



CRÉDIT AGRICOLE SA
DIRECTION DES RISQUES DU GROUPE
GROUPE DE RECHERCHE OPÉRATIONNELLE¹

Modélisation du risque de crédit

DEA de Statistique et Modèles aléatoires en économie et finance
Université Paris 7 — Université Paris 1

David KURTZ² & Thomas B. PIGNARD

Mots-clefs : Risque de crédit, modèles structurels, modèles à forme réduite, copules, produits dérivés de crédit.

¹ Le Centorial, 18 rue du 4 septembre, 75002 Paris.

²david.kurtz@creditlyonnais.fr

Hey, teacher, leave those kids alone!

Pink Floyd, "The Wall" (1979)

REMERCIEMENTS

Merci à Thierry Roncalli et Romain Camus pour leur aide.

Table des matières

1	Introduction	9
1.1	Zoologie des risques	10
1.2	Enjeux	11
1.3	Modèles du risque de crédit	11
2	Produits dérivés de crédit	15
2.1	Le marché des produits dérivés de crédit	15
2.2	Les obligations risquées	17
2.3	Credit default swap	18
2.4	Les Basket Default Swaps	21
2.5	Collateralized Debt Obligations	23
3	Introduction aux modèles structurels	27
3.1	Modèle de Merton	27
3.2	Modèles de premier instant de passage	35
4	Modèles à forme réduite	41
4.1	Préliminaires mathématiques	41
4.2	Évaluation des actifs risqués	47
5	Défauts corrélés	53
5.1	Produits dérivés sur un panier de crédits	54
5.2	Corrélation dans les modèles structurels	56
5.3	Corrélation des intensités	57
5.4	Corrélation des instants de défaut	58
6	Modèle Hybride : le cas des obligations convertibles	65
6.1	Caractéristiques des obligations convertibles	65
6.1.1	Définition	65
6.1.2	Utilisation des obligations convertibles	66

6.1.3	Obligations convertibles et options	66
6.1.4	Les déterminants du contrat	67
6.1.5	Influence des variables de marché sur le prix	67
6.2	Modélisation du risque de crédit	68
6.3	Implémentation des modèles	69
6.3.1	Les arbres multinomiaux	70
6.3.2	Les équations aux dérivées partielles	71
6.3.3	Synthèse des obligations convertibles	72
A	Théorie générale des processus	77
A.1	Généralités	77
A.2	Temps d'arrêt et tribus associées	78
A.3	Les tribus optionnelles et prévisibles	79
A.4	Temps d'arrêt prévisibles	79
A.5	Classification des temps d'arrêt	80
B	Formule d'Itô avec sauts	83
B.1	Processus à variation finie	83
B.2	Les semimartingales simples	87
C	Rappel sur les diffusions	93
C.1	Definitions	93
C.2	Diffusions avec sauts	94
C.3	Diffusion affine avec des sauts	96
D	Les fonctions copules	99
D.1	Definitions et exemples	99
D.2	Les copules archimédiennes	101
E	Démonstration du Lemme 3.1	105

CHAPITRE 1

Introduction

Ces notes constituent le support d'un cours de DEA concernant la modélisation du risque de crédit. Les deux principales sources qui ont servi de base à l'écriture de ce texte sont [11] et [2].

Le but de ce cours est de présenter quelques outils et concepts qui peuvent servir de base à la modélisation du risque de crédit ; il est plus spécifiquement orienté vers la valorisation des produits dont la valeur dépend du risque de crédit auquel il est exposé. Les modèles et techniques présentés peuvent néanmoins être utilisés aussi bien pour la mesure que pour l'évaluation du risque de crédit ¹. Les avancées théoriques autour de la modélisation du risque de crédit est l'un des facteurs qui ont rendu possible l'essor d'un marché des produits dérivés de crédit.

Le développement du marché des produits dérivés de crédit, qui a atteint cette année un encours total d'environ 2300 milliards de dollars, a révolutionné en retour la gestion du risque de crédit ainsi que l'ingénierie financière qui lui est liée. Les gestionnaires de portefeuille et les investisseurs disposent désormais d'instruments financiers permettant le transfert efficace du risque de crédit. D'après Alan Greenspan, la diffusion² des risques que génère ce marché des produits dérivés de crédit augmenterait la résistance de l'économie mondiale aux chocs systémiques comme l'aurait prouvé son bon comportement face aux récentes faillites d'Enron et de WorldCom.

Dans ce chapitre introductif et après avoir rappelé quels sont les principaux risques auxquels sont confrontés les institutions financières, nous expliquons quels sont les enjeux de la modélisation du risque de crédit. Nous présentons alors, brièvement, les deux grandes classes de modèles du risque de crédit : les modèles structurels et les modèles à forme réduite.

¹Autrement dit, ces techniques servent aussi bien sous la probabilité historique que sous une probabilité risque-neutre.

²et donc la mutualisation

1.1 Zoologie des risques

On distingue traditionnellement quatre grands types de risques financiers [11] :

1. Le risque de marché.

Le risque de marché peut se définir comme le risque de perte lié aux variations des conditions de marché (prix, taux, taux de change, volatilités, etc...)

2. Le risque de crédit.

Le risque de crédit est défini comme le risque de perte lié à l'évolution de la qualité de la signature d'un émetteur. On peut distinguer deux types de risque de crédit : le risque de contrepartie et le risque de référence. Pour un émetteur donné, ce risque peut se matérialiser sous la forme

- du changement de sa note (*upgrade* ou *downgrade*) telle celle émise par les grandes agences de notations Moody's et Standard & Poor's,
- d'une variation de son spread de crédit,
- d'un événement de crédit (*credit event*) tel le défaut de paiement ou la restructuration de sa dette.

Ces trois risques sont, bien évidemment, corrélés. Une augmentation brutale du niveau du spread émetteur augmente la probabilité d'un événement de crédit. De la même manière, un changement de notation influe fortement sur la probabilité du défaut d'un émetteur.

Lorsque A entre en relation avec une contrepartie B via un instrument financier, il peut être soumis au risque que B soit dans l'impossibilité d'honorer ses engagements. Par exemple, si A est en possession d'une obligation émise par B, il court le risque qu'à maturité B ne puisse lui rembourser le capital investi. On parle dans ce cas de *risque de contrepartie unilatéral* puisque B n'est pas soumis au risque de crédit de A. Si A et B sont les deux contreparties d'un swap, ils sont tout deux soumis au risque de contrepartie : on parle alors de *risque de contrepartie bilatéral*.

Supposons maintenant que la qualité de la signature des contreparties A et B soit de qualité infinie (de sorte que le risque de contrepartie bilatéral soit nul). Les parties A et B peuvent entrer dans un contrat qui fait intervenir le risque de crédit d'une troisième contrepartie C³. Le risque de crédit associé à C est appelé *risque de référence*. Nous verrons que le but des produits dérivés de crédit est le transfert de ce risque de référence.

La distinction entre ces risques n'est pas toujours aisée : le risque de crédit lié à la variation des spread de credit default swap peut être considéré comme un risque de marché. Les portefeuilles de produits dérivés OTC sont, bien évidemment, soumis aux risques de marché mais ils sont aussi exposés au risque de contrepartie.

3. Le risque de liquidité.

Il s'agit, pour une entreprise, du risque de ne pas pouvoir mobiliser à un instant donné assez de liquidités pour pouvoir faire face à ses engagements. Voir [11] pour un supplément d'information.

4. Le risque opérationnel.

³un *credit swap* dont le payoff dépend de l'occurrence d'un événement de crédit est un exemple de tel contrat.

Dans cette catégorie sont regroupés, par exemple, les risques de fraude, d'erreurs des opérateurs, de pannes des systèmes, etc... Pour plus de détails, voir [26].

1.2 Enjeux

Le risque de crédit peut être défini, en première approximation, comme le risque de perte lié au changement de la qualité de la signature d'une contrepartie. Toutes les institutions financières (ainsi que tout les acteurs du marché) accumulent une grande quantité de risque de crédit : soit directement par l'intermédiaire de leurs portefeuilles de créances, soit indirectement sous la forme de risques de contrepartie dans leurs portefeuilles d'actifs et de produits dérivés OTC.

L'enjeu que représente la modélisation de ce risque est donc très important : il s'agit de pouvoir

- mesurer le risque de crédit contenu dans les portefeuilles,
- évaluer les instruments financiers sensibles au risque de crédit et, plus généralement, tout instrument exposé à ce risque (risque de contrepartie),

et ce à un niveau agrégé. Il est, en effet, important de pouvoir contrôler l'exposition au risque de crédit contrepartie par contrepartie ainsi que l'évolution de cette exposition par secteurs géographiques et industriels. De telles pratiques permettent, par exemple, de réduire le risque de concentration.

La production de résultats quantitatifs robustes permet alors à l'institution concernée

- d'allouer à chaque centre de profit un capital économique adéquat,
- d'évaluer la performance des centres de profit au regard des risques pris,
- de fournir des informations fiables sur son intégrité financière aux régulateurs, aux investisseurs et aux agences de notation,
- de diversifier et réduire le risque en imposant, par exemple, des limites à l'exposition au risque de crédit par contrepartie.

La mise en place de tels processus de gestion du risque correspond à l'évolution de la réglementation prudentielle et aux trois piliers du nouvel accord de Bâle (sur ce sujet, nous renvoyons le lecteur à [26] pour plus de détails).

1.3 Modèles du risque de crédit

Dans cette section, nous exposons les approches classiques de la modélisation du risque de défaut à partir de l'étude du concept central d'obligation zéro-coupon risquée (c'est-à-dire soumise au risque de défaut). Il existe deux grandes familles de modèles d'évaluation de la dette risquée :

- les modèles structurels,
- les modèles à forme réduite.

Nous mentionnons aussi les modèles de corrélation d'instant de défaut qui sont nécessaires à l'évaluation des produits dérivés exotiques de crédit ainsi que les modèles dit hybrides dont le but est de permettre l'évaluation précise du risque de contrepartie des portefeuilles de produits dérivés OTC.

Dans la suite, nous nous plaçons dans le cadre de l'évaluation risque-neutre des actifs financiers : nous supposons donnés un espace probabilisé filtré $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})$ sur lequel est défini le processus des taux d'intérêt instantanés $(r_t)_{t \geq 0}$ et une probabilité risque-neutre \mathbb{P}^* . Rappelons qu'alors la valeur des actifs contingents⁴ est calculée comme l'espérance sous cette probabilité de ses flux futurs actualisés au taux sans risque.

Une obligation zéro-coupon (sans risque) est un actif qui paye 1 à sa maturité T . La valeur $B(t, T)$ de cet actif est

$$B(t, T) = \mathbb{E}_t^* \left[\exp \left(- \int_t^T r(s) \, ds \right) \right].$$

Si le risque de contrepartie de l'émetteur du zéro-coupon n'est pas nul, l'évaluation du zéro-coupon doit tenir compte de la possibilité du défaut de celui-ci : deux nouveaux risques entrent en jeu

- l'instant du défaut,
- la perte en cas de défaut (*Loss Given Default*).

La perte en cas de défaut s'exprime en terme d'un taux de recouvrement δ éventuellement aléatoire et d'une hypothèse de recouvrement. Noter que ces hypothèses de recouvrement, que nous présentons maintenant, s'étendent immédiatement à d'autres actifs que les zéro-coupons risqués. Nous notons $D(t, T)$ la valeur en t du zéro-coupon risqué de maturité T et τ l'instant du défaut de l'émetteur de ce titre.

L'hypothèse de recouvrement la plus courante s'appelle *fractional recovery of par value* et consiste en le recouvrement à l'instant du défaut d'une fraction δ du nominal du titre. Dans ce cas, on a

$$D(t, T) = \mathbb{E}_t^* \left[e^{-\int_t^T r(s) \, ds} I_{\{\tau > T\}} + e^{-\int_t^\tau r(s) \, ds} \delta I_{\{t < \tau \leq T\}} \right].$$

Si, en cas de défaut, la fraction recouvrée l'est à maturité, on parle de *fractional recovery of Treasury value*. Dans ce cas, on a

$$D(t, T) = \mathbb{E}_t^* \left[e^{-\int_t^T r(s) \, ds} (I_{\{\tau > T\}} + \delta I_{\{t < \tau \leq T\}}) \right].$$

Une autre hypothèse classique dite *fractional recovery of market value* suppose qu'à l'instant de défaut le propriétaire du titre perçoit une portion δ de la valeur précédant immédiatement le défaut. Dans ce cas

$$D(t, T) = \mathbb{E}_t^* \left[e^{-\int_t^T r(s) \, ds} I_{\{\tau > T\}} + e^{-\int_t^\tau r(s) \, ds} \delta D(\tau-, T) I_{\{t < \tau \leq T\}} \right],$$

où $D(\tau-, T)$ ⁵ est la valeur avant défaut de l'instrument considéré.

Modèles structurels

L'approche structurelle de l'évaluation de la dette risquée a été initiée par Merton (1974) [21] et est basée sur la modélisation de l'évolution du bilan de l'entreprise. Dans

⁴c'est-à-dire dont la valeur est contingente à l'évolution des conditions de marché.

⁵Si f est une fonction à valeurs réelles, nous notons $f(t-)$ la limite $\lim_{s \uparrow t, s < t} f(s)$.

ce cadre, le défaut intervient si l'émetteur de la dette est dans l'impossibilité d'honorer ses engagements et les zéro-coupons risqués apparaissent comme des produits dérivés sur la valeur de l'entreprise émettrice de cette dette. La qualité d'une signature est alors fonction de trois variables fondamentales : sa valeur, la variabilité de ses actifs⁶ et son levier d'endettement (qui est le rapport entre le niveau de sa dette et sa valeur totale).

Les modèles structurels sont largement utilisés par les praticiens. Pour s'en convaincre, il suffit de mentionner l'entreprise Moody's KMV⁷ qui a développé une offre complète de services financiers basés sur un tel modèle. Ainsi, ils proposent à leurs clients des outils d'analyse et leur fournissent des données (telles des estimations de probabilités de défaut) obtenues à partir de leur modèle [4].

Dans le chapitre 3, nous étudions le modèle de Merton ainsi que son extension aux modèles dits *de premier instant de passage*. Nous mettrons en avant les liens existant entre le risque equity et le risque de crédit.

Modèles à forme réduite

Cette approche de l'évaluation de la dette risquée qui remonte à [24] a été popularisée par de nombreux travaux et en particulier l'article [10].

Dans les modèles dits *à forme réduite*, et contrairement aux modèles structurels, le défaut est considéré comme un événement imprévisible dont la loi est gouvernée par un processus stochastique appelé *intensité d'arrivée* ou *taux de hasard*. L'exemple le plus simple d'un tel modèle est celui où l'instant de défaut τ est défini comme le premier instant d'arrivée d'un processus de Poisson d'intensité λ ⁸. Dans ce cas, on a

$$\mathbb{P}[\tau > t] = e^{-\lambda t}, \quad \mathbb{E}[\tau] = \frac{1}{\lambda}, \quad \mathbb{P}[\tau \in (t, t + \Delta t) | \tau > t] = \lambda \Delta t + o(\Delta t).$$

En pratique, on peut faire dépendre cette intensité de défaut d'un certain nombre de variables économiques (tels les taux d'intérêt) et/ou de variables liées à l'entreprise (telle sa notation). Noter que ces modèles sont couramment utilisés pour l'évaluation des produits dérivés de crédit.

Dans le chapitre 4, nous généralisons considérablement le modèle poissonnien : nous traitons le cas général où l'intensité de défaut est un processus stochastique en donnant un sens précis à l'expression

$$\mathbb{P}_t[\tau \in (t, t + dt) | \tau > t] = \lambda_t dt.$$

Nous verrons que l'intensité de défaut λ est intimement liée au spread émetteur court-terme s . Plus précisément, nous montrerons que sous l'hypothèse de recouvrement *fractional recovery of market value*, on a

$$s(t) = (1 - \delta)\lambda(t),$$

⁶Par exemple, cette volatilité est plus forte pour une entreprise travaillant dans le secteur des nouvelles technologies que pour une entreprise des secteurs traditionnels.

⁷www.moodyskmv.com.

⁸Autrement dit, τ suit une loi exponentielle de paramètre $\lambda > 0$.

où δ est le taux de recouvrement de l'obligation considérée. Cette égalité fondamentale est connue sous le nom d'égalité du triangle.

Défauts corrélés

L'évaluation des produits dérivés de crédit exotiques⁹ tels les CDO¹⁰ et les TDP¹¹ nécessitent la modélisation de la structure de dépendance qui existe entre les instants de défaut (τ_1, \dots, τ_N) de N entreprises.

Il existe, pour l'essentiel, deux manières de modéliser cette dépendance :

- en choisissant une structure de co-dépendance pour les rendements instantanés des processus de valeur des entreprises,
- en spécifiant la corrélation entre les intensités des instants de défaut.

Ces questions seront abordées en détail dans le chapitre 5. Nous y présentons aussi quelques applications concrètes de ces techniques.

Modèles hybrides

L'évaluation du risque de contrepartie ou celle de produits mixtes¹² nécessite la mise en place de modèles prenant en compte l'évolution jointe de différents facteurs de risque : les taux d'intérêt, les taux de change, les cours des actions, l'intensité de défaut des entreprises. Nous donnons un exemple de tels modèles dans le chapitre 6 consacré aux obligations convertibles.

⁹Par opposition aux produits vanilla que sont les Credit Default Swap.

¹⁰Collateralized Debt Obligations.

¹¹Tranche Default Product.

¹²c'est-à-dire dont la valeur dépend de plusieurs facteurs de risques.

CHAPITRE 2

Produits dérivés de crédit

Dans ce chapitre, nous présentons les principaux dérivés de crédit. Nous y décrivons leurs *payoffs* et donnons quelques indications sur leur intérêt financier.

Après un rapide état des lieux du marché des dérivés de crédit, nous décrivons les caractéristiques du Credit Default Swap (CDS), qui constitue le produit à la vanille de ce marché. Les CDS sont à la base de produits exotiques plus complexes, et sont utilisés à la fois en tant que produits de couverture (diversification du risque, couverture d'un risque en préservant la relation commerciale, réduction du risque de concentration, transfert du risque de crédit (balance sheet CDO),...), et d'investissement (long d'un risque de crédit à l'aide d'un instrument hors-bilan, effets de levier, création de position de crédit synthétique (CDO synthétiques),...). Enfin, nous présentons les produits de corrélation les plus courants : les n^{th} -to default, et les Collateralized Debt Obligations (CDO).

Dans la suite, r désigne le taux sans risque, B le processus $t \mapsto \exp(\int_0^t r(s) ds)$ et τ l'instant de défaut d'une entreprise.

2.1 Le marché des produits dérivés de crédit

Le marché des produits dérivés connaît une croissance exponentielle depuis le début des années 1990. Pour fixer les idées, précisons tout de suite que le nominal total des encours sur produits dérivés de crédit est de 2306 milliards de dollars (Risk Magazine 2003) en augmentation de plus de 50% par rapport à l'année dernière. La standardisation des CDS est devenue une réalité grâce aux nouvelles normes et définitions mises en place par l'ISDA¹ [16]. L'année 2003 a été marquée par les faits suivants : généralisation de l'utilisation des produits synthétiques, accroissement de la liquidité sur les produits de corrélation (cotation bid-ask de tranches synthétiques), croissance

¹International Swap & Derivatives Association.

du marché des credit default swaptions (credit option). Ils restent, cependant, de nombreux problèmes à résoudre. Citons, par exemple, le besoin de liquidité sur le court et le long terme de la courbe de crédit qui se fait parfois sentir et les méthodologies de calibration de taux de recouvrement qui demandent à être améliorées.

L'essentiel de l'encours se répartit sur : les credit default swaps (73%) et les produits sur paniers de crédit (22%), en particulier, les n^{th} -to-default swap à hauteur de 0.3% et toutes les transactions synthétiques tels les CDO (Collateralized Debt Obligations) et les TDP (Tranche Default Product). Le reste de l'encours est constitué de Credit Linked Notes, de Total Rate of Return et de Spread Option.

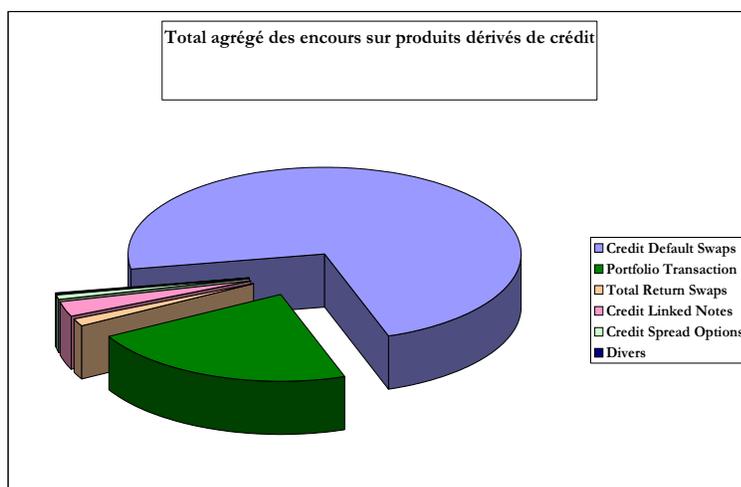


FIG. 2.1 – Répartition des encours sur produits dérivés de crédit.

Le marché vanilla (celui des credit default swaps) est essentiellement concentré sur l'Amérique du nord et l'Europe. Le point de plus grande liquidité est celui des CDS de maturité comprise entre 4 et 6 ans sur des signatures *investment grade*. Cette répartition géographique est intéressante compte tenu du fait qu'elle ne correspond pas à celle des obligations (la proportion d'obligations européennes est significativement plus faible). La croissance de ce marché est en partie due à la demande toujours plus forte de produits permettant de couvrir les positions synthétiques (CDO).

Les banques sont les principales utilisatrices de produits dérivés de crédit. Ceci est dû à leur utilisation massive des CDS pour couvrir leurs portefeuilles de créances et leurs positions synthétiques². L'activité de couverture engendrée par l'émission de CDO synthétiques a, pour la première fois, permis de satisfaire à la demande d'achat de protection provenant des gestionnaires de portefeuille de créances. Les principaux investisseurs dans les positions synthétiques restent les compagnies d'assurance : elles détiennent 65% des TDP et 81% des CDO de bilan³. Les *hedge funds* participent désormais activement à ce marché : ils sont, par exemple, régulièrement acheteurs de

²Une position synthétique est une position de crédit courte ou longue (c'est-à-dire acheteuse ou vendeuse de protection) obtenue à partir d'un portefeuille de CDS donc sans prêt effectif de capital.

³Ces produits sont émis par les banques qui veulent couvrir le risque de crédit de leur portefeuille de créances en vue de réduire leur charge en capital économique.

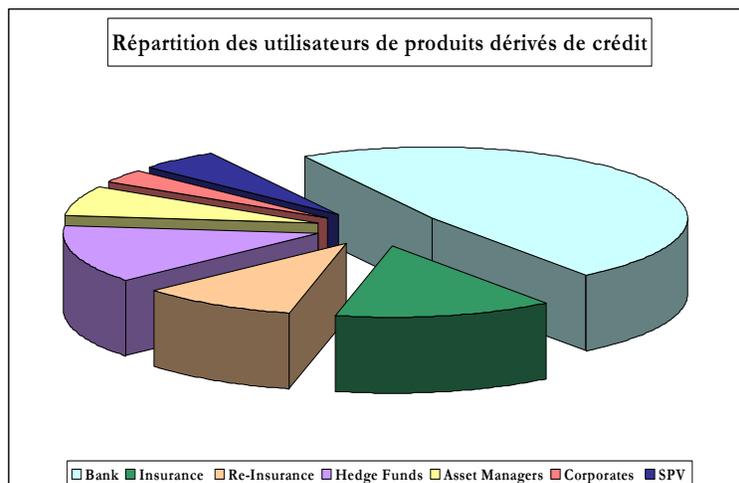


FIG. 2.2 – Répartition des utilisateurs de produits dérivés de crédit.

CDS pour leur activité d'arbitrage d'obligations convertibles.

2.2 Les obligations risquées

Outre les prêts traditionnellement accordés aux entreprises, une part importante du marché de la dette risquée consiste en obligations émises par les entreprises et certains états. Contrairement aux prêts, les obligations s'échangent sur les marchés organisés et sont ainsi soumis au risque de marché (risque de taux), au risque de crédit (risque de contrepartie) et, dans une certaine mesure, au risque de liquidité.

Le rendement de ces obligations est, en général, supérieur au rendement d'obligations identiques dont le risque de contrepartie est considéré comme nul (par exemple, les obligations du Trésor pour les pays de l'OCDE). La différence de rendement ou *spread* est une prime demandée par le marché pour prendre en charge les risques de contrepartie et de liquidité⁴ inhérents à l'instrument. On distingue, pour un émetteur donné, le *spread* calculé à partir d'obligations à taux fixe de celui coté dans les obligations à taux variables.

Noter que les obligations peuvent présenter des caractéristiques particulières : elles peuvent être rappelés par l'émetteur avant maturité (callable bond), être convertible en actions (convertible bond).

⁴En pratique, il est difficile de modéliser le risque de liquidité. On considère souvent, en première approximation, qu'il est pris en compte dans le *spread* de crédit.

2.3 Credit default swap

Description du produit

Le credit default swap (CDS) est le plus simple des produits dérivés de crédit et doit être considéré comme la brique de base (ou le sous-jacent) des produits dérivés plus exotiques.

Le CDS permet le transfert de risque de crédit de référence d'une entreprise C (*entité de référence*) entre deux contrepartie A et B. Dans le contrat standard, l'une des parties en question, disons A, achète une protection contre le risque de perte en cas de défaut de l'entité de référence C. Ce défaut est déclenché par un événement de crédit formel spécifié dans le contrat. Cet événement peut être la faillite de l'entreprise, un défaut de paiement ou la restructuration de sa dette.

La protection est valable jusqu'à la maturité du swap. En échange de cette protection, l'acheteur A verse périodiquement (en général, tous les 3 mois) au vendeur B une prime et ce jusqu'au défaut de C ou jusqu'à maturité du swap. La jambe du swap correspondante est appelé *premium leg*.

Si le défaut intervient avant la maturité du swap, le vendeur de protection effectue un paiement à l'acheteur de protection. Ce paiement équivaut à la différence entre le nominal de la dette couverte par le swap et le taux de recouvrement observé à l'instant du défaut. Ce paiement peut être effectué selon deux modalités : *physical settlement* ou *cash settlement*. Dans le premier cas, l'acheteur de protection A livre au vendeur de protection B un nombre d'obligations émises par C correspondant au nominal du swap et reçoit en retour le nominal du swap payé en cash. Dans le deuxième cas, un paiement en cash qui correspond au pair moins le taux de recouvrement est effectué par le vendeur de protection B vers l'acheteur de protection A. Ce taux de recouvrement est calculé à partir de cotations obtenues quelques temps après que le défaut se soit produit.

La *prime* (aussi appelé spread ou marge) qui annule la valeur du CDS est dite à *la monnaie* (fair margin ou fair spread ou simplement spread).

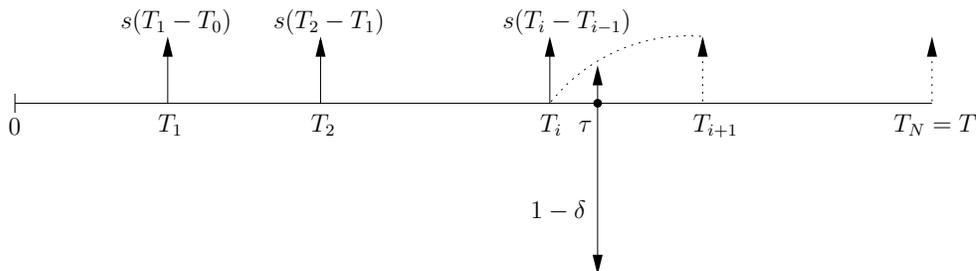


FIG. 2.3 – Un credit default swap.

Soient T_0 la date d'entrée dans le swap, T sa maturité, $T_0 < T_1 < \dots < T_N = T$ les dates de paiements, δ le taux de recouvrement en cas de défaut et s la valeur de

la marge. Le payoff (payé en T) correspondant à la jambe fixe s'écrit

$$\sum_{i=0}^{N-1} \left\{ (T_{i+1} - T_i) \frac{B_T}{B_{T_{i+1}}} I_{\{\tau > T_i\}} + (\tau - T_i) \frac{B_T}{B_\tau} I_{\{T_{i+1} \geq \tau > T_i\}} \right\}.$$

tandis que le payoff associé à la jambe variable s'écrit

$$(1 - \delta) \frac{B_T}{B_\tau} I_{\{\tau \leq T\}}.$$

Évaluation de la marge d'un CDS

Un raisonnement simple d'absence d'opportunité d'arbitrage permet d'obtenir une première approximation du spread (fair margin) d'un credit default swap. Nous notons C l'entité de référence.

Considérons les deux portefeuilles suivants :

$$P_1 \left\{ \begin{array}{l} \text{long d'une obligation à taux variable émise par } C \text{ de spread } U \\ \text{court d'une obligation à taux variable sans risque} \end{array} \right\}$$

$$P_2 \left\{ \text{court d'un credit default swap sur } C \text{ de spread } S \right\}$$

Nous supposons que toutes les obligations et le CDS considérés ont même maturité, même dates de tombée de coupon et même nominal. Nous supposons aussi que le défaut ne peut intervenir qu'immédiatement après une tombée de coupon.

Les cash-flows générés par le portefeuille P_1 sont décrits dans la figure 2.4. Ils correspondent à une position acheteuse de protection sur un CDS de spread U . Les deux portefeuilles P_1 et P_2 étant à coût d'entrée nul, on a nécessairement par absence d'opportunité d'arbitrage $U = S$.

Ainsi, en première approximation, la fair margin d'un CDS est égale au spread d'une obligation à taux variable (FRN = floating rate note) de même maturité et ayant les mêmes dates de tombée de coupons.

REMARQUE. L'égalité précédente n'est pas toujours observée dans la pratique. Ceci est dû au fait que les hypothèses implicites à notre raisonnement⁵ ne sont pas toujours vérifiées.

REMARQUE. Ce genre de raisonnement n'est valide que pour déterminer la marge à la monnaie d'un CDS. Pour déterminer sa valeur en cours de vie (NPV = net present value) qui n'est plus nécessairement égale à 0, il faut avoir recours à un modèle.

EXEMPLE. (Pricing d'un CDS.)

Nous considérons un CDS de maturité T . Nous supposons que le taux de recouvrement en cas de défaut δ , le taux d'intérêt sans risque r et le taux de défaut $\lambda > 0$ sont constants. L'instant de défaut τ suit une loi exponentielle de paramètre λ :

$$\mathbb{P}[\tau > t] = e^{-\lambda t}.$$

⁵telle l'absence de coût de transaction (spread bid-ask) sur le marché du cash et des obligations à taux fixes

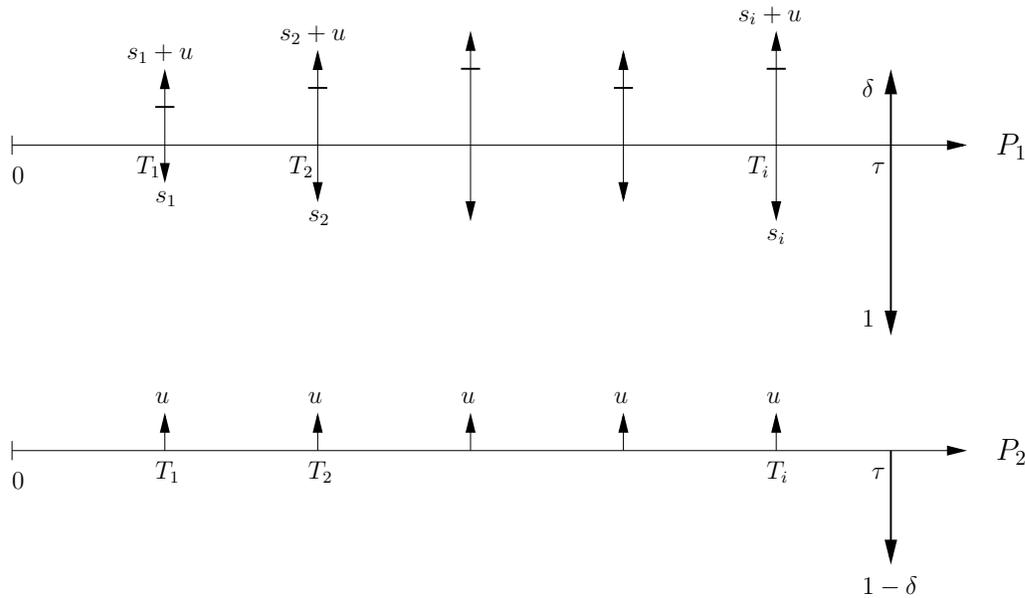


FIG. 2.4 – Synthèse d'un CDS.

On cherche à calculer la marge s^* qui annule la valeur du CDS à l'origine. En considérant que la prime est payée jusqu'au défaut, et sans tenir compte du coupon couru, la valeur de la jambe fixe s'écrit :

$$\begin{aligned}
 JF(s) &= s \sum_{i=0}^{N-1} (T_{i+1} - T_i) \mathbb{E} \left[e^{-rT_{i+1}} \mathbb{1}_{\{\tau > T_{i+1}\}} \right] \\
 &= s \sum_{i=0}^{N-1} (T_{i+1} - T_i) e^{-(r+\lambda)T_{i+1}} \\
 &\simeq s \int_0^t e^{-(r+\lambda)u} du \\
 &= s \frac{1 - e^{-(r+\lambda)T}}{r + \lambda} .
 \end{aligned}$$

En supposant que le flux variable est payé à l'instant τ d'occurrence du défaut, la valeur de la jambe variable est :

$$\begin{aligned}
 JV &= (1 - \delta) \mathbb{E} \left[e^{-r\tau} \mathbb{1}_{\{\tau \leq T\}} \right] \\
 &= (1 - \delta) \int_0^T \lambda e^{-(r+\lambda)u} du \\
 &= (1 - \delta) \frac{\lambda}{r + \lambda} \left(1 - e^{-(r+\lambda)T} \right).
 \end{aligned}$$

Le *fair spread* s^* est tel que $JF(s^*) = JV$ soit

$$s^* = (1 - \delta)\lambda.$$

Cette égalité est connue sous le nom d'*égalité du triangle*

EXERCICE. (Credit Default Swap et risque de contrepartie.)

On considère 3 contreparties A , B et C . La contrepartie A entre dans un credit default swap avec la contrepartie B sur l'entité de référence C . Nous supposons que A est acheteur de protection ; que le taux sans risque est constant et égal à r ; que le taux de recouvrement associé à l'entité C est R_C ; que les intensités de défaut des firmes B et C valent λ_B et λ_C respectivement. On suppose de plus que la distribution jointe risque-neutre des défauts (τ_B, τ_C) des entités B et C est modélisée de la manière suivante : il existe trois variables aléatoires indépendantes V, T_B, T_C suivant des lois exponentielles de paramètres λ, l_B et l_C respectivement telles que

$$\tau_B = \min(V, T_B), \quad \tau_C = \min(V, T_C).$$

(1) Quelles conditions faut-il imposer aux paramètres pour que les hypothèses sur les intensités de défaut soient vérifiées ?

(2) Calculer la survie jointe des instants de défauts (τ_B, τ_C) .

(3) On suppose que le flux associé à la jambe de défaut est payé à la fin de la période en cours pourvu que B ne fasse pas défaut d'ici là et que les flux associés à la jambe de prime soient payés à la fin de chaque période pourvu que ni B , ni C n'aient fait défaut. Calculer le fair spread. Donner en une expression approchée et interpréter ce résultat.

(4) *Question subsidiaire.* Calculer la copule associée à la loi du couple (τ_B, τ_C) .

Nous allons à présent nous intéresser aux produits écrits sur des paniers de crédits : les n^{th} -to-default et les Collateralized Debt Obligations.

2.4 Les Basket Default Swaps

L'intérêt premier des paniers d'actifs risqués est de pouvoir utiliser leur corrélation pour restructurer le risque de crédit d'un portefeuille à l'aide de plusieurs nouveaux titres par le truchement d'un mécanisme de redistribution. La taille des portefeuilles considérés peut varier de cinq à plusieurs centaines de sous-jacents de crédit.

En général, le mécanisme de redistribution réalloue les pertes dans l'ordre de leur arrivée : certains titres supportent les premières pertes, et sont donc plus risqués que d'autres supportant des pertes plus lointaines. Ces mécanismes exposent les investisseurs à la tendance qu'ont les émetteurs à faire défaut ou à survivre ensemble. On modélise généralement cette co-dépendance des instants de défaut par une corrélation, c'est pourquoi on parle de produits de corrélation.

Le produit de corrélation le plus simple est le Basket Default Swap (n^{th} -2-Default), dont la composition typique comporte de 5 à 10 crédits. Ce produit est dit *idiosynchratique*, car l'identité du défaut, et donc ses caractéristiques doivent être connues :

1. le principe du swap est identique à celui du Credit Default Swap ;
2. seule change la définition de l'événement de crédit qui devient **le défaut du $n^{\text{ième}}$ crédit du panier.**

Nous allons à présent étudier deux cas particuliers à titre d'exemple : les First et Second-to-Default.

First-2-Default $(n = 1)$

Dans ce cas, le vendeur de protection touche un spread plus important que celui payé par n'importe quel émetteur du panier. En effet, ce spread est proportionnel à la probabilité d'arrivée de l'événement de crédit multipliée par la perte en cas de défaut. Or, en supposant que le panier est homogène (i.e. que la perte $1 - \delta_i = 1 - \delta$, où les δ_i et δ représentent les taux de recouvrement, est identique pour chacun des actifs i), on remarque que :

$$\mathbb{P}[\tau_1 \leq T] \geq \mathbb{P}[\tau_i \leq T], \quad \forall i.$$

L'investisseur obtient un *effet de levier* en augmentant le risque de crédit du produit. Le risque reste cependant limité au "pire cas", c'est à dire au risque de l'émetteur ayant la plus mauvaise qualité de crédit. L'évolution du spread des First-to-Default en fonction de la corrélation est négative.

Second-2-Default $(n = 2)$

Pour ces produits, l'événement de crédit est le deuxième défaut du panier. Le risque pour l'investisseur étant moindre que dans le cas du First-2-Default, son rendement l'est également. Contrairement au first-to-Default, le spread d'un Second-2-Default augmente quand la corrélation augmente.

Par extension, beaucoup d'autres produits sont envisageables (Third-2-Default, etc...)

Déterminants du spread d'un produit sur basket

1. **valeur de n** : un FTD est par exemple plus risqué qu'un STD donc son spread est plus élevé ;
2. **nombre de crédits** : le risque d'un Basket Default Swap augmente avec le nombre de crédits ;
3. **qualité du crédit des émetteurs** ;
4. **maturité** : l'effet de la maturité dépend de la forme des courbes de crédit des émetteurs ;
5. **taux de recouvrement** : son effet sur le prix est moindre car un taux de recouvrement plus élevé est compensé (à spread donné) par une probabilité de défaut plus forte (inégalité du triangle) ;
6. **corrélation de défaut** : si la corrélation augmente, la tendance des crédits à survivre ou à faire défaut ensemble augmente. L'effet sur le prix est donc fondamental, mais complexe.

Ce dernier point souligne l'importance des corrélations de défaut entre les crédits du panier. En effet, les Basket Default Swaps sont essentiellement des produits

de corrélation. Le spread d'un panier dépend donc sensiblement des hypothèses de corrélation que l'on fait.

En général, on suppose par exemple que les crédits d'un même secteur géographique et industriel sont plus corrélés que ceux de secteurs différents. On suppose aussi que le risque systémique est plus fort que le risque idiosyncratique, c'est à dire que l'on suppose que ces défauts sont corrélés positivement. Ces deux risques se reflètent dans les hypothèses de modélisation les plus couramment utilisées. Le chapitre 5 sera consacré aux outils de modélisation de cette corrélation.

2.5 Collateralized Debt Obligations

Les Collateralized Debt Obligations ou CDO sont des produits obligataires adossés à des dettes, résultant d'un mécanisme relativement complexe d'ingénierie financière appelé *titrisation* (securitization). À partir d'un panier de titres de dette (de 50 à 10000 créances), l'émetteur synthétise des actifs obligataires. Les CDO se distinguent selon la nature de la dette sous-jacente : s'il s'agit de produits obligataires, on parle de "Collateralized Bond Obligations" ou CBO. Dans le cas où le panier est constitué uniquement de titres de prêts, on parle de "Collateralized Loan Obligations" ou CLO. Bien entendu, dans le cas général, le panier est mixte. Depuis sa création dans le milieu des années 1990, le marché des CDO n'a cessé de se développer. En 2000, il dépassait les 100 Milliards de dollars d'émission. Nous présentons les enjeux du processus de titrisation ainsi que ses mécanismes, puis les techniques récentes liées à la génération synthétique de tranches utilisées en trading de corrélation.

Titrisation

La titrisation est une technique de gestion de bilan consistant à créer des produits obligataires à partir d'un ensemble de créances par le truchement d'une entité juridique particulière appelée *Special Purpose Vehicle* (SPV). Un établissement financier possédant un ensemble de créances ou de crédits auprès d'emprunteurs individuels ou institutionnels peut ainsi choisir de les titriser. Pour ce faire, il crée un SPV juridiquement indépendant, à qui il *vend* ses créances. Cette étape est fondamentale, car elle lui permet, d'une part, de transférer le risque de crédit au SPV, et d'autre part de retirer les créances titrisées de son bilan. Enfin, le SPV émet les CDO (voir figure 2.5). Les CDO émis comportent différentes tranches d'investissement, chacune d'elle possédant une qualité de crédit, et donc une notation différente :

- la tranche *junior* ou *equity* supporte les premières pertes sur l'ensemble de créances. Il s'agit donc d'un produit très risqué, payant un spread très élevé à l'investisseur. Il s'agit d'un produit purement spéculatif ;
- la tranche intermédiaire, dite *mezzanine* supporte les pertes au delà de la tranche *equity*, c'est un produit moyennement risqué, offrant un spread intéressant ;
- la tranche *senior* supporte les pertes restantes, si elles ont lieu. Elle est la moins soumise au risque de crédit, et offre donc un coupon faible.

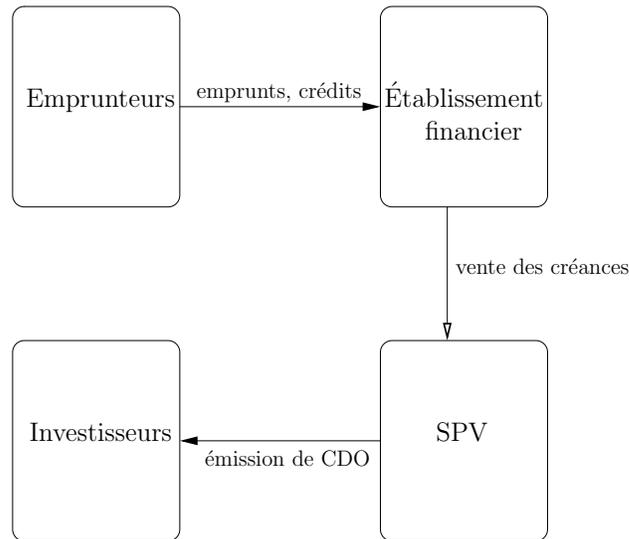


FIG. 2.5 – Le mécanisme de Titrisation

Comme chacune de ces tranches porte sur un ensemble de crédits, leur valorisation fait intervenir la corrélation entre les défauts du panier. Nous l'aborderons plus en détails au chapitre 5. Il est à noter que l'établissement financier peut parfois conserver une partie du risque de crédit, ce qui améliore la notation des tranches. Alternativement, il peut aussi faire appel à une compagnie d'assurance externe pour augmenter la qualité du crédit. On parle alors de *credit enhancement*.

L'intérêt de la titrisation est double : tout d'abord, elle permet d'économiser des fonds propres et ainsi d'améliorer leur rentabilité. En effet, étant donné que les créances titrisées peuvent être sorties du bilan, l'exigence de fonds propres sera moindre. De plus, la titrisation offre un accès simple et économique au marché financier à des entreprises de faible notation, qui devraient autrement se refinancer à des coûts prohibitifs. Elle permet l'assainissement de leur structure de capital.

Les produits synthétiques

Devant le développement impressionnant du marché des CDO et la demande de produits de corrélation de plus en plus forte de la part des investisseurs, des techniques d'ingénierie financière récentes ont donné naissance au concept des *CDO synthétiques*. De l'utilisation originelle dans les stratégies de gestion de fonds propres, les CDO deviennent peu à peu des produits d'investissement spéculatifs.

Le principe de la titrisation synthétique est de constituer des produits tranchés à partir, non plus d'un ensemble de crédits ou créances, mais d'un ensemble de CDS. Ceci revient à dire que l'organisme émetteur crée une exposition au risque de crédit et de corrélation en prenant des positions sur un ensemble de CDS. La titrisation synthétique peut donner naissance à deux types de produits :

1. les CDO synthétiques en *full capital structure* sont des CDO classiques à ceci près qu'ils sont écrits sur un ensemble de CDS. Comme pour les CDO classiques,

le risque de défaut est entièrement transféré aux investisseurs ;

2. les *Tranche Default Products* (TDP) consistent en l'émission d'une seule tranche de CDO synthétique. Dans ce cas, l'émetteur reste exposé au risque de corrélation, ce qui impose une gestion dynamique de cette corrélation par les *traders*, les positions deltas équivalentes étant évaluées à partir des spreads du marché et d'un modèle de co-dépendance.

Le développement des produits synthétiques, et des TDP en particulier, ouvre des perspectives prometteuses dans le domaine du trading de corrélation, et renforce l'importance de la modélisation.

Introduction aux modèles structurels

Les modèles structurels sont des modèles du risque de crédit où une entreprise donnée est considérée en cessation de paiement lorsque la valeur de ses actifs ne suffit plus à faire face à sa dette. Dans cette approche initiée par Merton (1974) [21], la valeur de la dette est évaluée à l'aide de la théorie des options : l'action de l'entreprise et sa dette y apparaissent comme des produits dérivés sur la valeur totale de ses actifs.

La popularité de ce modèle a permis la diffusion des idées de Merton et a transformé en quelques décennies la vision du risque de crédit. Le modèle développé par Moody's KMV [4] pour estimer les probabilités de défaut ou le modèle *Credit Grade* de JP Morgan [22] sont de parfaites illustrations de cette tendance. Les modèles “à la Merton” sont toujours largement utilisés pour valoriser la dette risquée, modéliser le *spread* de crédit, évaluer la qualité d'une signature, établir des liens entre le risque *equity* et le risque de crédit, etc.

Après avoir présenté le modèle de base tel que l'introduit Merton dans son article fondateur, nous décrivons, suivant Hull, Nelken et White [15], comment il peut être implémenté. Nous étendons ensuite le modèle de Merton dans la direction des modèles dits “de premier instant de passage” et traitons l'exemple du modèle *Credit Grade* de JP Morgan. Les exemples retenus mettent en évidence le lien entre risque de crédit et risque equity.

3.1 Modèle de Merton

Dans le modèle de Merton, l'action et la dette d'une entreprise sont considérées comme des produits dérivés sur sa valeur de marché et peuvent donc être évaluées dans le cadre de la théorie des options. Pour pouvoir appliquer cette théorie, il nous faut faire l'hypothèse de complétude suivante : il existe une classe d'actifs (comprenant la *cash*) permettant de répliquer les futurs *cash-flows* de l'entreprise. Cette hypothèse nous permet de justifier la valorisation des actifs de l'entreprise par absence d'oppor-

tunité d'arbitrage.

Soit par A la valeur totale des actifs de l'entreprise (A est la valeur de marché de l'ensemble des futurs *cash-flows* générés par l'activité de l'entreprise) et nous supposons que ce processus satisfait à l'équation différentielle stochastique suivante (mouvement brownien géométrique) :

$$\frac{dA_t}{A_t} = \mu dt + \sigma dB_t, \quad (\mu, \sigma > 0),$$

où B est un mouvement brownien unidimensionnel. Nous supposons aussi que les propriétaires de cette entreprise ont choisi de structurer son capital sous la forme d'actions (*pure equity*) et d'un unique zéro-coupon de maturité T et de nominal L (*debt*)¹. Le bilan d'une telle entreprise est résumé dans le tableau suivant :

Actif	Passif
Actifs $A_t = E_t + D_t$	Equity E_t Debt D_t

Produits dérivés sur la valeur des actifs de la firme

À maturité, si la valeur de l'entreprise est inférieure à la somme L due aux détenteurs d'obligations (zéro-coupon), nous considérons que la firme fait défaut. Dans ce cas, elle passe aux mains des détenteurs de sa dette qui ne récupèrent qu'une portion A_T/L de leur capital initial. Ainsi, les détenteurs d'obligations reçoivent $\min(A_T, L)$ à maturité tandis que les actionnaires perçoivent le reliquat $(A_T - L)_+$.

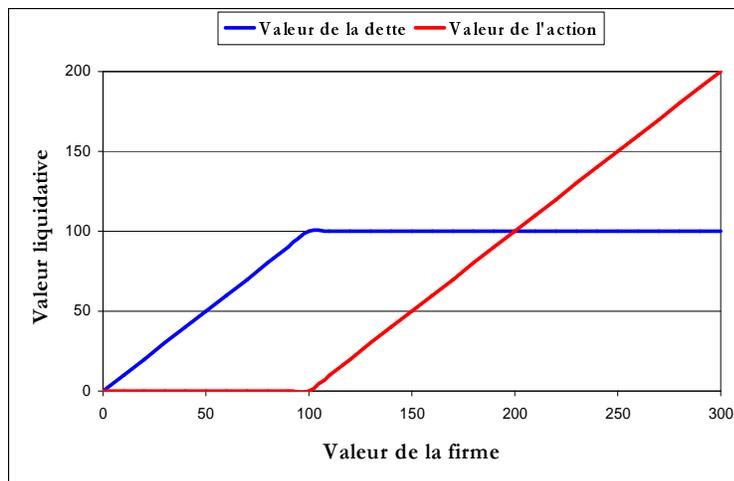


FIG. 3.1 – Valeurs liquidatives pour une entreprise de dette 100.

Notons E_t et D_t les valeurs respectives en t des actions et de la dette et $B(t, T)$ la valeur en t d'un zéro-coupon sans risque de maturité T . D'après la théorie des options

¹Cette hypothèse simplificatrice n'est que très rarement vérifiée. En pratique, la structure de capital d'une firme est infiniment plus complexe et peut comprendre, par exemple, des portions de dette convertible en action.

et en vertu de l'hypothèse de complétude, les processus E et D peuvent s'écrire sous la forme de l'espérance, sous la probabilité risque neutre, de leur valeur finale actualisée. Autrement dit, si nous avons

$$\begin{aligned} E_t &= B(t, T) \mathbb{E}_t^* [(A_T - L)_+] = A_t - D(t, T), \\ D_t &= B(t, T) \mathbb{E}_t^* [\min(A_T, L)] = B(t, T)L - B(t, T) \mathbb{E}^* [(L - A_T)_+]. \end{aligned}$$

Posons $L_t = B(t, T)L$. Appliquer la formule de Black & Scholes permet alors d'obtenir la formule de Merton pour la valeur de la dette risquée :

$$(3.1) \quad \boxed{D_t = A_t N(-d_1) + L_t N(d_2)},$$

où

$$\begin{aligned} d_1 &= \frac{\ln(A_t/L) + (r + \sigma^2/2)(T - t)}{\sigma\sqrt{T - t}}, \\ d_2 &= d_1 - \sigma\sqrt{T - t}, \end{aligned}$$

et

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt.$$

La valeur de la dette apparaît comme celle d'un zéro-coupon au taux sans risque de nominal L auquel on retranche un put (appelé *put-to-default*) sur la valeur de la firme de strike L et de maturité T . La valeur de actions apparaît, quant à elle, comme celle d'un call sur la valeur de la firme de strike L et de maturité T .

Probabilités de défaut et Loss Given Default

La formule (3.1) peut se réécrire

$$D(t, T) = L_t \mathbb{P}_t^* [A_T \geq L] + \mathbb{E}_t^* [B(t, T) A_T \mathbb{1}_{\{L > A_T\}}] = L_t \mathbb{P}_t^* [A_T \geq L] + A_t \bar{\mathbb{P}}_t [A_T < L],$$

où $\bar{\mathbb{P}}$ est la mesure martingale équivalente pour le numéraire A , c'est-à-dire la probabilité définie par

$$\frac{d\bar{\mathbb{P}}}{d\mathbb{P}} = \frac{1}{A_0} \frac{A_T}{B(0, T)}.$$

Les probabilités de défaut conditionnelles

$$p_t^* := \mathbb{P}_t^* [A_T < L], \quad \bar{p}_t^* := \bar{\mathbb{P}}_t [A_T < L],$$

sont données par les formules suivantes

$$p_t^* = N(-d_2), \quad \bar{p}_t^* = N(-d_1).$$

En utilisant ces notations, la formule de Merton devient

$$D_t = L_t(1 - p_t^*) + L_t p_t^* \delta_t^* = L_t(1 - p_t^* w_t^*),$$

où δ_t est le taux de recouvrement (Recovery Rate) en cas de défaut défini par

$$\delta_t^* = \frac{\mathbb{E}_t^* [A_T \mathbb{1}_{\{A_T < L\}}]}{L \mathbb{P}_t^* [A_T < L]} = \frac{A_t N(-d_1)}{L_t N(-d_2)},$$

et où $w_t^* = 1 - \delta_t^*$ est la perte en cas de défaut (Loss Given Default).

Distance-to-default

La probabilité de défaut s'interprète naturellement comme une fonction de la *distance-au-défaut* (distance-to-default) définie par

$$X_t = \frac{\ln A_t - \ln L}{\sigma} = \frac{1}{\sigma} \ln \frac{A_t}{L}.$$

Plus précisément, si $m^* = (r - \sigma^2/2)/\sigma$, on peut écrire que

$$p_t^* = \mathbb{P}^*[X_T \leq 0 | X_t] = N(u^*(t, T)),$$

où

$$u^*(t, T) = \frac{-X_t - m^*(T - t)}{\sqrt{T - t}}.$$

La distance au défaut est la distance entre la valeur des actifs de la firme et le niveau L de la dette mesuré en nombre d'écart-types. Il est facile de vérifier (exercice!) que X est un mouvement brownien avec dérive de drift m^* .

Levier d'endettement

Introduisons la variable

$$l_t = \frac{L_t}{A_t}$$

appelée *levier d'endettement* (leverage ratio) de l'entreprise. Ce ratio mesure le taux d'endettement de la firme. Il permet de réinterpréter les formules de Merton.

Les formules de Merton peuvent ainsi s'exprimer sous la forme

$$D_t = L_t \left(\frac{N(-h_1)}{l_t} + N(h_2) \right),$$

$$E_t = A_t \left(N(h_1) - \frac{N(h_2)}{l_t} \right),$$

où

$$h_1 = \frac{\ln(1/l_t) + \sigma^2(T - t)/2}{\sigma\sqrt{T - t}},$$

$$h_2 = h_1 - \sigma\sqrt{T - t}.$$

Les trois paramètres A_t , l_t et σ sont des paramètres-clés de l'analyse du risque de crédit d'une entreprise. Pour plus de détails sur les liens existant entre ces paramètres et le risque de crédit d'une firme voir [4]

Spread implicite

Le spread de crédit $S(t, T)$ peut se définir comme la différence de rendement (yield) entre une dette sans risque (par exemple une obligation du Trésor) et la dette émise par une entreprise (corporate bond). Ainsi, si $D(t, T)$ désigne la valeur d'un zéro-coupon risqué, on a

$$S(t, T) = Y^d(t, T) - Y(t, T),$$

où

$$Y(t, T) = -\frac{\ln B(t, T)}{T - t}, \quad Y^d(t, T) = -\frac{\ln D(t, T)}{T - t}.$$

Étant donné un modèle, nous appellerons *spread implicite*, le spread induit par ce modèle. Remarquons que dans le cas du modèle de Merton

$$S(t, T) = -\frac{L}{T - t} \ln\left(\frac{N(-h_1)}{l_t} + N(h_2)\right).$$

Le spread de crédit ne dépend ainsi que de la variabilité (volatilité) des actifs de la firme, du levier d'endettement et de la distance à la maturité. Lorsque $t \rightarrow T$, on observe le comportement suivant (exercice !)

$$\lim_{t \rightarrow T} S(t, T) = \infty \mathbb{1}_{\{A_T < L\}} + 0 \mathbb{1}_{\{A_T \geq L\}}.$$

Une première implémentation du modèle de Merton

La difficulté majeure de toute implémentation d'un modèle financier est sa calibration : comment peut-on estimer de manière fiable les paramètres apparaissant dans les formules précédentes ? En pratique, l'on peut estimer le paramètre L en évaluant la dette de l'entreprise. Les paramètres A_t et σ n'étant pas directement observables ; néanmoins, ils peuvent être estimés implicitement à partir des données de la volatilité (historique) σ_E de l'action et de sa valeur spot E_t .

Rappelons que la valeur de l'action est donnée par la formule

$$(3.2) \quad E_t = A_t N(d_1) - L_t N(d_2).$$

La formule d'Itô montre alors que E_t est un processus de volatilité

$$(3.3) \quad \sigma_E(t) = \frac{A_t}{E_t} N(d_1) \sigma.$$

Par suite, si l'on connaît σ_E et E_t , on connaît, en principe, la valeur des variables A_t et σ_A . Il suffit pour cela de résoudre en A_t et σ le système

$$\begin{cases} E_t &= A_t N(d_1) - L_t N(d_2), \\ \sigma_E(t) &= \frac{A_t}{E_t} N(d_1) \sigma. \end{cases}$$

La relation (3.3) n'ayant lieu qu'instantanément les résultats obtenus par cette méthode risquent d'être instables [4].

Spread implicite et smile de volatilité

Nous présentons maintenant une nouvelle méthode pour implémenter le modèle de Merton, d'après Hull, Nelken et White [15]. Il s'agit d'utiliser l'information contenue dans le smile de volatilité implicite pour calibrer le modèle de Merton et en particulier estimer le spread implicite.

Cette implémentation permet de classer les émetteurs de dette risquée selon la qualité de leur signature. L'intérêt d'un tel classement est important : d'une part, il permet de comparer le risque de crédit entre plusieurs firmes et d'autre part, il peut servir de base à une estimation des probabilités de défaut à horizon donné. Enfin, ce modèle permet de relier risque equity et risque de crédit.

Dans le modèle de Merton, un put de maturité $\tau < T$ et de strike K apparaît comme une option composée (voir [14]) sur la valeur de l'entreprise. Ainsi, la valeur du put sur l'action est donnée par la formule suivante :

$$P = L_0 M(-a_2, d_2; -\sqrt{\tau/T}) - A_0 M(-a_1, d_1; -\sqrt{\tau/T}) + KB(0, \tau)N(-a_2),$$

avec

$$a_1 = \frac{\ln(A_0/A_\tau^* B(0, \tau)) + \sigma^2 \tau / 2}{\sigma \sqrt{\tau}},$$

$$a_2 = a_1 - \sigma \sqrt{\tau},$$

où M désigne la fonction de répartition d'une gaussienne bivariée :

$$M(x, y; \rho) = \frac{1}{2\pi} \int_{-\infty}^x du \int_{-\infty}^{\frac{y-\rho u}{\sqrt{1-\rho^2}}} dv e^{-(u^2+v^2)/2}, \quad (-1 \leq \rho \leq +1),$$

et A_τ^* représente la valeur de la firme telle qu'à l'instant τ , l'on ait

$$E_\tau^* := A_\tau^* N(d_1) - L_\tau N(d_2) K = K.$$

Autrement dit, A_τ^* est la valeur en dessous de laquelle le put sur l'action sera exercé.

Notons ν la volatilité implicite du put et posons

$$\alpha = \frac{A_\tau^*}{A_0 B(0, \tau)}, \quad \kappa = \frac{K}{E_0 B(0, T)}.$$

Le paramètre κ est souvent appelé moneyness de l'option (lorsque $\kappa = 1$, l'option est à la monnaie du forward). Le paramètre α est la moneyness du point de vue de la valeur de la firme. Par définition, la volatilité implicite ν du put est solution de l'équation

$$L_0 M(-a_2, d_2; -\sqrt{\tau/T}) - A_0 M(-a_1, d_1; -\sqrt{\tau/T}) + \kappa E_0 N(-a_2) \\ = \kappa E_0 N(-d_2^*) - E_0 N(-d_1^*).$$

où

$$\begin{aligned} d_1^* &= \frac{\ln(1/\kappa) + \nu^2\tau/2}{\nu\sqrt{\tau}}, & d_2^* &= d_1^* - \nu\sqrt{\tau}, \\ a_1 &= \frac{\ln(1/\alpha) + \sigma^2\tau/2}{\sigma\sqrt{\tau}}, & a_2 &= a_1 - \sigma\sqrt{\tau}. \end{aligned}$$

En utilisant alors l'équation

$$(3.4) \quad E_0 = A_0 \left(N(h_1) - l_0^{-1} N(h_2) \right),$$

il vient

$$(3.5) \quad \begin{aligned} l_0^{-1} M(-a_2, d_2; -\sqrt{\tau/T}) - M(-a_1, d_1; -\sqrt{\tau/T}) + \kappa N(-a_2) [N(h_1) - l_0^{-1} N(h_2)] \\ = \kappa [N(-d_2^*) - N(-d_1^*)] [N(h_1) - l_0^{-1} N(h_2)]. \end{aligned}$$

En utilisant une nouvelle fois l'équation (3.4), nous pouvons aussi écrire que

$$\kappa E_0 e^{rT} = E^*(\tau, T) = A_\tau^* \left[N(h_{1,\tau}) - \frac{1}{\alpha l_0} N(h_{2,\tau}) \right],$$

où

$$\begin{aligned} h_{1,\tau} &= \frac{\ln(1/(\alpha l_0)) + \sigma^2(T-t)/2}{\sigma\sqrt{T-t}}, \\ h_{2,\tau} &= h_{1,\tau} - \sigma\sqrt{T-t}. \end{aligned}$$

de sorte que

$$(3.6) \quad \kappa = \frac{\alpha N(d_{1,\tau}) - l_0^{-1} N(h_{2,\tau})}{N(h_1) - l_0^{-1} N(h_2)},$$

Pour un jeu de paramètres (l_0, σ, T) du modèle de Merton et une maturité d'option $\tau (< T)$, les équations (3.5) et (3.6) définissent une relation implicite de la forme

$$\nu = \text{Fonction}(\kappa),$$

qui conduit à un smile de volatilité. Les propriétés de cette relation peuvent se résumer en quelques faits stylisés :

- (1) le spread de crédit est une fonction croissante de la volatilité implicite ;
- (2) la relation entre la pente (skew) du smile de volatilité et le spread de crédit est plus complexe : pour des spreads faibles, la pente est une fonction croissante du spread ; pour des niveaux de spread plus élevés, la variation de la pente devient négligeable ;
- (3) enfin, la pente est une fonction croissante de la volatilité implicite à la monnaie.

Ces résultats suggèrent une nouvelle méthode pour implémenter le modèle de Merton : étant données deux volatilités implicites et une valeur pour T , on peut résoudre les équations (3.5) et (3.6) pour obtenir l_0 et σ . Cette implémentation permet donc d'estimer directement le spread de crédit et la probabilité risque-neutre de défaut à partir du smile de volatilité tout en évitant le recours à la relation instantanée (3.3).

Ces résultats peuvent ensuite être utilisés comme un indicateur du risque de crédit. Les auteurs ont testé cette approche en comparant le classement des spreads obtenus par cette méthode avec les spreads du marché des CDS et ont montré que ces classements sont proches. Cette méthode permet donc, par exemple, d'étudier la qualité de la signature d'entreprises dont le marché de CDS n'existe pas ou n'est pas très liquide.

Limites du modèle et extensions

La principale limitation du modèle de Merton est que les spreads implicites court-terme sont nécessairement proches de zéro contrairement aux spreads réellement observés sur les marchés de capitaux. Cette propriété du modèle est liée à la continuité de la filtration brownienne (ou, de manière équivalente, à la prévisibilité de l'instant de défaut). Duffie et Lando expliquent cette apparente contradiction par le fait que l'asymétrie d'information qui existe entre les actionnaires et les détenteurs de la dette n'est pas prise en compte par ce modèle. Ils ont montré [8] que si ces derniers ne disposent que d'une information comptable partielle (ou "bruitée") l'instant de défaut leur apparaît comme totalement inaccessible. Ces résultats permettent de justifier l'utilisation des modèles à forme réduite dont nous aborderons l'étude dans le prochain chapitre.

Le modèle de Merton peut être étendu dans différentes directions : Nous verrons dans la section 3.2 consacrée aux modèles de premier instant de passage comment l'on peut incorporer des défaut se produisant à un instant quelconque de l'intervalle $[0, T]$; Vasiček [30] propose un modèle de Merton prenant en compte l'échelonnement de la dette ; Longstaff & Schwartz [20] introduisent un modèle de Merton intégrant des taux stochastiques.

Notons enfin que si l'on souhaite tenir compte des imperfections de marché (tels l'incomplétude ou le contrôle que peuvent exercer actionnaires et détenteurs de la dette sur la conduite de l'entreprise), la théorie s'éloigne sensiblement de la théorie des options. Sur ce sujet, l'on pourra se référer à Leland [19].

EXERCICE. (Extension du modèle de Merton au cas où la valeur des actifs de la firme est un processus à sauts.)

Soit N un processus de Poisson de paramètre λ et $U = \{U_i, i \geq 1\}$ une suite de v.a. i.i.d. de moyenne ν . Nous supposons que les tribus $\sigma(W)$, $\sigma(N)$ et $\sigma(U)$ sont mutuellement indépendantes.

Nous considérons une entreprise dont la dynamique risque-neutre de la valeur de ses actifs est décrite par l'EDS

$$dV_t = V_{t-}((r - \lambda\nu) dt + \sigma_V dW_t + d\pi_t),$$

où W est un mouvement brownien standard et π est le processus de saut $\pi_t = \sum_{i=1}^{N_t} U_i$ et r est la taux sans risque. Dans la suite, \mathcal{F}_t désigne la tribu $\sigma(W_s, \pi_s; s \leq t)$.

(1) Vérifier que $\tilde{\pi}_t = \pi_t - \lambda\nu t$ est une martingale relativement à la filtration \mathcal{F} .

(2) Montrer que $V_t^* = e^{-rt}V_t$ satisfait à

$$(3.7) \quad dV_t^* = V_{t-}^* (\sigma_V dW_t + d\tilde{\pi}_t).$$

En déduire que V^* est une martingale.

(3) Vérifier que

$$X_t = V_0 e^{\tilde{\pi}_t + \sigma_V W_t - \frac{1}{2}\sigma_V^2 t} \prod_{0 < u \leq t} (1 + \Delta\tilde{\pi}_u) e^{-\Delta\tilde{\pi}_u} = V_0 e^{\sigma_V W_t - (\sigma_V^2/2 + \lambda\nu)t} \prod_{i=1}^{N_t} (1 + U_i)$$

est solution de (3.7). On admettra l'unicité des solutions de (3.7) de sorte que $V^* = X$.

(4) Nous supposons désormais que $\ln(1 + U_i)$ suit une loi normale de moyenne μ et de variance σ^2 . Montrer que $\nu = e^{\mu + \sigma^2/2} - 1$.

(5) On considère le bond risqué de maturité T et de nominal L dont le payoff est $L(I_{\{V_T \geq L\}} + V_T/L)I_{\{V_T < L\}}$.² Montrer que la probabilité de défaut du bond s'écrit

$$\mathbb{P}[V_T < L] = \sum_{i=0}^{+\infty} e^{-\lambda T} \frac{(\lambda T)^i}{i!} N(-d_2(i)),$$

où

$$d_2(i) = \frac{\ln(V_0/L) + \mu_i(T)}{\sigma_i(T)}, \quad \mu_i(T) = (r - \sigma_V^2/2 - \lambda\nu) + i\mu, \quad \sigma_i^2(T) = \sigma_V^2 T + i\sigma^2.$$

(6) Vérifier que la valeur du bond risqué est

$$D(0, T) = LB(0, T) \left(1 - \sum_{i=0}^{+\infty} e^{-\lambda T} \frac{(\lambda T)^i}{i!} \left[N(-d_2(i)) + \frac{V_0}{L} e^{\mu_i(T) + \sigma_i^2(T)T/2} N(-d_1(i)) \right] \right),$$

où

$$d_1(i) = d_2(i) + \sigma_i(T).$$

3.2 Modèles de premier instant de passage

L'une des limites du modèle de Merton réside dans le fait que le défaut de l'émetteur ne peut intervenir qu'à la maturité de la dette. Dans les modèles de premier instant de passage, au contraire, l'instant de défaut est (un temps d'arrêt) de la forme

$$\tau = \inf \{t > 0; A_t < Bar(t)\}$$

où Bar est une barrière qui peut être aléatoire et A est une variable de type "valeur de la firme". Dans les modèles de ce type, il est possible de spécifier une grande variété d'hypothèses de recouvrement en cas de défaut (voir [2], chapitre 3). Après quelques préliminaires mathématiques, nous présenterons un exemple de modèle de premier instant de passage : le modèle *Credit Grade* développé par JP Morgan.

²Ce sont les hypothèses habituelles du modèle de la firme selon Merton.

Préliminaires mathématiques

Nous commençons ce paragraphe par un lemme bien connu de la théorie des diffusions (voir [17]). Dans la suite, Y désigne un mouvement brownien avec dérive de sorte que

$$Y_t = y_0 + \mu t + \sigma B_t, \quad (y_0 > 0, \mu \in \mathbb{R}, \sigma > 0),$$

où B est un mouvement brownien unidimensionnel.

LEMME 3.1

$$\mathbb{P} \left[\inf_{0 \leq u \leq s} (Y_u - y_0) \geq y \right] = N \left(\frac{-y + \mu s}{\sigma \sqrt{s}} \right) - e^{\frac{2\mu}{\sigma^2} y} N \left(\frac{y + \mu s}{\sigma \sqrt{s}} \right).$$

Une preuve de ce lemme est donnée en appendice E

Considérons le temps d'arrêt

$$\tau = \inf \{ t > 0; Y_t < 0 \}$$

Appliquer le lemme précédent à $Y - y_0$ suffit à prouver :

PROPOSITION 3.2 *La variable aléatoire τ est distribué selon la loi gaussienne inverse. Plus précisément,*

$$\mathbb{P}[\tau \leq s] = \mathbb{P}[\tau < s] = N(h_1(s)) + e^{-\frac{2\nu}{\sigma^2} y_0} N(h_2(s)),$$

où

$$h_1(s) = -\frac{y_0 + \nu s}{\sigma \sqrt{s}}, \quad h_2(s) = -\frac{y_0 - \nu s}{\sigma \sqrt{s}}.$$

EXEMPLE. Nous considérons une entreprise dont la valeur A est décrite comme dans le modèle de Merton et nous supposons que l'instant de défaut de la firme se présente sous la forme

$$\tau = \inf \{ t > 0; A_t < v \} \text{ où } v \text{ est un réel inférieur à } A.$$

Dans ce cas, nous avons :

$$\mathbb{P}[\tau > s | A_t] = N \left(\frac{\ln(v/A_t) - \nu(s-t)}{\sigma \sqrt{s-t}} \right) + \left(\frac{v}{A_t} \right)^{2a} N \left(\frac{\ln(v/A_t) + \nu(s-t)}{\sigma \sqrt{s-t}} \right),$$

où

$$a = \frac{\nu}{\sigma^2} = \frac{r - \sigma^2/2}{\sigma^2}.$$

Modèle *Credit Grade*

Dans ce paragraphe, nous présentons un modèle de premier instant de passage développé par JP Morgan (*Credit Grade*, voir [22]). Nous avons choisi de présenter ce modèle car il nous semble qu'il contient des intuitions puissantes sur ce que sont les déterminants du risque de défaut et les liens qui existent entre le risque de crédit et le risque equity.

Description du modèle

Nous supposons que la “valeur” V d'une entreprise est décrite par un processus satisfaisant à l'EDS

$$\frac{dV_t}{V_t} = \sigma dW_t, \quad (\sigma > 0),$$

où W est un mouvement brownien unidimensionnel et $\sigma > 0$ est la volatilité de V . Ici, V n'est pas réellement la valeur de la firme mais plutôt un indice mesurant l'évolution temporelle de la qualité du crédit de l'entreprise. Dans ce modèle, le défaut est défini comme le premier instant où V atteint une barrière LD où

- (1) D est le ratio debt-per-share,
- (2) L est une grandeur aléatoire représentant le taux de recouvrement moyen global en cas de défaut.

La variable L est supposée lognormale de moyenne \bar{L} et d'écart-type λ de sorte que

$$LD = \bar{L}De^{\lambda Z - \lambda^2/2}$$

où Z suit une loi normale centrée réduite.

La moyenne \bar{L} et l'écart-type λ sont estimés historiquement en utilisant des données de taux de recouvrement telles celles fournies par Standard & Poor's. Dans [22], les auteurs mentionnent les valeurs $\bar{L} = 0.5$ et $\lambda = 0.3$ obtenues à partir des données de défaut de 300 entreprises américaines (hors institutions financières) entre 1987 et 1997. Le ratio debt-per-share D est obtenu en divisant le nominal de la dette globale par le nombre d'actions émises par l'entreprise.

Pour une valeur V_0 donnée, l'instant τ de défaut est donc

$$\tau = \inf\{t > 0; V_t < LD\},$$

et si l'on pose

$$X_t = \sigma W_t - \lambda Z - \frac{\sigma^2}{2}t - \frac{\lambda^2}{2},$$

de sorte que $X_t \sim N(-1/2A_t^2, A_t^2)$ où $A_t = \sigma^2 t + \lambda^2$, cette formule peut se réécrire

$$\tau = \inf\{t > 0; X_t < \ln(\bar{L}D/V_0) - \lambda^2\}.$$

Pour appliquer les formules présentées dans le paragraphe précédent (afin d'obtenir une formule fermée pour la structure par terme de probabilité de défaut), les auteurs

proposent de remplacer X_t avec un mouvement brownien Y_t de loi $N(\mu t, \theta^2 t)$, où $\theta^2 t = A_t^2 = \sigma^2 t + \lambda^2$ et $\mu t = -A_t^2/2$. D'après le lemme 3.1, l'on a

$$\mathbb{P}\left[\inf_{s \leq t} Y_s > y\right] = N\left(\frac{\mu t - y}{\theta \sqrt{t}}\right) - e^{\frac{2\mu y}{\theta^2}} N\left(\frac{\mu t + y}{\theta \sqrt{t}}\right)$$

et en posant $y = \ln(\bar{L}D/V_0) - \lambda^2$, l'on obtient la formule

$$P(t) := \mathbb{P}[\tau > t] = N\left(-\frac{A_t}{2} + \frac{\ln d}{A_t}\right) - dN\left(-\frac{A_t}{2} - \frac{\ln d}{A_t}\right),$$

où

$$d = \frac{V_0 e^{\lambda^2}}{\bar{L}D}.$$

REMARQUE. L'on peut, au prix d'une intégration supplémentaire, éviter l'approximation ($X \sim Y$) précédente. En effet, l'on a

$$\begin{aligned} \mathbb{P}[\tau > t] &= \mathbb{P}[\sigma W_s - \sigma^2 s/2 > \ln(\bar{L}D/V_0) + \lambda Z - \lambda^2/2, \forall s \leq t] \\ &= \int_{-\infty}^{+\infty} \frac{dz}{\sqrt{2\pi}} \mathbb{P}[\sigma W_s - \sigma^2 s/2 > F(z), \forall s \leq t] \\ &= \int_{-\infty}^{+\infty} \frac{dz}{\sqrt{2\pi}} \left\{ N\left(\frac{-F(z) - \sigma^2 t/2}{\sigma \sqrt{t}}\right) - e^{-F(z)} N\left(\frac{F(z) - \sigma^2 t/2}{\sigma \sqrt{t}}\right) \right\} \end{aligned}$$

où

$$F(z) = \ln\left(\frac{\bar{L}D}{V_0}\right) - \frac{\lambda^2}{2} + \lambda z.$$

REMARQUE. L'introduction dans un modèle structurel d'une barrière aléatoire implique que la probabilité de défaut instantanée n'est plus nulle. Ceci conduit à un spread court-terme non nul et permet de résoudre l'un des problèmes inhérents aux modèles à la Merton.

Probabilité de défaut et spread de crédit

Dans ce paragraphe, nous désignons par spread de crédit la valeur de la marge d'un Credit Default Swap qui annule sa valeur au moment de l'entrée dans le swap. Autrement dit, le spread est donné par la formule

$$s(t) = (1 - R) \frac{\int_{]0,t]} DF(0, s) P(ds) + 1 - P(0)}{\int_0^t DF(0, s) P(s) ds}.$$

Il faut bien prendre garde au fait que R est le taux de recouvrement spécifique au titre couvert par le CDS et n'est donc pas nécessairement égal à \bar{L} qui est un taux de

recouvrement moyen global. Typiquement, le taux de recouvrement pour une dette *unsecured* sera plus faible que \bar{L} , alors que celui lié à une dette *secured* sera plus grand que \bar{L} .

REMARQUE. Si l'on note

$$p(t) = -\frac{1}{t} \ln P(t),$$

le taux de défaut moyen et si l'on suppose que

- (1) $p(t) \simeq p$ constant,
- (2) $P(0) \simeq 1$,

alors l'approximation suivante est justifiée

$$s(t) \simeq (1 - R)p.$$

Calibration du modèle sur les données de marché

Nous cherchons maintenant à calibrer ce modèle sur des données de marché observables (données equity). Soient S et σ_S respectivement la valeur de l'action et la volatilité de l'action de l'entreprise considérée. La méthode proposée par JP Morgan consiste à examiner les conditions aux bords à long terme sur une expression de type distance-au-défaut pour les deux régimes extrêmes

- près du défaut c'est-à-dire $S \rightarrow 0$,
- loin de la barrière c'est-à-dire $S \gg LD$.

Soit donc η la distance-au-défaut mesuré en écart-type de V et définie par

$$\eta = \frac{\ln V - \ln(LD)}{\sigma} = \frac{1}{\sigma} \ln \frac{V}{LD}.$$

Reprendre le raisonnement qui nous a conduit à la formule 3.3 permet d'exprimer la distance-au-défaut sous la forme

$$(3.8) \quad \eta = \frac{V}{\sigma_S S} \frac{\partial S}{\partial V} \ln \left(\frac{V}{LD} \right).$$

Nous allons établir les conditions aux bords pour η . Près du défaut ($S \rightarrow 0$), nous avons

$$V \simeq V(S=0) + \frac{\partial V}{\partial S} S + o(S^2) = LD + \frac{S}{\alpha} + o(S^2)$$

En utilisant l'équation précédente et (3.8), on montre que lorsque $S \rightarrow 0$

$$\eta = \frac{1}{\sigma_S} + o(S).$$

Lorsque $S \gg LD$, on suppose que $S/V \rightarrow 1$ (ce qui est cohérent avec le comportement d'un modèle de Merton standard). Par suite,

$$\eta \simeq \frac{1}{\sigma_S} \ln \left(\frac{S}{LD} \right), \quad (S \gg LD).$$

L'expression la plus simple pour η qui satisfasse simultanément à ces deux conditions aux bords est

$$\eta = \frac{S + LD}{S\sigma_S} \ln\left(\frac{S + LD}{LD}\right).$$

En comparant l'équation précédente avec (3.8) nous sommes conduits à

$$V = S + LD,$$

et donc

$$V_0 = S_0 + \bar{L}D.$$

Finalement, l'on obtient

$$\sigma = \sigma_S^* \frac{S^*}{S^* + \bar{L}D},$$

pour une valeur de l'action égale à S^* et sa volatilité (historique ou implicite) correspondante.

La formule suivante pour la probabilité (risque neutre) de défaut ne fait plus intervenir que des variables observables

$$P(t) = N\left(-\frac{A_t}{2} + \frac{\ln d}{A_t}\right) - dN\left(-\frac{A_t}{2} - \frac{\ln d}{A_t}\right),$$

où

$$A_t^2 = \left(\sigma_S^* \frac{S^*}{S^* + \bar{L}D}\right)^2 t + \lambda^2, \quad d = \frac{S^* + \bar{L}D}{\bar{L}D} e^{\lambda^2}.$$

CHAPITRE 4

Modèles à forme réduite

Nous présentons dans ce chapitre les bases de la modélisation du risque de défaut à l'aide de modèles à forme réduite.

Après avoir posé les fondations mathématiques des modèles à forme réduite, nous plaçons dans le cadre des modèles doublement stochastiques. Conditionnellement à l'information de marché, les instants de défaut y apparaissent comme le premier instant de saut d'un processus de Poisson. Nous verrons que cette propriété est au coeur de la tractabilité analytique de cette approche. Nous terminons ce chapitre par des exemples d'applications lorsque l'intensité de défaut est un processus affine (cf. appendice C).

4.1 Préliminaires mathématiques

Dans cette section inspirée de [2], [7] et de [29], nous présentons les définitions et résultats mathématiques qui sous-tendent la théorie des modèles à intensité. Elle se veut élémentaire et c'est pourquoi nous n'entrons pas dans toutes les subtilités techniques. Pour plus de détails sur la théorie générale des processus, nous renvoyons à l'annexe A.

Dans la suite, $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{G} = (\mathcal{G}_t)_{t \geq 0})$ désigne un espace probabilisé satisfaisant aux conditions habituelles. Toutes les filtrations considérées dans la suite sont, elles aussi, supposées vérifier les conditions habituelles.

Définitions

Nous introduisons la définition suivante qui est centrale dans la suite :

DÉFINITION 4.1 Soient τ un temps d'arrêt de la filtration \mathbb{G} et N le processus $t \mapsto N_t = \mathbb{1}_{\{\tau \leq t\}}$. On dit que le processus \mathbb{G} -adapté et positif $\lambda = (\lambda_t)_{t \geq 0}$ est l'intensité (ou

taux de hasard) de τ si

$$N_t - \int_0^t \mathbb{1}_{\llbracket 0, \tau \rrbracket}(s) \lambda_s ds = N_t - \int_0^{\tau \wedge t} \lambda_s ds$$

est une (\mathbb{G}, \mathbb{P}) -martingale.

REMARQUE. Si λ est déterministe, ceci revient à dire que τ est le premier instant de saut d'un processus de Poisson d'intensité λ .

Intuitivement, cette définition signifie que, conditionnellement à la réalisation de l'événement $\{\tau > t\}$, on a

$$\mathbb{P}[\tau \in (t, t + dt] | \mathcal{G}_t] = \lambda_t dt.$$

Le résultat suivant [1] formalise mathématiquement cette dernière égalité.

THÉORÈME 4.2 Soient $(\epsilon_n)_{n \geq 0}$ une suite de réels strictement positifs décroissant vers 0 et Y_n une version mesurable du processus $Y_n(t) = \epsilon_n^{-1} \mathbb{P}[t < \tau \leq t + \epsilon_n | \mathcal{G}_t]$. S'il existe deux processus λ et y tels que, pour tout t , la limite $\lambda_t = \lim_n Y_n(t)$ existe et

$$|Y_n(s) - \lambda_s| \leq y_s, \forall s \leq t \quad \text{et} \quad \int_0^t y_s ds < +\infty \text{ p.s.},$$

alors λ est l'intensité de τ .

Nous utiliserons ce résultat dans le prochain chapitre de la manière suivante : si la survie conditionnelle $S(t, T) = \mathbb{P}[\tau > T | \mathcal{G}_t]$ est un processus suffisamment régulier, alors l'intensité de τ est donnée par

$$\lambda_t = - \frac{\partial}{\partial T} S(t, T) \Big|_{T=t}.$$

REMARQUE. Par analogie avec la définition des taux courts, on peut interpréter λ comme un spread instantané :

$$r_t = - \frac{\partial}{\partial T} B(t, T) \Big|_{T=t}.$$

EXERCICE. Soit τ une variable aléatoire exponentielle de paramètre $\lambda > 0$: par définition, on a $\mathbb{P}[\tau > t] = \exp(-\lambda t)$. Montrer que dans la plus petite filtration dans laquelle τ un temps d'arrêt, τ admet pour intensité le processus constant égal à λ .

Soit τ un temps d'arrêt tel que $\mathbb{P}[0 < \tau < +\infty] > 0$. Si la filtration \mathbb{G} est engendrée par un mouvement brownien (disons W), le temps d'arrêt τ ne peut pas admettre d'intensité. En effet, dans le cas contraire, on pourrait trouver un processus prévisible H tel que

$$M_t = N_t - \int_0^{\tau \wedge t} \lambda(s) ds = \int_0^t H_s dW_s,$$

ce qui est absurde car sur l'événement $\{0 < \tau < +\infty\}$, on aurait alors $\Delta M_\tau = 0$. En fait, pour qu'un temps d'arrêt admettent une intensité, il faut qu'il soit totalement inaccessible (c'est-à-dire qu'il arrive comme une surprise totale). A contrario, les temps d'arrêt d'une filtration brownienne sont prévisibles et sont donc annoncés par une série de signes avant-coureurs.

Nous nous intéressons désormais à la situation où la variable aléatoire τ est un temps d'arrêt d'une filtration plus grande que celle dans laquelle est définie son taux de hasard. Nous supposons donc que $\mathbb{F} = (\mathcal{F})_{t \geq 0}$ est une filtration telle que $\mathcal{F}_t \subset \mathcal{G}_t$.

L'utilisation de ce cadre peut se justifier de la manière suivante : supposons que l'instant de défaut est défini par un modèle structurel brownien. Pour les agents ayant accès à l'information de cette filtration (disons le conseil d'administration de l'entreprise et ses managers), l'instant de défaut est un temps d'arrêt prévisible qui n'admet pas d'intensité. Or, si l'on fait l'hypothèse que les informations comptables nécessaires à la calibration du modèle structurel précédent sont "bruitées"¹, les agents de marché n'ont accès, en terme d'information, qu'à une sous-filtration de la première dans laquelle l'instant de défaut est totalement inaccessible et possède une intensité [8]. Les modèles à forme réduite rendent compte de l'asymétrie d'information qui existe entre l'agent du marché non-informé et le management d'une entreprise.²

DÉFINITION 4.3 *Soit τ un temps d'arrêt de \mathbb{G} d'intensité λ . On dit que τ est doublement stochastique relativement à la filtration \mathbb{F} si λ est un processus \mathbb{F} -adapté³ et si pour tout $t \leq s$, on a*

$$\mathbb{P}[\tau > s | \mathcal{F}_s \vee \mathcal{G}_t] = \exp\left(-\int_t^s \lambda_u du\right) \mathbb{1}_{\{\tau > t\}}.$$

Cette définition signifie que conditionnellement à \mathcal{F}_∞ , τ est le premier instant de saut d'un processus de Poisson $\tilde{N} = (\tilde{N}_t, t \geq 0)$ par rapport à \mathbb{G} d'intensité $(\lambda_t, t \geq 0)$. En effet

$$\begin{aligned} \mathbb{P}[\tau > s | \mathcal{F}_\infty \vee \mathcal{G}_t] &= \mathbb{P}\left[\tilde{N}_s - \tilde{N}_t = 0; \tilde{N}_t = 0 \mid \mathcal{F}_\infty \vee \mathcal{G}_t\right] \\ &= \mathbb{1}_{\{\tilde{N}_t=0\}} \mathbb{P}\left[\tilde{N}_s - \tilde{N}_t \mid \mathcal{F}_\infty\right] \\ &= \mathbb{1}_{\{\tau > t\}} e^{-\int_s^t \lambda_u du}, \end{aligned}$$

et donc

$$\begin{aligned} \mathbb{P}[\tau > s | \mathcal{F}_s \vee \mathcal{G}_t] &= \mathbb{E}\left[\mathbb{1}_{\{\tau > t\}} e^{-\int_s^t \lambda_u du} \mid \mathcal{F}_s \vee \mathcal{G}_t\right] \\ &= \mathbb{1}_{\{\tau > t\}} e^{-\int_s^t \lambda_u du}. \end{aligned}$$

Dans la pratique, nous partons d'une filtration $\tilde{\mathbb{F}}$ qui est, en général, la filtration engendrée par un processus de Markov (disons, un processus affine) représentant l'information de marché hors instant de défaut et nous construisons une v.a. $\tilde{\tau}$ et une filtration $\tilde{\mathbb{G}}$ de la manière suivante (processus de Cox) :

¹voire incomplètes, voire truquées...

²le délit d'initié protège, en principe, le marché des abus potentiels lié à l'utilisation en vue d'un enrichissement personnel de cette asymétrie d'information.

³en fait, prévisible.

Soient λ un processus $\tilde{\mathbb{F}}$ -adapté et positif, γ le processus $t \mapsto \exp(-\int_0^t \lambda_s ds)$ et U une v.a. de loi uniforme sur $[0, 1]$ indépendante de $\tilde{\mathcal{F}}_\infty$ ⁴. Si nous définissons la v.a.

$$\tilde{\tau} = \inf\{t \geq 0; \gamma(t) \leq U\}$$

et la tribu

$$\tilde{\mathcal{G}}_t = \tilde{\mathcal{F}}_t \vee \tilde{\mathcal{H}}_t$$

où $\tilde{\mathcal{H}}_t = \sigma(\tau \wedge t)$, alors il est facile de vérifier que $\tilde{\tau}$ est un temps d'arrêt de $\tilde{\mathcal{G}}$ d'intensité λ (exercice!). Nous montrerons dans la prochaine sous-section que le temps d'arrêt $\tilde{\tau}$ est aussi doublement stochastique relativement à $\tilde{\mathbb{F}}$.

Calculs d'espérances conditionnelles

Nous détaillons une série de calculs d'espérances conditionnelles valable dans un contexte assez général. Ces résultats vont nous permettre de démontrer que les processus de Cox sont doublement stochastiques mais aussi de prouver une formule qui interviendra de manière récurrente dans la suite.

Soient $\mathbb{F} \subset \mathbb{G}$ deux filtrations et τ une v.a. telle que si \mathbb{H} désigne la filtration $\mathcal{H}_t = \sigma(\tau \wedge t)$, alors on a $\mathbb{G} = \mathbb{F} \vee \mathbb{H}$. Dans la pratique, \mathbb{G} sera la filtration du trader de CDS, convertibles, sur la firme en question, tandis que \mathbb{F} désignera la filtration de marché hors défaut de la firme considérée. Nous aurons besoin du lemme fondamental suivant sur les tribus :

LEMME 4.4 *Soit \mathcal{T} une sous-tribu de \mathcal{G}_∞ . Pour tout t , on a*

$$\mathcal{H}_t \cap \mathcal{T} \subset \{A \in \mathcal{G}_\infty; \exists B \in \mathcal{T}, A \cap \{\tau > t\} = B \cap \{\tau > t\}\},$$

autrement dit $\mathcal{H}_t \cap \mathcal{T}$ est inclus dans l'ensemble des éléments de \mathcal{G}_∞ dont la restriction à $\{\tau > t\}$ est dans \mathcal{T} .

PREUVE. Notons \mathcal{G}_t^* le membre de droite dans l'égalité ci-dessus. Il est clair que \mathcal{G}_t^* est une sous-tribu de \mathcal{G}_∞ . Il suffit donc de prouver que $\mathcal{H}_t \subset \mathcal{G}_t^*$ et que $\mathcal{T} \subset \mathcal{G}_t^*$. Pour ce faire, il suffit encore de prouver que si $A = \{\tau \leq u\}$ ($u \leq t$) ou $A \in \mathcal{T}$, alors il existe $B \in \mathcal{T}$ tel que $A \cap \{\tau > t\} = B \cap \{\tau > t\}$. Dans le premier cas, on peut choisir $B = \emptyset$, dans le second $B = A$. ■

PROPOSITION 4.5 *Soient \mathcal{T} une sous-tribu de \mathcal{G}_∞ et Y une v.a. \mathcal{G}_∞ -mesurable et intégrable. On a*

$$\mathbb{E}[Y \mathbb{1}_{\{\tau > t\}} | \mathcal{H}_t \vee \mathcal{T}] = \mathbb{1}_{\{\tau > t\}} \frac{\mathbb{E}[\mathbb{1}_{\{\tau > t\}} Y | \mathcal{T}]}{\mathbb{P}[\tau > t | \mathcal{T}]},$$

ce qui revient à dire que si l'on se restreint à $\{\tau > t\}$, la tribu \mathcal{H}_t ne contient plus d'information.

⁴On pourrait dire que c'est dans l'aléa de cette variable aléatoire que se concentre l'information non disponible.

PREUVE. Puisque le membre de droite dans l'égalité précédente est $\mathcal{H}_t \vee \mathcal{T}$ -mesurable, il suffit de prouver que pour tout $A \in \mathcal{H}_t \vee \mathcal{T}$, on a

$$\int_A \mathbb{1}_C Y \mathbb{P}[C | \mathcal{F}_t] d\mathbb{P} = \int_A \mathbb{1}_C \mathbb{E}[\mathbb{1}_C Y | \mathcal{F}_t] d\mathbb{P}.$$

où $C = \{\tau > t\}$. D'après le lemme précédent, il existe $B \in \mathcal{T}$ tel que $A \cap C = B \cap C$ et par suite, on peut écrire que

$$\begin{aligned} \int_A \mathbb{1}_C Y \mathbb{P}[C | \mathcal{T}] d\mathbb{P} &= \int_{A \cap C} Y \mathbb{P}[C | \mathcal{T}] d\mathbb{P} = \int_{B \cap C} Y \mathbb{P}[C | \mathcal{T}] d\mathbb{P} \\ &= \int_B \mathbb{1}_C Y \mathbb{P}[C | \mathcal{T}] d\mathbb{P} = \int_B \mathbb{E}[\mathbb{1}_C Y | \mathcal{T}] \mathbb{P}[C | \mathcal{T}] d\mathbb{P} \\ &= \int_B \mathbb{E}[\mathbb{1}_C \mathbb{E}[\mathbb{1}_C Y | \mathcal{T}] | \mathcal{T}] d\mathbb{P} = \int_{B \cap C} \mathbb{E}[\mathbb{1}_C Y | \mathcal{T}] d\mathbb{P} \\ &= \int_{B \cap A} \mathbb{E}[\mathbb{1}_C Y | \mathcal{T}] d\mathbb{P} = \int_A \mathbb{1}_C \mathbb{E}[\mathbb{1}_C Y | \mathcal{T}] d\mathbb{P}, \end{aligned}$$

d'où le résultat. ■

Nous allons démontrer que le processus de Cox défini dans la sous-section précédente est doublement stochastique. Nous reprenons pour un instant les notations définies précédemment. Il s'agit de prouver que

$$\mathbb{P}[\tilde{\tau} > s | \tilde{\mathcal{F}}_s \vee \tilde{\mathcal{G}}_t] = \exp\left(-\int_t^s \lambda_u du\right) \mathbb{1}_{\{\tilde{\tau} > t\}}, \quad (t \leq s).$$

Les hypothèse du lemme précédent étant vérifiées ($\tilde{\mathcal{G}} = \tilde{\mathbb{F}} \vee \tilde{\mathbb{H}}$), il s'applique à nos filtrations pour donner d'une part

$$\begin{aligned} \mathbb{P}[\tilde{\tau} > s | \tilde{\mathcal{F}}_s \vee \tilde{\mathcal{G}}_t] &= \mathbb{P}[\tilde{\tau} > s | \tilde{\mathcal{F}}_s \vee \tilde{\mathcal{H}}_t] \\ &= \mathbb{1}_{\{\tilde{\tau} > t\}} \frac{\mathbb{P}[\tilde{\tau} > s | \tilde{\mathcal{F}}_s]}{\mathbb{P}[\tilde{\tau} > t | \tilde{\mathcal{F}}_s]}. \end{aligned}$$

D'autre part et puisque par hypothèse U est indépendant de $\tilde{\mathcal{G}}_\infty$ et donc de $\tilde{\mathcal{F}}_s$, on peut aussi écrire que pour tout $t \leq s$

$$\mathbb{P}[\tilde{\tau} > t | \tilde{\mathcal{F}}_s] = \exp\left(-\int_0^t \lambda_u du\right).$$

Le résultat désiré est alors immédiat.

Nous sommes désormais en mesure de prouver le principal théorème de ce paragraphe. Celui-ci est fondamental et il interviendra à de nombreuses reprises dans la suite.

THÉORÈME 4.6 *Soient τ un temps d'arrêt de \mathbb{G} d'intensité λ et doublement stochastique relativement à \mathbb{F} et Z un processus \mathbb{F} -prévisible borné. Alors, pour tout $s \geq t$, on a*

$$\mathbb{E}[\mathbb{1}_{\{t < \tau \leq s\}} Z_\tau | \mathcal{G}_t] = \mathbb{1}_{\{\tau > t\}} e^{\Lambda(t)} \mathbb{E}\left[\int_{]t, s]} Z_u \lambda_u e^{-\Lambda_u} du \middle| \mathcal{F}_t\right],$$

$$\text{où } \Lambda(t) = \int_0^t \lambda_s ds.$$

Noter que ce théorème s'applique aux processus de Cox.

REMARQUE. Ce résultat très puissant nous dit que les calculs dans la filtration étendue (du trader) peuvent se ramener (dans le cadre des modèles doublement stochastiques) à des calculs dans la filtration de marché (où les calculs sont simples, par exemple lorsque \mathbb{F} est engendrée par un processus affine).

REMARQUE. (admise) La tribu \mathbb{F} -prévisible est la tribu sur $\Omega \times \mathbb{R}_+$ engendrée par

1. les processus $(X_t, t \geq 0)$ continus et \mathbb{F} -adaptés ;
2. les ensembles de la forme $\begin{cases} A \times]s, t], s < t, A \in \mathcal{F}_s \\ A \times 0, A \in \mathcal{F}_0. \end{cases}$

PREUVE. Nous commençons par prouver le résultat pour un processus \mathbb{F} -prévisible Z de la forme $\mathbb{1}_{]u,v]} Y_u$ où Y_u est une v.a. \mathcal{F}_u -mesurable. Dans ce cas, le résultat se déduit des calculs suivants :

$$\begin{aligned} \mathbb{E}[\mathbb{1}_{\{t < \tau \leq s\}} Z_\tau | \mathcal{G}_t] &= \mathbb{E}[\mathbb{1}_{\{\tau > t\}} \mathbb{1}_{\{u \vee t < \tau \leq s \wedge v\}} Y_u | \mathcal{H}_t \vee \mathcal{F}_t] \\ &= \mathbb{1}_{\{\tau > t\}} \frac{\mathbb{E}[\mathbb{1}_{\{u \vee t < \tau \leq s \wedge v\}} Y_u | \mathcal{F}_t]}{\mathbb{P}[\tau > t | \mathcal{F}_t]} \\ &= \mathbb{1}_{\{\tau > t\}} e^{\Lambda t} \mathbb{E}[Y_u (\mathbb{1}_{\{\tau > u \vee t\}} - \mathbb{1}_{\{\tau > s \wedge v\}}) | \mathcal{F}_t] \\ &= \mathbb{1}_{\{\tau > t\}} e^{\Lambda t} \mathbb{E}[Y_u (\mathbb{E}[\mathbb{1}_{\{\tau > u \vee t\}} | \mathcal{F}_{u \vee t}] - \mathbb{E}[\mathbb{1}_{\{\tau > s \wedge v\}} | \mathcal{F}_{s \wedge v}])] | \mathcal{F}_t] \\ &= \mathbb{1}_{\{\tau > t\}} e^{\Lambda t} \mathbb{E}[Y_u (e^{-\Lambda_{u \vee t}} - e^{-\Lambda_{s \wedge v}}) | \mathcal{F}_t] \\ &= \mathbb{1}_{\{\tau > t\}} e^{\Lambda(t)} \mathbb{E}\left[\int_{]t,s]} Z_w de^{-\Lambda_w} \middle| \mathcal{F}_t\right] \\ &= \mathbb{1}_{\{\tau > t\}} e^{\Lambda(t)} \mathbb{E}\left[\int_{]t,s]} Z_w \lambda_w e^{-\Lambda_w} dw \middle| \mathcal{F}_t\right]. \end{aligned}$$

Par linéarité, le résultat reste vrai pour un Z de la forme $\sum_{i=1}^N \mathbb{1}_{]t_i, t_{i+1}]} Y_{t_i}$ où les Y_{t_i} sont des v.a. \mathcal{F}_{t_i} -mesurables bornées. Pour prouver le cas général, il suffit alors de remarquer qu'un processus \mathbb{F} -prévisible borné Z peut être approché uniformément par une suite de processus prévisible de la forme précédente. ■

Un théorème de représentation

Nous terminons cette section par l'énoncé d'un théorème qui précise la structure de la filtration \mathbb{G} lorsque \mathbb{F} est une filtration brownienne. Nous supposons donc ici que \mathbb{F} est la filtration engendrée par un mouvement brownien B (éventuellement multi-dimensionnel).

THÉORÈME 4.7 Soient X une v.a. et M la \mathbb{G} -martingale $M_t = \mathbb{E}[X | \mathcal{G}_t], t \in [0, T]$.

Alors, il existe un processus \mathbb{F} -prévisible ξ et un processus \mathbb{G} -prévisible ζ tels que

$$M_t = M_0 + \int_0^t \xi_s dB_s + \int_{]0, \tau \wedge t]} \zeta_u (\lambda_u du - dN_u).$$

PREUVE. Voir [2]. ■

4.2 Évaluation des actifs risqués

Dans cette section, nous nous intéressons à l'évaluation des actifs soumis au risque de défaut. Nous nous plaçons dans le cadre de l'absence d'opportunité d'arbitrage. Nous supposons donc que tous les actifs considérés sont définis sur un espace probabilisé filtré $(\Omega, \mathcal{A}, \mathbb{P}^*, \mathbb{G} = (\mathcal{G}_t)_{t \geq 0})$. Nous supposons de plus qu'il existe un temps d'arrêt τ de la filtration \mathbb{G} , d'intensité λ , doublement stochastique relativement à une filtration $\mathbb{F} \subset \mathbb{G}$ telle que $\mathbb{G} = \mathbb{F} \vee \mathbb{H}$, où $\mathcal{H}_t = \sigma(\tau \wedge t)$. Ce temps τ représente l'instant de défaut d'une entité de référence fixée. L'hypothèse sur les filtration signifie que toutes les informations de marché autres que l'instant de défaut τ sont contenues dans la filtration \mathbb{F} .

La probabilité \mathbb{P}^* est une probabilité risque-neutre relativement au processus de taux d'intérêt $r = (r_t)_{t \geq 0}$ que nous choisirons \mathbb{F} -adapté. Ceci entraîne, en particulier, que les prix des actifs contingents à évaluer se calculent comme l'espérance sous \mathbb{P}^* des flux futurs réactualisés à l'aide du taux sans risque r .

Le cas du fractional recovery of par value

Nous considérons un produit donnant droit en T à un flux F représenté par une v.a. \mathcal{G}_T -mesurable et payant un flux W_τ (taux de recouvrement) si l'entité de référence fait défaut avant la maturité T du titre où W est un processus \mathbb{F} -prévisible. La valeur en t d'un tel actif s'écrit

$$p(t, T) = \mathbb{E}^* \left[e^{-\int_0^T r_s ds} F \mathbb{1}_{\{\tau > t\}} + e^{-\int_0^\tau r_s ds} W_\tau \mathbb{1}_{\{t < \tau \leq T\}} \middle| \mathcal{G}_t \right],$$

or, il découle du théorème 4.6 que

$$\begin{aligned} \mathbb{E}^* \left[e^{-\int_0^T r_s ds} F \mathbb{1}_{\{\tau > t\}} \middle| \mathcal{G}_t \right] &= \mathbb{1}_{\{\tau > t\}} e^{\int_0^t \lambda_u du} \mathbb{E} \left[e^{-\int_t^T r_u du} F \mathbb{1}_{\{\tau > t\}} \middle| \mathcal{F}_t \right] \\ &= \mathbb{1}_{\{\tau > t\}} e^{\int_0^t \lambda_u du} \mathbb{E} \left[e^{-\int_t^T (r_u + \lambda_u) du} F \middle| \mathcal{F}_t \right], \end{aligned}$$

et

$$\begin{aligned} \mathbb{E}^* \left[e^{-\int_0^\tau r_s ds} W_\tau \mathbb{1}_{\{t < \tau \leq T\}} \middle| \mathcal{G}_t \right] &= \mathbb{1}_{\{\tau > t\}} e^{\int_0^t \lambda_u du} \mathbb{E} \left[\int_{(t, T]} W_s \lambda_s e^{-\int_t^s (r_u + \lambda_u) du} ds \middle| \mathcal{F}_t \right] \\ &= \mathbb{1}_{\{\tau > t\}} e^{\int_0^t \lambda_u du} \int_{(t, T]} \phi(t, s) ds, \end{aligned}$$

où

$$\phi(t, s) = \mathbb{E}^* \left[e^{\int_0^s (r_u + \lambda_u) du} W_s \lambda_s \middle| \mathcal{F}_t \right],$$

et finalement, la valeur de l'actif s'écrit

$$p(t, T) = \mathbb{1}_{\{\tau > t\}} \left(\mathbb{E}^* \left[e^{-\int_0^T (r_s + \lambda_s) ds} F \middle| \mathcal{F}_t \right] + e^{\int_0^t \lambda_s ds} \int_{]t, T]} \phi(t, s) ds \right).$$

REMARQUE. Le calcul de la première espérance met en lumière le fait que la valeur d'un flux de la forme $F \mathbb{1}_{\{\tau > t\}}$ se calcule comme celle d'un flux de la forme F avec un taux d'actualisation $r + \lambda$. Ceci constitue une première interprétation de l'intensité de défaut comme un spread émetteur instantané.

Le cas du fractional recovery of market value

Une hypothèse de recouvrement particulièrement intéressante est l'hypothèse dite *Market Value Recovery* (MVR). Soient une variable aléatoire \mathcal{F}_T -mesurable F et un processus \mathbb{F} -prévisible l à valeurs dans $(0, 1)$, nous considérons l'actif contingent W comme le processus \mathbb{F} -prévisible payant le flux suivant :

- F en T si l'entité de référence n'a pas fait défaut à maturité ;
- $W_t = (1 - l_t)p(t-, T)$ si le défaut intervient en $t < T$ où $p(t, T)$ représente la valeur en t de cet actif.

Par absence d'opportunité d'arbitrage, la valeur $p(t, T)$ en t de cet actif doit satisfaire à l'équation

$$p(t, T) = \mathbb{E} \left[e^{-\int_t^T r_s ds} F \mathbb{1}_{\{\tau > t\}} + e^{-\int_t^T r_s ds} l_\tau p(\tau-, T) \mathbb{1}_{\{t < \tau \leq T\}} \middle| \mathcal{G}_t \right].$$

La valeur p s'exprime donc sous la forme d'une équation différentielle stochastique backward (EDSB). Nous admettrons dans la suite que cette équation admet une unique solution. Pour plus d'informations sur la théorie des EDSB, le lecteur pourra se référer à [13] ou au cours de M. Chaleyat-Maurel.

PROPOSITION 4.8 *Soit*

$$V_t = \mathbb{E} \left[e^{-\int_t^T (r_s + (1-l_s)\lambda_s) ds} F \middle| \mathcal{F}_t \right],$$

nous supposons que, presque sûrement, $\Delta V_\tau = 0$. Alors $U_t = V_t \mathbb{1}_{\{\tau > t\}}$ est l'unique solution de l'équation différentielle stochastique backward

$$U_t = \mathbb{E} \left[e^{-\int_t^T r_s ds} F \mathbb{1}_{\{\tau > t\}} + e^{-\int_t^T r_s ds} l_\tau U_{\tau-} \mathbb{1}_{\{t < \tau \leq T\}} \middle| \mathcal{G}_t \right].$$

PREUVE. Voir [2] ou l'article fondateur [10]. ■

REMARQUE. Pour valoriser un actif soumis au risque de défaut dans le cadre du fractional recovery of market value, il suffit d'ajouter au taux sans risque r le *spread instantané* $s_t = (1 - l_t)\lambda_t$. Ceci est constituée une autre version de la relation du triangle déjà obtenue dans le cadre du pricing des CDS :

$$S = (1 - R)\lambda,$$

où R désigne le taux de recouvrement et λ le taux de défaut.

EXERCICE. (Intensité de défaut et Modèle Affine de Base.)

Soit N un processus de Poisson de paramètre λ et $U = \{U_i, i \geq 1\}$ une suite de v.a. i.i.d. de loi ν à support inclus dans \mathbb{R}_+ . On considère le processus h dont la dynamique est décrite par l'EDS

$$dh_t = \kappa(\bar{h} - h_t) dt + \gamma\sqrt{h_t} dW_t + dJ_t$$

où W est un mouvement brownien standard et J est le processus de saut $J_t = \sum_{i=0}^{N_t} U_i$. On désigne par \mathcal{F} la filtration engendrée par le processus h .

(1) Soit L l'opérateur défini par

$$Lf(x) = \frac{\gamma^2}{2} x f''(x) + \kappa(\bar{h} - x) f'(x) + \lambda \int (f(x+z) - f(x)) \nu(dz).$$

Vérifier que pour toute fonction f de classe C^2 , bornée et à dérivés bornés, le processus

$$f(h_t) - \int_0^t (Lf)(h_{s-}) ds$$

est une martingale.

(2) On admet qu'il existe deux fonctions de classe C^1 α_u et β_u telles que

$$\mathbb{E} \left[e^{uh_T - \int_t^T h_s ds} \middle| \mathcal{F}_t \right] = e^{\alpha_u(T-t) + h_t \beta_u(T-t)}, \quad (0 \leq t \leq T).$$

Montrer que α_u et β_u sont solutions des équations de Riccati

$$\begin{cases} \alpha' &= -\bar{h}\kappa\beta - \lambda(\theta(\beta) - 1) \\ \beta' &= 1 - \kappa\bar{h}\beta - \frac{\gamma^2}{2}\beta^2, \end{cases}$$

avec $\alpha(0) = 0$, $\beta(0) = u$ et $\theta(c) = \mathbb{E}[\exp(cU_1)]$ ($c \in \mathbb{C}$).

(3) On suppose que h est l'intensité risque-neutre de l'instant de défaut τ d'une entreprise et que le taux sans risque est constant égal à r . Calculer la marge à la monnaie (break even spread) d'un credit default swap dont la jambe fixe prend en compte les flux intervenant avant le défaut et le coupon couru, et la jambe variable est payée dès l'occurrence du défaut.

EXERCICE. (Examen juin 2004.)

Le but de ce problème est d'établir les formules de base de la valorisation des actifs soumis au risque de défaut dans un modèle où taux sans risque et intensité de défaut sont des processus corrélés.

Partie I : Préliminaires. Soit $(\Omega, \mathcal{A}, \mathbb{P}, \mathbb{F} = (\mathcal{F}_t)_{t \geq 0})$ un espace probabilisé filtré sur lequel vit un \mathbb{F} -mouvement brownien W . On considère le processus x solution de l'EDS

$$dx_t = (\kappa_t - \alpha x_t) dt + \sigma dW_t, \quad x_0 > 0$$

où κ est une fonction déterministe strictement positive et $\alpha, \sigma \in (0, +\infty)$.

(1) Vérifier que $y_t := e^{\alpha t} x_t$ satisfait à l'EDS

$$dy_t = e^{\alpha t} (\kappa_t dt + \sigma dW_t), \quad y_0 = x_0.$$

(2) Montrer que, pour tout $s \geq t$, $z(s, t) = x_s - e^{-\alpha(s-t)} x_t$ est une gaussienne indépendante de la tribu \mathcal{F}_t , dont on déterminera la variance.

(3) En déduire l'existence de deux fonctions A et B (que l'on ne cherchera pas à calculer) telles que

$$\mathbb{E}[e^{-\int_t^T x_s ds} | \mathcal{F}_t] = \exp(A(t, T) - x_t B(t, T)) =: \psi(t, x_t).$$

(4) Montrer que les fonctions A et B sont solutions des EDO

$$\begin{aligned} \frac{\partial B}{\partial t}(t, T) &= \alpha B(t, T) - 1, & B(T, T) &= 0, \\ \frac{\partial A}{\partial t}(t, T) &= -\frac{\sigma^2}{2} B^2(t, T) + \kappa_t B(t, T), & A(T, T) &= 0. \end{aligned}$$

[Indication : le processus $M_t = e^{-\int_0^t x_s ds} \psi(t, x_t)$ est une \mathbb{F} -martingale.]

(5) En déduire que

$$\begin{aligned} A(t, T) &= A(t, T; \alpha, \kappa, \sigma) = \frac{\sigma^2}{2} \int_t^T B^2(s, T) ds - \int_t^T \kappa_s B(s, T) ds, \\ B(t, T) &= B(t, T; \alpha) = \frac{1 - e^{-\alpha(T-t)}}{\alpha}. \end{aligned}$$

Partie II. Tous les processus et v.a. considérés dans la suite vivent sur un espace probabilisé complet $(\Omega, \mathcal{A}, \mathbb{P})$. Soient $B = (B_t^1, B_t^2; t \geq 0)$ un mouvement brownien standard et $W = (W_t^1, W_t^2; t \geq 0)$ le processus défini par

$$W^1 = B^1, \quad W^2 = \rho B^1 + \sqrt{1 - \rho^2} B^2 \quad (\rho \in [-1, +1]),$$

de sorte que W soit un mouvement brownien tel que

$$d\langle W^1, W^2 \rangle_t = \rho dt.$$

On désigne par $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ la filtration engendrée par B . On considère une économie dans laquelle les dynamiques risque-neutre des taux sans risque r et de l'intensité de défaut λ de l'entité de référence sont de la forme

$$\begin{aligned} dr_t &= (k_t - ar_t) dt + \sigma dW_t^1, \\ d\lambda_t &= (\bar{k}_t - \bar{a}\lambda_t) dt + \bar{\sigma} dW_t^2. \end{aligned}$$

Les fonctions strictement positives et déterministes k et \bar{k} ainsi que les paramètres a, \bar{a}, σ et $\bar{\sigma}$ sont supposés connus. On considère une v.a. τ à valeurs dans $[0, +\infty]$ telle que, si \mathbb{G} est la filtration engendrée par \mathbb{F} et τ , à savoir $\mathcal{G}_t = \mathcal{F}_t \vee \sigma(\tau \wedge t)$, alors

- (i) τ admet λ pour \mathbb{G} -intensité,
- (ii) τ est doublement stochastique relativement à \mathbb{F} .

On admettra que, dans ce cas, la propriété suivante est vérifiée : tout \mathbb{F} -mouvement brownien est un \mathbb{G} -mouvement brownien.

L'espace probabilisé filtré $(\Omega, \mathcal{A}, \mathbb{P}, \mathbb{G})$ est la base du modèle d'évaluation des actifs de sorte que la tribu \mathcal{G}_t modélise l'information disponible à l'instant t et τ représente l'instant de défaut de l'entité de référence.

(1) Montrer que la valeur $B(t, T)$, en t , d'une obligation zéro-coupon sans risque de maturité T s'écrit

$$\exp(A(t, T; a; k; \sigma) - r_t B(t, T; a)).$$

Quelle est la volatilité de cette obligation zéro-coupon ?

(2) Montrer que la probabilité de survie de l'entité de référence au delà de l'instant T , conditionnellement à l'information disponible à l'instant t , s'écrit

$$I_{\{\tau > t\}} \exp(A(t, T; \bar{a}; \bar{k}; \bar{\sigma}) - \lambda_t B(t, T; \bar{a})).$$

(3) Soit \mathbb{P}_T la probabilité forward-neutre de la maturité T dans le monde sans risque de défaut. Autrement dit, \mathbb{P}_T est défini par les relations

$$\frac{d\mathbb{P}_T}{d\mathbb{P}} \Big|_{\mathcal{F}_t} = \frac{1}{e^{\int_0^t r_s ds}} \frac{B(t, T)}{B(0, T)}, \quad (0 \leq t \leq T).$$

- (i) Montrer que le processus $B_t^T = (B_t^1 + \sigma \int_0^t B(s, T; a) ds, B_t^2)$ est un \mathbb{P}_T -mouvement brownien standard.
- (ii) En déduire que la valeur $\bar{B}(t, T)$, en t , d'une obligation zéro-coupon risqué de maturité T et de taux de recouvrement nul s'écrit

$$I_{\{\tau > t\}} B(t, T) \exp(A(t, T; \bar{a}; \tilde{k}; \bar{\sigma}) - \lambda_t B(t, T; \bar{a}))$$

où

$$\tilde{k}_t = \bar{k}_t - \rho \bar{\sigma} \sigma B(t, T; a).$$

(4) Montrer que la valeur $\bar{B}_{RMV}(t, T)$, en t , d'une obligation zéro-coupon risqué de maturité T soumise à l'hypothèse "Fractional Recovery of Market Value" et de taux de recouvrement égal à $R \in (0, 1)$ s'écrit

$$I_{\{\tau > t\}} B(t, T) \exp(A'(t, T) - (1 - R) \lambda_t B(t, T; \bar{a}))$$

où

$$A'(t, T) = \frac{(1 - R)^2 \bar{\sigma}^2}{2} \int_t^T B(s, T; \bar{a})^2 ds - (1 - R) \int_t^T B(s, T; \bar{a}) \tilde{k}_s ds.$$

(5) Montrer que la valeur, en t , d'un titre rapportant 1 à l'instant du défaut, si celui-ci se produit avant la maturité T , s'écrit

$$I_{\{\tau > t\}} \int_t^T e(t, s) ds$$

où

$$e(t, s) = \mathbb{E}[\lambda(s)e^{-\int_t^s (r_u + \lambda_u) du}], \quad (s \geq t).$$

(6) Soit \mathbb{P}'_T la probabilité définie par les relations

$$\frac{d\mathbb{P}'_T}{d\mathbb{P}} \Big|_{\mathcal{F}_t} = \frac{1}{e^{\int_0^t (r_s + \lambda_s) ds}} \frac{\bar{P}(t, T)}{\bar{P}(0, T)}, \quad (0 \leq t \leq T)$$

où $\bar{P}(t, T) = \mathbb{E}[e^{-\int_t^T (r_s + \lambda_s) ds} | \mathcal{F}_t]$.

(i) Montrer que le processus

$$\tilde{B}_t^T = \left(B_t^1 + \int_0^t (\sigma B(s, T; a) + \rho \bar{\sigma} B(s, T; \bar{a})) ds, B_t^2 + \sqrt{1 - \rho^2} \int_0^t \bar{\sigma} B(s, T; \bar{a}) ds \right)$$

est un \mathbb{P}'_T -mouvement brownien.

(ii) En déduire que

$$e(t, s) = \bar{B}(t, T) \left(\lambda_t e^{-\bar{a}(T-t)} + \int_t^T e^{-a(T-s)\tilde{k}'(s)} ds \right)$$

où

$$\tilde{k}'(t) = \bar{k}(t) - \rho \bar{\sigma} \sigma B(t, T; a) - \bar{\sigma}^2 B(t, T; \bar{a}).$$

CHAPITRE 5

Défauts corrélés

Dans ce chapitre, nous décrivons plusieurs méthodes pour définir la loi jointe d'une famille d'instants de défauts.

Une entreprise donnée possède nécessairement des liens économiques avec d'autres acteurs (fournisseurs, clients, débiteurs...), de telle sorte que son défaut peut affecter la qualité du crédit d'autres entreprises. Il s'avère donc crucial de modéliser cette relation de dépendance (souvent appelée corrélation) entre les défauts lorsque l'on s'intéresse à des situations faisant intervenir le risque de défaut de nombreux émetteurs. De telles situations sont courantes : citons les problèmes d'évaluation et couverture des produits dérivés de crédit sur panier (Nth-to-default, CDO), ceux de la mesure du risque de crédit d'un portefeuille de créance et des calculs d'exposition au risque de contrepartie d'un portefeuille de produits dérivés OTC.

De manière générale, on distingue deux grands principes :

1. On suppose en général que le principal facteur de risque de crédit est un facteur économique global. La qualité du crédit des entreprises d'un même secteur économique (Europe, États-Unis, Asie) est donc a priori corrélée positivement.
2. La dégradation de la qualité de crédit d'une firme donnée peut entraîner la dégradation de la santé financière des entreprises partenaires d'un même secteur industriel. Là encore, la corrélation est donc positive.

Dans la pratique, la co-dépendance s'observe effectivement sur les marchés par la dégradation du spread de crédit des entreprises partenaires lors du défaut d'une entreprise particulière.

L'enjeu est d'élaborer des modèles qui rendent compte de ces phénomènes tout en restant aisément interprétables. Nous nous restreindrons ici à la modélisation la plus courante à l'aide des copules.

Nous présentons tout d'abord des méthodes de modélisation du défaut d'un ensemble de crédits. Nous nous intéressons ensuite à la généralisation des deux approches précédentes, structurelle et à forme réduite, dans le cadre des défauts multiples. Enfin,

nous traitons le problème de la calibration sur sauts de spread.

5.1 Produits dérivés sur un panier de crédits

On considère ici des produits dérivés écrits sur un panier de k crédits possédant les caractéristiques suivantes :

- $R_i \in (0, 1)$ est le taux de recouvrement du $i^{\text{ème}}$ crédit ;
- N_i est le nominal du $i^{\text{ème}}$ crédit ;
- τ_i est l'instant de défaut du $i^{\text{ème}}$ crédit ;

On fait l'hypothèse selon laquelle les instants de défaut sont indépendants du taux court :

$$(r_s)_{s \geq 0} \perp (\tau_1, \dots, \tau_k)$$

et l'on suppose connue la distribution risque neutre

$$F_i(t) = \mathbb{P}^*[\tau_i \leq t]$$

de chacun des instants de défaut. Par exemple, dans le cas d'une intensité λ_i de défaut constante, le spread de CDS s_i s'écrit

$$s_i = (1 - R_i)\lambda_i,$$

ce qui conduit à la distribution suivante :

$$F_i(t) = e^{-\lambda_i t} = \exp\left(-\frac{s_i}{1 - R_i} t\right).$$

Dans la suite, on notera

$$\min(\tau_i) = \tau_{(1)} < \tau_{(2)} < \dots < \tau_{(k)} = \max(\tau_i)$$

le vecteur des statistiques d'ordre du vecteur $\tau = (\tau_1, \dots, \tau_k)$.

Nous allons distinguer deux types de produits dérivés sur panier de crédits : les n^{th} -to-default et les CDO.

n^{th} -to-default

Un n^{th} -to-default offre une protection contre le $n^{\text{ième}}$ défaut d'un panier de crédits. $n \in \{1, \dots, k\}$ est appelé *ordre de défaut*. Nous supposons pour simplifier que le panier est homogène, i.e.

$$\forall i \in \{1, \dots, k\}, N_i = N, R_i = R.$$

Une prime (spread) d'assurance est payée jusqu'à l'occurrence du $n^{\text{ième}}$ défaut. Le vendeur de protection paye alors à l'acheteur du n^{th} -to-default la quantité L définie par :

$$L = (1 - R_i)N \quad \text{si } \tau_{(n)} = \tau_i.$$

La valorisation d'un n^{th} -to-default nécessite l'évaluation des jambes fixe et variable sous la forme des espérances suivantes :

$$NPV(JF) = N \sum_i s(T_{i+1} - T_i) \mathbb{E}^* \left[e^{-\int_0^{T_i} r_s ds} \mathbb{1}_{\{\tau_{(n)} > T_i\}} \right],$$

et sous l'hypothèse d'indépendance $r \perp (\tau_1, \dots, \tau_k)$:

$$\begin{aligned} NPV(JF) &= N \sum_i s(T_{i+1} - T_i) B(0, T_i) \mathbb{P}^* [\tau_{(n)} > T_i] \\ &\simeq SN \int_0^T B(0, t) Q_n(t) dt \end{aligned}$$

où $Q_n(t)$ désigne la survie du $n^{\text{ième}}$ crédit à faire défaut au temps t . En ce qui concerne la jambe variable :

$$NPV(JV) = \mathbb{E}^* \left[e^{-\int_0^{\tau_{(n)}} r_s ds} L \mathbb{1}_{\{\tau_{(n)} \leq T\}} \right],$$

et sous l'hypothèse d'homogénéité, $L = (1 - R)N$:

$$\begin{aligned} NPV(JV) &= (1 - R)N \int_0^T B(0, t) \mathbb{P}^* [\tau_{(n)} \in dt] \\ &\simeq (1 - R)N \int_0^T B(0, t) Q_n(t) dt, \end{aligned}$$

ce qui permet d'en déduire l'expression du spread immédiatement.

Collateralized Debt Obligations (CDO)

Dans cette partie, on note $0 \leq A < B \leq 1$ et $N = \sum_i N_i$ le nominal du panier. L'acheteur de protection sur la tranche A - B paye à chaque période un spread correspondant au capital restant sur la tranche et reçoit en échange une protection contre toutes les pertes comprises entre $A \times N$ et $B \times N$. La perte cumulée L_t à l'instant t s'écrit :

$$L_t = \sum_{i=1}^k N_i (1 - R_i) \mathbb{1}_{\{\tau_i \leq t\}},$$

et donc la perte cumulée sur la tranche A - B est

$$L_t(A, B) = (L_t - A)_+ - (L_t - B)_+.$$

On en déduit le pourcentage du capital $C_t^{A,B}$ restant à l'instant t :

$$C_t^{A,B} = 1 - \frac{(L_t - A)_+ - (L_t - B)_+}{B - A}.$$

On peut donc calculer le prix du CDO : l'acheteur de protection verse $SN(B-A)C_{T_i}^{A,B}$ à l'instant T_i donc la jambe fixe s'exprime comme :

$$\begin{aligned} NPV(JF) &= \sum_i (T_{i+1} - T_i) SN(B-A) \mathbb{E}^*[C_{T_i}^{A,B}] \\ &\simeq SN(B-A) \int_0^T B(0,t) Q(A,B,t) dt, \end{aligned}$$

où $Q(A,B,t) = \mathbb{E}^*[C_{T_i}^{A,B}]$ est la survie sur la tranche $A-B$. Le vendeur de protection verse à l'instant du $i^{\text{ème}}$ défaut $L_{\tau(i)}^{(A,B)} - L_{\tau(i)-}^{(A,B)}$, donc la jambe variable vaut :

$$\begin{aligned} NPV(JV) &= \mathbb{E}^* \left[\int_0^T B(0,t) dL_t^{A,B} \right] \\ &= \int_0^T B(0,t) d\mathbb{E}^*[L_t^{A,B}] \\ &= -(B-A) \int_0^T B(0,t) Q(A,B,t) dt. \end{aligned}$$

Tout le problème se résume donc au calcul de

$$Q(A,B,t) = \mathbb{E}^*[C_t^{A,B}]$$

pour tout t , ce qui nécessite de modéliser la loi jointe des (τ_1, \dots, τ_k) connaissant les lois marginales, i.e. la loi des τ_i , ($1 \leq i \leq k$).

L'outil privilégié pour résoudre ce type de problème est la *fonction copule* (voir l'annexe D).

5.2 Corrélation dans les modèles structurels

Dans cette partie, on suppose que chaque défaut est déclenché par un mécanisme à la Merton, i.e. $\forall i \in \{1, \dots, k\}, \exists X_i \sim N(0,1)$ telles que

$$\tau_i \leq t \Leftrightarrow X_i \leq B_i(t).$$

X_i peut s'interpréter comme le rendement normalisé de valeur de la firme. Pour que l'hypothèse sur la loi des τ_i soit vérifiée, il faut que

$$F_i(t) = \mathbb{P}[X_i \leq B_i(t)] = \phi(B_i(t)),$$

soit

$$B_i(t) = \phi^{-1}(F_i(t)),$$

où

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$

EXEMPLE. (Modèle à un facteur)

On suppose que les variables X_i partagent un facteur commun X , décrivant l'état

de l'économie mondiale. Un tel facteur est dit *facteur systémique* car il modélise un risque global, par opposition aux *facteurs idiosyncratiques* Y_i modélisant un risque de crédit spécifique à l'émetteur. En d'autres termes, $\exists X, Y_1, \dots, Y_n$ iid suivant une loi $N(0, 1)$ et $\rho_i \in [0, 1]$ tels que

$$X_i = \rho_i X + \sqrt{1 - \rho_i^2} Y_i.$$

ρ_i est une mesure de la dépendance de la qualité du crédit aux facteurs économiques globaux.

REMARQUE. Conditionnellement au facteur X , les instants de défaut sont indépendants. En conséquence, la loi jointe peut s'écrire :

$$\begin{aligned} & \mathbb{P}[\tau_1 \leq t_1, \dots, \tau_k \leq t_k] \\ &= \mathbb{E}[\mathbb{P}[\tau_1 \leq t_1, \dots, \tau_k \leq t_k | X]] \\ &= \mathbb{E}[\mathbb{P}[X_1 \leq B_1(t), \dots, X_k \leq B_k(t) | X]] \\ &= \mathbb{E}[\mathbb{P}[X_1 \leq B_1(t) | X] \dots \mathbb{P}[X_k \leq B_k(t) | X]] \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathbb{P}[X_1 \leq B_1(t) | X = x] \dots \mathbb{P}[X_k \leq B_k(t) | X = x] e^{-\frac{x^2}{2}} dx, \end{aligned}$$

où

$$\mathbb{P}[X_i \leq B_i(t) | X = x] = \Phi\left(\frac{\phi^{-1}(F_i(t)) - \rho_i x}{\sqrt{1 - \rho_i^2}}\right).$$

REMARQUE. Cet exemple est en fait un cas de modélisation avec la *copule gaussienne*. Dans la pratique, le choix de la copule à utiliser est guidé par l'hypothèse que l'on fait sur la dépendance jointe des rendements des actifs des entreprises. Des études récentes tendent à prouver que cette dépendance n'est pas effectivement gaussienne, mais de type student (dont le degré de liberté est estimé à 8 ou 9). Ces études justifieraient la pratique courante qui consiste à utiliser la copule de student.

5.3 Corrélation des intensités

Dans cette section, $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{G} = (\mathcal{G}_t)_{t \geq 0})$ désigne un espace probabilisé satisfaisant aux conditions habituelles. Nous introduisons la définition suivante qui généralise la définition unidimensionnelle introduite au chapitre précédent (cf. définition 4.3).

DÉFINITION 5.1 Une famille (τ_1, \dots, τ_N) de \mathbb{G} -temps d'arrêt d'intensité λ_i est dite doublement stochastique relativement à la filtration \mathbb{F} si les processus λ_i sont \mathbb{F} -adapté et si pour tout $t_1, \dots, t_N \in [t, s]$, on a

$$\mathbb{P}[\tau_1 > t_1, \dots, \tau_N > t_N | \mathcal{F}_s \vee \mathcal{G}_t] = \prod_{i=1}^N \exp\left(-\int_t^{t_i} \lambda_i(u) du\right) \mathbb{1}_{\{\tau_i > t\}}.$$

EXERCICE. Montrer que dans ce cadre le \mathbb{G} -temps d'arrêt $\tau = \tau_1 \wedge \dots \wedge \tau_N$ est doublement stochastique relativement à la filtration \mathbb{F} d'intensité $\lambda = \sum_{i=1}^N \lambda_i$.

Ici encore, nous partons d'une filtration $\tilde{\mathbb{F}}$ représentant l'information de marché hors instant de défaut et nous construisons une famille de v.a. $\tilde{\tau}_1, \tilde{\lambda}_N$ et une filtration $\tilde{\mathbb{G}}$ de la manière suivante :

Soient λ_i des processus $\tilde{\mathbb{F}}$ -adapté et positif, γ_i les processus $t \mapsto \exp(-\int_0^t \lambda_i(s) ds)$ et les U_i des v.a. indépendantes de loi uniforme sur $[0, 1]$ et indépendante de $\tilde{\mathcal{F}}_\infty$. Si nous définissons les v.a.

$$\tilde{\tau}_i = \inf\{t \geq 0; \gamma_i(t) \leq U_i\}$$

et les tribus

$$\tilde{\mathcal{G}}_t = \tilde{\mathcal{F}}_t \vee \tilde{\mathcal{H}}_t^1 \vee \dots \vee \tilde{\mathcal{H}}_t^1,$$

où $\tilde{\mathcal{H}}_t^i = \sigma(\tau_i \wedge t)$, alors il est facile de vérifier que les $\tilde{\tau}_i$ sont des temps d'arrêt de $\tilde{\mathbb{G}}$ d'intensité λ_i (exercice!). On peut montrer en utilisant des arguments similaires à ceux déjà employé dans le chapitre précédent (exercice!) que les $\tilde{\tau}_i$ sont, de plus, doublement stochastique relativement à \mathbb{F} .

5.4 Corrélation des instants de défaut

L'approche que nous présentons dans cette section est due à Schonbücher et Schubert [29]. Dans ce modèle, l'on "corrèle" directement les instants de défaut à l'aide d'une copule. Cette technique va permettre d'éclairer l'impact de la modélisation sur les trajectoires des intensités.

Dans la suite $(\Omega, \mathcal{A}, \mathbb{P})$ est l'espace probabilisé de base sur lequel nous travaillons. Il est supposé suffisamment riche pour contenir toutes les v.a. que nous introduirons par la suite.

Description du modèle

On suppose donnés une filtration \mathbb{F} contenant l'information sur le marché autre que les défauts (*background filtration*) et une famille de processus \mathbb{F} -adaptés et positifs λ_i . Nous appelons *pseudo-intensité* ces processus pour des raisons qui seront bientôt claires. Soit $U = (U_1, \dots, U_N)$ une v.a. indépendantes de \mathcal{F}_∞ de loi déterminée par une copule C (voir annexe D pour plus de détails sur ces objets).

Pour tout $1 \leq i \leq N$, nous définissons le temps aléatoire τ_i en posant

$$\tau_i = \inf\{t \geq 0; \gamma_i(t) \leq U_i\}$$

où $\gamma_i(t) = \exp(-\int_0^t \lambda_i(s) ds)$. Il est facile (exercice!) de calculer la fonction de survie des τ_i :

$$\mathbb{P}[\tau_1 > t_1, \dots, \tau_N > t_N] = \mathbb{E}[C(\gamma_1(t_1), \dots, \gamma_N(t_N))].$$

Nous introduisons aussi les filtrations suivantes :

$$\begin{aligned}\mathbb{G}_i &= \mathbb{F} \vee \mathbb{H}_i, & \tilde{\mathbb{G}}_i &= \mathcal{F}_\infty \vee \mathbb{G}_i, \\ \mathbb{G} &= \bigvee_{i=1}^N \mathbb{G}_i, & \tilde{\mathbb{G}} &= \bigvee_{i=1}^N \tilde{\mathbb{G}}_i,\end{aligned}$$

où $\mathbb{H}_i = (\sigma(\tau_i \wedge t))_{t \geq 0}$.

Le point important à comprendre (et qui est aussi la raison pour laquelle nous avons appelé les processus λ_i pseudo-intensité), est que l'intensité des τ_i dépend de manière cruciale de la filtration de référence considérée. Par exemple, le processus λ_i est bien la \mathbb{G}_i -intensité de τ_i (exercice!), mais ce résultat n'est plus vrai relativement à la filtration qui nous intéresse ici c'est-à-dire \mathbb{G} .

Dynamique de l'intensité

Nous allons calculer la \mathbb{G} -intensité des instants τ_i induite par le modèle et déterminer sa dynamique. Pour ce faire, nous allons utiliser le théorème 4.2 et nous commençons par la remarque suivante qui simplifiera nos calculs : si l'on note $P_i(t, T)$ et $\tilde{P}_i(t, T)$ la survie conditionnelle de τ_i relative à la filtration \mathbb{G} et $\tilde{\mathbb{G}}$ respectivement, alors la \mathbb{G} -intensité h_i de τ_i peut se calculer de la manière suivante :

$$h_i(t) = \left. \frac{\partial}{\partial T} P_i(t, T) \right|_{T=t}, \quad \tilde{h}_i(t) = \left. \frac{\partial}{\partial T} \tilde{P}_i(t, T) \right|_{T=t}.$$

$$P_i(t, T) = \mathbb{E}[\tau_i > T | \mathcal{G}_t], \quad \tilde{P}_i(t, T) = \mathbb{E}[\tau_i > T | \tilde{\mathcal{G}}_t].$$

Avant de démontrer le lemme à la base de ces calculs, nous introduisons quelques notations : si d est un sous-ensemble de $\{1, \dots, N\}$, on note C_d la dérivée $(\prod_{k \in d} \partial_k)C$; plus généralement, si d_1 et d_2 sont deux sous-ensembles de $\{1, \dots, N\}$, on note $C_{d_1+d_2}$ la dérivée $(\prod_{k_1 \in d_1} \partial_{k_1} \prod_{k_2 \in d_2} \partial_{k_2})C$; si $d_2 = \{j\}$, on écrira simplement C_{d_1+j} ; on désigne par $d(t)$ l'ensemble aléatoire $\{k; \tau_k \leq t\}$. Si $u = (u_1, \dots, u_N)$ est un vecteur et v un nombre réels, on note $u \wedge v$ et (u_{-i}, v) les vecteurs $(u_1 \wedge v, \dots, u_N \wedge v)$ et $(u_1, \dots, u_{i-1}, v, u_{i+1}, \dots, u_N)$ respectivement.

LEMME 5.2 *La survie conditionnelle $\tilde{P}_i(t, T)$ de τ_i est donnée par la formule suivante*

$$\tilde{P}_i(t, T) = \mathbb{1}_{\{\tau_i > t\}} \frac{C_{d(t)}(\gamma_{-i}(t \wedge \tau), \gamma_i(T))}{C_{d(t)}(\gamma(t \wedge \tau))}.$$

où $\gamma(t) = \left(e^{\int_0^t \lambda_1(s) ds}, \dots, e^{\int_0^t \lambda_k(s) ds} \right)$

PREUVE. Pour montrer ce lemme, on commence par remarquer que

$$\begin{aligned}\tilde{P}_i(t, T) &= \mathbb{P}[\tau_i > T | \mathcal{F}_\infty \vee \mathcal{H}_t] \\ &= \sum_{d; d \ni i} \mathbb{1}_{\{d(t)=d\}} \mathbb{P}[\tau_i > T | \mathcal{F}_\infty, \tau_j > t, \forall j \notin d, U_k = \gamma(\tau_k), \forall k \in d].\end{aligned}$$

Il suffit alors pour conclure de remarquer que pour tout d tel que $i \notin d$, on a

$$\begin{aligned} \mathbb{P}[\tau_i > T | \mathcal{F}_\infty, \tau_j > t, \forall j \notin d, U_k = \gamma(\tau_k), \forall k \in d] &= \\ \frac{\mathbb{P}[\gamma_i(T) > U_i, \gamma_j(t) > U_j, \forall j \notin d - \{i\}, U_k = \gamma(\tau_k), \forall k \in d | \mathcal{F}_\infty]}{\mathbb{P}[\gamma_j(t) > U_j, \forall j \notin d, U_k = \gamma(\tau_k), \forall k \in d | \mathcal{F}_\infty]} &= \\ \frac{C_d(\gamma_{-i}(t \wedge \tau); \gamma_i(T))}{C_d(\gamma(t \wedge \tau))}, & \end{aligned}$$

en vertu de l'indépendance de U et de \mathcal{F}_∞ . ■

REMARQUE. Ce lemme signifie que la survie conditionnelle dépend du nombre de sauts (défauts) déjà arrivés avant l'instant t . En effet, cette formule prend en compte l'information sur la structure de dépendance qui arrive au fur et à mesure que les défauts se produisent.

EXEMPLE. (*Application 1*)

De la proposition précédente, on déduit de façon immédiate que sur $\{\tau_{(1)} > t\}$ (pas de défaut)

$$\tilde{h}_i(t) = \lambda_i(t) \gamma_i(t) \frac{C_i(\gamma(t))}{C(\gamma(t))} = h_i(t)$$

car $h_i(t) = \mathbb{E}[\tilde{h}_i(t) | \mathcal{G}_t] = \tilde{h}_i(t)$.

EXEMPLE. (*Application 2*)

On va chercher à déterminer l'intensité de l'émetteur i sachant qu'en t , l'émetteur $j \neq i$ à déjà fait défaut. Autrement dit, on se place sur l'événement $\{\tau_{(1)} = \tau_j \leq t\}$. On a alors :

$$\begin{aligned} \tilde{P}_i(t, T) &= \frac{C_j(\gamma_{-i}(t \wedge \tau), \gamma_i(T))}{C_j(\gamma(t \wedge \tau))} \\ \tilde{h}_i(t) &= h_i(t) \\ h_i(t) &= \lambda_i(t) \gamma_i(t) \frac{C_{ij}(\gamma(t))}{C_j(\gamma(t))}. \end{aligned}$$

La proposition suivante décrit la dynamique de l'intensité induite par le modèle. Une fois un modèle de recouvrement choisi, nous pourrons en déduire la dynamique du spread court terme. Ce résultat nous ouvrira aussi la voie d'une méthode de calibration de ce type de modèle.

PROPOSITION 5.3 *Pour tout $1 \leq i \leq N$, on a*

$$\frac{dh_i}{h_i} = \frac{d\lambda_i}{\lambda_i} - (h_i \Delta_{ii} + \lambda_i) dt - dN_i + \sum_{j \notin d: j \neq i} \Delta_{ij} (dN_j - h_j) dt,$$

où $N_i(t) = \mathbb{1}_{\{\tau \leq t\}}$ et

$$\Delta_{ij} = \Delta_{ij}(t, d(t)) = \frac{C_{d(t)+i+j}C_{d(t)}}{C_{d(t)+i}C_{d(t)+j}}(\gamma(\tau \wedge t)) - 1.$$

Le terme $d\lambda_i/\lambda_i$ est le terme de risque de diffusion de h_i : λ_i et h_i ont la même volatilité. Le terme suivant est un terme correctif exprimant le fait que $h_i \neq \lambda_i$. Le terme $-dN_i$ est le saut à 0 si l'émetteur i fait défaut. Le terme de somme représente l'influence sur l'intensité de τ_i du comportement des autres émetteurs : il prend en compte l'information contenue dans les défauts potentiels et la survie de toutes les entités de référence du panier. Cette information est résumée dans la matrice $\Delta = (\Delta_{ij})_{ij}$. Cette matrice Δ dépend pour une grande part des caractéristiques de la copule.

PREUVE. Nous supposons que λ_i est une diffusion. Au vu du lemme précédent, h_i saute à chaque défaut et tombe à 0 en cas de défaut de l'émetteur i . Soient $\tau_{(k)} < \tau_{(k+1)}$ deux instants de défauts consécutifs tels que $\tau_i > \tau_{(k)}$. Le saut de l'intensité h_i en $t = \tau_{(k+1)}$ se calcule explicitement :

$$\frac{\Delta h_i(t)}{h_i(t-)} = \begin{cases} -1, & \text{si } \tau_{(k+1)} = \tau_i, \\ \frac{C_{d+i+j}}{C_{d+j}} - \frac{C_{d+i}}{C_d}, & \text{si } \tau_{(k+1)} = \tau_j, i \neq j. \end{cases}$$

Pour obtenir la dynamique du terme de diffusion de h_i sur l'intervalle $]\tau_{(k)}, \tau_{(k+1)}[$, on applique la formule d'Itô à

$$h_i = \lambda_i f(\gamma), \text{ où } f(\gamma) = \gamma_i \frac{C_{d+i}(\gamma)}{C_d(\gamma)},$$

pour obtenir

$$\begin{aligned} dh_i &= \lambda_i df(\gamma) + f(\gamma) d\lambda_i \\ &= \lambda_i f_i d\gamma_i + \lambda_i \sum_{j \in d: j \neq i} f_j(\gamma) d\gamma_j + h_i \frac{d\lambda_i}{\lambda_i} \\ &= h_i \left(\frac{d\lambda_i}{\lambda_i} + \lambda_i dt + h_i \left(\frac{C_{d+i+i}C_d}{C_{d+i}C_{d+i}} - 1 \right) dt + \sum_{j \in d: j \neq i} \left(\frac{C_{d+i+j}C_d}{C_{d+i}C_{d+j}} - 1 \right) h_j dt \right) dt. \end{aligned}$$

Pour trouver la dynamique de h_i sur l'intervalle $]\tau_{(k)}, \tau_{(k+1)}[$, il suffit d'intégrer le saut en $\tau_{(k+1)}$; il vient alors

$$\frac{dh_i}{h_i} = \frac{d\lambda_i}{\lambda_i} - (h_i \Delta_{ii} + \lambda_i) dt - dN_i + \sum_{j \in d: j \neq i} \Delta_{ij} (dN_j - h_j) dt,$$

et le résultat s'en déduit immédiatement. ■

Calibration sur sauts de spread

Nous allons nous servir de ce résultat pour expliquer comment calibrer la structure de corrélation du modèle sur un a priori de sauts du spread court-terme.

Supposons qu'aucun défaut n'a eu lieu juste avant t et qu'en t l'émetteur j fasse défaut. Si l'on note $h_i^{-j}(t)$ la valeur juste après cet événement de crédit, on déduit du résultat précédent que

$$\frac{h_i^{-j}(t)}{h_i(t)} = 1 + \Delta_{ij}(t, \emptyset).$$

et dans le cadre du modèle de recouvrement MVR la même relation est vrai pour le spread court-terme (et est approximativement vrai pour le modèle de recouvrement dit RFV recovery of face value). Par conséquent, étant donné un a priori du trader sur la valeur du saut de spread de l'émetteur i en cas de défaut de l'émetteur j , il est possible dans les bons cas, de choisir les paramètres de la fonction copule redonnant cet a priori.

Commençons par examiner la nature de ce saut de spread lorsque la copule utilisée est archimédienne. Autrement dit, on suppose qu'il existe une fonction ϕ , appelée générateur (qui est, par exemple, la transformée de Laplace inversible d'une v.a. $Y > 0$: $\phi(s) = \mathbb{E}[e^{-sY}]$, $\psi = \phi^{-1}$) telle que

$$C(u) = \phi\left(\sum_{i=1}^N \psi(u_i)\right).$$

Par exemple, la copule de Gumbel a pour générateur la fonction $\phi(s) = (-\ln(x))^\theta$ et correspond à une v.a. de loi α -stable de paramètre $\alpha = 1/\theta$. La copule de Clayton a, quant à elle, pour générateur la fonction $\phi(s) = (1+s)^{-1/\theta}$ et correspond à une v.a. Y de loi gamma de paramètre $1/\theta$.

Dans le cadre des copules archimédiennes, on peut montrer (exercice!) que, si tous les instants de défaut sont postérieurs à t ,

$$\begin{aligned} h_i(t) &= \lambda_i \gamma_i \frac{\psi'(\gamma_i)}{C(\gamma) \psi'(C(\gamma))} \\ \frac{h_i^{-j}(t)}{h_i(t)} &= -\frac{C(\gamma_t) \psi''(C(\gamma_t))}{\psi'(\gamma_t)}. \end{aligned}$$

Et ces formules, nous conduisent à (exercice!)

$$\begin{aligned} h_i &= \lambda_i \left(\frac{\Lambda_i}{\|\Lambda\|_\theta} \right)^{\theta-1}, \\ \frac{h_i^{-j}(t)}{h_i(t)} &= 1 + \frac{\theta - 1}{\|\Lambda\|_\theta} \end{aligned}$$

dans le cas de la copule de Gumbel et à

$$h_i = \lambda_i \left(\frac{C(\gamma)}{\gamma_i} \right)^\alpha,$$

$$\frac{h_i^{-j}(t)}{h_i(t)} = 1 + \alpha$$

dans le cas de la copule de Clayton.

Il y a ici deux remarques importantes à faire. La première c'est que les copules archimédiennes n'autorisent pas de structure asymétriques de dépendance : lorsqu'une entité de référence fait défaut l'intensité de défaut de tout les autres noms du panier est multipliée par le même facteur. Dans le cas de la copule de Gumbel, ce facteur est indépendant du moment ou le saut à lieu et cette propriété peut s'avérer intéressante comme nous le verrons dans la suite.

Il apparaît nécessaire de généraliser le concept de copules archimédienne pour pouvoir autoriser des changements asymétriques dans le spread court-terme en cas de défaut d'une firme. C'est pourquoi, suivant [28], on peut utiliser des copules archimédiennes généralisées (cf. Appendice D).

EXERCICE. (Évaluation de CDO par Monte-Carlo.)

On considère un panier de crédit contenant N noms. On note τ_i l'instant de défaut du i ème nom ; N_i son nominal ; R_i son taux de recouvrement et λ_i son intensité risque-neutre supposée constante. Le taux sans risque est constant égal à r . On suppose que la structure de dépendance de la loi jointe des τ_i est caractérisée par la copule gaussienne associée à une matrice de corrélation ρ . On considère un CDO sur ce panier de points d'attachement $0 \leq A < B \leq 1$.

(1) Exprimer sous forme d'espérance la valeur présente des deux jambes du CDO en supposant que la jambe de prime prend en compte tous les flux et que la jambe de protection est payée à la fin de la période au cours de laquelle le défaut à lieu.

(2) Proposer un algorithme de simulation par Monte-Carlo pour calculer la valeur des deux jambes du CDO.

EXERCICE. (Détermination de la loi du premier et du second instant de défaut).

On considère un panier de crédit contenant N noms. On note τ_i l'instant de défaut du i ème nom ; N_i son nominal ; $R_i = R$ son taux de recouvrement supposé constant. Le taux sans risque est constant égal à r .

On suppose de plus qu'il existe une v.a. X (facteur de risque) telle que conditionnellement à X les τ_i forment une famille de v.a. indépendantes et l'on pose

$$p_i(t, X) = \mathbb{P}[\tau_i \leq t | X], \quad q_i(t, X) = \mathbb{P}[\tau_i > t | X].$$

On désigne par N_t la v.a. $\sum_{i=1}^N I_{\{\tau_i \leq t\}}$ et on considère sa fonction génératrice des moments

$$\psi_{N_t}(u) = \mathbb{E}[e^{Nu}] = \sum_{l=0}^N \mathbb{P}[N_t = l] u^l.$$

(1) Montrer que

$$\psi_{N_t}(u) = \mathbb{E} \left[\prod_{i=1}^N (q_i(t, X) + p_i(t, X)u) \right].$$

(2) En déduire une expression intégrale pour les probabilités $\mathbb{P}[N_t = 0]$ et $\mathbb{P}[N_t = 1]$.

(3) On suppose désormais que les τ_i sont des variables exponentielles d'intensités respectives λ_i et que leur structure de dépendance est décrite par la copule associée à un modèle gaussien à 1 facteur à corrélation constante. Autrement dit, on suppose que

$$\tau_i \leq t \iff X_i = \rho X + \sqrt{1 - \rho^2} Y_i \leq B_i(t),$$

où les v.a. X, Y_1, \dots, Y_N sont des gaussiennes standard indépendantes. En utilisant les résultats précédents, calculer des formules semi-explicites pour les jambes d'un First-2-Default (jambe de prime de tenant uniquement compte des flux intervenant avant le défaut, et jambe de protection payée à la fin de la période en cours).

Modèle Hybride : le cas des obligations convertibles

Ce chapitre constitue une application des chapitres précédents à partir de l'exemple des obligations convertibles. Ces produits prennent une importance de plus en plus considérable et certains *hedge funds* fondent même toute leur stratégie d'investissement sur cet unique actif. Après une brève description des caractéristiques de ces titres, et de leurs sources de risque, nous présentons la structure des modèles du risque de crédit adaptés à ces produits, ainsi que les techniques numériques intervenant dans l'implémentation de tels modèles. La modélisation des obligations convertibles permet de répliquer artificiellement le produit, dans une optique de couverture des risques.

6.1 Caractéristiques des obligations convertibles

6.1.1 Définition

Une obligation convertible est un titre de créance similaire à une obligation traditionnelle à ceci près qu'elle peut être échangée contre un nombre prédéterminé d'autres titres (habituellement l'action de l'entreprise émettrice). Le détenteur d'un tel produit reçoit un coupon payé périodiquement jusqu'à la conversion de l'obligation ou sa maturité. Si l'obligation n'est pas convertie avant la maturité, l'investisseur reçoit le nominal, comme dans le cas d'une obligation traditionnelle. Après conversion, l'investisseur possède des actions : il reçoit des dividendes mais plus de coupon ni, de remboursement du nominal. Le prix de conversion étant prédéterminé, tout se passe donc comme si l'émetteur avait vendu un call sur l'action sous-jacente. Comme ce call possède une valeur pour l'investisseur, celui-ci accepte de recevoir un coupon moins important que dans le cas d'une obligation standard. Les obligations convertibles sont souvent considérées comme de la dette subordonnée, en ce sens que si l'entreprise émettrice fait défaut avant la maturité, les obligations traditionnelles ont priorité sur les obligations convertibles, qui elles mêmes ont priorité sur les actions. Les obligations convertibles constituent donc des produits hybrides complexes, à la limite des

produits de taux, d'actions et de crédit.

6.1.2 Utilisation des obligations convertibles

Les entreprises émettent des obligations convertibles car celles-ci permettent de lever des fonds importants de façon économique, la présence de l'option de conversion persuadant les investisseurs d'accepter un taux de coupon moins élevé que ceux en vigueur. Les obligations convertibles sont aussi utilisées au cours d'opérations financières comme par exemple les Leverage Buy-Out (une prise de contrôle particulière d'une société qui consiste à en financer l'achat par un endettement, pour exploiter l'effet de levier), comme motivation pour les nouveaux dirigeants : si l'entreprise fait du profit et que le prix de ses actions augmente, le prix des obligations convertibles augmentera aussi.

6.1.3 Obligations convertibles et options

Les obligations convertibles sont des produits dérivés complexes, dont le prix dépend à la fois de la valeur de l'actif et du passif de l'entreprise. Pour simplifier l'étude de leur comportement, on les décompose parfois en produits financiers plus simples. Il existe deux représentations de base, reliées par la relation de parité call-put :

- OC = Obligation + Call d'échange de l'obligation contre l'action ;
- OC = Action + Put d'échange de l'action contre l'obligation + Swap d'échange des dividendes contre les coupons jusqu'à maturité.

La première représentation peut être interprétée dans le cadre des marchés de taux d'intérêts, l'émetteur vendant une obligation, et un call pour réduire le coût de la dette. La seconde représentation est plus cohérente dans le cadre des marchés d'actions, l'investisseur possédant une action, un put de protection, et un swap pour assurer un paiement de coupons constant au lieu de dividendes aléatoires.

La convertibilité de l'obligation a lieu jusqu'à maturité pour les produits à la vanille, donc les détenteurs d'obligations convertibles vont exercer l'option de façon à optimiser leur richesse. Cependant, pour la plupart des produits convertibles, la possibilité d'exercer l'option n'a lieu qu'après une période prédéterminée. D'autres produits fixent des dates ou des périodes de conversion. De plus, les obligations convertibles incluent parfois des clauses dites *callable* et *puttable* : l'obligation est dite callable si l'émetteur se réserve la possibilité de racheter le produit à un prix fixé au préalable. Le détenteur peut alors choisir de convertir le produit en actions, ou bien est obligé de le revendre à l'émetteur. En pratique, cette clause est utilisée pour forcer la conversion lorsque la valeur du produit converti est un peu au dessus de la valeur de l'obligation. Un produit callable possède un degré d'optionnalité moindre que pour une obligation convertible à la vanille et est donc moins cher. A contrario, une obligation convertible puttable donne le droit à l'investisseur de revendre le produit à l'acheteur à un prix prédéterminé. De tels produits sont donc plus chers.

6.1.4 Les déterminants du contrat

Le prix des obligations convertibles dépend de nombreuses caractéristiques définissant le comportement du contrat. Nous donnons ici les principales :

1. le *nominal* est la valeur faciale de l'obligation ;
2. le *coupon* est le taux d'intérêt annuel versé par le produit ;
3. la *fréquence des coupons* représente le nombre de coupons payés par an ;
4. le *ratio de conversion* est le nombre d'actions sous-jacentes contre lequel l'obligation convertible peut être échangée ;
5. le *prix de conversion* représente le prix payé pour chaque action à la conversion, c'est à dire le quotient du nominal par le ratio de conversion ;
6. la *parité* représente le ratio de conversion multiplié par le prix des actions ;
7. la *première date de conversion* est la première date à laquelle le détenteur est autorisé à convertir s'il existe une période de prescription ;
8. l'existence d'un *call émetteur* (caractère callable) ;
9. l'existence d'un *put investisseur* (caractère puttable).

Tous ces paramètres sont généralement définis dans le contrat et sont pris en compte pour établir le prix initial de l'obligation convertible. Cependant, pendant la vie de l'obligation, ce prix va être influencé par les variables de marché suivantes :

1. le prix des actions sous-jacentes ;
2. la volatilité des actions ;
3. le taux de dividende ;
4. le taux d'emprunt de l'action (taux de REPO) ;
5. le taux sans risque ;
6. le spread de crédit de l'émetteur.

Les quatre premiers facteurs sont tous reliés au prix de l'action sous-jacente, donc pour simplifier, le prix des obligations convertibles est influencé par trois sources de risque : le prix de l'action, le taux sans risque, et le spread de crédit de l'émetteur.

6.1.5 Influence des variables de marché sur le prix

Comme les paramètres intrinsèques n'affectent plus le prix une fois que le contrat est établi, les variations du prix durant la vie de l'obligation convertible résultent des variations d'un ou plusieurs paramètres de marché. Il est possible d'étudier historiquement la dépendance des variations du prix par rapport à celles de variables de marché. Les tendances générales sont résumées dans le tableau suivant :

Param. intrinsèques	Var. Prix	Param. de marché	Var. Prix
Maturité	↘	Prix de l'action	↗
Coupon	↗	Taux sans risque	↘
Nominal	↗	Spread de l'émetteur	↘
Prix de conversion	↘	Volatilité de l'action	↗
Existence d'un call	↘	Dividendes	↘
Existence d'un put	↗	Taux de REPO	↘

Eu égard à la dépendance complexe du prix des obligations convertibles aux trois sources de risque (taux, action, et crédit), l'évaluation de tels produits nécessite une modélisation du risque de crédit mettant en jeu les techniques introduites aux chapitres précédents.

6.2 Modélisation du risque de crédit

Dans cette section, nous considérons un espace probabilisé filtré $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{G} = (\mathcal{G}_t)_{t \geq 0})$ et nous nous plaçons dans le cadre doublement stochastique (voir chapitre 4) dans lequel l'instant de défaut τ possède une \mathbb{G} -intensité λ , \mathbb{F} -prévisible où \mathbb{F} est une sous-filtration de la filtration \mathbb{G} générée par le processus de diffusion de l'action. Pour simplifier le problème, nous allons supposer que la variation des taux d'intérêts est déterministe (il est en effet d'usage de considérer que les variation de taux agissent au second ordre face à l'influence du risque de défaut). Nous nous intéressons donc uniquement à l'influence du prix de l'action et du risque de crédit sur le prix de l'obligation convertible. Ces deux sources de risque ne peuvent être considérées comme indépendantes : supposons en effet que la valeur du spot chute, ceci peut présager de l'imminence d'un possible défaut, et en conséquence, le spread de crédit aura tendance à augmenter. De façon empirique, on remarque que l'évolution du spread de crédit à cinq ans en fonction de la valeur du spot possède la forme représentée à la figure 6.1.

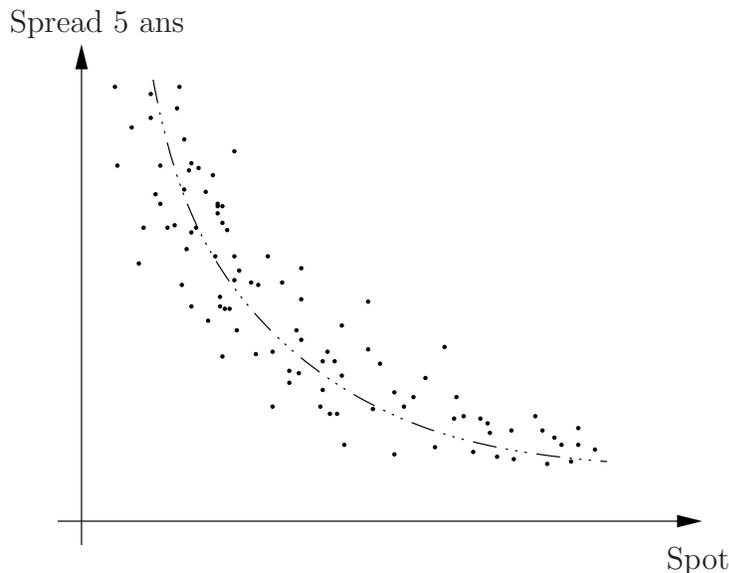


FIG. 6.1 – Évolution du spread à 5 ans en fonction du spot.

Nous allons donc utiliser un modèle hybride, c'est à dire que l'intensité de défaut λ est fonction convexe de S :

$$\lambda_t = \lambda(t, S_t),$$

prenant une forme similaire à celle représentée sur la figure 6.2.

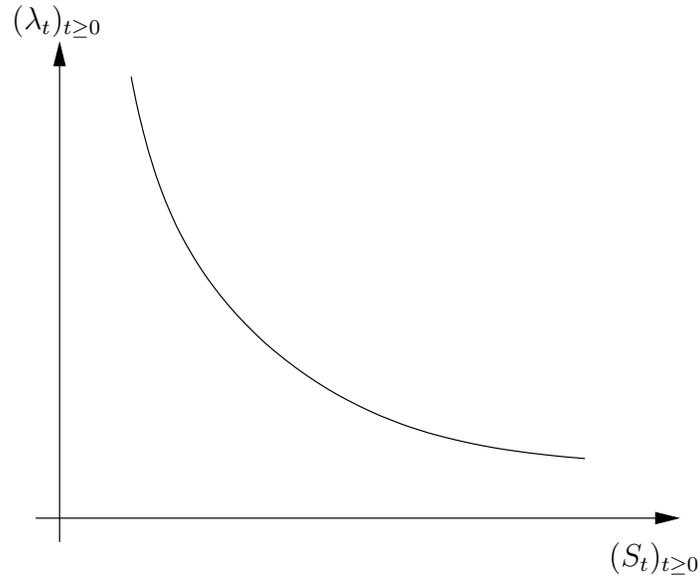


FIG. 6.2 – Relation intensité-spot.

On trouve dans la littérature plusieurs formes paramétriques différentes pour λ dont

$$\begin{aligned}\lambda(t, S_t) &= a + \frac{b}{S_t} \\ \lambda(t, S_t) &= c - d \ln(S_t) \\ \lambda(t, S_t) &= e + f e^{-g S_t}.\end{aligned}$$

La première paramétrisation se retrouve dans le modèle de Takahashi et al (2001) notamment, la deuxième par Bloch et Miralles (2002), et la troisième par Arvanitis et Gregory (2001). Ces formes paramétriques permettent de capturer l'une des caractéristiques du prix des obligations convertibles : quand le prix de l'action S est très élevé par rapport au prix de l'obligation elle-même, l'obligation convertible se comporte comme un call sur S . En revanche, lorsque la valeur de l'action S est très faible, l'intensité λ augmente et avec elle la probabilité de défaut, donc le prix de l'obligation chute. (voir [11]). Le graphe du prix de l'obligation convertible, en fonction du prix de l'action possède donc la forme décrite sur la figure 6.3.

6.3 Implémentation des modèles

Considérons une obligation convertible de nominal N et de maturité T écrite sur un sous-jacent S ne versant pas de dividende, de taux de recouvrement δ , et de ratio

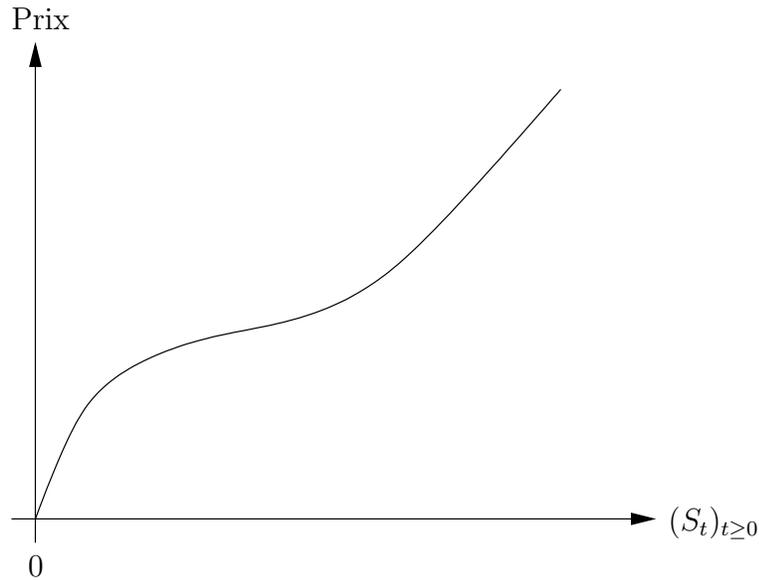


FIG. 6.3 – Prix de l’obligation convertible en fonction du prix de l’action.

de conversion κ . L’obligation est supposée verser continûment un taux de coupon constant c . Nous supposons que la valeur de l’action S suit l’équation de diffusion avec saut suivante :

$$\begin{aligned} dS_t &= r_t S_t dt + \sigma S_t dW_t - S_{t-} (dN_t - \lambda_t dt) \quad \text{sur } \{t < \tau\} \\ &= (r_t + \lambda_t) S_t dt + \sigma S_t dW_t - S_{t-} dN_t \quad \text{sur } \{t < \tau\}, \end{aligned}$$

où W est un mouvement Brownien standard, et $N_t = \mathbb{1}_{\{\tau \leq t\}}$ un processus de Cox d’intensité λ . Comme précédemment, nous faisons l’hypothèse que l’intensité de défaut λ est une fonction déterministe de t et de S .

6.3.1 Les arbres multinomiaux

L’idée à la base des modèles de résolution par arbre est la discrétisation en temps de la dynamique de l’action. On divise la période considérée en petits intervalles de temps de longueur Δt , et de calculer le prix à l’instant initial par rétropropagation dans l’arbre de la matrice des valeurs finales.

On fait l’hypothèse que le défaut ne peut avoir lieu qu’à un noeud de l’arbre (voir figure 6.4).

En un noeud donné de l’arbre, la firme peut faire défaut avec une probabilité $1 - e^{-\lambda_t \Delta t}$ – auquel cas la valeur de l’obligation convertible tombe à δ – ou survivre avec une probabilité $e^{-\lambda_t \Delta t}$. Comme dans un arbre binomial classique, le rendement de l’action peut être positif avec une probabilité $e^{-\lambda_t \Delta t} p$ pendant le petit intervalle de temps Δt et prendre une valeur μ_+ ou être négatif avec une probabilité $e^{-\lambda_t \Delta t} (1 - p)$ de valeur μ_- . La valeur de l’obligation convertible au temps $t - 1$ est égale au maximum de la valeur convertie, i.e. κS , et de la valeur V de l’obligation calculée par

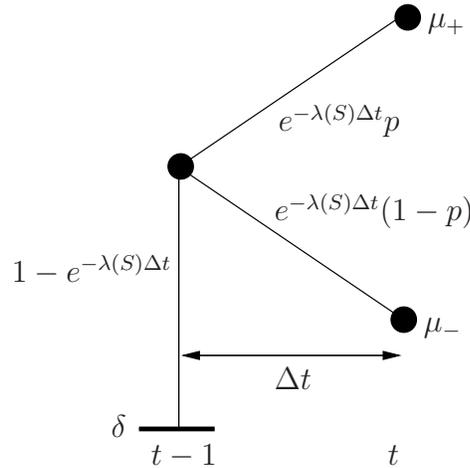


FIG. 6.4 – Rétropropagation et risque de défaut.

rétropropagation, soit

$$V(S, t-1) = \max \left(\kappa S, e^{-(r_t-c)\Delta t} \left(e^{\lambda(S,t)\Delta t} (pV(S\mu_+, t) + (1-p)V(S\mu_-, t)) + (1 - e^{-\lambda(S,t)\Delta t})\delta \right) \right).$$

En chaque noeud de l'arbre, la valeur de S est déterminée par la valeur au pas précédent et le rendement. On en déduit la valeur de λ grâce à la formule paramétrique, et celle du prix de l'obligation convertible via l'expression ci-dessus.

Reste à initialiser la matrice des valeurs du prix à maturité. Pour chaque valeur de S_T , le prix final est égal au maximum de l'ensemble des flux obligataires et de la valeur convertie soit

$$V(S, T) = \max(\kappa S, N e^{cT}).$$

Cet algorithme peut être étendu facilement en ajoutant des clauses callable ou puttable, des périodes de non-convertibilité, des dividendes, etc...

6.3.2 Les équations aux dérivées partielles

Sachant que le prix V de l'obligation convertible est fonction de S et de t , Nous allons utiliser la formule d'Itô avec sauts pour trouver l'équation aux dérivées partielles qu'il vérifie. Considérons un portefeuille π composé d'une obligation convertible, et d'une quantité α de sous-jacent S :

$$\pi = V + \alpha S,$$

par application de la formule d'Itô :

$$d\pi(t, S_t) = \frac{\partial V}{\partial t} dt + \left(\frac{\partial V}{\partial S} + \alpha \right) dS + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} d\langle S, S \rangle + \Delta V + \alpha \Delta S.$$

On peut donc supprimer le risque lié à la variation Brownienne en choisissant $\alpha = -\frac{\partial V}{\partial S}$ et donc

$$\begin{aligned} d\pi(t, S_t) &= \frac{\partial V}{\partial t} dt + \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial S^2} dt + (\delta - 1)V dN_t - \frac{\partial V}{\partial S} dN_t. \\ &= \frac{\partial V}{\partial t} dt + \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial S^2} dt + \left((\delta - 1)V - \frac{\partial V}{\partial S} \right) (dN_t - \lambda_t dt) \\ &\quad + \left((\delta - 1)V - \frac{\partial V}{\partial S} \right) \lambda_t dt. \end{aligned}$$

Les sauts étant compensés, le portefeuille est sans risque, donc par absence d'opportunité d'arbitrage, il rapporte le taux sans risque r :

$$\mathbb{E} \left[\frac{d\pi}{\pi} \middle| \mathcal{F}_t \right] = (r_t - c) dt,$$

donc finalement

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial S^2} - \lambda_t \frac{\partial V}{\partial S} + r_t \frac{\partial V}{\partial S} S = ((r_t - c) + (1 - \delta)\lambda_t)V,$$

les conditions à la limite étant identiques à celles utilisées dans le cas des arbres. Ce système peut ensuite être résolu par des méthodes numériques de différences ou éléments finis.

6.3.3 Synthèse des obligations convertibles

Outre la valorisation des produits, la modélisation des obligations convertibles revêt un intérêt tout particulier en matière de couverture des risques. En effet, l'identification des sources de risque va permettre de répliquer le produit par une combinaison d'autres actifs.

Les obligations convertibles sont souvent considérées en première approximation comme un produit dérivé sur l'action sous-jacente. Nous avons vu en effet qu'elle pouvait être globalement décomposée comme une obligation classique ainsi qu'une option d'échange. Cette vision néglige cependant les risques de crédit et de taux liés au produit. En effet, comme le montre la figure 6.3, le prix de l'obligation convertible ne se comporte pas de la même façon que celui d'un dérivé action. Une couverture en *delta* classique par achat du sous-jacent n'immuniserait pas le portefeuille aux fortes variations de S à la baisse (risque de défaut) et aux changements de volatilité. En pratique, les traders répliquent l'effet de volatilité en couvrant en *vega*, c'est à dire en prenant une position opposée sur des options ayant la même sensibilité à la volatilité. Enfin, le risque de crédit peut être répliqué par des CDS digitaux, à savoir un contrat payant 1 au temps t si $\tau \leq t$. Ces produits ne sont malheureusement pas du tout liquides sur le marché, et sont eux même répliqués par des CDS classiques dans la pratique. Le risque de taux étant, pour sa part, couvert par des swaps.

L'un des principaux problèmes liés à cette approche pour la gestion des risques liée aux obligations convertibles est la grande maturité de ce type de contrat : selon Jim Vinci, directeur des risques du fond d'arbitrage d'obligations convertibles *Paloma*

Partners. “One of the final pieces would be if dealers start offering more longer-dated equity options – the replication possibility in convertibles would become even better”. En effet, à l’heure actuelle, la liquidité des options sur action de longue maturité est insuffisante.

Bibliographie

- [1] AVEN (T.). A theorem for determining the compensator of a counting process. *Scandinavian Journal of Statistics*, 12(1), pp. 69-72 (1985).
- [2] BIELECKI (T. R.), RUTKOWSKI (M.). Credit Risk : Modeling, Valuation and Hedging. Springer (2002).
- [3] BOMFIM (A. N.). Credit Derivatives and Their Potential to Synthesize Riskless Assets. *The Journal of Fixed Income* December 2002, Vol. 12, No. 3, pp. 6-16 (2002).
- [4] CROSBIE (P. J.). Modeling Default Risk. KMV Corporation, San Francisco, www.kmv.com (1997).
- [5] DELLACHERIE (C.) MEYER (P.A.). *Probabilités et potentiels, chapitres I à IV*, Hermann, Paris (1975).
- [6] DUFFIE (D.), FILIPOVIĆ (D.) & SCHACHERMAYER (W.). Affine Processes and Applications in Finance. *Annals of Applied Probability*, Vol.13, No 3, August 2003 (2003).
- [7] DUFFIE (D.). A Short Course on Credit Risk Modeling with Affine Processes. Working Paper. Stanford University and Scuola Normale Superior (2002). Disponible sur http://fibonacci.dm.unipi.it/finance/duffie_notes.pdf.
- [8] DUFFIE (D.), LANDO (D.). The Term Structure of Credit Spreads with Incomplete Accounting Information. *Econometrica* 69, pp. 633-664 (2001).
- [9] DUFFIE (D.), PAN (J.), SINGLETON (K.). Transform Analysis and Asset Pricing for Affine Jump Diffusions. *Econometrica*, Vol. 68, pp. 1343-1376 (2000).
- [10] DUFFIE (D.), SINGLETON (K. J.). Modeling Term Structure of Defaultable Bonds. *The Review of Financial Studies* Special 1999 Vol. 12, No. 4, pp. 687-720 (1999).
- [11] DUFFIE (D.), SINGLETON (K. J.). *Credit Risk*. Princeton University Press (2003).

- [12] ÉMERY (M.). Classical Probability Theory : An Outline of Stochastic Integrals and Diffusions. Quantum Probability Communications, QP-PQ, Vol. XI, World Scientific (1999). Disponible sur <http://www-irma.u-strasbg.fr>
- [13] EL KAROUI (N.) & MAZLIAK (L.). Backward Stochastic Differential Equations. *Pitman Research Notes in Mathematic Series* (1997).
- [14] GESKE (R.). The valuation of compound options. *J. Finan. Econom.* 7, pp. 63-81 (1979).
- [15] HULL (J. C.), NELKEN (I.), WHITE (A.). Merton's Model, Credit Risk, and Volatility Skews. Working Paper. University of Toronto (2003).
- [16] ISDA 2003 Credit Derivatives Definitions. www.isda.org.
- [17] KRYLOV (N. V.). *Introduction to the theory of diffusion processes*. American Mathematical Society, Providence.
- [18] LAURENT (J.-P.). I will survive. Risk, June 2003. www.risk.net (2003).
- [19] LELAND (H. E.). Corporate debt value, bond covenants, and optimal capital structure. *J. Finance* 49, pp. 1213-1252.
- [20] LONGSTAFF (F. A.), SCHWARTZ (E. S.). A simple approach to valuing risky fixed and floating rate debt. *J. Finance* 50, pp. 789-819 (1995).
- [21] MERTON (R. C.). On the pricing of corporate debt :The risk structure of interest rates. *J. Finance* 29, pp. 449-470 (1974).
- [22] PAN (G.). Equity to credit pricing. Risk, November 2001. www.risk.net (2001).
- [23] PROTTER (P.). *Stochastic Integration and Differentiel Equations*. Springer (1990).
- [24] PYE (G.). Gauging the Default Premium. *Financial Analyst's Journal*, January-February, pp.49-50 (1974).
- [25] Credit Derivatives Survey. Risk, February 2003. www.risk.net (2003).
- [26] RONCALLI (T.). *La gestion des risques financiers*. Cours du DESS de Gestion des risques et des actifs de l'université d'Evry. gro.creditlyonnais.fr.
- [27] RUDIN (W.). *Real and Complex Analysis*. McGraw-Hill, New York, third edition (1987).
- [28] SCHONBÜCHER (P.), ROGGE (E.). Modeling Dynamic Portfolio Credit Risk. Working Paper. ETH Zürich (2002). Disponible sur www.schonbucher.de.
- [29] SCHONBÜCHER (P.), SCHUBERT (D.). Copula-Dependent Default Risk in Intensity Models. Working Paper. University of Bonn (2000). Disponible sur www.schonbucher.de.
- [30] VASICEK (O.). Credit Valuation. Working Paper. KMV Corporation. www.kmv.com (1984).

Théorie générale des processus

Cette annexe se veut une présentation très succincte de la théorie générale des processus et du calcul stochastique. Pour plus de détails et les preuves, le lecteur pourra consulter [5]. Dans toute la suite $(\Omega, \mathcal{F}, \mathbb{P})$ désigne un espace probabilisé complet.

A.1 Généralités

L'objet fondamental de la théorie s'appelle *filtration* et consiste en la donnée d'une famille croissante $(\mathcal{F}_t)_{t \geq 0}$ de sous-tribu de \mathcal{F} :

$$\mathcal{F}_s \subset \mathcal{F}_t, \quad \forall s \leq t.$$

Nous supposons que les hypothèses suivantes sont vérifiées :

- (1) La tribu \mathcal{F}_0 contient les ensembles \mathbb{P} -négligeables,
- (2) $\mathcal{F}_t = \bigcap_{s > t} \mathcal{F}_s$ (continuité à droite).

Les conditions (1) et (2) s'appellent les *conditions usuelles* de la théorie générale des processus.

Un *processus* sur cet espace est une famille $(X_t)_{t \geq 0}$ de v.a. à valeurs réelles. Si l'application $(t, \omega) \mapsto X_t(\omega)$ est mesurable, on dit que le processus $(X_t)_{t \geq 0}$ est *mesurable*. Si pour tout $t \geq 0$, la v.a. X_t est \mathcal{F}_t -mesurable, on dit que le processus est *adapté*. Si pour \mathbb{P} -presque tout ω la trajectoire $t \mapsto X_t(\omega)$ est continue (respectivement continue à droite, continue à gauche, pourvue de limite à droite,...) on dit que le processus $(X_t)_{t \geq 0}$ est continu (respectivement càd, càg, làd, ...). Un processus continu à droite et pourvue de limites à gauches est dit *càd-làg*. Un tel processus est mesurable.

Soient $X = (X_t)_{t \geq 0}$ et $Y = (Y_t)_{t \geq 0}$ deux processus. On dit que X et Y sont des *versions* l'un de l'autre ou que Y est une *modification* de X si $X_t = Y_t$ *p.s.* pour tout $t \geq 0$. On dit que X et Y sont *indistinguables* si $X_t = Y_t$ *p.s.* Une partie A de $\mathbb{R}_+ \times \Omega$ est dite *évanescence* si sa projection sur Ω est une partie \mathbb{P} -négligeable. Un processus X est dit *évanescent* s'il est indistinguishable du processus nul.

A.2 Temps d'arrêt et tribus associées

Une v.a. positive T est appelée un *temps d'arrêt* de la filtration $(\mathcal{F}_t)_{t \geq 0}$ si

$$\{T \leq t\} \in \mathcal{F}_t$$

pour tout $t \geq 0$. La filtration étant supposé continue à droite, nous pouvons remplacer l'ensemble $\{T \leq t\}$ par $\{T < t\}$ dans la définition précédente.

On vérifie que si S et T sont des temps d'arrêt, il en est de même des v.a. $S \wedge T$ et $S \vee T$. Si (S_n) est une suite de temps d'arrêt telle que $S_n \uparrow S$ (respectivement $S_n \downarrow S$), alors S est un temps d'arrêt.

Soit T un temps d'arrêt. La tribu \mathcal{F}_T des *événements antérieurs* à T est l'ensemble des $A \in \mathcal{F}_\infty$ tels que la v.a. $T_A = T \mathbb{1}_A + \infty \mathbb{1}_{A^c}$ soit un temps d'arrêt. On vérifie que

$$\mathcal{F}_T = \{A \in \mathcal{F}_\infty ; A \cap \{T \leq t\} \in \mathcal{F}_t, \forall t \geq 0\} .$$

La tribu \mathcal{F}_{T-} des *événements strictement antérieurs* à T est la tribu engendrée par \mathcal{F}_0 et les événements de la forme

$$A \cap \{t < T\} \quad , A \in \mathcal{F}_t \quad (t \geq 0) .$$

Les résultats de base sur ces différentes notions sont regroupés dans le théorème suivant.

THÉORÈME A.1 *Soient S et T deux temps d'arrêt.*

- (1) $\mathcal{F}_{S-} \subset \mathcal{F}_S$ et S est \mathcal{F}_{S-} -mesurable,
- (2) Pour tout $A \in \mathcal{F}_S$, $A \cap \{S \leq T\} \in \mathcal{F}_T$ et $A \cap \{S < T\} \in \mathcal{F}_{T-}$. En particulier, $\{S \leq T\}$ et $\{S = T\}$ appartiennent à $\mathcal{F}_S \cap \mathcal{F}_T$ et $\{S < T\}$ appartient à $\mathcal{F}_S \cap \mathcal{F}_{T-}$,
- (3) Pour tout $A \in \mathcal{F}_S$, $A \cap \{S = +\infty\} \in \mathcal{F}_{S-}$,
- (4) Si $S \leq T$, alors $\mathcal{F}_S \subset \mathcal{F}_T$ et $\mathcal{F}_{S-} \subset \mathcal{F}_{T-}$. Si de plus, $S < T$ partout sur $\{0 < T < +\infty\}$, alors $\mathcal{F}_S \subset \mathcal{F}_{T-}$,
- (5) Soit $(S_n)_{n \in \mathbb{N}}$ une suite monotone de temps d'arrêt. Si $S_n \uparrow S$, alors $\mathcal{F}_{S-} = \bigvee_{n \in \mathbb{N}} \mathcal{F}_{S_n-}$. Si de plus, $S_n < S$ sur $\{0 < S < +\infty\}$, alors $\mathcal{F}_{S-} = \bigvee_{n \in \mathbb{N}} \mathcal{F}_{S_n}$. Si $S_n \downarrow S$, alors $\mathcal{F}_S = \bigcap_{n \in \mathbb{N}} \mathcal{F}_{S_n}$.

Nous présentons maintenant l'exemple fondamental de temps d'arrêt. On dit qu'un processus $(X_t)_{t \geq 0}$ est *progressivement mesurable* ou *progressif* si l'application

$$[0, t] \times \Omega \rightarrow \mathbb{R} \quad , (s, \omega) \mapsto X_s(\omega)$$

est $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$ -mesurable. En particulier, un processus progressivement mesurable est adapté. Les processus adaptés et continus (respectivement càd, càg,...) sont progressivement mesurables. On note $\mathcal{P}rog$ la tribu sur $\mathbb{R}_+ \times \Omega$ engendrée par les processus progressifs. Les éléments de cette tribu sont appelés *ensembles progressifs*. Soit A une partie de $\mathbb{R}_+ \times \Omega$. On appelle *début de A* la fonction positive définie sur Ω par

$$D_A(\omega) = \inf \{t \geq 0 ; (t, \omega) \in A\} \quad ,$$

avec la convention $\inf \emptyset = +\infty$.

THÉORÈME A.2 *Le début D_A d'un ensemble progressif A est un temps d'arrêt de la filtration $(\mathcal{F}_t)_{t \geq 0}$.*

A.3 Les tribus optionnelles et prévisibles

Soient U, V deux applications définies sur Ω et à valeurs dans $\overline{\mathbb{R}}_+$ telles que $U \leq V$. On pose

$$\llbracket U, V \llbracket = \{(t, \omega) \in \mathbb{R}_+ \times \Omega; U(\omega) \leq t < V(\omega)\} .$$

Remarquer que $\llbracket U, V \llbracket$ est un sous-ensemble de $\mathbb{R}_+ \times \Omega$. Nous définissons de la même manière les intervalles de la forme $\llbracket U, V \rrbracket,]U, V], \dots$. En particulier, $\llbracket U \rrbracket = \llbracket U, U \rrbracket$ est le graphe de U .

DÉFINITION A.3 *La tribu optionnelle est la tribu \mathcal{O} sur $\mathbb{R}_+ \times \Omega$ engendrée par les processus adaptés et càd-làg. La tribu prévisible est la tribu \mathcal{P} sur $\mathbb{R}_+ \times \Omega$ engendrée par les processus adaptés et càg sur $(0, +\infty)$.*

EXEMPLE. Si S et T sont deux temps d'arrêt tels que $S \leq T$, les ensembles $\llbracket 0, T \rrbracket$ et $\llbracket S, +\infty \rrbracket$ sont optionnels (car adaptés et càd-làg). Donc, $\llbracket S, T \llbracket$ est optionnel. En prenant $T = S + 1/n$ et en faisant tendre $n \rightarrow +\infty$, nous prouvons que $\llbracket S \rrbracket$ est optionnel. Ainsi, tous les intervalles stochastiques déterminés par S et T sont optionnels. Plus généralement, le processus $Z \mathbb{1}_{\llbracket S, T \llbracket}$, où Z est \mathcal{F}_S -mesurable, est optionnel. Il suffit, en effet, de le vérifier pour $Z = \mathbb{1}_A (A \in \mathcal{F}_S)$ et dans ce cas $Z \mathbb{1}_{\llbracket S, T \llbracket} = \mathbb{1}_{\llbracket S_A, T_A \llbracket}$. L'intervalle $\llbracket S, T \rrbracket$ est prévisible (car adapté et càg).

Nous résumons dans les deux théorèmes suivants les propriétés fondamentales des tribus \mathcal{O} et \mathcal{P} .

THÉORÈME A.4

- (1) *Les ensembles optionnels sont progressifs,*
- (2) *Si T est un temps d'arrêt et X un processus progressif, alors $X_T \mathbb{1}_{\{T < +\infty\}}$ est une v.a. \mathcal{F}_T -mesurable. Inversement, si Y est une v.a. \mathcal{F}_T -mesurable alors il existe un processus optionnel X tel que $Y \mathbb{1}_{\{T < +\infty\}} = X_T \mathbb{1}_{\{T < +\infty\}}$,*
- (3) *La tribu optionnelle est engendrée par les intervalles stochastiques de la forme $\llbracket S, +\infty \rrbracket$, où S est un temps d'arrêt,*
- (4) *La tribu optionnelle est engendrée par les processus adaptés et continus à droite.*

THÉORÈME A.5

- (1) *Les ensembles prévisibles sont optionnels,*
- (2) *Si T est un temps d'arrêt et si X est un processus prévisible, alors la variable aléatoire $X_T \mathbb{1}_{\{T < +\infty\}}$ est \mathcal{F}_{T-} -mesurable. Inversement, si Y est une v.a. \mathcal{F}_{T-} -mesurable alors il existe un processus prévisible X tel que $Y \mathbb{1}_{\{T < +\infty\}} = X_T \mathbb{1}_{\{T < +\infty\}}$.*
- (3) *La tribu prévisible est engendrée par les ensembles de la formes*

$$\{0\} \times A \quad (A \in \mathcal{F}_0) \text{ et } (s, t] \times A \quad (0 < s < t, A \in \mathcal{F}_s).$$

A.4 Temps d'arrêt prévisibles

DÉFINITION A.6 *Un temps d'arrêt T est dit prévisible s'il existe une suite $(T_n)_{n \in \mathbb{N}}$ de temps d'arrêt telle que*

- (1) $T_n \uparrow T$,
- (2) $T_n < T$ sur $\{T > 0\}$, pour tout n .

Dans ce cas, on dit que la suite $(T_n)_{n \in \mathbb{N}}$ annonce T .

Sous les conditions habituelles une suite (T_n) annonce un temps d'arrêt T dès qu'elle l'annonce presque sûrement. En effet, si (T_n) annonce T sur A , où A^c est négligeable, alors $A \in \mathcal{F}_0$ et les temps d'arrêt

$$T'_n = T_n \mathbb{1}_{A^c} + (T - 1/n)^+ \mathbb{1}_A$$

annoncent T partout.

La signification intuitive de la définition est claire : un phénomène physique est prévisible s'il existe une suite de signes avant-coureurs annonçant son apparition.

THÉORÈME A.7 Soient S un temps d'arrêt prévisible, $A \in \mathcal{F}_{S-}$ et T un temps d'arrêt. Alors $A \cap \{S \leq T\} \in \mathcal{F}_{T-}$. En particulier, $\{S < T\}, \{S = T\}, \{S \leq T\}$ appartiennent à \mathcal{F}_{T-} .

THÉORÈME A.8

- (1) Si S, T sont des temps d'arrêt prévisibles, il en est de même des temps $S \wedge T$ et $S \vee T$,
- (2) Soit (S_n) une suite monotone de temps d'arrêt prévisibles. Si $S_n \uparrow S$, alors S est un temps d'arrêt prévisible. Si $S_n \downarrow S$ et si pour tout ω il existe $n(\omega)$ tel que $S_n(\omega) = S(\omega), \forall n \geq n(\omega)$, alors S est un temps d'arrêt prévisible,
- (3) Si S et T sont des temps d'arrêt prévisibles, il en est de même de $S_{\{S < T\}}$,
- (4) Si T est un temps d'arrêt prévisible, T_A est un temps d'arrêt prévisible si et seulement si $A \in \mathcal{F}_{T-}$.

THÉORÈME A.9 La tribu sur $\mathbb{R}_+ \times \Omega$ engendrée par les intervalles stochastiques de la forme $\llbracket S, T \llbracket$, où S et T sont des temps d'arrêt prévisibles, est la tribu prévisible \mathcal{P} .

A.5 Classification des temps d'arrêt

On se propose de classer les temps d'arrêt selon leur situation "géométrique" par rapport aux temps d'arrêt prévisibles. Plus précisément, nous introduisons la définition suivante :

DÉFINITION A.10 Un temps d'arrêt T est dit accessible s'il existe une suite $(T_n)_{n \in \mathbb{N}}$ de temps d'arrêt prévisibles telle que

$$\llbracket T \llbracket \subset \bigcup_{n \in \mathbb{N}} \llbracket T_n \llbracket, \text{ à un ensemble évanescent près,}$$

ce qui s'écrit encore $\mathbb{P}[\bigcup_n \{T = T_n < +\infty\}] = \mathbb{P}[T < +\infty]$.

Un temps d'arrêt T est dit totalement inaccessible si pour tout temps d'arrêt prévisible S , on a

$$\llbracket T \llbracket \cap \llbracket S \llbracket = \emptyset, \text{ à un ensemble évanescent près,}$$

ce qui s'écrit encore $\mathbb{P}[S = T < +\infty] = 0$.

Les temps d'arrêt prévisibles sont accessibles. Un temps d'arrêt à la fois accessible et totalement inaccessible est *p.s.* infini. Un temps d'arrêt totalement inaccessible est *p.s.* strictement positif.

Soit T un temps d'arrêt et considérons l'ensemble $\mathcal{S}[T]$ des suites croissantes de temps d'arrêt majorées par T . Pour toute suite (S_n) de $\mathcal{S}[T]$, posons

$$\mathcal{A}[(S_n)] = \{\lim_n \uparrow S_n = T; S_n < T, \forall n\} \cup \{T = 0\} .$$

Si l'on oublie $\{T = 0\}$, $\mathcal{A}[(S_n)]$ est l'ensemble sur lequel la suite (S_n) annonce T . De plus, nous avons

$$\mathcal{A}[(S_n)] = \{T = 0\} \cup \bigcap_n \{S_n < T\} \setminus \{S < T\} ,$$

ce qui prouve que $\mathcal{A}[(S_n)]$ est un élément de la tribu \mathcal{F}_{T-} . Soit $\mathcal{A}[T]$ un représentant de la réunion essentielle de tous les $\mathcal{A}[(S_n)]$ et $\mathcal{J}[T] = \mathcal{A}[T]^c$. Remarquons que $\mathcal{A}[T]$ contient *p.s.* les ensembles $\{T = 0\}$ et $\{T = +\infty\}$.

THÉORÈME A.11

- (1) *Un temps d'arrêt T est accessible si et seulement si $\mathcal{A}[T] = \Omega$ p.s.*
- (2) *Un temps d'arrêt T est totalement inaccessible si et seulement si $\mathcal{A}[T] = \{T = +\infty\}$ p.s.*
- (3) *Le temps d'arrêt $T_{\mathcal{A}[T]}$ est accessible, le temps d'arrêt $T_{\mathcal{J}[T]}$ est totalement inaccessible et cette décomposition est unique, en ce sens que si U et V sont deux temps d'arrêt tels que U soit accessible, V totalement inaccessible, $U \wedge V = T$, $U \vee V = +\infty$, alors $T_{\mathcal{A}[T]} = U$ p.s. et $T_{\mathcal{J}[T]} = V$ p.s.*

ANNEXE B

Formule d'Itô avec sauts

Le but de cet appendice est d'énoncer une version de la formule d'Itô pour des processus avec des sauts. Pour plus de détails sur ces questions le lecteur pourra consulter l'ouvrage [23] ou le survey [12]. Dans la suite $(\Omega, \mathcal{A}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})$ désigne un espace probabilisé filtré satisfaisant aux conditions habituelles.

Avant d'aller plus loin rappelons quelques définitions classiques : une fonction $F : \mathbb{R}_+ \mapsto \mathbb{R}$ est dite *continue à droite* si $\lim_{s \downarrow t} F(s) = F(t)$, *continue à gauche* si $\lim_{s \uparrow t} F(s) = F(t)$; on dit que F admet une limite gauche si la limite $F(t-) := \lim_{s \uparrow t} F(s)$ existe. Le saut ΔF de F à l'instant t est définie par $\Delta F(t) := F(t) - F(t-)$. Dans la suite, toutes les fonctions (et par extension tout les processus) considérés seront supposés *càd-làg* c'est-à-dire continus à droite et admettant des limites à gauche en tout point.

B.1 Processus à variation finie

Nous commençons ce paragraphe par quelques rappels sur les fonctions à variation finie (pour plus de détails, on peut consulter l'excellent [27]).

Fonctions à variation finie

On note Θ l'ensemble des subdivisions $\{0 = t_0 < t_1 < \dots < t_N = t\}$. Une fonction à valeurs réelles F est dite à *variation finie sur* \mathbb{R}_+ , si pour tout $t \geq 0$, sa variation totale

$$\begin{aligned} \int_{(0,t]} |df_s| &:= \sup_{\Theta} \sum_{i=1}^N |f(t_i) - f(t_{i-1})| \\ &= \lim_{n \rightarrow +\infty} \sum_{k \geq 0} |f((k+1)2^{-n} \wedge t) - f(k2^{-n} \wedge t)| \end{aligned}$$

est une quantité finie.

EXEMPLE. Une fonction croissante est une fonction à variation finie, en effet :

$$\begin{aligned} \int_{(0,t]} |df_s| &= \lim_{n \rightarrow \infty} \sum_{k \geq 0} |f((k+1)2^{-n} \wedge t) - f(k2^{-n} \wedge t)| \\ &= \lim_{n \rightarrow \infty} \sum_{k \geq 0} f((k+1)2^{-n} \wedge t) - f(k2^{-n} \wedge t) \\ &= \lim_{n \rightarrow \infty} (f(t) - f(0)) \\ &= f(t) - f(0) < \infty. \end{aligned}$$

EXEMPLE. Soit g appartenant à $L^1_{\text{loc}}(\mathbb{R}_+)$, alors $f(t) = f(0) + \int_0^t g(s) ds$ est à variation finie et $\int_{(0,t]} |df_s| = \int_0^t |g(s)| ds$. La démonstration de ce résultat dans le cas où g est continue repose sur le théorème des accroissements finis (exercice).

EXEMPLE. Soient $(t_n)_{n \geq 0} \subseteq \mathbb{R}_+^*$ et $(\Delta_n)_{n \geq 0} \subseteq \mathbb{R}$ telle que $\sum_{n \geq 0} |\Delta_n| \leq \infty$, alors

$$f(t) = \sum_{n \geq 0} \Delta_n \mathbb{1}_{\{t_n \leq t\}}$$

est à variation finie. Montrer que dans ce cas, sa variation est donnée par la fonction

$$\int_{(0,t]} |df_s| = \sum_{n \geq 0} |\Delta_n| \mathbb{1}_{\{t_n \leq t\}}.$$

La mesure de Radon correspondante est la mesure $\mu = \sum_n \Delta_n \delta_{t_n}$, où δ_t est la masse de Dirac au point t .

On dit dans ce cas que f est *purement atomique* ou *purement de saut*.

REMARQUE. On note parfois $V(f)(t) = \int_{(0,t]} |df_s|$ la variation totale de la fonction f .

PROPOSITION B.1

1. Les fonctions à variation finie forment un espace vectoriel.
2. Si f est à variation finie, alors

$$f = f(0) + f^+ + f^-$$

où f^+ et f^- sont deux fonctions croissantes qui s'écrivent de la façon suivante :

$$\begin{aligned} f^+(t) &= \frac{1}{2} \left(f(t) - f(0) + \int_{(0,t]} |df_s| \right) \\ f^-(t) &= \frac{1}{2} \left(f(t) - f(0) - \int_{(0,t]} |df_s| \right). \end{aligned}$$

PREUVE. Montrons que f^+ est croissante. Soit $t_1 \leq t_2$ appartenant au domaine de définition de f^+ , on a alors :

$$\begin{aligned} 2(f^+(t_2) - f^+(t_1)) &= f(t_2) - f(t_1) + V(f)(t_2) - V(f)(t_1) \\ &\geq f(t_2) - f(t_1) + |f(t_2) - f(t_1)| \geq 0. \end{aligned}$$

■

Identification de f avec une mesure signée

Les fonctions à variation finie sont en bijection avec les mesures signées (différence de deux mesures finies sur \mathbb{R}^+) : la donnée de la fonction à variation finie f est équivalente à celle de la mesure μ_f définie par $\mu_f(a, b] = f(b) - f(a)$. Plus précisément, on peut définir pour toute fonction g continue à support compact :

$$\int_{(0, \infty)} g(t) f(dt) := \int_{(0, \infty)} g(t) \mu_f(dt).$$

Le membre de gauche est appelé *intégrale au sens de Riemann-Stieljes*, et est définie comme la limite

$$\lim_{n \rightarrow \infty} \sum_{k \geq 0} g(k2^{-n}) (f((k+1)2^{-n}) - f(k2^{-n})).$$

REMARQUE. La mesure μ_f possède un atome (masse de Dirac) en t_0 si et seulement si $\Delta f(t_0) \neq 0$. En effet, $\mu_f(\{t_0\}) = f(t_0) - f(t_0-) = \Delta f(t_0)$.

Décomposition de f en parties continue et atomique

Soit f une fonction à variation finie, on peut écrire f sous la forme

$$f = f_0 + f^c + f^d$$

où $f^d(t) := \sum_{0 < s \leq t} \Delta f(s)$ (la somme des sauts de f), est une fonction purement discontinue, et $f^c = f - f_0 - f^d$ est continue.

EXERCICE. Montrer que $\left| \int g df \right| \leq \int |g| dV(f)$.

PREUVE. Modulo l'utilisation d'un lemme de classe monotone, il suffit de prouver le résultat pour $g = \mathbb{1}_{(a,b]}$, et dans ce cas, on a :

$$\left| \int g df \right| = \left| \int_{(a,b]} df \right| = |f(b) - f(a)| \leq \int_{(a,b]} |df|.$$

■

EXERCICE. Montrer que $\sum_{0 \leq s \leq t} |\Delta f(s)| \leq \int_0^t |df_s| < \infty$.

EXERCICE. Montrer que l'ensemble des sauts d'une fonction à variation finie est au plus dénombrable

EXERCICE. Montrer que $k : t \mapsto \int_{(0,t]} g df$ est à variation finie et $\Delta k(t) = g(t) \Delta f(t)$.

Formule de changement de variable

Nous allons généraliser la formule d'intégration par partie aux fonctions à variation finie.

LEMME B.2 (INTÉGRATION PAR PARTIE.) *Soient $f, g : [0, \infty) \mapsto \mathbb{R}_+$ à variation finie, alors la fonction produit fg est à variation finie et on a*

$$f(t)g(t) = f(0)g(0) + \int_{(0,t]} f_{s-} dg_s + \int_{(0,t]} g_{s-} df_s + [f, g]_t$$

où $[f, g]_t$ représente le crochet ou covariation quadratique de f et g et est défini par la relation :

$$[f, g]_t = \sum_{0 < s \leq t} \Delta f(s) \Delta g(s).$$

REMARQUE. La covariation quadratique de f et g est bien définie et est elle-même une fonction à variation finie (exercice, indication : utiliser une inégalité de type Cauchy-Schwarz).

PREUVE. Soient μ et ν les mesures associées à f et à g respectivement. Le membre de gauche de l'égalité à démontrer n'est autre que $(\mu \otimes \nu)((0, t] \times (0, t])$. Soient $T_1 = \{(x, y); 0 < x < y \leq t\}$, $T_2 = \{(x, y); 0 < y < x \leq t\}$ et $D = \{(x, y); x = y\}$. Il est facile de montrer que le membre de droite de l'égalité précédente n'est autre que $(\mu \otimes \nu)(T_1) + (\mu \otimes \nu)(T_2) + (\mu \otimes \nu)(D)$ et le résultat s'en déduit immédiatement. ■

Ce résultat étant acquis, il est aisé de démontrer la proposition suivante :

PROPOSITION B.3 (FORMULE DE CHANGEMENT DE VARIABLE) *Soient f une fonction de classe C^1 et F une fonction à variation finie. On a*

$$f(F_t) - f(F_0) = \int_{(0,t]} f'(F_{s-}) F(ds) + \sum_{0 < s \leq t} [\Delta f(F_s) - f'(F_{s-}) \Delta F_s],$$

ou sous forme différentielle en notant $g_t = f(F_t)$:

$$dg_t = f'(F_{t-}) dF_t + (\Delta g_t - f'(F_{t-}) \Delta F_t).$$

Ce résultat se généralise aisément au cadre multidimensionnel : soient $f \in C^1(\mathbb{R}^d)$ et F^1, \dots, F^d à variation finie :

$$\begin{aligned} f(F_t^1, \dots, F_t^d) &= f(F_0^1, \dots, F_0^d) + \int_{(0,t]} \sum_i \frac{\partial f}{\partial x_i}(F_{s-}) dF_s^i \\ &\quad + \sum_{0 < s \leq t} \left(\Delta f(F_s^1, \dots, F_s^d) - \sum_i \frac{\partial f}{\partial x_i}(F_{s-}) \Delta F_s^i \right) \end{aligned}$$

Un moyen mnémotechnique pour retenir cette formule est d'écrire que les sauts sont les mêmes à gauche et à droite de l'égalité.

PREUVE. Le résultat est vrai pour $f(x) = x$. Par intégration par partie, il est vrai pour $xf(x)$, donc pour les polynômes, et pour les fonctions continues par des arguments de densité. Voir [23]. ■

Nous terminons ce paragraphe par une définition cruciale dans ce qui suit

DÉFINITION B.4 *Un processus à variation finie est un processus stochastique A adapté à la filtration \mathbb{F} et tel que, pour presque sûrement tout ω , la trajectoire $t \mapsto A_t(\omega)$ est une fonction à variation finie.*

Les deux exemples précédents peuvent se généraliser à cette situation stochastique : si B est un processus adapté tel que, presque sûrement, $\int_0^t |B_s| ds < \infty$, alors le processus $A_t = \int_0^t B_s ds$ est un processus à variation finie. Soient $(t_n)_{n \geq 0}$ une suite de réels et $(A_n)_{n \geq 0}$ une suite de variables aléatoires telles que A_n soit \mathcal{F}_{t_n} -mesurable pour tout n . Si, pour tout $t \geq 0$, $\sum_n |A_n| \mathbb{1}_{\{t_n \leq t\}} < \infty$ presque sûrement, alors le processus $A_t = \sum_n A_n \mathbb{1}_{\{t_n \leq t\}}$ est un processus à variation finie. Cet exemple peut encore se généraliser au cas où les t_n sont des temps d'arrêt de la filtration \mathbb{F} .

B.2 Les semimartingales simples

Dans cette partie, nous nous plaçons dans le cadre d'un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathbb{P}, ((\mathcal{F}_t)_{t \geq 0}))$ tel que

- (i) \mathcal{F}_0 contient les \mathbb{P} -négligeables (complétude) ;
- (ii) $\mathcal{F}_t = \cap_{s > t} \mathcal{F}_s$ (continuité à droite).

REMARQUE. Considérons deux processus mesurables $X = (X_t)_{t \geq 0}$ et $Y = (Y_t)_{t \geq 0}$, la condition (i) entraîne que si X et Y sont deux versions (voir Annexe A) l'un de l'autre et que X est adapté à la filtration $(\mathcal{F}_t)_{t \geq 0}$, alors Y est lui aussi adapté à cette filtration.

REMARQUE. L'hypothèse (ii) nous permettra de supposer que toutes les martingales considérées sont càd-làg. En effet, si M est une \mathcal{F}_t -martingale et que (ii) est vérifiée, il existe une version \widetilde{M} de M dont les trajectoires sont càd-làg. (cf. Théorème fondamental de la théorie des martingales)

Soit H un processus vérifiant des conditions de mesurabilité et d'intégrabilité à préciser et X une semimartingale "simple" de la forme

$$X = X_0 + A + M,$$

où A désigne un processus à variation finie et $M = \int K dW$ une martingale locale brownienne. L'objet de cette section est définir l'intégrale stochastique $\int H dX$ telle que si X est un martingale locale, alors $\int H dX$ est également une martingale locale. C'est pour préserver cette propriété (vraie si $A = 0$, ce qui est le cas de l'intégrale stochastique standard), que l'on impose des conditions de mesurabilité sur H , en particulier, nous allons imposer à H d'être un processus prévisible (cf. Annexe A).

DÉFINITION B.5 Soient A un processus à variation finie et H un processus prévisible, on dit que H est intégrable par rapport à A si

$$\forall t \int_{(0,t]} |H_s| dA_s < \infty,$$

dans ce cas, $\int_{(0,t]} |H_s| dA_s$ est simplement l'intégrale de Stieljes considérée trajectoire par trajectoire.

PROPOSITION B.6 (ADMISE) Soit A un processus à variation finie. Si A est un martingale (locale) et si H est un processus prévisible et (localement) borné, alors $\int H dA$ est une martingale locale.

Intégration par rapport aux semimartingales simples

DÉFINITION B.7 On appelle semimartingale simple un processus de la forme

$$X_t = X_0 + A_t + M_t,$$

où

- (i) $X_0 \in \mathbb{R}$;
- (ii) A est un processus à variation finie ;
- (iii) M est de la forme $\int K \cdot dW$ où W est un mouvement brownien (d -dimensionnel) par rapport à $(\mathcal{F}_t)_{t>0}$ et K est un processus prévisible localement de carré intégrable i.e. $\int_0^t \|K_s\|^2 ds < \infty$ pour tout $t \geq 0$.

DÉFINITION B.8 Soit $X = X_0 + A + K$ une semimartingale simple. Un processus H prévisible est dit intégrable par rapport à X si et seulement si

- (i) $\forall t \int_{(0,t]} |H_s| |dA_s| < \infty$ p.s. (l'intégrale usuelle existe pour toutes les trajectoires) ;
- (ii) $\forall t \int_{(0,t]} |H_s|^2 \|K_s\|^2 ds < \infty$ p.s. (existence de l'intégrale brownienne).

Dans ce cas, on pose

$$\int_{(0,t]} H_s dX_s := \int_{(0,t]} H_s dA_s + \int_{(0,t]} H_s K_s \cdot dW_s,$$

ce qui définit une intégration par rapport aux semimartingales simples.

Crochets droit et oblique de deux semimartingales simples

Soient X et Y deux semimartingales simples telles que $X = X_0 + A + \int H dW$ et $Y = Y_0 + B + \int K dW$, on définit le *crochet droit* de X et Y en t par :

$$\begin{aligned} [X, Y] &= \left\langle \int H dW, \int K dW \right\rangle + [A, B] \\ &= \int HK dt + [A, B]. \end{aligned}$$

Le crochet droit représente la covariation quadratique des semimartingales simples.

Grâce à cette définition, nous pouvons généraliser à la fois la formule d'intégration par parties pour les fonctions à variation finie et celle (classique) pour les processus d'Itô :

LEMME B.9 (INTÉGRATION PAR PARTIES) *Soient X et Y deux semimartingales simples, alors $Z = XY$ est aussi une semimartingale simple et l'on a :*

$$Z_t = Z_0 + \int_{(0,t]} X_{s-} dY_s + \int_{(0,t]} Y_{s-} dX_s + [X, Y]_t.$$

PREUVE. Il suffit de démontrer le résultat dans les trois cas suivants :

1. $X = A, Y = B$;
2. $X = \int H dW, Y = \int K dW$;
3. $X = A, Y = \int K dW$.

Les cas 1. et 2. sont immédiats : il s'agit respectivement du lemme sur les processus à variation finie et de la théorie classique d'Itô. Reste donc à prouver

$$A_t M_t = \int_{(0,t]} M_s dA_s + \int_0^t A_s K_s dW_s,$$

où $M_t = \int_0^t K_s dW_s$. Modulo l'utilisation d'une procédure de localisation, on peut supposer que sur $[0, t]$

- (i) A est à variation bornée : $\int_{(0,t]} |dA_s| \leq cte$;
- (ii) M est bornée : $\sup_{s \leq t} |M_s| \leq cte$.

REMARQUE. Localiser le problème signifie trouver une suite de temps d'arrêts $(T_n)_{n \geq 0} \uparrow \infty$ telle que

$$A_t^n = A_{t \wedge T_n}, M_t^n = M_{t \wedge T_n}$$

vérifient les propriétés considérées. Il suffit ici de choisir

$$T_n = \inf_{t \geq 0} \left\{ \int_{(0,t]} |dA_s| > n \text{ ou } \sup_{s \leq t} |M_s| > n \right\}.$$

Une fois le résultat prouvé dans ce cas, il suffit de faire tendre n vers $+\infty$.

Soit $t_k^n = k2^{-n} \wedge t$, on écrit que

$$A_t M_t = \sum_{k \geq 0} A_{t_k^n} \Delta M_{t_k^n} + \sum_{k \geq 0} M_{t_k^n} \Delta A_{t_k^n} + \sum_{k \geq 0} \Delta A_{t_k^n} \Delta M_{t_k^n}$$

avec $\Delta M_{t_k^n} = M_{t_{k+1}^n} - M_{t_k^n}$ et $\Delta A_{t_k^n} = A_{t_{k+1}^n} - A_{t_k^n}$, et on étudie séparément la convergence de chacune de ces trois sommes. Tout d'abord,

$$\sum_{k \geq 0} A_{t_k^n} \Delta M_{t_k^n} = \int_0^t \tilde{A}_s^n dM_s,$$

avec $\tilde{A}_s^n = \sum_{k \geq 0} A_{t_k^n} \mathbb{1}_{(t_k^n, t_{k+1}^n]}$, et d'après le théorème de Lebesgue, on a

$$\mathbb{E} \left[\int_0^t (\tilde{A}_s^n - A_s)^2 K_s^2 ds \right] \rightarrow 0,$$

donc $\sum_{k \geq 0} A_{t_k^n} \Delta M_{t_k^n} \rightarrow_{L^2} \int_0^t A_s dM_s$.

En ce qui concerne la deuxième somme, on a immédiatement par définition de l'intégrale de Riemann-Stieltjes

$$\sum_{k \geq 0} M_{t_k^n} \Delta A_{t_k^n} \rightarrow_{p.s.} \int_{(0,t]} M_s dA_s,$$

et enfin, on a

$$\sum_{k \geq 0} \Delta A_{t_k^n} \Delta M_{t_k^n} \rightarrow_{p.s.} 0$$

car

$$\sum_{k \geq 0} \Delta A_{t_k^n} \Delta M_{t_k^n} \leq \sup \{ |M_u - M_s|; u, s \leq t, |u - s| \leq 2^{-n} \} \int_{(0,t]} dA_s,$$

ce qui tend bien vers 0 car la borne supérieure tend vers 0 par continuité de M et l'intégrale est à variation finie.

Ceci achève la démonstration ■

Formule de changement de variable

Nous allons maintenant généraliser la formule de changement de variable au cas des semimartingales simples.

PROPOSITION B.10 *Soit X une semimartingale simple et f une fonction C^2 , alors $f(X)$ est aussi une semimartingale simple et l'on a la relation suivante :*

$$\begin{aligned} f(X_t) = f(X_0) + \int_{(0,t]} f'(X_{s-}) dX_s + \frac{1}{2} \int_{(0,t]} f''(X_{s-}) d\langle X^c, X^c \rangle_s \\ + \sum_{0 < s \leq t} (\Delta f(X_s) - f'(X_{s-}) \Delta X_s). \end{aligned}$$

Ce résultat s'étend au cadre multidimensionnel de la façon suivante :

PROPOSITION B.11 *Soit $X = (X_1, \dots, X_n)$ une semimartingale simple et f une fonction C^2 , alors $f(X)$ est aussi une semimartingale simple et l'on a la relation suivante :*

$$\begin{aligned} f(X_t) = f(X_0) + \sum_i \int_{(0,t]} \partial_i f(X_{s-}) dX_s^i + \frac{1}{2} \sum_{i,j} \int_{(0,t]} \partial_{ij} f(X_{s-}) d\langle X_i^c, X_j^c \rangle_s \\ + \sum_{0 < s \leq t} (\Delta f(X_s) - \sum_i \partial_i f(X_{s-}) \Delta X_s^i). \end{aligned}$$

PREUVE. La démonstration est analogue à celle effectuée dans le cadre des fonctions à variation finie, à partir de la formule d'intégration par parties. ■

On peut alors démontrer le lemme suivant :

LEMME B.12 Soient $X = X_0 + A + M$ et $Y = Y_0 + B + N$ deux semimartingales simples,

$$X = Y \Rightarrow A + X_0 = B + Y_0, M = N$$

PREUVE.

$$X = Y \Rightarrow A - B + X_0 - Y_0 = \int (K - H) dW.$$

Le membre de gauche, disons φ , est un processus à variation finie et celui de droite est une martingale continue. Le processus φ est donc une martingale continue et est à variation finie. Par suite, et en vertu de la formule d'intégration par parties, on peut écrire :

$$\mathbb{E}[(\varphi_t - \varphi_0)^2] = \mathbb{E}[\varphi_t^2 - \varphi_0^2] = \mathbb{E}\left[\int_0^t N_s dN_s\right] = 0.$$

Le processus φ est donc constant et vaut 0 en 0 donc c'est le processus nul. ■

Application : autour du processus de Poisson

Soit $(N_t)_{t \geq 0}$ un processus de comptage, i.e. tel que

$$\left\{ \begin{array}{l} \Delta N_t \in \{0, 1\} \\ N \text{ est constant entre deux sauts } N_t < \infty \forall t \text{ p.s.} \end{array} \right. ,$$

Nous faisons l'hypothèse selon laquelle $\forall t, N_t < \infty$ presque sûrement.

DÉFINITION B.13 Soit $\lambda : \mathbb{R}_+ \rightarrow \mathbb{R}_+^*$. Un processus de comptage N est un processus de Poisson d'intensité λ si

$$\forall s \in (0, t], \mathcal{L}(N_t - N_s | N_u, u \leq s) = \mathcal{P}\left(\int_0^t \lambda(u) du\right).$$

REMARQUE. On rappelle que $X \sim \mathcal{P}(\lambda) \Leftrightarrow \mathbb{P}[X = k] = e^{-\lambda} \frac{\lambda^k}{k!}, k \in \mathbb{N}$. La propriété précédente entraîne que le processus de Poisson est à accroissements indépendants : $0 \leq u_1 < u_2 < v_1 < v_2 < \infty \Rightarrow N_{u_2} - N_{u_1} \perp N_{v_2} - N_{v_1}$.

PROPOSITION B.14 Soit N un processus de Poisson, les assertions suivantes sont équivalentes :

- (i) N est un processus de Poisson d'intensité λ ;
- (ii) $\mathbb{E}[N_{t+h} - N_t | \mathcal{F}_t^N] = \lambda_t h + o(h)$
- (iii) $N_t - \int_0^t \lambda_u du$ est une \mathcal{F}^N -martingale.

PREUVE.

(i) \Rightarrow (ii) est trivial

(ii) \Rightarrow (iii) : supposons que $\mathbb{E}[N_{t+h} - N_t | \mathcal{F}_t^N] = \lambda_t h + o(h)$ et montrons que, pour $s \leq t$,

$$\mathbb{E}[N_t - N_s | \mathcal{F}_s^N] = \int_s^t \lambda_u du.$$

$$\begin{aligned} \mathbb{E}[N_t - N_s | \mathcal{F}_s^N] &= \sum_{k=1}^n \mathbb{E} \left[N_{s+\frac{k}{n}(t-s)} - N_{s-\frac{k-1}{n}(t-s)} \middle| \mathcal{F}_s^N \right] \\ &= \sum_{k=1}^n \mathbb{E} \left[N_{s+\frac{k}{n}(t-s)} - N_{s-\frac{k-1}{n}(t-s)} \middle| \mathcal{F}_{s+\frac{k-1}{n}(t-s)}^N \middle| \mathcal{F}_s^N \right] \\ &= \sum_{k=1}^n \mathbb{E} \left[\lambda_{s+\frac{k-1}{n}(t-s)} \frac{1}{n} + o\left(\frac{1}{n}\right) \middle| \mathcal{F}_s^N \right] \\ &= \sum_{k=1}^n \frac{1}{n} \lambda_{s+\frac{k-1}{n}(t-s)} + o(1) \xrightarrow{n \rightarrow \infty} \int_s^t \lambda_u du + o(1). \end{aligned}$$

(iii) \Rightarrow (i) : d'après la formule de changement de variable, on a

$$\begin{aligned} f(N_t - N_s) &= 1 + \int (s, t] f'(N_{u-} - N_s) dN_u \\ &\quad + \sum_{s < u \leq t} \left(f(1 + N_{u-} - N_s) - f(N_{u-}) - f'(N_{u-} - N_s) \Delta N_u \right) \\ &= 1 + \sum_{s < u \leq t} f'(N_u - N_s) \Delta N_u + f(1 + N_{u-} - N_s) \\ &\quad - f(N_{u-} - N_s) - f'(N_u - N_s) \Delta N_u \\ &= 1 + \int (s, t] (f(1 + N_{u-} - N_s) - f(N_{u-} - N_s)) dN_s. \end{aligned}$$

Pour $f(x) = e^{icx}$, l'expression devient :

$$\begin{aligned} e^{ic(N_t - N_s)} &= 1 + (e^{ic} - 1) \int_{(s,t]} e^{ic(N_{u-} - N_s)} dN_u \\ &= 1 + (e^{ic} - 1) \int_{(s,t]} e^{ic(N_{u-} - N_s)} (dN_u - \lambda_u du) + (e^{ic} - 1) \int_{(s,t]} e^{ic(N_{u-} - N_s)} \lambda_u du, \end{aligned}$$

et comme $dN_u - \lambda_u du$ est une martingale, on a

$$\begin{aligned} \mathbb{E} \left[e^{ic(N_t - N_s)} \middle| \mathcal{F}_s^N \right] &= 1 + \mathbb{E} \left[(e^{ic} - 1) \int_{(s,t]} e^{ic(N_{u-} - N_s)} \lambda_u du \middle| \mathcal{F}_s^N \right] \\ &= 1 + (e^{ic} - 1) \int_s^t \mathbb{E} \left[e^{ic(N_{u-} - N_s)} \middle| \mathcal{F}_s^N \right] \lambda_u du. \end{aligned}$$

Et en posant $g(t) = \mathbb{E} \left[e^{ic(N_t - N_s)} \middle| \mathcal{F}_s^N \right]$, on obtient pour g une équation différentielle usuelle qui se résout en

$$g(t) = \exp \left((e^{ic} - 1) \int_s^t \lambda_u du \right),$$

qui est bien la fonction caractéristique d'un loi de Poisson de paramètre $\int_0^t \lambda_u du$. ■

Rappel sur les diffusions

Dans cette partie, nous nous intéressons aux propriétés de base des processus de diffusion, et à leurs extensions aux diffusions avec des sauts. Un cas particulièrement intéressant en finance est celui des diffusions affines qui seront étudiées en fin de chapitre.

C.1 Définitions

Point de vue des caractéristiques locales

Une *diffusion* est un processus continu $(X_t, t \geq 0)$ à valeurs dans \mathbb{R}^d , défini sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})$ tel qu'il existe deux fonctions $\mu : \mathbb{R}^d \rightarrow \mathbb{R}^d$ et $a : \mathbb{R}^d \rightarrow S_+^{d \times d}$ (où $S_+^{d \times d}$ désigne l'ensemble des matrices symétriques définies positives sur $d \times d$), tels que

1. $\mathbb{E}[dX_t | \mathcal{F}_t] = \mu(X_t) dt + o(dt)$
2. $\mathbb{E}[(dX_t^i - \mu^i dt)(dX_t^j - \mu^j dt) | \mathcal{F}_t] = a^{i,j}(X_t) dt + o(dt)$.

On sait construire de tels processus à l'aide de la théorie des équations différentielles stochastiques : Si μ et a sont des fonctions suffisamment régulières, i.e. vérifient les conditions de Lipschitz, alors on peut construire X comme solution de l'équation différentielle stochastique

$$dX_t = \mu(X_t) dt + \sigma(X_t) dW_t$$

où σ vérifie $\sigma \sigma^* = a$.

REMARQUE. Les propriétés 1. et 2. caractérisent la loi du processus.

Point de vue des processus de Markov

On peut également voir le processus X défini précédemment comme un processus de Markov de générateur infinitésimal

$$L = \frac{1}{2} \sum_{i,j=1}^n a^{i,j} D_i D_j + \sum_{i=1}^n \mu^i D_i$$

où D_i désigne l'opérateur de dérivation par rapport à la variable x_i

$$D_i = \frac{\partial}{\partial x_i}.$$

Point de vue des problèmes de martingales

Enfin, il est possible de définir une diffusion comme un processus X tel que, pour toute fonction $f \in C_b^2$, le processus

$$M_t = f(X_t) - f(X_0) - \int_0^t (Lf)(X_s) ds$$

est une martingale, où l'opérateur L désigne le générateur défini ci-dessus.

Ces définitions sont équivalentes, et il est possible de les généraliser dans le cas des processus à sauts.

C.2 Diffusions avec sauts

DÉFINITION C.1 Une diffusion avec sauts est un processus càd-làg $(X_t, t \geq 0)$ défini sur $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})$ et à valeurs dans \mathbb{R} tel qu'il existe $\mu : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $a : \mathbb{R}^d \rightarrow S_+^{d \times d}$, $\lambda : \mathbb{R}^d \rightarrow \mathbb{R}_+^*$ et ν une mesure sur \mathbb{R}^d tels que

- (i) $\mathbb{E}[dX_t | \mathcal{F}_t] = \mu(X_t) dt + o(dt)$
- (ii) $\mathbb{E}[(dX_t^i - \mu^i dt)(dX_t^j - \mu^j dt) | \mathcal{F}_t] = a^{i,j}(X_t) dt + o(dt)$
- (iii) $\mathbb{E}[dN_t | \mathcal{F}_t] = \lambda(X_{t-}) dt + o(dt)$
- (iv) $\mathbb{E}[f(\Delta X_t) | \mathcal{F}_t, \Delta X_t \neq 0] = \int f(x) \nu(dx)$,

où N_t est le processus de comptage des sauts de X :

$$N_t := \sum_{0 < s \leq t} \mathbb{1}_{\{\Delta X_s \neq 0\}}.$$

REMARQUE. Les deux premières conditions sont directement héritées de la définition classique des diffusions (sans saut) tandis que l'on peut interpréter les deux dernières conditions de la façon suivante :

1. (iii) signifie que, conditionnellement à l'ensemble de l'information disponible au temps t , les sauts du processus $(X_t, t \geq 0)$ se produisent suivant un processus de Poisson d'intensité $\lambda(X_{t-}), t > 0$;

2. la condition (iv) est équivalente à

$$(iv') \quad \mathcal{L}(\Delta X_t | \mathcal{F}_t, \Delta X_t \neq 0) = \nu.$$

Ceci signifie que la loi ν de l'amplitude de chacun des sauts est indépendante des autres paramètres du processus.

Un tel processus se construit comme :

1. Une diffusion avec des sauts

$$dX_t = \mu(X_t) dt + \sigma(X_t) dW_t + dJ_t,$$

où J est un processus purement de saut d'intensité $\lambda(X_{t-})$ et d'amplitude ν ;

2. un processus de Markov de générateur

$$(Lf)(x) = \frac{1}{2} \sum_{i,j} a^{i,j} D_{i,j} + \sum_i \mu^i D_i + \lambda(x) \int (f(x+z) - f(x)) \nu(dz).$$

Le lien entre (1) et (2) s'écrit :

PROPOSITION C.2 *Pour toute fonction f appartenant à C_b^2 , le processus*

$$M_t = f(X_t) - f(X_0) - \int_0^t (Lf)(X_{s-}) ds$$

est une martingale.

PREUVE. D'après la formule d'Itô

$$\begin{aligned} df(X_t) &= f'(X_{t-}) dX_t + \frac{1}{2} f''(X_{t-}) d\langle X^c, X^c \rangle_t + \Delta f(X_t) - f'(X_{t-}) \Delta(X_t) \\ &= f'(X_{t-}) \mu(X_t) dt + f'(X_{t-}) \sigma(X_t) dW_t + \frac{1}{2} f''(X_{t-}) \sigma^2(X_t) dt \\ &\quad + f'(X_{t-}) dJ_t + \Delta f(X_t) - f'(X_{t-}) \Delta(X_t) \\ &= (Lf)(X_{t-}) dt + \Delta f(X_t) - \lambda(X_{t-}) \int (f(z + X_{t-}) - f(X_{t-})) \nu(dz) \\ &= (Lf)(X_{t-}) dt + (f(X_t + \Delta X_t) - f(X_{t-})) dN_t \\ &\quad - \lambda(X_{t-}) \int (f(z + X_{t-}) - f(X_{t-})) \nu(dz), \end{aligned}$$

donc

$$\begin{aligned} &\mathbb{E}[df(X_t) - (Lf)(X_{t-}) dt | \mathcal{F}_t] \\ &= \mathbb{E}[(f(X_t + \Delta X_t) - f(X_{t-})) dN_t - \lambda(X_{t-}) \int (f(z + X_{t-}) - f(X_{t-})) \nu(dz) | \mathcal{F}_t] \\ &= \mathbb{E}[(f(X_t + \Delta X_t) - f(X_{t-})) dN_t | \mathcal{F}_t] - \lambda(X_{t-}) \int (f(z + X_{t-}) - f(X_{t-})) \nu(dz) \\ &= \mathbb{E}[dN_t | \mathcal{F}_t] \int (f(z + X_{t-}) - f(X_{t-})) \nu(dz) \\ &\quad - \lambda(X_{t-}) \int (f(z + X_{t-}) - f(X_{t-})) \nu(dz) \\ &= 0, \end{aligned}$$

ce qui achève la preuve. ■

La théorie des diffusions offre un cadre de travail très général, qui ne permet pas toujours de faire les calculs explicitement. C'est pourquoi, dans la pratique, on considère souvent des diffusions affines avec sauts. Ces processus sont commodes pour leur tractabilité analytique qui en font des outils de modélisation incontournables. Pour plus de détails sur ces processus, nous renvoyons le lecteur à Pan et Filipovic ainsi qu'aux articles [9] et [6].

C.3 Diffusion affine avec des sauts

On suppose que μ, a, λ sont des fonctions affines de X , autrement dit qu'il existe $k_0 \in \mathbb{R}^d, k_1 \in \mathbb{R}^{d \times d}, H_0 \in \mathbb{R}^{d \times d}, H_1 \in \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ linéaire, et $l_0 \in \mathbb{R}, l_1 \in \mathbb{R}^d$ tels que

$$\begin{aligned}\mu(x) &= k_0 + k_1 x \\ a(x) &= H_0 + H_1 x \\ \lambda(x) &= l_0 + l_1 x.\end{aligned}$$

On note θ la transformée de Laplace de ν . Autrement dit, pour tout $c \in \mathbb{Z}$, on a

$$\theta(c) = \int e^{cz} \nu(dz).$$

PROPOSITION C.3 Soient X une diffusion affine avec des sauts et $R : \mathbb{R}^d \rightarrow \mathbb{R}, y \mapsto l_0 + l_1 y$ une application linéaire, alors

$$\mathbb{E} \left[\exp \left(- \int_t^T R(X_s) ds \right) e^{u \cdot X_T} \middle| \mathcal{F}_t^X \right] = e^{\alpha_u(T-t) + \beta_u(T-t) \cdot X_t},$$

où α_u et β_u sont solution des équations différentielles suivantes :

$$\begin{cases} \alpha' = l_0 - \beta^* \cdot k_0 - \frac{1}{2} \beta^* H_0 \beta - l_0 (\theta(\beta) - 1) \\ \beta' = l_1 - k_1^* \cdot \beta - \frac{1}{2} \beta^* H_1 \beta - l_1 (\theta(\beta) - 1), \end{cases}$$

avec les conditions aux bords

$$\alpha_u(T, T) = 0, \beta_u(T, T) = u.$$

On est donc ramené au problème de résolution d'un système de deux équations différentielles ordinaires. En pratique, soit il existe des solutions explicites, soit on utilise un algorithme numérique de type Runge-Kutta.

PREUVE. On admet que ϕ s'écrit bien sous cette forme. Nous allons vérifier que les équations différentielles ordinaires sont satisfaites. Pour simplifier les notations, nous nous limiterons à un cadre de dimension 1, l'extension aux dimensions supérieures étant immédiate. Posons

$$\phi(t, X_t) = \mathbb{E} \left[e^{\int_t^T R(X_s) ds} e^{u \cdot X_T} \middle| \mathcal{F}_t^X \right] = e^{\alpha_u(T-t) + \beta_u(T-t) \cdot X_t}.$$

Il est clair que $e^{-\int_0^t R(X_s) ds} \phi(t, X_t)$ est une martingale, et de plus

$$\mathbb{E} \left[\frac{d\phi(t, T)}{\phi(t, T)} \middle| \mathcal{F}_t \right] = R(X_t) dt.$$

Par application de la formule d'Itô :

$$\begin{aligned} d\phi(t, X_t) &= \frac{\partial \phi}{\partial t} dt + \frac{\partial \phi(t, X_{t-})}{\partial X} dX_t + \frac{1}{2} \frac{\partial^2 \phi}{\partial X^2} d\langle X^c, X^c \rangle + \Delta \phi - \frac{\partial \phi}{\partial X} \Delta X \\ &= \frac{\partial \phi}{\partial t} dt + \frac{\partial \phi}{\partial X} (\mu dt + \sigma(X_t) dW_t + dJ_t) + \frac{1}{2} \frac{\partial^2 \phi}{\partial X^2} \sigma^2(X_t) dt \\ &\quad + \Delta \phi - \frac{\partial \phi}{\partial X} \Delta X. \end{aligned}$$

Or, d'après la forme affine de ϕ , on a

$$\phi(t, X_{t-} + z) = \phi(t, X_{t-}) e^{z\beta(T-t)},$$

et donc

$$\begin{aligned} \frac{d\phi(t, X_t)}{\phi(t, X_{t-})} &= \left(\alpha' + \beta' X_{t-} + \beta(k_0 + k_1 X_t) + \frac{\beta^2}{2} (H_0 + H_1 X_{t-}) \right. \\ &\quad \left. + (l_0 + l_1 X_{t-}) \int (e^{\beta z} - 1) \nu(dz) \right) dt + \sigma(X_t) dW_t \\ &\quad + \left(\frac{\Delta(\phi(t, X_t))}{\phi(t, X_{t-})} - (l_0 + l_1 X_{t-}) \int \frac{\phi(t, X_{t-} + z) - \phi(t, X_{t-})}{\phi(t, X_{t-})} \nu(dz) dt \right). \end{aligned}$$

Les deux derniers termes de la somme étant martingales, on a d'après l'expression de l'espérance

$$\begin{aligned} \alpha' - l_0 + \beta k_0 + \frac{\beta^2}{2} H_0 \beta' X_{t-} + l_0(\theta(\beta) - 1) \\ + (\beta' - l_1 + \beta k_1 + \frac{\beta^2}{2} H_1 + l_1(\theta(\beta) - 1)) X_{t-} = 0, \end{aligned}$$

ce qui est vrai pour toutes les valeurs du processus X . On en déduit donc immédiatement les deux équations

$$\begin{cases} \alpha' = l_0 - \beta \cdot k_0 - \frac{1}{2} \beta^* H_0 \beta - l_0(\theta(\beta) - 1) \\ \beta' = l_1 - k_1^* \beta - \frac{1}{2} \beta^* H_1 \beta - l_1(\theta(\beta) - 1), \end{cases}$$

les conditions aux bords étant la conséquence de

$$\phi(T, X_T) = e^{\alpha(T,T) + \beta(T,T)X_T} = e^{0 + uX_T}.$$

■

COROLLAIRE C.4 Soient X une diffusion affine avec des sauts et $R : \mathbb{R}^d \rightarrow \mathbb{R}, y \mapsto l_0 + l_1 y$ une application linéaire.

$$\mathbb{E} \left[\exp \left(- \int_t^T R(X_s) ds \right) X_T e^{u \cdot X_T} \middle| \mathcal{F}_t^X \right] = (A_u(T-t) + B_u(T-t) \cdot X_t) e^{\alpha_u(T-t) + \beta_u(T-t) \cdot X_t},$$

où A_u et B_u sont solution des équations différentielles suivantes :

$$\begin{cases} -A' = k_0^* \cdot B + \beta^* H_0 B + l_0(\nabla \theta)(\beta) \cdot B \\ -B' = k_1^* \cdot B + \beta^* H_1 B + l_1(\nabla \theta)(\beta) \cdot B, \end{cases}$$

avec les conditions aux bords

$$A_u(T, T) = 0, \quad B_u(T, T) = 1.$$

PREUVE. Il suffit de reprendre la preuve précédente en dérivant par rapport à u . ■

Les fonctions copules

Les *fonctions copules* ou plus simplement *copules* sont un outil mathématique permettant de modéliser la dépendance entre des variables aléatoires connaissant leurs lois marginales. Dans cette partie, nous définissons les fonctions copules, et présentons les exemples les plus couramment utilisés en finance. Le résultat fondamental de la théorie étant le théorème de Sklar. Nous nous intéressons ensuite au cas particulier des copules archimédiennes qui revêtent une importance particulière pour les calculs.

D.1 Définitions et exemples

DÉFINITION D.1 *La fonction $C : [0, 1]^N \rightarrow [0, 1]$ est une copule de dimension N s'il existe un vecteur aléatoire à marges uniformes $U = (U_1, \dots, U_N)$ tel que*

$$C(u_1, \dots, u_N) = \mathbb{P}[U_1 \leq u_1, \dots, U_N \leq u_N].$$

Autrement dit, C est la fonction de répartition d'un vecteur aléatoire à marges uniformes.

REMARQUE. On rappelle que, si X est une variable aléatoire et F_X sa fonction de répartition,

$$Z = F_X(X) \sim U(0, 1),$$

où U est une variable aléatoire de loi uniforme sur $[0, 1]$. Soient X_1, X_2, \dots, X_N des variables aléatoires, de fonctions de répartitions F_1, \dots, F_n , alors la fonction de répartition jointe F est définie par

$$F(x) = \mathbb{P}[X_1 \leq x_1, \dots, X_N \leq x_N].$$

Le théorème fondamental est le suivant :

THÉORÈME D.2 (SKLAR)

Avec les notations précédentes, il existe une fonction copule C de dimension N telle que

$$F(x) = C(F_1(x_1), \dots, F_N(x_N)).$$

Si les fonctions F_i sont continues, C est unique, sinon, elle est uniquement déterminée sur $\text{Im}(F_1) \times \dots \times \text{Im}(F_N)$.

REMARQUE. D'après le théorème de Sklar, une distribution multidimensionnelle continue F est uniquement déterminée par

- ses lois marginales ;
- sa structure de dépendance décrite par une fonction copule C .

On dit alors que F a pour copule C . Dans la suite, on s'intéressera uniquement à des distributions continues, ce qui permet d'en déduire l'unicité de la fonction copule associée.

PROPOSITION D.3 La fonction copule de $X = (X_1, \dots, X_N)$ est la fonction de répartition du vecteur aléatoire $(F_1(X_1), \dots, F_N(X_N))$.

PREUVE. Notons \bar{C} la fonction de répartition du vecteur $(F_1(X_1), \dots, F_N(X_N))$,

$$\begin{aligned} F(x) &= \mathbb{P}[X_1 \leq x_1, \dots, X_N \leq x_N] \\ &= \mathbb{P}[F_1(X_1) \leq F(x_1), \dots, F_N(X_N) \leq F(x_N)] \\ &= \bar{C}(F_1(X_1), \dots, F_N(X_N)), \end{aligned}$$

d'où le résultat par unicité. ■

PROPOSITION D.4 Soient f_1, \dots, f_N des fonction croissantes et X_1, \dots, X_N des variables aléatoires. Pour tout i appartenant à $\{1, \dots, N\}$, on pose $Y_i = F_i(X_i)$. Alors la copule des Y_i est la même que celle est X_i . En d'autres termes, la structure de dépendance est invariante par transformation monotone.

PREUVE. La preuve est laissée au lecteur ■

EXEMPLE. La copule indépendante représente la structure de dépendance associée à une famille de variables aléatoires uniformes indépendantes

$$C(u_1, \dots, u_N) = u_1 \dots u_N.$$

EXEMPLE. La copule gaussienne est associée à la distribution normale multidimensionnelle. Soit $X = (X_1, \dots, X_N)$ un vecteur gaussien standard (i.e pour tout i appartenant à $\{1, \dots, N\}$, $\mathbb{E}[X_i] = 1$, $\text{Var}(X_i) = 0$), de matrice de corrélation $(\rho_{ij})_{i,j \in \{1, \dots, N\}}$. Notons ϕ_ρ la fonction de répartition d'une gaussienne multivariée, on a par définition :

$$C_\rho(\phi(u_1), \dots, \phi(u_N)) = \phi_\rho(u_1, \dots, u_N).$$

La copule gaussienne est donc définie de façon implicite par la relation

$$C_\rho(u_1, \dots, u_N) = \phi_\rho(\phi^{-1}(u_1), \dots, \phi^{-1}(u_N)).$$

EXERCICE. Montrer que la copule gaussienne bivariée admet la représentation intégrale suivante :

$$C(u_1, u_2, \rho) = \int_0^{u_1} \phi\left(\frac{\phi^{-1}(u_2) - \rho\phi^{-1}(u)}{\sqrt{1-\rho^2}}\right) du.$$

REMARQUE. Bien qu'il n'existe pas d'expression explicite pour cette fonction copule, il est très aisé de simuler des variables aléatoires de lois marginales données ayant une structure de dépendance décrite par cette fonction copule. Cette propriété explique en partie la popularité de cette structure de dépendance.

REMARQUE. Dans les cas où la matrice des corrélations est de la forme

$$\text{cor}(X_i, X_j) = \rho_i \rho_j,$$

on peut trouver des variables aléatoires gaussiennes indépendantes X, Y_1, \dots, Y_N telles que

$$X_i = \rho_i X + \sqrt{1 - \rho_i^2} Y_i \quad (\text{modèle à un facteur}).$$

Dans ce cas, la fonction copule correspondante s'écrit (exercice) :

$$C(u_1, u_2, \rho) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi\left(\frac{\phi^{-1}(u_1) - \rho_1 x}{\sqrt{1 - \rho_1^2}}\right) \dots \phi\left(\frac{\phi^{-1}(u_N) - \rho_N x}{\sqrt{1 - \rho_N^2}}\right) dx.$$

EXEMPLE. De la même manière que pour la copule gaussienne, il est possible de définir la *copule de student*.

D.2 Les copules archimédiennes

DÉFINITION D.5 On appelle copule archimédienne une fonction copule telle qu'il existe une fonction inversible $\psi : [0, 1] \rightarrow \mathbb{R}_+ \cup \{\infty\}$ vérifiant

$$C(u_1, \dots, u_N) = \psi^{-1}\left(\sum_{i=1}^N \psi(x_i)\right).$$

Dans ce cas, pour tout x appartenant à $\mathbb{R}_+ \cup \{\infty\}$, on a $\psi'(x) < 0$, $\psi''(x) > 0$.

EXERCICE. Soit C une copule archimédienne de générateur ψ , montrer que

$$\frac{\partial C}{\partial u_i}(u) = \frac{\psi'(u_i)}{\psi'(C(u_1, \dots, u_N))},$$

et

$$\frac{\partial C}{\partial u_i \partial u_j}(u) = -\psi'(u_i)\psi'(u_j) \frac{\psi''(C(u))}{\psi'(C(u))^3}.$$

EXEMPLE. (*Copule de Gumbel*)

$$\psi(x) = -\ln(x)^\theta, \theta \geq 1$$

$$C(u_1, \dots, u_N) = \exp\left(-\left[\sum_{i=1}^N -\ln x_i^\theta\right]\right).$$

EXEMPLE. (*Copule de Clayton*)

$$\psi(x) = \frac{\left(\frac{1}{x^\alpha} - 1\right)}{\alpha}, \alpha \geq 0$$

$$C(u_1, \dots, u_N) = \left(1 - N + \sum_{i=1}^N \frac{1}{u_i^\alpha}\right)^{-\frac{1}{\alpha}}.$$

EXERCICE. Soient Y une variable aléatoire à valeurs dans $(0, +\infty)$ et ψ sa transformée de Laplace définie par

$$\psi(u) = \mathbb{E}[e^{-uY}], \quad (u \geq 0).$$

(1) Montrer que ψ réalise une bijection de $[0, +\infty)$ sur $[0, 1]$.

On note ϕ l'inverse ψ^{-1} de la transformée de Laplace de Y . On considère une famille X_1, \dots, X_N de variables aléatoires i.i.d. de loi uniforme sur $(0, 1)$. On suppose que les X_i sont indépendantes de la variable Y . On pose

$$U_i = \psi\left(-\frac{1}{Y} \ln X_i\right), \quad (1 \leq i \leq N).$$

(2) Montrer que les U_i sont des v.a. de loi uniforme sur $(0, 1)$ dont la distribution jointe est décrite par la copule archimédienne de générateur ϕ ; autrement dit, que

$$\mathbb{P}[U_1 \leq u_1; \dots; U_N \leq u_N] = \psi\left(\sum_{i=1}^N \phi(u_i)\right).$$

(3) Dédurre de ce qui précède un algorithme pour simuler un vecteur de v.a. de loi uniforme dont la loi jointe est caractérisée par la copule archimédienne de générateur ϕ .

Le cas des copules archimédiennes généralisés

Pour définir une copule archimédienne généralisée nous partons des données suivantes : un vecteur $Y \in \mathbb{R}^K$ de facteurs de risque qui sont des v.a. positives et une matrice $A \in \mathbb{R}^{N \times K}$ de poids. On pose alors pour tout $1 \leq i \leq N$ et $1 \leq k \leq K$

$$\tilde{Y}_i = \sum_{k=1}^K a_{ik} Y_k = a_i \cdot Y,$$

$$\tilde{\phi}_i(s) = \mathbb{E}[e^{-s\tilde{Y}_i}], \quad \tilde{\psi}_i = \tilde{\phi}_i^{-1},$$

$$\phi(s) = \mathbb{E}[e^{-sY}].$$

La copule archimédienne généralisée correspondante est alors définie par

$$C(u) = \phi\left(\sum_{i=1}^N a_{i1}\tilde{\psi}_i(u_i), \dots, \sum_{i=1}^N a_{iK}\tilde{\psi}_i(u_i)\right)$$

Le lemme suivant nous assure que C est bien une copule et fournit un moyen commode de simuler des v.a. de loi décrite par la copule C .

LEMME D.6 Soient $(X_i, 1 \leq i \leq N)$ des v.a. indépendantes de loi uniformes sur $[0, 1]$ et $(U_i, 1 \leq i \leq N)$ les v.a. définies par la formule suivante :

$$U_i = \tilde{\psi}_i(-\ln X_i/\tilde{Y}_i).$$

Alors, la distribution jointe des U_i est donnée par

$$\mathbb{P}[U_i \leq u_i, 1 \leq i \leq N] = \phi\left(\sum_{i=1}^N a_{i1}\tilde{\psi}_i(u_i), \dots, \sum_{i=1}^N a_{iK}\tilde{\psi}_i(u_i)\right).$$

et les lois marginales des U_i sont des lois uniformes sur $[0, 1]$.

PREUVE. Par définition, on a

$$\begin{aligned} \mathbb{P}[U_i \leq u_i, \forall i] &= \mathbb{P}[X_i \leq \exp(-\tilde{Y}_i\tilde{\psi}_i(u_i)), \forall i] \\ &= \mathbb{E}\left[\mathbb{P}[X_i \leq \exp(-\tilde{Y}_i\tilde{\psi}_i(u_i)), \forall i | \tilde{Y}_1, \dots, \tilde{Y}_N]\right] \\ &= \mathbb{E}\left[\exp\left(\sum_{i=1}^N \tilde{Y}_i\tilde{\psi}_i(u_i)\right)\right] \\ &= \mathbb{E}\left[\exp\left(\sum_{k=1}^K \left(\sum_{i=1}^N a_{ik}\tilde{\psi}_i(u_i)\right)Y_k\right)\right] \\ &= \phi\left(\sum_{i=1}^N a_{i1}\tilde{\psi}_i(u_i), \dots, \sum_{i=1}^N a_{iK}\tilde{\psi}_i(u_i)\right). \end{aligned}$$

Il reste à montrer que pour tout i , on a $\mathbb{P}[U_i \leq u_i] = u_i$. Pour le voir, on écrit en utilisant le même genre de conditionnement que précédemment que

$$\mathbb{P}[U_i \leq u_i] = \mathbb{E}\left[\exp(-\tilde{Y}_i\tilde{\psi}_i(u_i))\right] = \tilde{\phi}_i(\tilde{\psi}_i(u_i)) = u_i.$$

■

Dans la pratique, il est judicieux de choisir des facteurs tels que les transformées de Laplace précédentes soient faciles à calculer et à inverser. Un exemple simple consiste à choisir les facteurs de risques indépendants et dans une famille dont les lois sont stable par sommation.

Démonstration du Lemme 3.1

Dans cette partie, nous allons démontrer le résultat suivant :

$$\mathbb{P}\left[\inf_{0 \leq s \leq t} (\mu s + \sigma B_s) \geq y\right] = N\left(\frac{\mu t - y}{\sigma \sqrt{t}}\right) - e^{\frac{2\mu}{\sigma^2}y} N\left(\frac{y + \mu t}{\sigma \sqrt{t}}\right).$$

Posons $X_t = B_t + \mu t$, $M_t = \max_{0 \leq s \leq t} X_t$, et $m_t = \min_{0 \leq s \leq t} X_t$. La démonstration revient à prouver le lemme suivant :

LEMME E.1

$$\mathbb{P}[X_t \leq x; M_t \geq y] = e^{2\mu y} \mathbb{P}[X_t \geq 2y - x + 2\mu t] = e^{2\mu y} N\left(\frac{x - 2y - \mu t}{\sqrt{t}}\right).$$

En effet :

$$\begin{aligned} \mathbb{P}[M_t \geq y] &= \mathbb{P}[X_t \geq y] + \mathbb{P}[X_t < y; M_t \geq y] \\ &= N\left(\frac{\mu t - y}{\sigma \sqrt{t}}\right) + e^{2\mu y} N\left(\frac{-\mu t - y}{\sigma \sqrt{t}}\right), \end{aligned}$$

et donc

$$\begin{aligned} \mathbb{P}[m_t \geq y] &= \mathbb{P}\left[-\min_{0 \leq s \leq t} (B_s + \mu s) \leq -y\right] \\ &= \mathbb{P}\left[\max_{0 \leq s \leq t} (-B_s - \mu s) \leq -y\right] \\ &= \mathbb{P}\left[\max_{0 \leq s \leq t} (B_s - \mu s) \leq -y\right] \\ &= 1 - \mathbb{P}\left[\max_{0 \leq s \leq t} (B_s - \mu s) \geq -y\right] \\ &= 1 - N\left(\frac{y - \mu t}{\sigma \sqrt{t}}\right) - e^{2\mu y} N\left(\frac{y + \mu t}{\sigma \sqrt{t}}\right) \\ &= N\left(\frac{\mu t - y}{\sigma \sqrt{t}}\right) - e^{2\mu y} N\left(\frac{y + \mu t}{\sigma \sqrt{t}}\right). \end{aligned}$$

L'extension au cas général est alors immédiate en écrivant :

$$\mathbb{P}\left[\inf_{0 \leq s \leq t} (\mu s + \sigma B_s) \geq y\right] = \mathbb{P}\left[\inf_{0 \leq s \leq t} \left(\frac{\mu}{\sigma} s + B_s\right) \geq \frac{y}{\sigma}\right].$$

La démonstration du lemme E.1 repose sur le *Principe de réflexion* dont nous rappelons l'énoncé :

PROPOSITION E.2 (PRINCIPE DE RÉFLEXION) *Parmi les trajectoires browniennes qui atteignent le niveau y avant l'instant t , il y en a autant en dessous de x qu'au dessus de $2y - x$ en t . Autrement dit :*

$$\begin{aligned} \mathbb{P}[B_t \leq x; M_t \geq y] &= \mathbb{P}[B_t \geq 2y - x; M_t \geq y] \\ &= \mathbb{P}[B_t \geq 2y - x], \end{aligned}$$

ce qui se réécrit de la façon suivante :

$$\mathbb{E}[\mathbb{1}_{\{M_t \geq y\}} f(B_t)] = \mathbb{E}[\mathbb{1}_{\{M_t \geq y\}} f(2y - B_t)].$$

Nous allons à présent prouver le lemme. D'après le théorème de Girsanov il existe une probabilité $\bar{\mathbb{P}}$ équivalente à \mathbb{P} définie par sa densité de Radon-Nikodym

$$\frac{d\bar{\mathbb{P}}}{d\mathbb{P}} = e^{-\mu B_T - \frac{\mu^2}{2} T}$$

telle que $X_t = B_t + \mu t$ soit un $\bar{\mathbb{P}}$ -mouvement brownien, et une probabilité $\tilde{\mathbb{P}}$ équivalente à $\bar{\mathbb{P}}$ définie par la densité

$$\frac{d\tilde{\mathbb{P}}}{d\bar{\mathbb{P}}} = e^{-\mu X_T - \frac{\mu^2}{2} T}$$

telle que $W_t = X_t + \mu t$ soit un $\tilde{\mathbb{P}}$ -mouvement brownien. Soit $x \leq y$, on a

$$\begin{aligned} \mathbb{P}[X_t \leq x; M_t \geq y] &= \mathbb{E}[\mathbb{1}_{\{X_t \leq x; M_t \geq y\}}] \\ &= \bar{\mathbb{E}}[e^{\mu B_t + \frac{\mu^2}{2} t} \mathbb{1}_{\{X_t \leq x; M_t \geq y\}}] \\ &= \bar{\mathbb{E}}[e^{\mu X_t - \frac{\mu^2}{2} t} \mathbb{1}_{\{X_t \leq x; M_t \geq y\}}], \end{aligned}$$

et d'après le principe de réflexion :

$$\begin{aligned} \mathbb{P}[X_t \leq x; M_t \geq y] &= \bar{\mathbb{E}}[e^{\mu(2y - X_t) - \frac{\mu^2}{2} t} \mathbb{1}_{\{2y - X_t \leq x; M_t \geq y\}}] \\ &= e^{2\mu y} \bar{\mathbb{E}}[e^{-\mu X_t - \frac{\mu^2}{2} t} \mathbb{1}_{\{X_t \geq 2y - x\}}] \\ &= e^{2\mu y} \bar{\mathbb{P}}[X_t \geq 2y - x] \\ &= e^{2\mu y} \bar{\mathbb{P}}[W_t + \mu t \geq 2y - x + 2\mu t] \\ &= e^{2\mu y} \mathbb{P}[X_t \geq 2y - x + 2\mu t], \end{aligned}$$

car W est un $\tilde{\mathbb{P}}$ -mouvement brownien. Ceci achève la démonstration ■