P}} C_k, \quad (13.2.82)$$

where the constant  $C_k$  is defined by (13.2.20). When  $\mu = 0$  then:

$$\frac{X_k(h, R)}{h^2} \xrightarrow[h \rightarrow +\infty]{\mathcal{L}} 2C_k \int_0^1 (1-t) \mathbb{1}_{\{B_t < 0\}} dt, \quad (13.2.83)$$

where  $(B_t)_{t \geq 0}$  is the Brownian motion defined by (13.2.7).

2. In addition if  $\mu < 0$  then:

$$\left( \frac{X_1(h, R)}{h^2} - C_1 \right) h \xrightarrow[h \rightarrow +\infty]{a.s.} -2C_1 \left[ \frac{e^{-\tau t^*}}{\tau} + \int_0^{+\infty} \inf \left( 1, \frac{\gamma}{\alpha} R_t \right) dt \right], \quad (13.2.84)$$

where  $(R_t)_{t \geq 0}$  is the geometric Brownian motion defined by (13.2.7).

The proof of this proposition will be made in Section 13.3.3.

**Remark 13.2.17.**

1. Note that if  $\mu > 0$  then:

$$\int_0^{+\infty} \inf \left( 1, \frac{\gamma}{\alpha} R_t \right) dt = +\infty. \quad (13.2.85)$$

So when  $\mu > 0$  result (13.2.84) becomes:

$$\left( \frac{X_1(h, R)}{h^2} - C_1 \right) h \xrightarrow[h \rightarrow +\infty]{a.s.} +\infty. \quad (13.2.86)$$

In the case  $\mu > 0$  the renormalization factor is not sufficient to get a finite limit.

2. Result (13.2.84) implies in particular that (13.2.82) is true for  $k = 1$  and the convergence is almost surely.

Note that we can not deduce from Proposition 13.2.14 and 13.2.16, the convergence of  $\frac{P(h, R) - h^2 C_1}{h\sqrt{C_2}}$ , we need an additional result:

**Proposition 13.2.18.** *Let assumptions  $(A_1)$  to  $(A_9)$  hold. If  $\mu < 0$  it follows that:*

$$\left[ \frac{P(h, R) - X_1(h, R)}{\sqrt{X_2(h, R)}}, \left( \frac{X_1(h, R)}{h^2} - C_1 \right) h \right] \xrightarrow[h \rightarrow +\infty]{\mathcal{L}} \left[ G, -2C_1 \left[ \frac{e^{-\tau t^*}}{\tau} + \int_0^{+\infty} \inf \left( 1, \frac{\gamma}{\alpha} R_t \right) dt \right] \right]. \quad (13.2.87)$$

The proof of this proposition will be made in Section 13.3.4.

Let us come back to the proof of result (13.2.26). We have:

$$\frac{P(h, R) - h^2 C_1}{h\sqrt{C_2}} = \frac{P(h, R) - X_1(h, R)}{\sqrt{X_2(h, R)}} \sqrt{\frac{X_2(h, R)}{h^2 C_2}} + \left( \frac{X_1(h, R)}{h^2} - C_1 \right) \frac{h}{\sqrt{C_2}}. \quad (13.2.88)$$

Proposition 13.2.18 combined to (13.2.82) with  $k = 2$  allows us to write:

$$\frac{P(h, R) - h^2 C_1}{h\sqrt{C_2}} \xrightarrow[h \rightarrow +\infty]{\mathcal{L}} X^*,$$

where  $X^*$  is defined by (13.2.21). This established (13.2.26).

**Proof of formula (13.2.27) of Theorem 13.2.5:**

Firstly we give a technical lemma useful for the next:

**Lemma 13.2.19.**

$$\lim_{h \rightarrow +\infty} \int_0^1 dy \int_0^{hy} [\phi_t(hy) - \alpha] dt = -\alpha \frac{e^{-\tau t^*}}{\tau}. \quad (13.2.89)$$

This lemma will be proved in Section 13.3.5. Now we will focus on the proof of formula (13.2.27). We have:

$$\frac{P(h, R) - \mathbb{E}[P(h, R)]}{h\sqrt{C_2}} = \frac{P(h, R) - h^2 C_1}{h\sqrt{C_2}} + \frac{h}{\sqrt{C_2}} \left[ C_1 - \frac{\mathbb{E}[P(h, R)]}{h^2} \right]. \quad (13.2.90)$$

From (13.2.38) and using the change of variable  $s = hy$  it follows:

$$h \left[ C_1 - \frac{\mathbb{E}[P(h, R)]}{h^2} \right] = \lambda_2 \mathbb{E}[M_1] \int_0^1 dy \int_0^{hy} [\Gamma_1(t, hy) + \Gamma_2(t, hy)] dt, \quad (13.2.91)$$

where:

$$\Gamma_1(t, hy) := \alpha - \phi_t(hy) \Phi \left[ \frac{\ln \left[ \frac{\phi_t(hy)}{\gamma} \right] - \mu t}{\sigma \sqrt{t}} \right], \quad (13.2.92)$$

$$\Gamma_2(t, hy) := \gamma e^{(\mu + \frac{\sigma^2}{2})t} \Phi \left[ \frac{\ln \left[ \frac{\phi_t(hy)}{\gamma} \right] - (\mu + \sigma^2)t}{\sigma \sqrt{t}} \right]. \quad (13.2.93)$$

From (13.2.54) it comes that:

$$\lim_{h \rightarrow +\infty} \Gamma_1(t, hy) = \alpha - \alpha \Phi \left[ \frac{\ln \left[ \frac{\alpha}{\gamma} \right] - \mu t}{\sigma \sqrt{t}} \right]$$

$$= \alpha \Phi \left[ \frac{\mu t - \ln \left[ \frac{\alpha}{\gamma} \right]}{\sigma \sqrt{t}} \right], \quad (13.2.94)$$

$$\lim_{h \rightarrow +\infty} \Gamma_2(t, hy) = \gamma e^{(\mu + \frac{\sigma^2}{2})t} \Phi \left[ \frac{\ln \left[ \frac{\alpha}{\gamma} \right] - (\mu + \sigma^2)t}{\sigma \sqrt{t}} \right]. \quad (13.2.95)$$

From (13.3.6) and (13.3.21) we have:

$$\alpha(1 - e^{-\tau t^*}) \leq \phi_t(hy) \leq \alpha. \quad (13.2.96)$$

So:

$$\begin{aligned} \Gamma_1(t, hy) &= \alpha - \phi_t(hy) \left( 1 - \Phi \left[ \frac{\mu t - \ln \left[ \frac{\phi_t(hy)}{\gamma} \right]}{\sigma \sqrt{t}} \right] \right) \\ &= \alpha - \phi_t(hy) + \phi_t(hy) \Phi \left[ \frac{\mu t - \ln \left[ \frac{\phi_t(hy)}{\gamma} \right]}{\sigma \sqrt{t}} \right] \\ &\leq \alpha - \phi_t(hy) + \alpha \Phi \left[ \frac{\mu t - \ln \left[ \frac{\alpha(1 - e^{-\tau t^*})}{\gamma} \right]}{\sigma \sqrt{t}} \right], \end{aligned} \quad (13.2.97)$$

$$\Gamma_2(t, hy) \leq \gamma e^{(\mu + \frac{\sigma^2}{2})t} \Phi \left[ \frac{\ln \left[ \frac{\alpha}{\gamma} \right] - (\mu + \sigma^2)t}{\sigma \sqrt{t}} \right]. \quad (13.2.98)$$

By combining (13.2.91), (13.2.92), (13.2.93), (13.2.94), (13.2.95), (13.2.97), (13.2.98), (13.2.51), (13.2.52), and (13.2.89) it follows:

$$\begin{aligned} \lim_{h \rightarrow +\infty} h \left[ C_1 - \frac{\mathbb{E}[P(h, R)]}{h^2} \right] &= \lambda_2 \mathbb{E}[M_1] \int_0^{+\infty} \left( \alpha \Phi \left[ \frac{\mu t - \ln \left[ \frac{\alpha}{\gamma} \right]}{\sigma \sqrt{t}} \right] \right. \\ &\quad \left. + \gamma e^{(\mu + \frac{\sigma^2}{2})t} \Phi \left[ \frac{\ln \left[ \frac{\alpha}{\gamma} \right] - (\mu + \sigma^2)t}{\sigma \sqrt{t}} \right] \right) dt. \end{aligned} \quad (13.2.99)$$

So combining (13.2.90), (13.2.99), and (13.2.26) we obtain (13.2.27).

## 13.3 Technicals proofs

### 13.3.1 Proof of Lemma 13.2.12

Recall that:

$$\Phi(z) \sim \frac{1}{\sqrt{2\pi}|z|} e^{-\frac{z^2}{2}}, \quad z \rightarrow -\infty. \quad (13.3.1)$$

Formula (13.2.51) is a direct consequence of (13.3.1). Now we prove formula (13.2.52). We distinguish two cases:

1. If  $b + c^2 \leq 0$  then  $b + \frac{c^2}{2} < 0$  and formula (13.2.52) is verified because  $\Phi(x) \leq 1$  for all  $x \in \mathbb{R}$ .

2. If  $b + c^2 > 0$  we use that:

$$\Phi(z) \leq \frac{1}{\sqrt{2\pi}|z|} e^{-\frac{z^2}{2}}, \quad z < 0. \quad (13.3.2)$$

So:

$$\begin{aligned} e^{(b+\frac{c^2}{2})t} \Phi\left[\frac{a - (b + c^2)t}{c\sqrt{t}}\right] &\leq \frac{c\sqrt{t}}{\sqrt{2\pi}|a - (b + c^2)t|} \exp\left[\left(b + \frac{c^2}{2}\right)t - \frac{1}{2tc^2}[a - (b + c^2)t]^2\right] \\ &\leq \frac{c\sqrt{t}}{\sqrt{2\pi}|a - (b + c^2)t|} \exp\left[-\frac{b^2}{2c^2}t + \frac{a}{c^2}(b + c^2) - \frac{a^2}{2tc^2}\right]. \end{aligned} \quad (13.3.3)$$

This established (13.2.52).

### 13.3.2 Proof of Proposition 13.2.14

**Notation 13.3.1.** *It is convenient to adopt the following notations:*

$$f_k := f(T_k, D_k, r_{T_k}), \quad k \geq 1, \quad (13.3.4)$$

$$N(A) := \sum_{j \geq 1} \mathbf{1}_{\{(T_j, D_j) \in A\}}, \quad (13.3.5)$$

where  $f$  is defined by (13.2.4),  $A$  is a subset of  $\{(t, s), 0 \leq t \leq s\}$ , and  $(r_t, t \geq 0)$  is a continuous function in  $]0, +\infty[$ .

The proof of Proposition 13.2.14 is based on the following lemma:

**Lemma 13.3.2.**

1. For all  $t, s, u \geq 0 : t < s$  we have:

$$f(t, s, u) \leq \phi_t(s) \leq \alpha, \quad (13.3.6)$$

where  $f$  and  $\alpha$  are defined by (13.2.4) and (13.2.6).

2. We have:

$$\mathbb{E}[P(h, r)] = X_1(h, r), \quad (13.3.7)$$

$$\mathbb{V}[P(h, r)] = X_2(h, r). \quad (13.3.8)$$

3. If:

$$\lim_{t \rightarrow +\infty} r_t = 0, \quad (13.3.9)$$

then:

$$\lim_{h \rightarrow +\infty} \int_{A_h} [f(t, s, r_t)]^2 dt ds = +\infty. \quad (13.3.10)$$

The proof of this lemma will be made at the end of this section. Now we prove Proposition 13.2.14. Formulas (13.2.67) and (13.2.68) are proved by (13.3.7) and (13.3.8). From Theorem 4.1 of [50] we have:

$$d_W \left[ \frac{P(h, r) - \mathbb{E}[P(h, r)]}{\mathbb{V}[P(h, r)]}, G \right] \leq \frac{\sqrt{X_4(h, r)}}{\mathbb{V}[P(h, r)]}, \quad (13.3.11)$$

where  $d_W$  is the Wasserstein distance defined by (13.2.64): this combined with (13.3.8) prove formula (13.2.66). Now we show that the right term in (13.3.11) converges to 0 when  $h \rightarrow +\infty$ . From (13.2.60), (13.3.8), and (13.3.6) it follows:

$$\begin{aligned} \frac{X_4(h, r)}{[\mathbb{V}[P(h, r)]]^2} &= \frac{X_4(h, r)}{[X_2(h, r)]^2} \\ &= \frac{\mathbb{E}[(M_1)^4] \int_{A_h} [f(t, s, r_t)]^4 dt ds}{\lambda_2 [\mathbb{E}[(M_1)^2] \int_{A_h} [f(t, s, r_t)]^2 dt ds]^2}. \end{aligned}$$

Since:

$$[f(t, s, r_t)]^4 \leq \alpha^2 [f(t, s, r_t)]^2, \quad (13.3.12)$$

it comes that:

$$\frac{X_4(h, r)}{[\mathbb{V}[P(h, r)]]^2} \leq \frac{\alpha^2 \mathbb{E}[(M_1)^4]}{\lambda_2 [\mathbb{E}[(M_1)^2]]^2 \int_{A_h} [f(t, s, r_t)]^2 dt ds}. \quad (13.3.13)$$

From (13.3.10), (13.3.13) and assumption (A<sub>9</sub>) it follows:

$$\frac{\sqrt{X_4(h, r)}}{\mathbb{V}[P(h, r)]} \xrightarrow{h \rightarrow +\infty} 0. \quad (13.3.14)$$

This established formula (13.2.70). By combining (13.3.14) and (13.3.11) it follows (13.2.71) is true. This achieves the proof of Proposition 13.2.14. Let us come back to the proof of Lemma 13.3.2.

*Proof of Lemma 13.3.2.*

1. From (13.2.4), since  $u \geq 0$  and  $\gamma \in ]0, 1[$  it follows:

$$f(t, s, u) \leq \phi_t(s), \quad (13.3.15)$$

where  $f$  and  $\phi$  are defined by (13.2.4) and (11.2.13). By definition of  $\phi_t(s)$  we can deduce from (13.3.15) that:

$$f(t, s, u) \leq \alpha, \text{ si } t \leq t^*. \quad (13.3.16)$$

When  $s > t^*$  the function  $t \rightarrow \phi_t(s)$  is decreasing over  $]t^*, s]$  and:

$$\lim_{t \rightarrow t^*} \phi_t(s) = \alpha.$$

We deduce that:

$$f(t, s, u) \leq \alpha, \text{ si } t > t^*. \quad (13.3.17)$$

This established (13.3.6).

2. Recall that  $N(A_h) \sim P(\Lambda(A_h))$  and given  $N(A_h) = n$  the RVs  $(T_1, D_1), \dots, (T_n, D_n)$  are i.i.d. with law  $\frac{1}{\Lambda(A_h)} \Lambda|_{A_h}$ . Then it follows:

$$\begin{aligned} \mathbb{E}[P(h, r)] &= \mathbb{E}\left[\sum_{j \geq 1} M_j f_j \mathbf{1}_{\{(T_j, D_j) \in A_h\}}\right] \\ &= X_1(h, r). \end{aligned}$$

This established formula (13.3.7). Now we calculate the variance of  $P(h, r)$ :

$$\begin{aligned} \mathbb{V}[P(h, r)] &= \mathbb{E}\left[\left(\sum_{j \geq 1} M_j f_j \mathbf{1}_{\{(T_j, D_j) \in A_h\}}\right)^2\right] - [X_1(h, r)]^2 \\ &= \sum_{n \geq 0} \mathbb{P}[N(A_h) = n] \left\{ \mathbb{E}\left[\sum_{j=1}^n (M_j f_j)^2\right] + \mathbb{E}\left[\sum_{1 \leq j \neq l \leq n} M_j f_j M_l f_l\right] \right\} \\ &\quad - [X_1(h, r)]^2 \\ &= \sum_{n \geq 0} \mathbb{P}[N(A_h) = n] \left\{ n \mathbb{E}[(M_1)^2] \mathbb{E}[(f_1)^2] + n(n-1) (\mathbb{E}[M_1] \mathbb{E}[f_1])^2 \right\} \\ &\quad - [X_1(h, r)]^2. \end{aligned} \quad (13.3.18)$$

In addition:

$$\begin{aligned} \mathbb{E}[(f_1)^k] &= \frac{\lambda_2}{\Lambda(A_h)} \int_{A_h} [f(t, s, r_t)]^k dt ds \\ &= \frac{\lambda_2}{\mathbb{E}[N(A_h)]} \int_{A_h} [f(t, s, r_t)]^k dt ds \\ &= \frac{X_k(h, r)}{\mathbb{E}[(M_1)^k]} \frac{1}{\mathbb{E}[N(A_h)]}. \end{aligned} \quad (13.3.19)$$

So:

$$\begin{aligned} \mathbb{V}[P(h, r)] &= \sum_{n \geq 0} \mathbb{P}[N(A_h) = n] \left\{ n \frac{X_2(h, r)}{\mathbb{E}[N(A_h)]} + n(n-1) \left[ \frac{X_1(h, r)}{\mathbb{E}[N(A_h)]} \right]^2 \right\} \\ &\quad - [X_1(h, r)]^2 \\ &= X_2(h, r) + \left[ \mathbb{E}[(N(A_h))^2] - \mathbb{E}[N(A_h)] \right] \left[ \frac{X_1(h, r)}{\mathbb{E}[N(A_h)]} \right]^2 - [X_1(h, r)]^2. \end{aligned}$$



Since:

$$\mathbb{V}[N(A_h)] = \mathbb{E}[N(A_h)] = \mathbb{E}[(N(A_h))^2] - (\mathbb{E}[N(A_h)])^2,$$

we deduce:

$$\mathbb{V}[P(h, r)] = X_2(h, r).$$

This established formula (13.3.8).

3. We have:

$$\begin{aligned} \int_{A_h} [f(t, s, r_t)]^2 dt ds &= \int_0^h ds \left[ \int_0^s [f(t, s, r_t)]^2 dt \right] \\ &\stackrel{s=hy}{=} h \int_0^1 dy \left[ \int_0^{+\infty} \mathbf{1}_{t \leq hy} \left( [\phi_t(hy) - r_t]_+ \right)^2 dt \right]. \end{aligned} \quad (13.3.20)$$

When  $s > t^*$  the function  $t \rightarrow \phi_t(s)$  is decreasing on  $]t^*, s]$ , we deduce:

$$\phi_s(s) \geq \alpha(1 - e^{-\tau t^*}) > 0, \quad 0 \leq t \leq s. \quad (13.3.21)$$

So:

$$\int_{A_h} [f(t, s, r_t)]^2 dt ds \geq h \int_0^1 dy \left[ \int_0^{+\infty} \mathbf{1}_{t \leq hy} \left( [\alpha(1 - e^{-\tau t^*}) - r_t]_+ \right)^2 dt \right]. \quad (13.3.22)$$

From (13.3.9) we have:

$$\int_0^{+\infty} [(a - r_t)_+]^2 dt = +\infty, \quad a > 0.$$

So:

$$\lim_{h \rightarrow +\infty} \int_{A_h} [f(t, s, r_t)]^2 dt ds = +\infty.$$

This achieves the proof of Lemma 13.3.2. □

### 13.3.3 Proof of Proposition 13.2.16

1. It is clear that proof of (13.2.82) and (13.2.83) of Proposition 13.2.16 is a direct consequence of the following lemma:

**Lemma 13.3.3.** *Let  $k \geq 1$ .*

(a) *For all  $h > 0$ :*

$$X_k(h, R) \stackrel{(d)}{=} \tilde{X}_k(h, R), \quad (13.3.23)$$

where function  $X_k(h, R)$  is defined by (13.2.60) and:

$$\tilde{X}_k(h, R) := h^2 \lambda_2 \mathbb{E}[(M_1)^k] \int_0^1 dy \left[ \int_0^y [f(hx, hy, \exp[\mu hx + \sigma \sqrt{h} B_x])]^k dx \right]. \quad (13.3.24)$$

(b) If  $\mu < 0$  then:

$$\lim_{h \rightarrow +\infty} \frac{\tilde{X}_k(h, R)}{h^2} = C_k \quad a.s., \quad (13.3.25)$$

where the constant  $C_k$  is defined by (13.2.20).

(c) If  $\mu = 0$  then:

$$\lim_{h \rightarrow +\infty} \frac{\tilde{X}_k(h, R)}{h^2} = 2C_k \int_0^1 (1-t) \mathbf{1}_{\{B_t < 0\}} dt, \quad a.s.. \quad (13.3.26)$$

The proof of this lemma will be made at the end of this section.

2. Now we want to prove result (13.2.84). We need to introduce the following notation:

$$\tilde{C}_1(h) := \frac{\lambda_2}{h} \mathbb{E}[M_1] \int_0^1 du \int_0^{hu} \phi_t(hu) dt. \quad (13.3.27)$$

The proof is based on the following lemma:

**Lemma 13.3.4.**

(a) If  $\mu < 0$  it comes that:

$$\left[ \frac{X_1(h, R)}{h^2} - \tilde{C}_1(h) \right] h \xrightarrow[h \rightarrow +\infty]{a.s.} -\lambda_2 \mathbb{E}[M_1] \int_0^{+\infty} \inf(\alpha, \gamma R_t) dt. \quad (13.3.28)$$

(b) In addition:

$$\lim_{h \rightarrow +\infty} \left[ \tilde{C}_1(h) - C_1 \right] h = -2 \frac{e^{-\tau t^*}}{\tau} C_1. \quad (13.3.29)$$

The proof of this lemma will be made at the end of this section. Now we want to prove formula (13.2.84). Since:

$$\left( \frac{X_1(h, R)}{h^2} - C_1 \right) h = \left[ \frac{X_1(h, R)}{h^2} - \tilde{C}_1(h) \right] h + \left[ \tilde{C}_1(h) - C_1 \right] h, \quad (13.3.30)$$

and using (13.3.28) and (13.3.29) we deduce the formula (13.2.84).

Let us come back to the proof of Lemma 13.3.3 and 13.3.4.

*Proof of Lemma 13.3.3.*

1. From formulas (13.2.60), (13.2.3), (13.2.8) and using the changes of variables  $s = hy$  and  $t = hx$  it comes:

$$\begin{aligned} \frac{X_k(h, R)}{\mathbb{E}[(M_1)^k]} &= \lambda_2 \int_0^h ds \left[ \int_0^s [f(t, s, R_t)]^k dt \right] \\ &= h^2 \lambda_2 \int_0^1 dy \left[ \int_0^y [f(hx, hy, \exp[\mu hx + \sigma B_{hx}])]^k dx \right]. \end{aligned} \quad (13.3.31)$$

But:

$$(B_{xh}, x \geq 0) \stackrel{(d)}{=} (\sqrt{h}B_x, x \geq 0). \quad (13.3.32)$$

So combining (13.3.31) and (13.3.32) it follows:

$$X_k(h, r) \stackrel{(d)}{=} \tilde{X}_k(h, R),$$

where  $\tilde{X}_k(h, R)$  is defined by (13.3.24). This establishes formula (13.3.23).

2. The previous change of variable allows us to study the asymptotic behaviour of  $\tilde{X}_k(h, R)$ . We want to apply the dominated convergence theorem. With (13.2.4) and (13.2.5) it follows:

$$f(hx, hy, \exp[\sigma\sqrt{h}B_x + \mu hx]) = [\phi_{hx}(hy) - \gamma \exp\{\sigma\sqrt{h}B_x + \mu hx\}]_+.$$

Because  $x > 0$  and from (13.2.42) we obtain:

$$\begin{aligned} \lim_{h \rightarrow +\infty} f(hx, hy, \exp[\sigma\sqrt{h}B_x + \mu hx]) \\ = \alpha[\mathbb{1}_{\mu < 0} + \mathbb{1}_{\{\mu=0, B_x < 0\}}] + \mathbb{1}_{\{\mu=0, B_x=0\}}(\alpha - \gamma)_+, \quad a.s.. \end{aligned}$$

In addition, inequality (13.3.6) allows us to apply the dominated convergence theorem. If  $\mu < 0$  we have:

$$\begin{aligned} \lim_{h \rightarrow +\infty} \frac{\tilde{X}_k(h, R)}{h^2} &= \lambda_2 \mathbb{E}[(M_1)^k] \int_0^1 dy \left[ \int_0^y \alpha^k dx \right] \\ &= \lambda_2 \mathbb{E}[(M_1)^k] \frac{\alpha^k}{2} = C_k. \end{aligned}$$

When  $\mu = 0$  we have:

$$\begin{aligned} \lim_{h \rightarrow +\infty} \frac{\tilde{X}_k(h, R)}{h^2} &= \lambda_2 \mathbb{E}[(M_1)^k] \int_0^1 dy \int_0^y [\alpha \mathbb{1}_{\{B_x < 0\}} dx + (\alpha - \gamma)_+ \mathbb{1}_{\{B_x = 0\}}]^k dx \\ &= \lambda_2 \mathbb{E}[(M_1)^k] \int_0^1 (1-x) [\alpha^k \mathbb{1}_{\{B_x < 0\}} dx + [(\alpha - \gamma)_+]^k \mathbb{1}_{\{B_x = 0\}}] dx \\ &= \lambda_2 \alpha^k \mathbb{E}[M_1] \int_0^1 (1-t) \mathbb{1}_{\{B_t < 0\}} dt \\ &= 2C_k \int_0^1 (1-t) \mathbb{1}_{\{B_t < 0\}} dt, \quad a.s., \end{aligned}$$

because  $\int_0^1 (1-t) \mathbb{1}_{\{B_t = 0\}} dt$  is a positive RV with mean equal to 0.

This achieves the proof of Lemma 13.3.3. □

*Proof of Lemma 13.3.4.*

1. From (13.2.60) and (13.3.27) it follows:

$$\begin{aligned}
 \left[ \frac{X_1(h, R)}{h^2} - \tilde{C}_1(h) \right] h &= \lambda_2 \mathbb{E}[M_1] \left[ \frac{1}{h} \int_0^h ds \int_0^s [\phi_t(s) - \gamma R_t]_+ dt \right. \\
 &\quad \left. - \int_0^1 du \int_0^{hu} \phi_t(hu) dt \right] \\
 &\stackrel{s=hu}{=} \lambda_2 \mathbb{E}[M_1] \int_0^1 du \int_0^{hu} [\phi_t(hu) - \gamma R_t]_+ - \phi_t(hu) dt \\
 &= -\lambda_2 \mathbb{E}[M_1] \int_0^1 du \int_0^{hu} \inf[\phi_t(hu), \gamma R_t] dt. \tag{13.3.33}
 \end{aligned}$$

So when  $\mu < 0$  it comes from (13.2.54) that:

$$\left[ \frac{X_1(h, R)}{h^2} - \tilde{C}_1(h) \right] h \xrightarrow[h \rightarrow +\infty]{a.s.} -\lambda_2 \mathbb{E}[M_1] \int_0^{+\infty} \inf(\alpha, \gamma R_t) dt,$$

because:

$$\int_0^{+\infty} \inf(\alpha, \gamma R_t) < +\infty.$$

This established formula (13.3.28).

2. From (13.3.27) and (13.2.20) it comes:

$$\begin{aligned}
 [\tilde{C}_1(h) - C_1] h &= \lambda_2 \mathbb{E}[M_1] \left[ \int_0^1 du \int_0^{hu} \phi_t(hu) dt - \frac{\alpha}{2} h \right] \\
 &= \lambda_2 \mathbb{E}[M_1] \left[ \int_0^1 du \int_0^{hu} [\phi_t(hu) - \alpha] dt \right]. \tag{13.3.34}
 \end{aligned}$$

Using result (13.2.89) it comes:

$$\lim_{h \rightarrow +\infty} [\tilde{C}_1(h) - C_1] h = -2 \frac{e^{-\tau t^*}}{\tau} C_1. \tag{13.3.35}$$

This establishes formula (13.3.29). □

### 13.3.4 Proof of Proposition 13.2.18

Let  $\phi_1$  et  $\phi_2$  two continuous and bounded functions. Then:

$$\begin{aligned}
 \Gamma_h &:= \mathbb{E} \left[ \phi_1 \left( \frac{P(h, R) - X_1(h, R)}{\sqrt{X_2(h, R)}} \right) \phi_2 \left( \left[ \frac{X_1(h, R)}{h^2} - C_1 \right] h \right) \right] \\
 &= \mathbb{E} \left[ \tilde{\phi}_1(h, R) \phi_2 \left( \left[ \frac{X_1(h, R)}{h^2} - C_1 \right] h \right) \right], \tag{13.3.36}
 \end{aligned}$$

where:

$$\tilde{\phi}_1(h, r) := \mathbb{E} \left[ \phi_1 \left( \frac{P(h, r) - X_1(h, r)}{\sqrt{X_2(h, r)}} \right) \right]. \tag{13.3.37}$$

In particular from (13.2.71) it follows:

$$\lim_{h \rightarrow +\infty} \tilde{\phi}_1(h, r) = \mathbb{E}[\phi_1(G)]. \quad (13.3.38)$$

Since  $\phi_2$  is bounded and using (13.2.84), (13.3.36), and (13.3.37) it comes that:

$$\Gamma_h \xrightarrow{h \rightarrow +\infty} \mathbb{E} \left[ \phi_1(G) \phi_2 \left( -2C_1 \left( \frac{e^{-\tau t^*}}{\tau} + \int_0^{+\infty} \inf \left( 1, \frac{\gamma}{\alpha} R_t \right) dt \right) \right) \right].$$

This achieves the proof of Proposition 13.2.18.

### 13.3.5 Proof of Lemma 13.2.19

From (13.2.5) and for  $h > t^*$  it comes:

$$\begin{aligned} \int_0^1 dy \int_0^{hy} [\phi_t(hy) - \alpha] dt &= \int_{\frac{t^*}{h}}^1 dy \int_{t^*}^{hy} \alpha \left[ \frac{e^{\tau hy} - e^{\tau(t-t^*)}}{e^{\tau hy} - 1} - 1 \right] dt \\ &= \alpha \int_{\frac{t^*}{h}}^1 dy \int_{t^*}^{hy} \frac{1 - e^{\tau(t-t^*)}}{e^{\tau hy} - 1} dt \\ &= \alpha \int_{\frac{t^*}{h}}^1 \frac{dy}{e^{\tau hy} - 1} \left[ (hy - t^*) - \frac{e^{-\tau t^*}}{\tau} [e^{\tau hy} - e^{\tau t^*}] \right] dt \\ &= \alpha \left[ \int_{\frac{t^*}{h}}^1 \frac{hy - t^* + \frac{1}{\tau}(1 - e^{-\tau t^*})}{e^{\tau hy} - 1} dy - \frac{e^{-\tau t^*}}{\tau} \left( 1 - \frac{t^*}{h} \right) \right] \\ &= \alpha \left[ \frac{1}{h} \int_{t^*}^h \frac{s - t^* + \frac{1}{\tau}(1 - e^{-\tau t^*})}{e^{\tau s} - 1} ds - \frac{e^{-\tau t^*}}{\tau} \left( 1 - \frac{t^*}{h} \right) \right]. \end{aligned} \quad (13.3.39)$$

Since:

$$\int_{t^*}^h \frac{s - t^* + \frac{1}{\tau}(1 - e^{-\tau t^*})}{e^{\tau s} - 1} ds < +\infty, \quad (13.3.40)$$

it follows that:

$$\lim_{h \rightarrow +\infty} \int_0^1 dy \int_0^{hy} [\phi_t(hy) - \alpha] dt = -\alpha \frac{e^{-\tau t^*}}{\tau}. \quad (13.3.41)$$

This achieves the proof of Lemma 13.2.19.



# Additional information concerning the solvency need calculation

---

In this chapter, we will give some numerical complements to the previous chapter and particularly to Theorems [13.2.5](#) and [13.2.6](#).

In Section [14.1](#) we will verify that the asymptotic mean (see formula [\(13.2.24\)](#)) is close to the theoretical mean (see formula [\(12.4.2\)](#)) and the empirical mean obtained with simulations (see formula [\(12.4.7\)](#)). We will analyze the influence of parameters  $h$  and  $\mu$  on the provision mean.

In Section [14.2](#) we will focus on the empirical distribution of the provision renormalized:

$$\frac{P(h, R) - \mathbb{E}[P(h, R)]}{h\sqrt{C_2}}. \quad (14.0.1)$$

To do this, we will use the Algorithm [12.1](#) to simulate the provision  $P(h, R)$ . We will point out the influence of  $\mu$  on the empirical distribution and we will show that for  $\mu$  “strongly” negative, the empirical distribution of renormalized provision seems to be Gaussian. We will show that the influence of the integral of a geometric Brownian motion (added to the Gaussian RV) is almost zero when  $\mu$  tends to minus infinity.

Finally in Section [14.3](#) we will define the empirical solvency need and we will give a numerical value for this quantity. Finally, we will observe how this amount fluctuates in function of  $\mu$ .

## 14.1 Asymptotic and theoretical means

In Chapter [13](#) we assumed that when  $h = 276$  months the behaviour of the renormalized provision (see [\(14.0.1\)](#)) is close to its asymptotic behaviour when  $h \rightarrow +\infty$ . We have to verify if this is the case in practice. Consider  $\mu = -0.07$  and plot curves:

- Asymptotic mean (blue curve) in function of  $h$  (see [\(13.2.24\)](#)), i.e.:

$$\left[ h, h^2 \frac{\hat{\lambda}_2 \hat{\alpha} \bar{m}}{2} \right], \quad (14.1.1)$$

where  $\hat{\lambda}_2$ ,  $\hat{\alpha}$ , and  $\bar{m}$  are defined in Section [12.4.2](#) (see Illustration [12.21](#)).

- Theoretical mean (black curve) in function of  $h$ , i.e.:

$$\left[ h, \mathbb{E}^{theo}[P(h, R)] \right], \quad (14.1.2)$$

where  $\mathbb{E}^{theo}[P(h, R)]$  is the theoretical mean of the provision evaluated at date 0 and concerning loans drawdown in year 0, i.e.:

$$\begin{aligned} \mathbb{E}^{theo}[P(h, R)] = \hat{\lambda}_2 \bar{m} \times \int_{A_h} \left\{ \hat{\phi}_t(s) \Phi \left[ \frac{\ln \left[ \frac{\hat{\phi}_t(s)}{\hat{\gamma}} \right] - \hat{\mu}t}{\hat{\sigma} \sqrt{t}} \right] \right. \\ \left. - \hat{\gamma} e^{(\hat{\mu} + \frac{\hat{\sigma}^2}{2})t} \Phi \left[ \frac{\ln \left[ \frac{\hat{\phi}_t(s)}{\hat{\gamma}} \right] - (\hat{\mu} + \hat{\sigma}^2)t}{\hat{\sigma} \sqrt{t}} \right] \right\}, \end{aligned} \quad (14.1.3)$$

where  $\hat{\gamma}$ ,  $\hat{\mu}$ ,  $\hat{\sigma}$ , and  $\hat{\phi}_t(s)$  are defined in Section 12.4.2.

- The yellow curve indicates the mean value for  $h = 276$  months.

We do not plot the empirical mean because we saw in Section 12.4.2 that it is very close to the theoretical mean. We will also give the ratio “Asymptotic mean over theoretical mean” in red. This ratio will be indicated on the second ordinate axis (right axis).

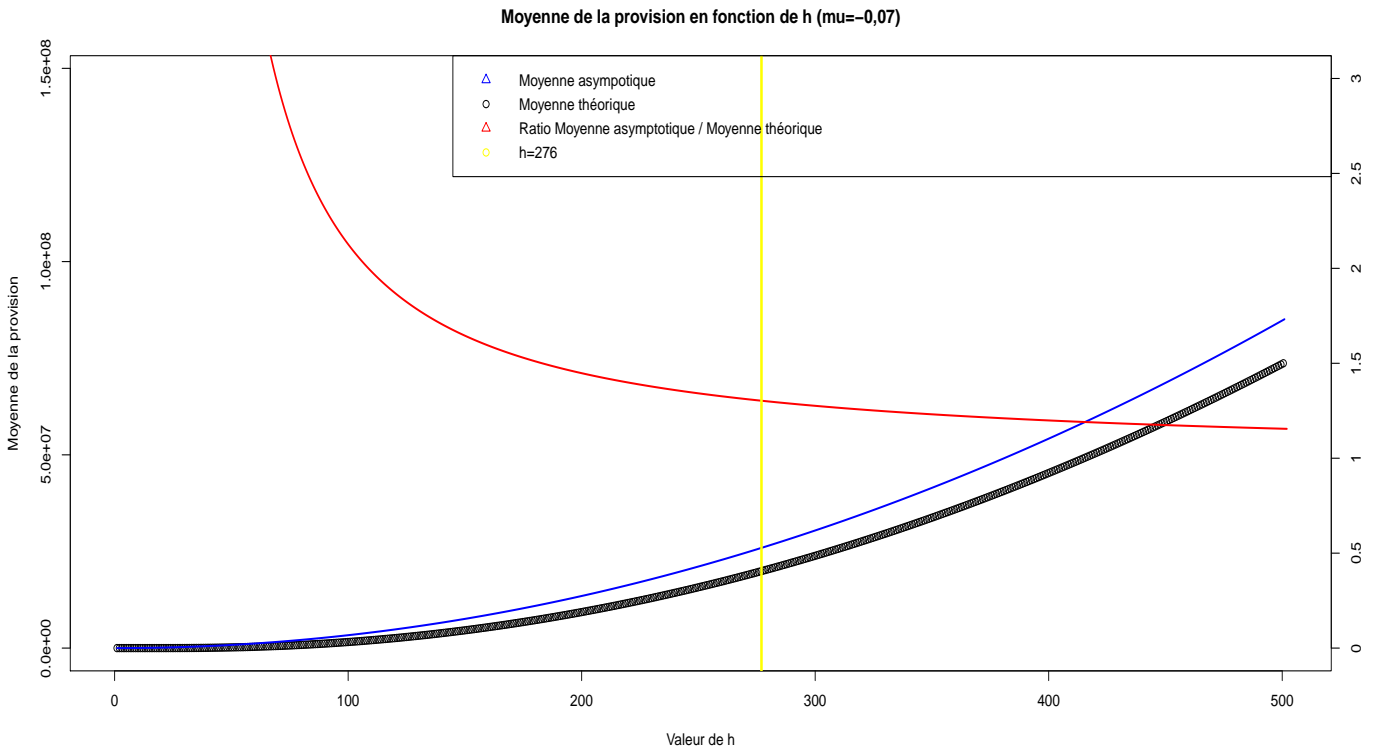


Illustration 14.1: Provision mean according  $h$ .

We can notice that when  $h$  increases, the asymptotic mean become clothier to the theoretical one. For  $h = 276$  months the ratio “asymptotic mean over theoretical mean ” is



around 1.2.

Fix  $h = 276$  months. We want to see how the asymptotic mean of provision (see (13.2.24)) and the theoretical mean (see (14.1.3)) fluctuate in function of  $\mu$ . To do this, we plot the following curves:

- Asymptotic mean in function of  $\mu$  (see (13.2.24)), i.e.:

$$\left[ \mu, \hat{\lambda}_2 \hat{\alpha} \bar{m} \hat{h}^2 \left( \frac{1}{2} \mathbf{1}_{\{\mu < 0\}} + \frac{1}{4} \mathbf{1}_{\{\mu = 0\}} \right) \right], \quad (14.1.4)$$

where  $\hat{\lambda}_2, \hat{\alpha}, \bar{m}$ , and  $\hat{h}$  are defined in Section 12.4.2.

- Theoretical mean (black) in function of  $\mu$ , i.e.:

$$\left[ \mu, \mathbb{E}^{theo}[P(h, R)] \right], \quad (14.1.5)$$

where  $\mathbb{E}^{theo}[P(h, R)]$  is defined by (14.1.3).

- The yellow curve represents mean values for the value of  $\mu$  estimated in Section 12.4.2, i.e.  $\mu = -0.07$ .

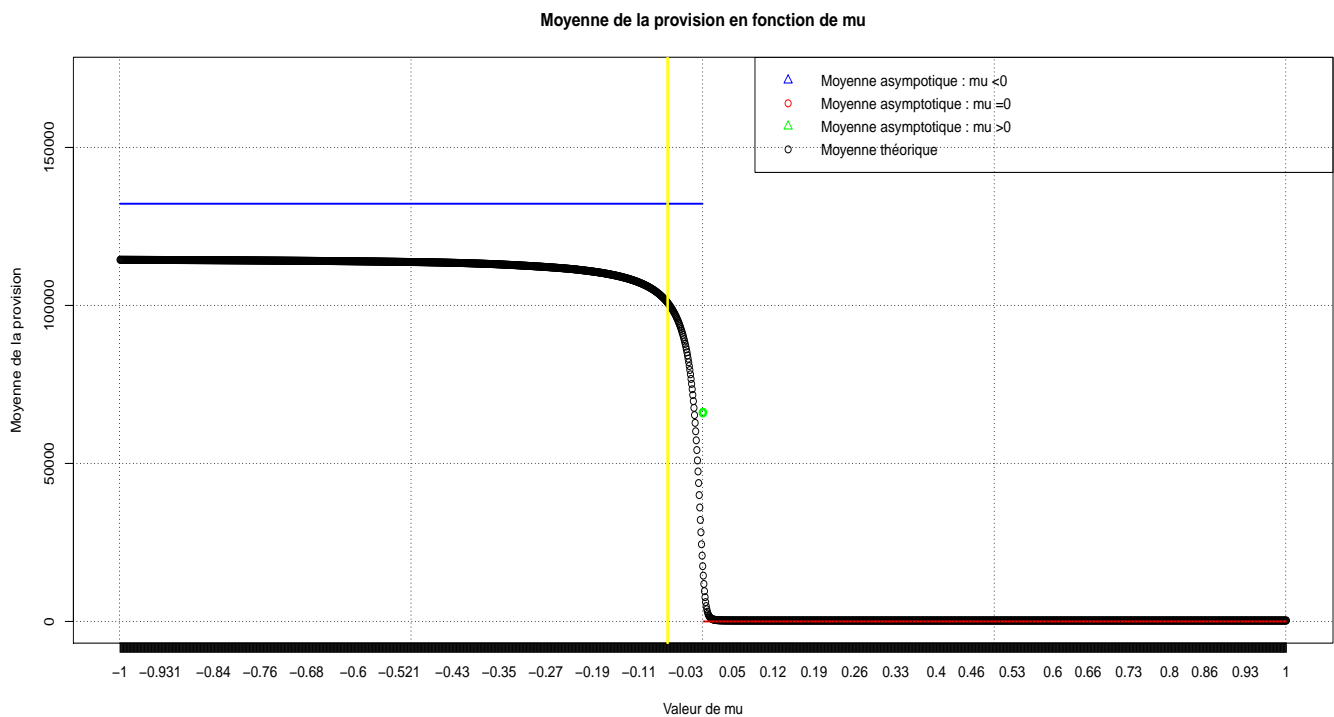


Illustration 14.2: Provision mean according  $\mu$ .

We find that:

1. Asymptotic mean is close to the theoretical mean.
2. Theoretical and asymptotic means vary differently in function of the sign of  $\mu$ . There are three different regimes:
  - (a) When  $\mu \leq 0$ : provision mean converges to a strictly positive constant,
  - (b) When  $\mu$  is close to 0: the mean of the provision approaches 0; this is a transient state,
  - (c) When  $\mu > 0$ : the provision mean tends to 0.

## 14.2 Empirical distribution of provision

In Theorem 13.2.5 we showed that the provision, properly renormalized, converges to the sum of two variables which one of them is a Gaussian, specifically:

$$\frac{P(h, R) - \mathbb{E}[P(h, R)]}{h\sqrt{C_2}} \xrightarrow[h \rightarrow +\infty]{\mathcal{L}} X^* + \kappa, \quad (14.2.1)$$

where coefficients  $C_k$  and  $\kappa$  are defined by (13.2.20) and (13.2.22) and:

$$X^* := G - \frac{2C_1}{\sqrt{C_2}} \left[ \frac{e^{-\tau t^*}}{\tau} + \int_0^{+\infty} \inf \left( 1, \frac{\gamma}{\alpha} R_t \right) dt \right]. \quad (14.2.2)$$

From (14.2.2) RV  $X^*$  is the sum of a constant and two RVs:

1. A centered and reduced Gaussian RV,
2. An integral of a function depending of a geometric Brownian motion.

In this section we want to show that the influence of the second RV is low when the variable  $\mu$  is negative. To do this, we consider the empirical distribution of:

$$P(h, R) - \mathbb{E}[P(h, R)]. \quad (14.2.3)$$

We want to see how this empirical distribution varies in function of  $\mu$ . Using Algorithm 12.1 it is possible to get a sample (size  $l = 5\,000$ ) of the provision  $P(h, R)$ . We denoted by  $(p_1^\mu, \dots, p_l^\mu)$  this sample. We plot the empirical histogram of:

$$\left( p_1^\mu - \frac{1}{l} \sum_{j=1}^l p_j^\mu, \dots, p_l^\mu - \frac{1}{l} \sum_{j=1}^l p_j^\mu \right), \quad (14.2.4)$$

for several value of  $\mu$  ( $\mu = -0.1, -0.07, -0.04, -0.01$ ):

- The blue curve represents the empirical density,
- The red curve represents a Gaussian density centered and with standard deviation  $\hat{h}\sqrt{\hat{C}_2}$ .

## 14.2 Empirical distribution of provision

We start by  $\mu = -0.01$ :

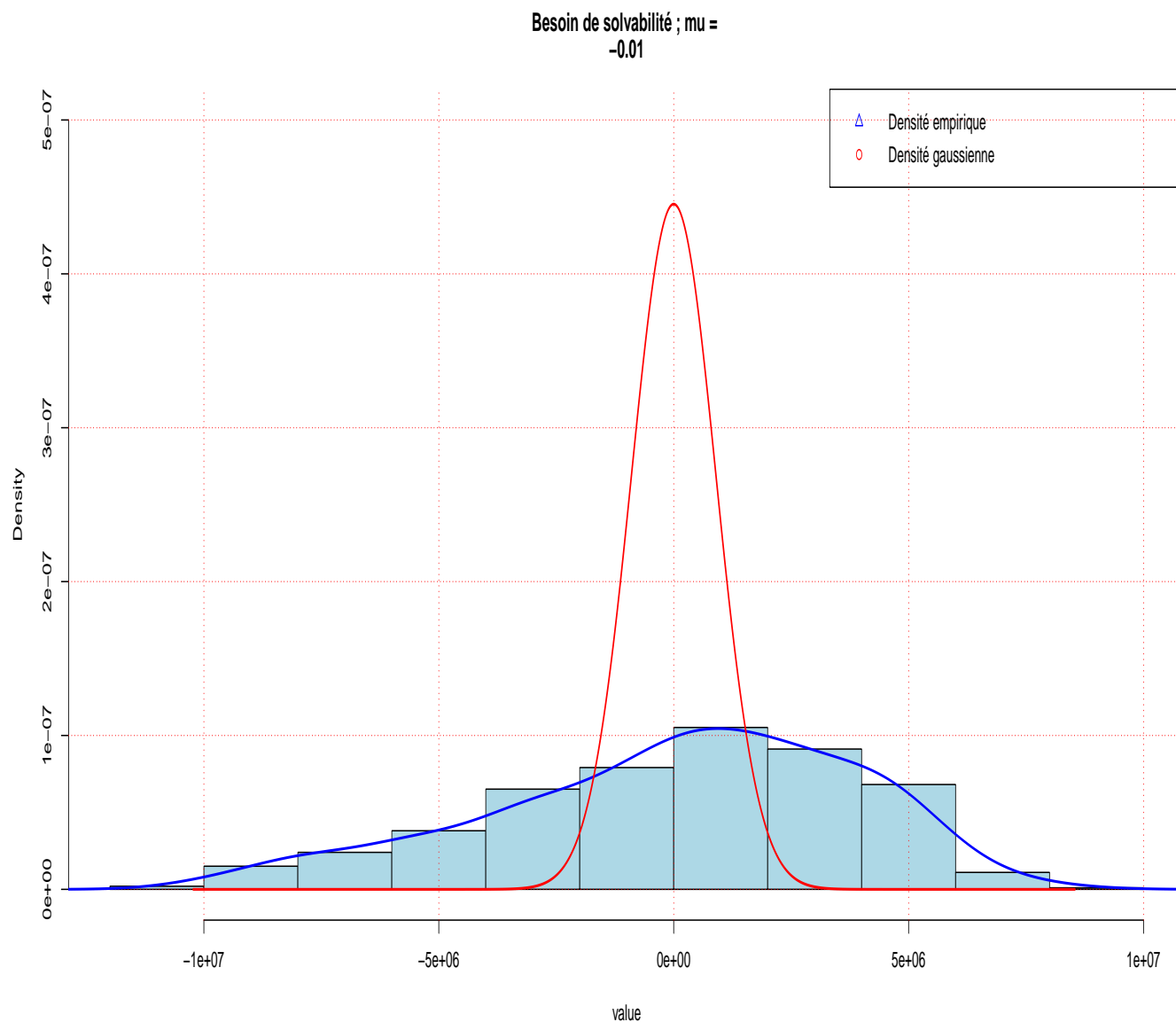


Illustration 14.3: Empirical distribution ;  $\mu = -0.01$ .

The empirical distribution does not appear Gaussian. Normality tests (Jarque- Bera, etc.) confirm this observation as the QQ-plot. Recall that a QQ-plot is a graph representing observed quantiles in function of Gaussian quantiles. On the next QQ-plot, we find empirical quantiles of the sample 14.2.4 on the x axis and quantiles of a centered Gaussian RV and standard deviation  $\hat{h}\sqrt{\hat{C}_2}$  on the y axis.

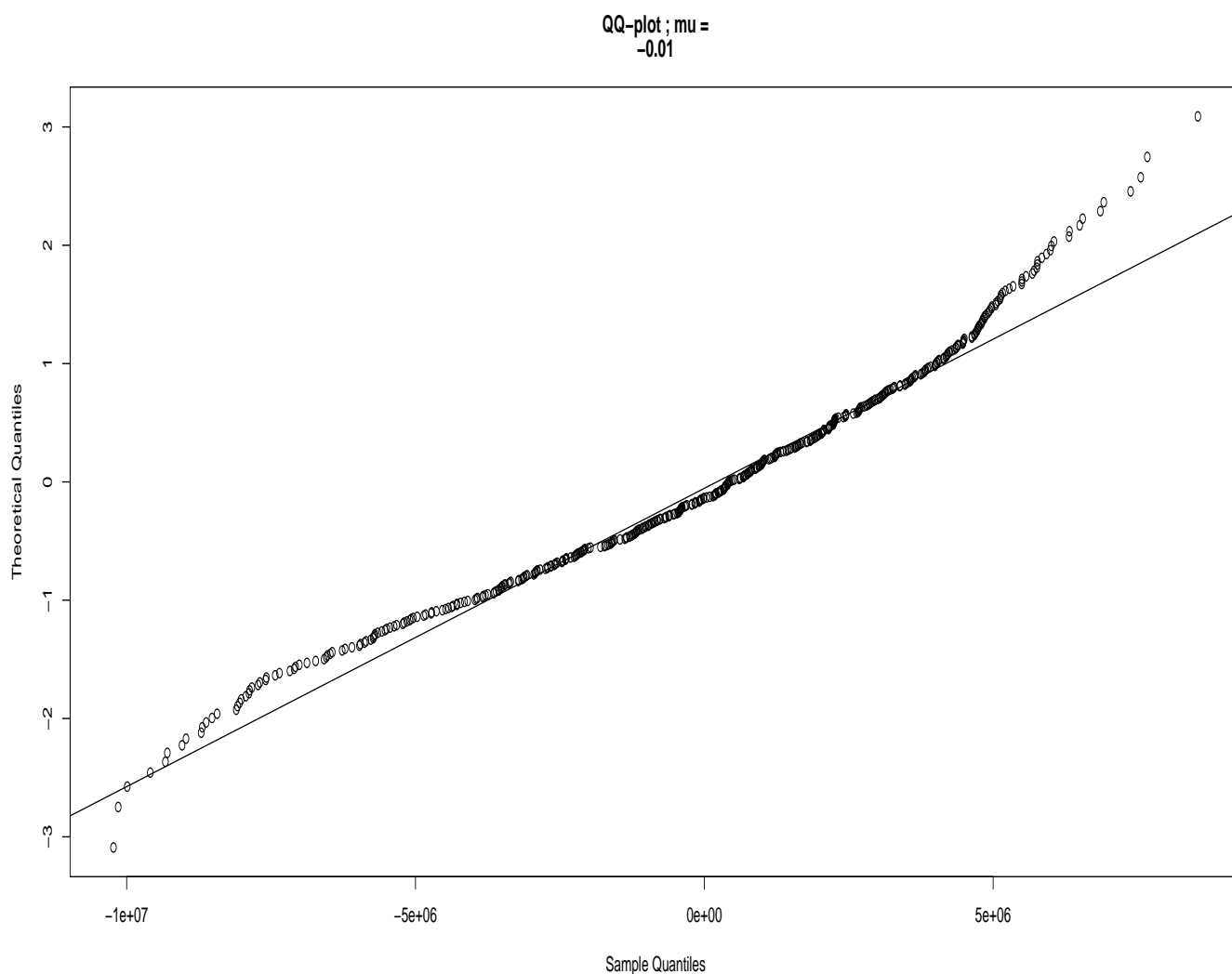


Illustration 14.4: QQ plot ;  $\mu = -0.01$ .

As the QQ-plot shows, the curve formed by empirical quantiles and the line of Gaussian quantiles does not match, so the distribution is not Gaussian.

Thus the influence of the additive term in  $X^*$ , i.e. the integral of a function of geometric Brownian motion, is not negligible for  $\mu = -0.01$ .

The same observations can be made for the case  $\mu = -0.04$ . Indeed:

14.2 Empirical distribution of provision

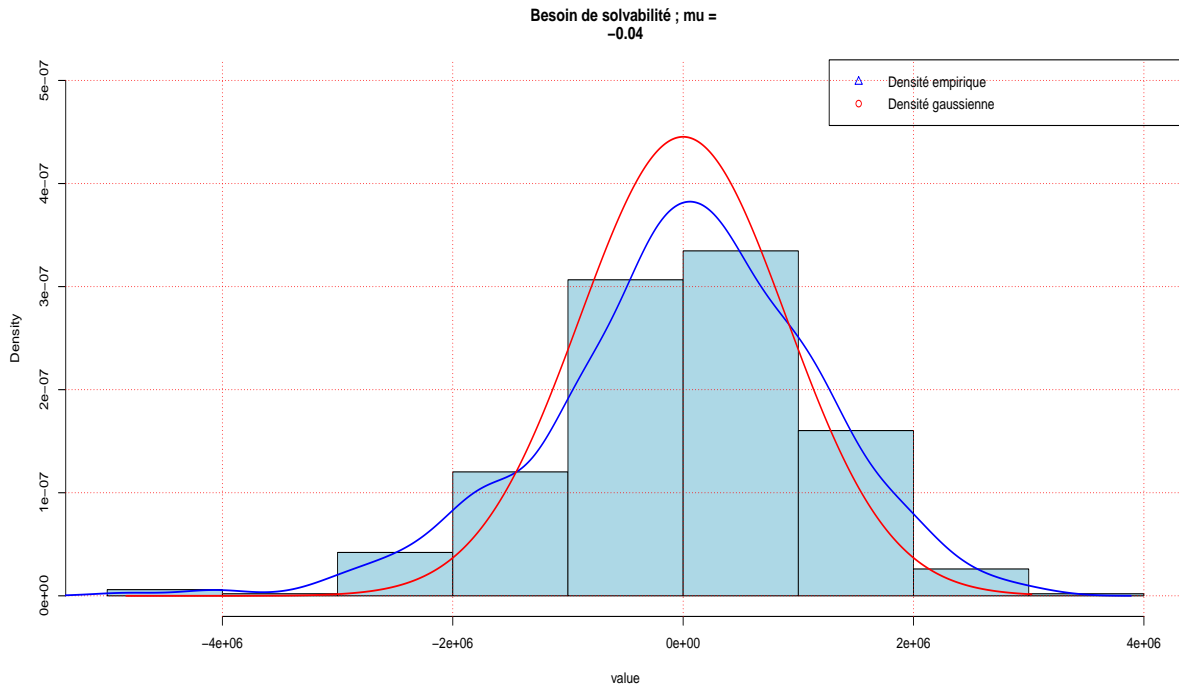


Illustration 14.5: Empirical distribution ;  $\mu = -0.04$ .

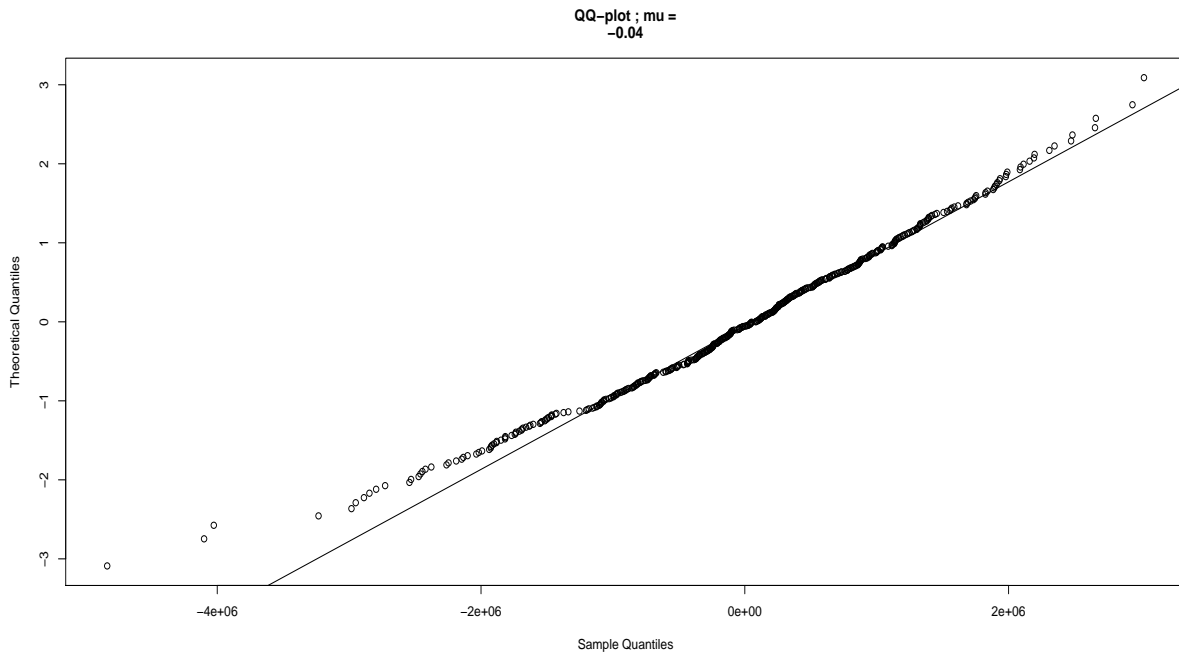


Illustration 14.6: QQ-plot ;  $\mu = -0.04$ .

In contrast, for  $\mu = -0.07$  and  $\mu = -0.1$  the density of centered provision is close to a Gaussian RV, as we can see on both following illustrations:

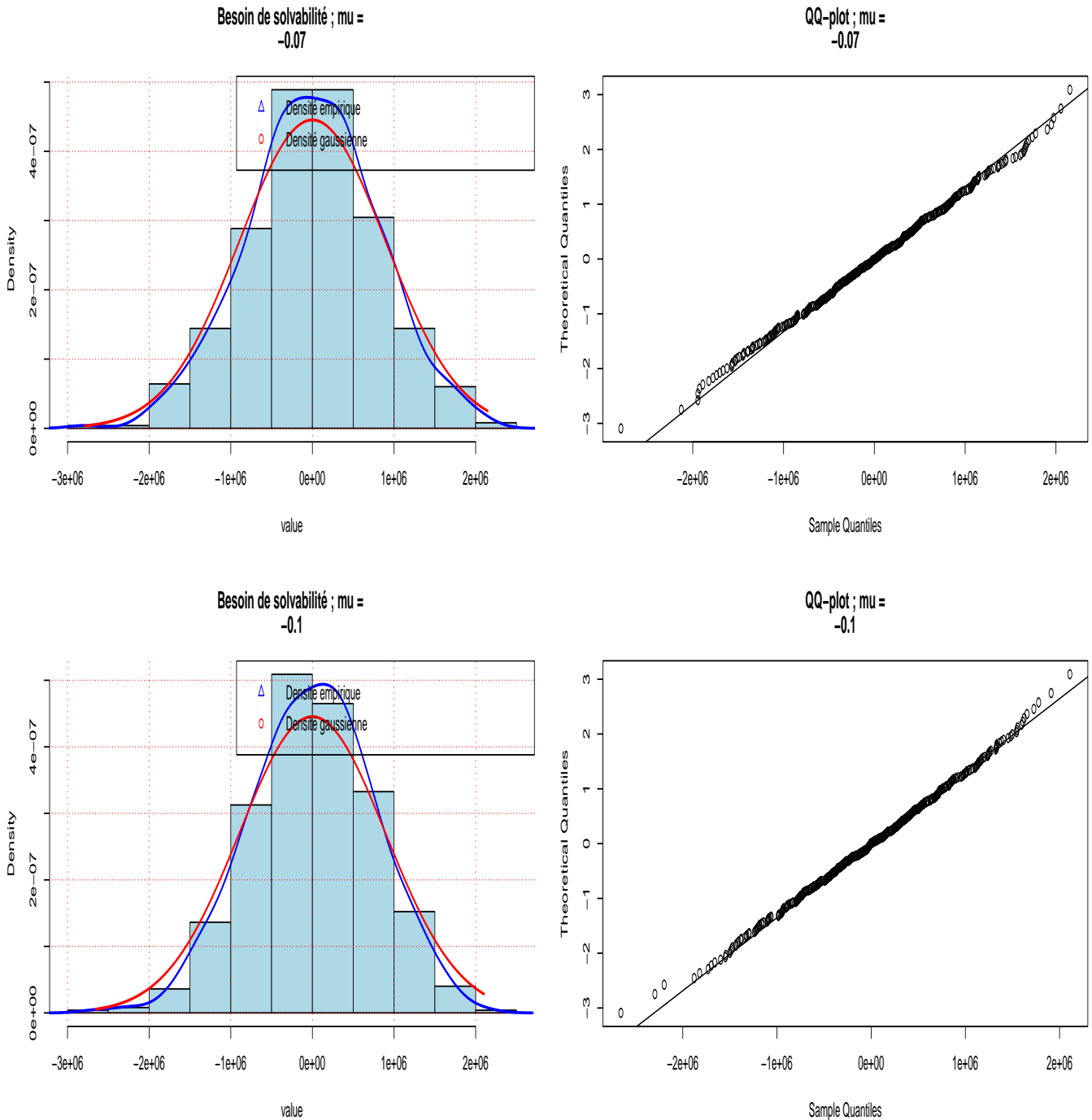


Illustration 14.7: Empirical distribution and QQ-plot ;  $\mu = -0.07$  and  $\mu = -0.1$ .

When  $\mu$  is negative (far from 0) the empirical distribution of  $P(h, R) - \mathbb{E}[P(h, R)]$  seems Gaussian.

Note that we hope that influence of  $\int_0^{+\infty} \inf(1, \frac{\gamma}{\alpha} R_t) dt$  in the limit law of  $P(h, R)$  is insignificant when  $\mu \rightarrow -\infty$ . Indeed:

$$\lim_{\mu \rightarrow -\infty} \int_0^{+\infty} \inf(1, \frac{\gamma}{\alpha} \exp[\mu t + \sigma B_t]) dt = 0. \quad (14.2.5)$$

So the additive term in the limit law of renormalized provision has no effect when  $\mu$  is very negative. The interest of the simulations is to point out that when  $\mu = -0.07$ , the influence of the integral of the Brownian motion is not significant.

### 14.3 Empirical solvency need

In this section we numerically evaluate the solvency need. As set out in Chapter 13, the asymptotic method does not allow us to calculate numerically the SN. Indeed the formula (13.2.28) is the following one:

$$SN^*(h) = h \left[ \mathcal{Q}_{X^*}(99.5\%) + \kappa \right] \sqrt{\frac{\lambda_2 \alpha^2 \mathbb{E}[(M_1)^2]}{2}}, \quad (14.3.1)$$

where  $\mathbb{E}[(M_1)^2]$  is the second moment order of the borrowed amount,  $\lambda_2, \alpha, \kappa$  are defined by (13.2.8), (13.2.6), (13.2.22), and  $\mathcal{Q}_{X^*}(99.5\%)$  (see (13.2.13)) is the 99.5% quantile of RV  $X^*$  which is defined by (14.2.2).

Because the law of  $\int_0^{+\infty} \inf(1, \frac{\gamma}{\alpha} R_t) dt$  is unknown, formula (14.3.1) does not allow for the numeric calculation of SN. Because  $h$  is large it is possible to make Monte Carlo simulation. Indeed we had seen in Chapters 11 and 12 that it is possible to simulate the provision: Algorithm 12.1 gives a method to make such simulations. We denoted by  $p^{\mu, l}$  the sample (size  $l = 5\ 000$ ) obtained with this algorithm, i.e.:

$$p^{\mu, l} = (p_1^\mu, \dots, p_l^\mu). \quad (14.3.2)$$

From Definition 13.2.2 and Remark 13.2.3 (see (13.2.14)), we define the empirical solvency need, denoted by  $SN_\mu^{emp}$ , as follows:

$$SN_\mu^{emp} := \inf \left\{ t : F_{p^{\mu, l}}(t) \geq 99.5\% \right\} - \frac{1}{l} \sum_{j=1}^l p_j^\mu. \quad (14.3.3)$$

where  $F_{p^{\mu, l}}(t)$  is the cumulative distributive function of the  $l$  sample.

We want to see how fluctuate the solvency need in function of  $\mu$ . To do this, we plot black curve  $(\mu, SN_\mu^{emp})$  for  $\mu \in [-1, 1]$ :

- $\mu$  is in the x axis,
- $BS_{\mu}^{emp}$  is in the y axis.
- The yellow vertical line represents the value of empirical solvency need for the parameter's value  $\mu$  estimated in Section 12.4.2, i.e.  $\mu = -0.07$ .

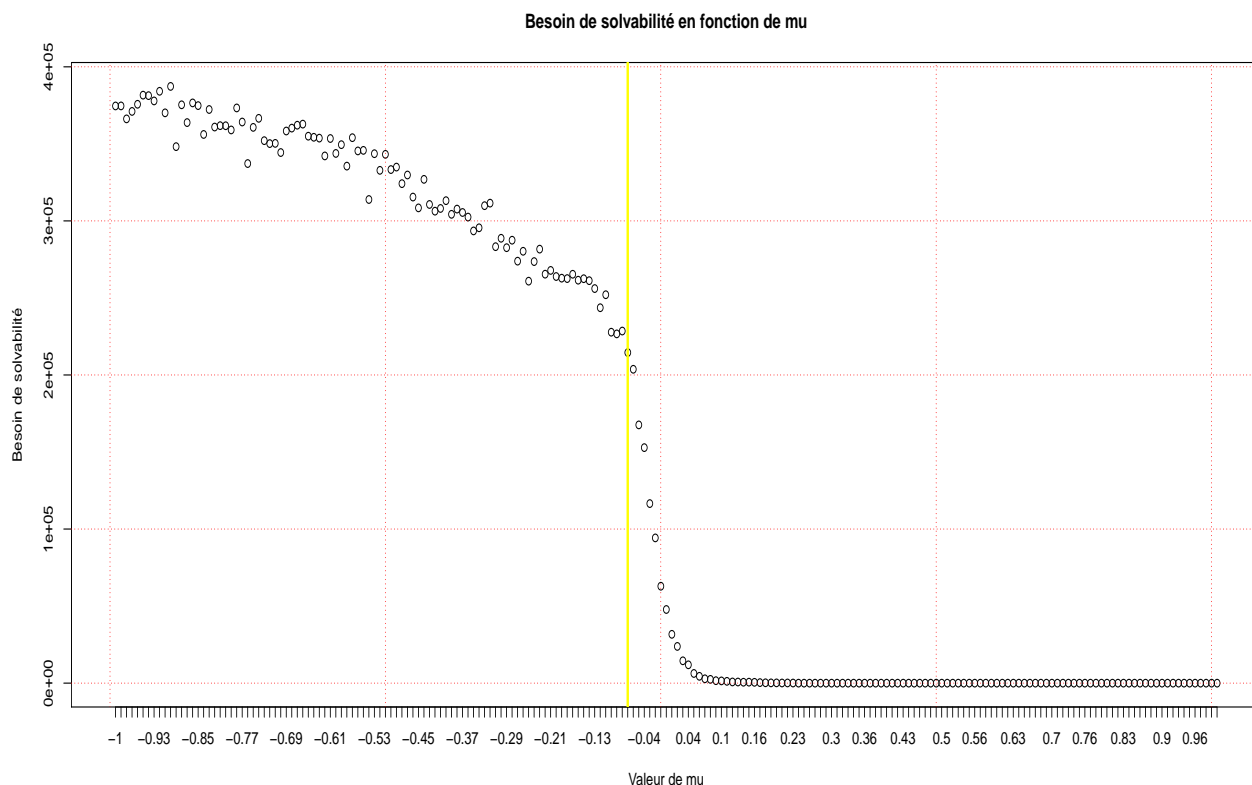


Illustration 14.8: Solvency need in function of  $\mu$ .

We distinguish three different regimes in the previous graph:

1. When  $\mu < 0$ : empirical solvency need approaches a positive constant when  $\mu$  becomes very negative.
2. When  $\mu$  is close to 0: empirical solvency need approaches 0. This is a transient state.
3. When  $\mu > 0$ : empirical solvency need requirement approaches 0 when  $\mu$  increases.





## Provisionnement en assurance non-vie pour des contrats à maturité longue et à prime unique - Application à la réforme Solvabilité 2

Nous considérons le cas d'un assureur qui doit indemniser une banque à la suite de pertes liées à un défaut de remboursement de ses emprunteurs. Les modèles couramment utilisés sont collectifs et ne permettent pas de prendre en compte les comportements individuels des emprunteurs. Dans une première partie nous définissons un modèle pour étudier le montant des pertes liées à ces défauts de paiement (provision) pour une période donnée. La quantité clé de notre modèle est le montant d'un défaut. Pour un emprunteur  $j$  et une date de fin de prêt  $T_j$ , ce montant vaut  $\max(S_{T_j}^j - R_{T_j}^j; 0)$ , où  $S_{T_j}^j$  est le montant dû par l'emprunteur et dépend de la durée et du montant du prêt, et  $R_{T_j}^j$  est le montant de la revente du bien immobilier financé par le prêt.  $R_{T_j}^j$  est proportionnel au montant emprunté; le coefficient de proportionnalité est modélisé par un mouvement Brownien géométrique et représente les fluctuations des prix de l'immobilier. La loi des couples (Date de fin du prêt, Durée du prêt) est modélisée par un processus ponctuel de Poisson. La provision  $P_h$ , où  $h$  est la durée maximale des contrats considérés, est alors définie comme la somme d'un nombre aléatoire de montants de défauts individuels. Nous pouvons ainsi calculer l'espérance et la variance de la provision mais aussi donner un algorithme de simulation. Il est également possible d'estimer les paramètres liés au modèle et de fournir une valeur numérique aux quantiles de la provision. Dans une deuxième partie nous nous intéresserons au besoin de solvabilité associé au risque de provisionnement (problématique imposée par la réforme européenne Solvabilité 2). La question se ramène à étudier le comportement asymptotique de  $P_h$  lorsque  $h \rightarrow +\infty$ . Nous montrons que  $P_h$ , convenablement normalisée, converge en loi vers une variable aléatoire qui est la somme de deux variables dont l'une est gaussienne.

**Mots clés:** *Risque de défaut; Provisionnement stochastique individuel; Processus ponctuel de Poisson; Mouvement Brownien géométrique; Temps de défaut; Quantile; Solvabilité 2 ; ORSA ; Besoin de solvabilité ; Convergence en loi ; Perpetual integral functional of Brownian motion.*

## Provisioning in non life insurance for contracts with long maturities and unique premium - Application to Solvency 2 reform

We consider an insurance company which has to indemnify a bank against losses related to a borrower defaulting on payments. Models normally used by insurers are collective and do not allow to take into account the personal characteristics of borrowers. In a first part, we defined a model to evaluate potential future default amounts (provision) over a fixed period. The amount of default is the key to our model. For a borrower  $j$  and an associated maturity  $T_j$ , this amount is  $\max(S_{T_j}^j - R_{T_j}^j; 0)$ , where  $S_{T_j}^j$  is the outstanding amount owed by the borrower and depends on the borrowed amount and the term of the loan, and  $R_{T_j}^j$  is the property sale amount.  $R_{T_j}^j$  is proportionate to the borrowed amount; the proportionality coefficient is modeled by a geometric Brownian motion and represents the fluctuation price of real estate. The couples (Maturity of the loan, Term of the loan) are modeled by a Poisson point process. The provision  $P_h$ , where  $h$  is the maximum duration of the loans, is defined as the sum of the random number of individual default amounts. We can calculate the mean and the variance of the provision and also give an algorithm to simulate the provision. It is also possible to estimate the parameters of our model and then give a numerical value of the provision quantile. In the second part we will focus on the solvency need due to provisioning risk (topic imposed by the European Solvency 2 reform). The question will be to study the asymptotic behaviour of  $P_h$  when  $h \rightarrow +\infty$ . We will show that  $P_h$ , well renormalized, converges in law to a random variable which is the sum of two random variables whose one is a Gaussian.

**Keys word:** *Default risk; Individual stochastic provisioning; Poisson point process; Geometrical Brownian motion; Time of default; Quantile; Solvency 2 ; ORSA ; Solvency need ; Convergence in law ; Perpetual integral functional of Brownian motion.*