

# Risk Management :

## 1/5 - Les différentes natures du risque et leur couverture

Matthieu Garcin

ESILV

25 janvier 2019

# Qu'est-ce que le risque ?

En général, c'est une exposition au hasard.

- En finance, on se concentre plutôt sur conséquences négatives.
- Qu'est-ce que le hasard ? Providence, événement fortuit, rencontre de deux chaînes causales à la Cournot, cygne noir ?
- Comment le modéliser ? Par des variables aléatoires (possible depuis 1933 et les travaux de Kolmogorov) ? des systèmes chaotiques ?

Plusieurs perception du risque :

- **Trader** : risque de pertes d'un portefeuille → mais une perte est envisageable.
- **Assureur** : risque de l'occurrence d'un événement (décès, incendie,...) traité de manière statistique → on prévoit une marge d'erreur dans la prime d'assurance pour avoir une probabilité proche de 1 d'avoir un gain si la clientèle est assez représentative de la population et grande.
- **Asset Liability Management** : risque que l'actif rémunère moins que ne coûte le passif → une perte n'est pas envisageable.

# Qu'est-ce que le risque ?

Ainsi, une optimisation de portefeuille en gestion d'actif ressemble à cela :

$$\min_{\omega} \{ \omega' \Sigma \omega \} \text{ sous contrainte que } \sum_i \omega_i r_i \geq r,$$

où les  $\omega$  sont des poids,  $\Sigma$  la matrice de covariance,  $r$  un rendement cible et les  $r_i$  des rendements attendus pour chaque actif  $i$ .

En ALM, le programme d'optimisation peut plutôt ressembler à cela :

$$\max_{\omega} \left\{ \mathbb{P} \left( \sum_i \omega_i R_i \geq r \right) \right\},$$

où  $R_i$  est une variable aléatoire ( $\mathbb{E}[R_i] = r_i$ ). Si les rendements forment un vecteur gaussien de moyenne  $\mu = (r_1, \dots, r_n)'$ , alors le problème revient à maximiser  $\Phi((-r + \omega' \mu) / \omega' \Sigma \omega)$ , où  $\Phi$  est la fonction de répartition gaussienne standard et où l'on voit apparaître la matrice de covariance.

# Qu'est-ce que le risque ?

Plusieurs natures de risque :

- Risque de marché,
- Risque de crédit (inclut risque de contrepartie),
- Risque de liquidité,
- Risque opérationnel : risque de pertes dues à mauvaises procédures (pas de contrôle, modèles mal spécifiés, erreurs, fraudes, pannes,...) ou à des risques externes (inondation, incendie,... mais aussi risque pays/politique/légal).

Aussi d'autres risques que nous n'évoquerons pas en détail et dont la classification est floue : risque météorologique (pour l'assureur mais aussi peut induire un risque marché si impact sur performances d'une entreprise). Autres risques qui peuvent avoir un impact indirect sur le risque de marché : risque de réputation (d'un émetteur), risque systémique.

# Qu'est-ce que le risque ?

D'après Kloman (1990) :

*Risk management is a discipline for living with the possibility that future events may cause adverse effects.*

Donc, le risque peut être utile (pas de rendement sans risque), mais on peut aussi vouloir le transférer, avec des produits dérivés notamment.

# Un peu d'histoire

- **17ème siècle** : utilisation d'options de vente et d'achat et vente à découvert (Amsterdam).
- **1952** : les titres ne sont plus jugés que sur leur espérance de rendement mais sur le couple risque/rendement (théorie du portefeuille, Markowitz). Le risque est l'écart type. Raffinements de la frontière efficiente ont suivi avec des ratios asymétriques au lieu du ratio de Sharpe par exemple etc.
- **1970** : abolition du système de Bretton-Woods, les taux de change ne sont plus fixes et deviennent facteur de risque à couvrir.
- **1973** : nouvelle méthode de valorisation et réplication des calls/puts avec la formule de Black-Scholes-Merton.
- **1973** : ouverture du Chicago Options Exchange (moins de 1000 options le premier jour, plus d'un million par jour en 1995).
- **1995** : faillite de la Barings à cause de straddles (short call et short put de même strike : prime à la stabilité) de Nick Leeson, révèle aux yeux de tous les effets de levier permis par les produits dérivés.
- **2007** : crise des subprimes, rôle de la titrisation dans le transfert du risque.

**Accord de Bâle I** (1988) : ratio de fonds propres réglementaire pour faire face essentiellement au risque de crédit. Les actifs sont pondérés selon leur risque (0% pour cash et treasuries, 20% pour mortgage-backed securities notées AAA, 50% pour les autres hypothèques, par exemple sur les prêts immobiliers, 100% pour les prêts corporate. Les banques internationales doivent détenir un capital d'au moins 8% de ces RWA (risk-weighted assets).

Critique : Le risque de crédit n'est pas pris en compte, seul le nominal du crédit compte.

Amendement de 1996 : incitation à calculer une VaR (value-at-risk, née en 1993).

**Accord de Bâle II** (2004) : 3 piliers (I : mesure du risque et exigence minimale de fonds propres ; II : système de supervision et contrôle des risques ; III : affichage public des données de pertes et des méthodes de gestion de risque).

Exigence de fonds propres : en complément du risque de crédit et de contrepartie apparaît le risque marché et le risque opérationnel. Les fonds propres doivent maintenant être supérieurs à 8% de 85% du risque de crédit +5% du risque de marché +10% du risque opérationnel. Le risque de crédit introduit probabilité de défaut et LGD (perte en cas de défaut).

Critique : dépendance aux agences de notation américaines (S& P, Moody's) ; accord signé par de nombreux pays (dont les USA) mais jamais appliqué par les Américains.



**Accord de Bâle III (2010)** : la crise de 2007 s'expliquerait par une croissance démesurée du hors bilan bancaire (produits dérivés en particulier) et pas assez de capitaux pour faire face au risque de liquidité.

Décisions :

- Mise en place de ratios de liquidité : le Liquidity Coverage Ratio est censé permettre aux banques de faire face à une crise de liquidité pendant un mois. Les réserves de liquidité (tout ce qui est considéré liquide : cash, treasuries refinançable en banque centrale) doivent être supérieures aux fuites dues à une brusque illiquidité. Selon le type d'actif est donc défini un taux de fuite (0% pour le cash, 50% pour les prêts aux particuliers, entre 25 et 75% pour les prêts aux entreprises, 100% pour les prêts interbancaires).
- Ajout d'un coussin au ratio de solvabilité pour prendre en compte le hors bilan.

**Basel IV Agreement (2019)** : ratios are sharper, use of the expected shortfall instead of the VaR, risk measures must include liquidity risk.

# Les conséquences de la régulation

- Pendant la crise des tulipes, interdiction de la vente à découvert a favorisé la croissance d'une bulle (seules les vues positives peuvent s'exprimer).
- En 1933, le Glass-Steagall Act interdisait aux banques commerciales américaines de souscrire à des assurances. Cela a indirectement favorisé l'émergence des banques d'investissement.
- Les accords de Bâle transforment le modèle de nombreuses banques qui originent mais détiennent moins. Le risque existe néanmoins et est transféré à d'autres investisseurs.

- Identifier, mesurer et limiter le risque.
- Enjeux pour mesurer le risque : peu de données (EVT permet d'extrapoler une VaR 99.9% même si historique de seulement 100 observations), mesurer à quelle échelle (aspects multifréquentiels), agréger les risques (aspect multidimensionnel), quelle mesure choisir (VaR, vol, autre?).
- Intégrer le risque (notamment de liquidité) dans des décisions d'investissement.
- Couvrir le risque avec des produits dérivés.

## 1 Risque de marché

- Risque actions
- Risque de taux
- Risque de change

## 2 Risque de crédit

- Valorisation sous risque de crédit
- Modèle structurel de crédit
- Dérivés de crédit

## 3 Risque de liquidité

- Notion de liquidité
- Mesures de liquidité
- Modèles d'impact

## 4 Risque opérationnel

- Aperçu général
- Risque de modèle
  - Surapprentissage et sélection de modèle
  - Méthodes de lissage
  - LASSO

# Couverture du risque de marché par des futures

On peut utiliser des futures pour couvrir un risque. Par exemple :

- si on possède un baril de pétrole à  $S_0 = 19\text{USD}$  en  $t_0$ , on s'expose à une variation du prix en  $t_1$  : on risque de perdre ou gagner  $S_1 - S_0$  et  $t_1$ .
- si on achète un futur sur le baril de maturité  $t_1$  en  $t_0$  au prix  $F_0 = 18.75\text{USD}$ , on récupère le baril au prix  $S_1$  (car le prix futur converge vers le prix spot, sinon arbitrage) en  $t_1$ .
- si on possède le baril et que l'on vend le futur, on gagne *in fine* le prix en  $t_1$ ,  $S_1 - F_1 = 0$  moins le prix en  $t_0$ ,  $-F_0$ , à savoir notre baril nous rapporte en  $t_1$  18.75USD.

Si le spot en  $t_1$  vaut 15USD ou 21USD, on touchera quand même 18.75USD.

Inconvénients :

- Les futures (côtés sur un marché organisé) engendrent des flux quotidiens (appels de marge) et nécessitent donc l'immobilisation d'une trésorerie, alors que les forwards (opération de gré à gré) n'engendrent qu'un flux terminal mais s'exposent davantage au risque de contrepartie.
- La couverture d'un risque n'est pas gratuite, il y a un coût (ici  $19 - 18.75$ ).

# Couverture du risque de marché par des futures

Le prix théorique du future est l'espérance du prix spot actualisé. Dans un monde à taux nul (ce qui ne change pas grand chose en fait car il faut emprunter à ce taux pour investir, ce qui empêche de faire un arbitrage) où le marché est efficient (pas d'opportunité d'arbitrage : l'espérance du spot en  $t_1$  est le spot en  $t_0$ ),  $F_0 = S_0$ . Cela permet de rouler des couvertures. En effet, la liquidité des futures n'est assurée que pour des maturités faibles, par exemple 3 mois. Exemple de couverture du risque de change (gré à gré, car marché organisé ne s'est jamais imposé) :

- en janvier, on achète une obligation 100GBP (obtenus en convertissant 125EUR à notre disposition) et on couvre le risque de change avec un contrat de change à terme en euros pour  $100\text{GBP}=125\text{EUR}$ . Concrètement :
  - on s'engage à donner 100GBP en avril en échange de 125EUR ;
  - on dépense 100GBP pour l'obligation, qui remboursera 100GBP plus un coupon (on ne couvre que le capital) en octobre.
- en avril, on roule la couverture : on reçoit 125EUR par le change à terme que l'on change contre 96.15GBP (nouveau taux de change :  $100\text{GBP}=130\text{EUR}$ ) et l'on donne 100GBP ; on s'engage ensuite à donner 96.15GBP en juillet en échange de 125EUR ;

# Couverture du risque de marché par des futures

- en juillet, on roule la couverture : on reçoit 125EUR par le change à terme que l'on change contre 113.64GBP (nouveau taux de change :  $100\text{GBP}=110\text{EUR}$ ) et l'on donne 96.15GBP ; on s'engage ensuite à donner 113.64GBP en octobre en échange de 125EUR ;
- en octobre, on reçoit 125EUR du change à terme et l'on donne 113.64GBP ; on reçoit aussi 100GBP et un coupon de l'obligation.

En résumé, les flux synthétiques (hors coupon) sont les suivants (attention à la nécessité de trésorerie chaque trimestre) :

- -125EUR en janvier
- -3.85GBP en avril
- +17.49GBP en juillet
- -13.64GBP + 125EUR en octobre

La somme de tous ces flux est bien nulle.

# Couverture du risque de marché par des futures

Des écarts à ce cas idéal existent, c'est le risque de base, la base étant l'écart entre spot et future  $S_t - F_t$  :

- l'actif est parfois différent du sous-jacent couvert (par exemple : future sur indice vs titre) ;
- décalage de dates entre maturité du future et opération voulue, dont la date peut être inconnue en  $t_0$  (on reçoit valeur future avant date idéale ou bien on doit vendre contrat à la date de l'opération, et la base est ajoutée positivement ou négativement à notre P&L) ;
- les décalages de dates peuvent apparaître même si la date de l'opération est connue en  $t_0$  car des contrats de la bonne maturité peuvent ne pas exister.



# Couverture du risque de marché par des futures

Si le but est simplement de minimiser la variance du prix de l'actif couvert (pertinent si risque de base avéré), on peut couvrir seulement une portion  $h$  de l'actif. On note

- $\Delta S$  la variation aléatoire de prix spot pendant la période de couverture ;
- $\Delta F$  la variation aléatoire de prix future pendant la période de couverture ;
- $\sigma_S$  l'écart type de  $\Delta S$  ;
- $\sigma_F$  l'écart type de  $\Delta F$  ;
- $\rho$  la corrélation entre  $\Delta S$  et  $\Delta F$ .

On suppose que l'on peut estimer historiquement  $\sigma_S$ ,  $\sigma_F$  et  $\rho$ . A la fin de la période de couverture, le profit est  $\Delta S - h\Delta F$ , dont la variance est

$$v = \sigma_S^2 + h^2\sigma_F^2 - 2\rho h\sigma_S\sigma_F$$

qui est minimale pour  $h$  valant  $h^*$  avec

$$h^* = \rho \frac{\sigma_S}{\sigma_F}.$$

Alors, en couvrant avec ce ratio, la variance passe de  $\sigma_S^2$  à  $(1 - \rho^2)\sigma_S^2$ .

## 1 Risque de marché

- Risque actions
- Risque de taux
- Risque de change

## 2 Risque de crédit

- Valorisation sous risque de crédit
- Modèle structurel de crédit
- Dérivés de crédit

## 3 Risque de liquidité

- Notion de liquidité
- Mesures de liquidité
- Modèles d'impact

## 4 Risque opérationnel

- Aperçu général
- Risque de modèle
  - Surapprentissage et sélection de modèle
  - Méthodes de lissage
  - LASSO

# VaR ou sensibilité ?

- Les grecques permettent d'exprimer la sensibilité d'un portefeuille à différents facteurs de risque.
- Ces facteurs sont limités pour une option sur action (le cours de l'action, le prix d'exercice, le temps avant maturité, la volatilité de l'action sous-jacente, le taux d'intérêt, les dividendes).
- Mais pour une banque entière, ils peuvent être nombreux. C'est pourquoi on leur préfère souvent la VaR, plus synthétique mais dont la philosophie est différente puisqu'elle remplace la sensibilité aux facteurs de risque par une hypothèse sur le comportement de ceux-ci.
- Attention, sensibilité est une mesure locale et peuvent mal décrire certains phénomènes non-linéaires quand intervient un choc.
- Concept d'antifragilité (Taleb) :
  - portefeuille fragile : évolue négativement avec choc ;
  - portefeuille robuste : insensible au choc ;
  - portefeuille "résilient" : évolue négativement avec choc puis revient à sa valeur d'avant choc ;
  - portefeuille antifragile : évolue positivement avec choc.

# Couverture risque action pour option

Portefeuille de valeur  $V_t$  investi dans  $\delta_t$  actions de prix  $S_t$ , alors que le reste épargné à taux  $r_t$ . Variation de valeur endogène :

$$dV_t = \delta_t dS_t + (V - \delta_t S_t) r_t dt.$$

$V_t$  est une fonction de  $t$  et  $S_t$  :  $V(t, S_t)$ . Si  $S_t$  est un processus d'Ito, c'est-à-dire

$$dS_t = \mu_t dt + \sigma_t dW_t,$$

avec  $W_t$  un mouvement brownien (cas particulier : modèle de Black-Scholes avec  $\mu_t = \mu S_t$  et  $\sigma_t = \sigma S_t$ , qui conduit à  $S_t = S_0 \exp(\sigma B_t + \mu t - \frac{1}{2} \sigma^2 t)$ ), on peut appliquer le lemme d'Ito si  $V$  est  $\mathcal{C}^2(\mathbb{R}_+ \times \mathbb{R}, \mathbb{R})$  :

$$d(V(t, S_t)) = \frac{\partial V}{\partial t}(t, S_t) dt + \frac{\partial V}{\partial x}(t, S_t) dS_t + \frac{1}{2} \frac{\partial^2 V}{\partial x^2}(t, S_t) \sigma_t^2 dt.$$

Ainsi, le résultat important de Black, Scholes et Merton concernant la couverture du risque action est que  $\delta_t = \frac{\partial V}{\partial x}(t, S_t)$  : si on veut couvrir le risque action d'un portefeuille quelconque, il faut vendre une quantité  $\frac{\partial V}{\partial x}(t, S_t)$  d'actions.

# Couverture risque action pour option

Dans le cas particulier d'un call européen avec  $S_t$  un mouvement brownien géométrique et avec une série d'hypothèses (absence d'opportunité d'arbitrage, temps continu, ventes à découvert autorisées, pas de coûts de transaction, ni de dividendes et actifs parfaitement divisibles), on obtient le prix du call comme l'espérance du payoff actualisé sous proba risque neutre :

$$C(S_0, K, r, T, \sigma) = S_0 \mathcal{N}(d_1) - Ke^{-rT} \mathcal{N}(d_2),$$

avec  $d_1 = \frac{1}{\sigma\sqrt{T}} \left[ \ln\left(\frac{S_0}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right) T \right]$  et  $d_2 = d_1 - \sigma\sqrt{T}$ . Les grecques (sensibilité aux facteurs de risque, qui peuvent être couverts par portefeuille répliquant où poids égal à ces grecques investi dans facteurs de risque, si disponibles) sont alors :

- $\delta = \partial C / \partial S = \mathcal{N}(d_1)$ ;
- $\Gamma = \partial^2 C / \partial S^2 = \mathcal{N}'(d_1) / S_0 \sigma \sqrt{T} > 0$ ;
- etc. avec le thêta  $\Theta = -\partial C / \partial T$  (donc dérivée sur vie restante), le vega  $\mathcal{V} = \partial C / \partial \sigma$ , le rho  $\rho = \partial C / \partial r$  et  $\partial C / \partial K$ .

En comparant les parties en  $dt$  de la formule d'Ito appliquée à  $C$  et de la formule de  $dV_t$  du portefeuille répliquant (donc  $V_t = C_t$ ), on a l'EDP :

$$\frac{1}{2}\sigma^2 S_0^2 \Gamma + r S_0 \delta - r C + \Theta = 0.$$

# Couverture d'un call européen

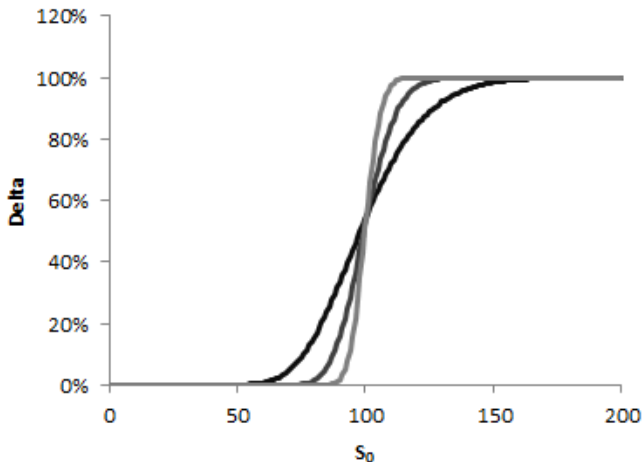


FIGURE – Delta pour  $\sigma\sqrt{T} = 0.05$  (gris clair), puis 0.1, puis 0.2 (noir).  $K = 100$ ,  $r = 0$ .

# Couverture d'un call européen

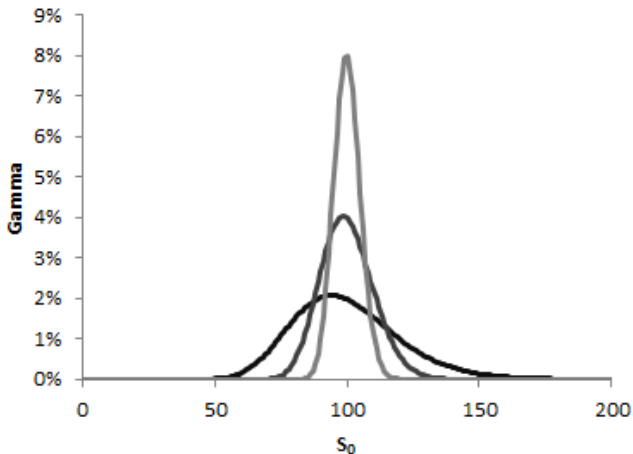


FIGURE – Gamma pour  $\sigma\sqrt{T} = 0.05$  (gris clair), puis 0.1, puis 0.2 (noir).  $K = 100$ ,  $r = 0$ .

# Couverture d'un call européen

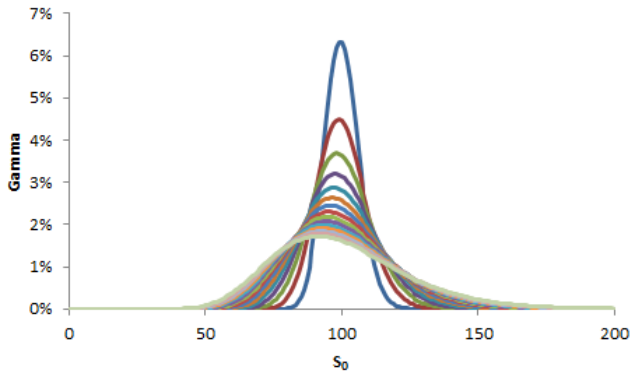


FIGURE – Gamma pour  $\sigma = 0.2$ ,  $K = 100$ ,  $r = 0$ .  $T$  varie de 0.1 (valeur plus élevée à  $S_0 = 100$ ) à 1.5.



# Couverture d'un call européen

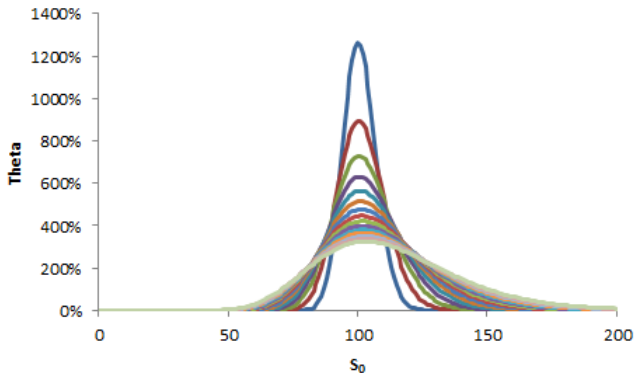


FIGURE – Theta pour  $\sigma = 0.2$ ,  $K = 100$ ,  $r = 0$ .  $T$  varie de 0.1 (valeur plus élevée à  $S_0 = 100$ ) à 1.5. Obtenu par EDP.

## 1 Risque de marché

- Risque actions
- Risque de taux
- Risque de change

## 2 Risque de crédit

- Valorisation sous risque de crédit
- Modèle structurel de crédit
- Dérivés de crédit

## 3 Risque de liquidité

- Notion de liquidité
- Mesures de liquidité
- Modèles d'impact

## 4 Risque opérationnel

- Aperçu général
- Risque de modèle
  - Surapprentissage et sélection de modèle
  - Méthodes de lissage
  - LASSO

# Sensibilité aux variations de taux

Le prix d'une obligation s'écrit

$$P = \sum_{i=1}^n c_i e^{-yt_i},$$

Où les  $c_i$  sont les flux futurs et  $y$  le taux actuariel de l'obligation en composition continue. La **duration (de McCauley)** est la date moyenne à laquelle la valeur de l'obligation est remboursée (temps pondérés par proportion de valeur) :

$$D = \frac{\sum_{i=1}^n t_i c_i e^{-yt_i}}{P}.$$

Si le taux actuariel change de  $\Delta y$ , le prix change de  $\Delta P$ , qui vaut, par approximation linéaire,  $\Delta P \approx \frac{\partial P}{\partial y} \Delta y$ , donc :

$$\frac{\Delta P}{P} \approx -D \Delta y.$$

# Sensibilité aux variations de taux

- Si le taux actuariel est en composition de  $m$  fois par an, le prix est :

$$P = \sum_{i=1}^n \frac{c_i}{(1 + y/m)^{mt_i}}.$$

Donc  $\frac{\partial P}{\partial y}$  devient égal à  $-(1 + y/m)^{-1} \sum_{i=1}^n t_i \frac{c_i}{(1 + y/m)^{mt_i}}$  et :

$$\frac{\Delta P}{P} \approx D^* \Delta y,$$

où  $D^* = D/(1 + y/m)$  est la **duration modifiée**, de limite  $D$  lorsque  $m \rightarrow \infty$ .

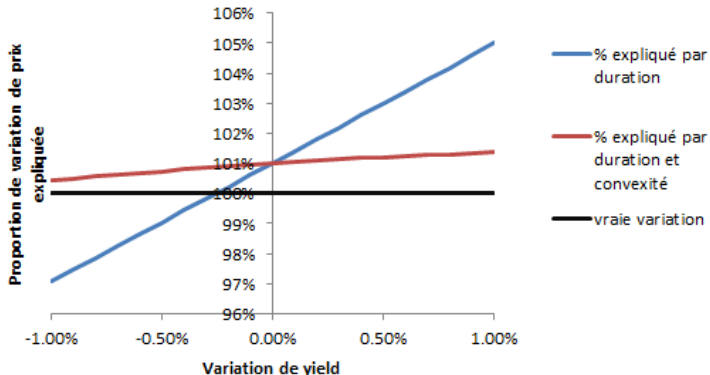
- On peut aussi ajuster la sensibilité avec la **convexité**,  $C$  :

$$C = \frac{1}{P} \frac{\partial^2 P}{\partial y^2} = \frac{\sum_{i=1}^n t_i^2 c_i e^{-yt_i}}{P}.$$

Alors :

$$\frac{\Delta P}{P} \approx -D \Delta y + \frac{1}{2} C (\Delta y)^2.$$

# Sensibilité aux variations de taux



**FIGURE** – Obligation 7 ans, coupon fixe 1% annuel, prix initial 100 (donc yield initial de 1%, duration 6.80 années, convexité 47.03 années<sup>2</sup>). Courbes non définies pour une variation de yield de 0 ; prolongement avec valeur 100% en 0 donne courbes discontinues.

# Sensibilité aux variations de taux

Si l'on décompose le taux actuariel en taux sans risque et *spread* (contenant risque de crédit, de liquidité, etc.), on peut décomposer en variation de taux et en variation de *spread*, définissant donc une duration (et une convexité) taux et une duration (et convexité) *spread*. Cela conduit à quelques remarques :

- On peut ne couvrir que le risque de taux par exemple et rester exposé au risque de *spread*.
- La couverture du risque de taux se fait en vendant des produits de taux *pur* (e.g. Schatz [2 ans], Bobl [5 ans], Bund [10 ans]) de même duration et convexité.
- Une telle couverture ne couvre en réalité que contre des translations de la courbe des taux : en effet, on écrit, en  $t_i$ ,  $r_i + s$  au lieu de la constante  $y$ , donc la duration taux s'entend comme

$$-\frac{1}{P} \frac{\partial}{\partial \varepsilon} \left( \sum_{i=1}^n c_i (1 + (r_i + \varepsilon) + s)^{-t_i} \right) \Bigg|_{\varepsilon=0}.$$

Une solution est le **gap management** qui consiste à se couvrir contre des variations de taux différentes sur plusieurs segments prédéfinis de maturités (avec des swaps, des futures, etc.).

# Sensibilité aux variations de taux

- ...
- Pour une obligation à taux variable,  $c_i = f_i$ , taux forward, égal par absence d'opportunité d'arbitrage à

$$f_i = -1 + (1 + r_i)^i / (1 + r_{i-1})^{i-1}.$$

On montre qu'une translation des taux a un impact identique sur  $f_i$ . Ainsi, une translation des taux va avoir un impact identique sur le rendement : la sensibilité du prix à une translation de taux est nulle. Donc la duration taux est nulle.

- La variation relative des *spreads* semble *plus invariante* que sa variation absolue, ce qui a conduit à introduire la DTS, avec l'approximation (si risque de taux est couvert par ailleurs) :  $(\Delta P/P) = -DTS(\Delta s/s)$ .

## 1 Risque de marché

- Risque actions
- Risque de taux
- Risque de change

## 2 Risque de crédit

- Valorisation sous risque de crédit
- Modèle structurel de crédit
- Dérivés de crédit

## 3 Risque de liquidité

- Notion de liquidité
- Mesures de liquidité
- Modèles d'impact

## 4 Risque opérationnel

- Aperçu général
- Risque de modèle
  - Surapprentissage et sélection de modèle
  - Méthodes de lissage
  - LASSO



# Couverture du risque de change

Couverture possible avec swap de devises :

- En  $t = 0$ , on donne 100 EUR et on reçoit  $100S_0$  GBP où  $S_0$  est le taux de change en  $t = 0$ .
- En  $t = 1$ , on paye le taux domestique britannique, disons  $2\% \times 100S_0$  GBP et on reçoit le taux européen, disons  $1\% \times 100$  EUR.
- En  $t = 2$  on paye le taux britannique (différent de  $t = 1$  si taux variable) et le principal de  $100S_0$  GBP contre le taux européen et le principal de 100 EUR ;

Si on a acheté une obligation britannique payant un coupon variable en GBP et que l'on entre dans le swap décrit, on a transformé l'obligation en un titre en EUR payant un coupon variable indexé sur le taux européen.

Les swaps sont des opérations de gré à gré, donc il y a un risque de contrepartie qui peut entrer en compte dans la valorisation. Sans cela, le prix est nul en  $t = 0$  et vaut en  $t$  la différence du prix des deux obligations constituant le swap dont l'une est convertie dans la devise de l'autre au taux de change spot  $S_t$ .

## 1 Risque de marché

- Risque actions
- Risque de taux
- Risque de change

## 2 Risque de crédit

- Valorisation sous risque de crédit
- Modèle structurel de crédit
- Dérivés de crédit

## 3 Risque de liquidité

- Notion de liquidité
- Mesures de liquidité
- Modèles d'impact

## 4 Risque opérationnel

- Aperçu général
- Risque de modèle
  - Surapprentissage et sélection de modèle
  - Méthodes de lissage
  - LASSO

# Notes de crédit

- Indépendance des agences de notation ?
- La probabilité de défaut historique est plus faible que celle qui est valorisée et donc observée dans le prix des obligations (prend en compte d'autres choses que le crédit comme la liquidité ? méfiance des investisseurs pour les notations ?), le ratio croissant de 1.2 pour du B à 16.8 pour du Aaa (source : Hull).
- IG de Aaa (ou AAA) à Baa (ou BBB) et HY en-dessous.
- CreditMetrics : migrations de ratings.

# Notes de crédit

**Table 8.2.** Probabilities of migrating from one rating quality to another within one year.  
Source: Standard & Poor's CreditWeek (15 April 1996).

Initial rating	Rating at year-end (%)							
	AAA	AA	A	BBB	BB	B	CCC	Default
AAA	90.81	8.33	0.68	0.06	0.12	0.00	0.00	0.00
AA	0.70	90.65	7.79	0.64	0.06	0.14	0.02	0.00
A	0.09	2.27	91.05	5.52	0.74	0.26	0.01	0.06
BBB	0.02	0.33	5.95	86.93	5.30	1.17	1.12	0.18
BB	0.03	0.14	0.67	7.73	80.53	8.84	1.00	1.06
B	0.00	0.11	0.24	0.43	6.48	83.46	4.07	5.20
CCC	0.22	0.00	0.22	1.30	2.38	11.24	64.86	19.79

FIGURE – Source : McNeil, Frey, Embrechts.

# Notes de crédit

- Tableaux de probabilités historique par note et maturité. Probabilité cumulée suit souvent loi log-logistique :  $F(x) = \frac{x^\beta}{\alpha^\beta + x^\beta}$ , avec  $x > 0$ ,  $\alpha > 0$  et  $\beta > 0$ .
- Probabilité de défaut non-conditionnelle entre  $s$  et  $t$  :  $F(t) - F(s)$ .
- Probabilité de défaut conditionnelle entre  $s$  et  $t$ , loi de Bayes :  $\text{proba}(\text{défaut en } t \text{ et pas en } s) / \text{proba}(\text{pas défaut en } s)$ , soit  $(F(t) - F(s)) / (1 - F(s))$ , supérieure à la proba non-conditionnelle.
- Intensité de défaut  $\lambda$  est la proba conditionnelle rapportée à  $t - s$ . Pour intervalle de temps faible, on obtient  $\lambda(t)dt = dF(t)/(1 - F(t))$ , donc

$$1 - F(t) = \exp\left(-\int_0^t \lambda(\tau)d\tau\right).$$

- Pour un titre de  $T$  années, avec  $F(t) = t/T$ , le défaut est certain. La proba non-conditionnelle attribue une probabilité de  $1/T$  chaque année. La proba conditionnelle est croissante et vaut pour l'année  $t$ ,  $(T + 1 - t)^{-1}$ , qui vaut 1 en  $T$ . Un proba conditionnelle annuelle de  $1/T$  équivaut à  $F(1) = 1/T$ ,  $F(2) = 2/T - 1/T^2$  et plus généralement à  $F(t) = 1/T + F(t - 1)(1 - 1/T)$ .

# Notes de crédit

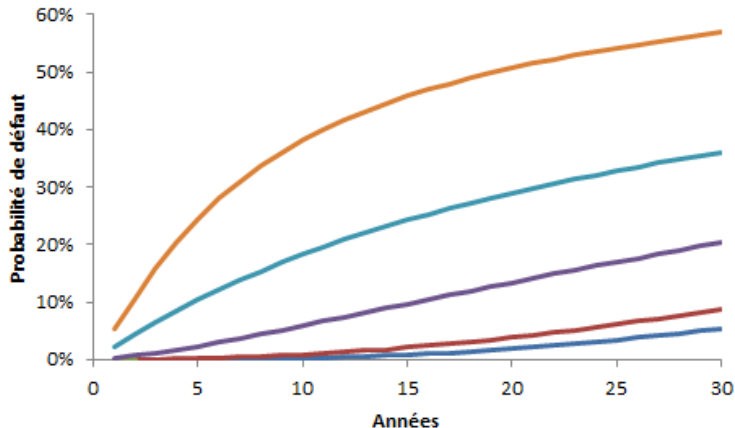


FIGURE – Probabilité de défaut cumulée pour notes AAA, AA, A, BBB, BB, B.

# Notes de crédit

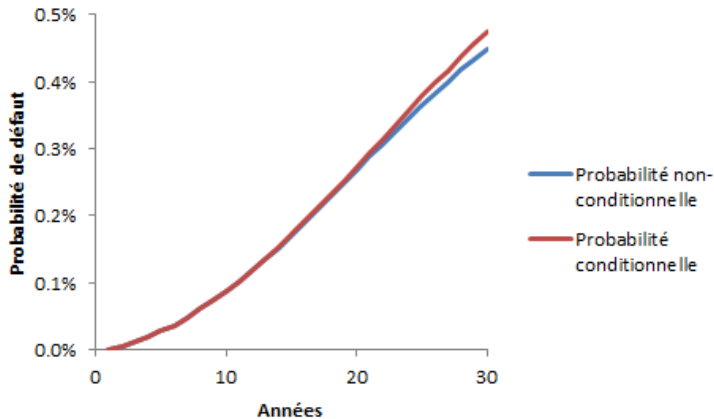


FIGURE – Densité de probabilité de défaut AAA.

# Notes de crédit

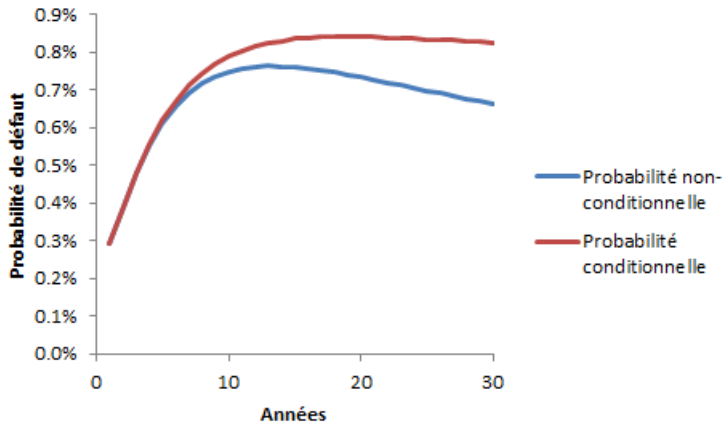


FIGURE – Densité de probabilité de défaut BBB.



# Notes de crédit

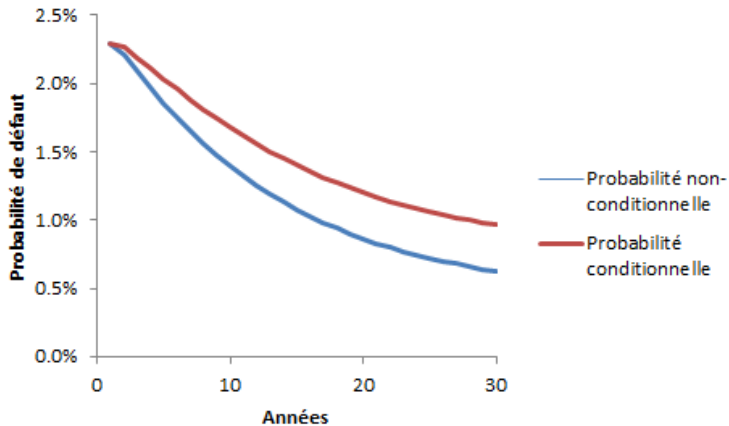


FIGURE – Densité de probabilité de défaut BB.

# Recouvrement

- Actions en justice pour recouvrement.
- Délai de recouvrement.
- Impact de la séniorité (1982-2003) :
  - senior garantie : 51.6%
  - senior non garantie : 36.1%
  - senior subordonnée : 32.5%
  - subordonnée : 31.1%
  - junior subordonnée : 24.5%
- Impact note de crédit (1982-2003) :  $R = 50.3\% - 6.3 \times P$ , où  $R$  est le taux de recouvrement moyen et  $P$  la proba de défaut moyenne.

# Valorisation d'une obligation

- Deux versions : risque de crédit au numérateur ou au dénominateur :

$$price_t = \sum_{i=1}^n \frac{\mathbb{E}_t[cashflow_{t_i}]}{(1 + r_{t_i})^{t_i - t}}$$

ou

$$price_t = \sum_{i=1}^n \frac{cashflow_{t_i}}{(1 + r_{t_i} + s)^{t_i - t}}$$

- L'intensité de défaut moyenne est égale à  $s/(1 - R)$ , où  $s$  est le spread et  $R$  le taux de recouvrement.

## 1 Risque de marché

- Risque actions
- Risque de taux
- Risque de change

## 2 Risque de crédit

- Valorisation sous risque de crédit
- Modèle structurel de crédit
- Dérivés de crédit

## 3 Risque de liquidité

- Notion de liquidité
- Mesures de liquidité
- Modèles d'impact

## 4 Risque opérationnel

- Aperçu général
- Risque de modèle
  - Surapprentissage et sélection de modèle
  - Méthodes de lissage
  - LASSO

# Modèle de Merton

- Entreprise de valeur  $V_t$  stochastique, financée par dette (zéro-coupon de valeur faciale  $B$  et de maturité  $T$ ) de valeur  $B_t$  et actions de valeur  $S_t$ .
- Sans taxes ni coûts de transactions :  $V_t = S_t + B_t$  pour tout  $t \leq T$ .
- Hypothèse pas dividendes ni émission nouvelle dette.
- Défaut si impossibilité de payer dette en  $T$ . Si  $V_T > B$ , pas de défaut et  $S_T = V_T - B$ . Sinon défaut, et  $S_T = 0$  et  $B_T = V_T$ . En résumé, action est call européen, alors que dette est valeur faciale moins put européen. :

$$S_T = (V_T - B)^+,$$

$$B_T = B - (B - V_T)^+.$$

- Conflit d'intérêts entre actionnaires (veulent projets risqués) et créanciers (veulent peu de volatilité).
- Modèle peu réaliste (une seule date de défaut possible, une seule obligation) mais bon point de départ.

# Modèle de Merton

In the Merton model it is assumed that under the real-world or physical probability measure  $P$  the process  $(V_t)$  follows a diffusion model (known as Black–Scholes model or geometric Brownian motion) of the form

$$dV_t = \mu_V V_t dt + \sigma_V V_t dW_t \quad (8.3)$$

for constants  $\mu_V \in \mathbb{R}$ ,  $\sigma_V > 0$ , and a standard Brownian motion  $(W_t)$ . Equation (8.3) implies that  $V_T = V_0 \exp((\mu_V - \frac{1}{2}\sigma_V^2)T + \sigma_V W_T)$ , and, in particular, that  $\ln V_T \sim N(\ln V_0 + (\mu_V - \frac{1}{2}\sigma_V^2)T, \sigma_V^2 T)$ . Under the dynamics (8.3) the default probability of our firm is readily computed. We have

$$P(V_T \leq B) = P(\ln V_T \leq \ln B) = \Phi\left(\frac{\ln(B/V_0) - (\mu_V - \frac{1}{2}\sigma_V^2)T}{\sigma_V \sqrt{T}}\right). \quad (8.4)$$

It is immediately seen from (8.4) that the default probability is increasing in  $B$ , decreasing in  $V_0$  and  $\mu_V$  and, for  $V_0 > B$ , increasing in  $\sigma_V$ , which is all perfectly in line with economic intuition.

## 1 Risque de marché

- Risque actions
- Risque de taux
- Risque de change

## 2 Risque de crédit

- Valorisation sous risque de crédit
- Modèle structurel de crédit
- Dérivés de crédit

## 3 Risque de liquidité

- Notion de liquidité
- Mesures de liquidité
- Modèles d'impact

## 4 Risque opérationnel

- Aperçu général
- Risque de modèle
  - Surapprentissage et sélection de modèle
  - Méthodes de lissage
  - LASSO

# Swap de défaut

- CDS permet de se couvrir contre le risque défaut (au risque de contrepartie près).
- Plus généralement, le nantissement (collatéralisation) actualisé par appels de marge permet de se prémunir contre ce risque.
- Deux possibilité pour le CDS : sous-jacent livré en cas de défaut (l'acheteur du CDS reçoit le nominal en cash), dénouement en cash (si prix milieu de fourchette à 35 EUR pour 100 EUR de nominal, l'acheteur du CDS reçoit 65 EUR). Le paiement a lieu au moment du défaut.
- L'acheteur du CDS paye au vendeur des montants convenus à intervalles réguliers (par exemple 90 bp/an).



# Valorisation d'un swap de défaut

Supposons une intensité de défaut constante de 2% pour une obligation de maturité 5 ans (source exemple : Hull).

Maturité	Proba non-cond.	Proba de survie
1	0.0200	0.9800
2	0.0196	0.9604
3	0.0192	0.9412
4	0.0188	0.9224
5	0.0184	0.9039

Supposons que le spread du CDS soit  $s$ . Le paiement à l'année 3 intervient seulement s'il n'y a pas de défaut, donc avec probabilité 0.9412. Pour un taux sans risque (LIBOR) de 5%, la valeur actuelle du paiement est

$$0.9412 \times s \times e^{-0.05 \times 3} = 0.8101s.$$

Alors, la valeur actualisée de l'ensemble des paiements est  $4.0704s$ .

# Valorisation d'un swap de défaut

En contrepartie, en supposant que le défaut intervient toujours en milieu d'année et avec un taux de recouvrement de 40%, l'acheteur de la protection reçoit un dédommagement avec probabilité 0.0192 en année 3. Sa valeur actuelle est

$$0.0192 \times 0.6 \times e^{-0.05 \times 2.5} = 0.0102.$$

Alors, la valeur actualisée du paiement total espéré est 0.0511. Il y a aussi un paiement résiduel d'une demi-année en cas de défaut, dont la valeur actualisée pour l'année 3 est :

$$0.0192 \times 0.5 \times s \times e^{-0.05 \times 2.5} = 0.0085s.$$

Pour l'ensemble du contrat, cela représente 0.0426s. Le spread du CDS de milieu de fourchette vérifie donc :

$$4.0704s + 0.0426s = 0.0511,$$

ce qui donne  $s = 0.0124$ , soit 124bp.

# Valorisation d'un swap de défaut

- La valorisation peut être plus précise en prenant en compte une occurrence du défaut à plusieurs date en non juste au milieu de l'année.
- Probabilité de défaut implicite.
- Cas d'une intensité non constante (obtenue par courbes de spreads CDS).
- Estimation du taux de recouvrement ( $R$ ) a place centrale dans valo.
- Swap digitaux comme CDS mais montant payé en cas de défaut (par exemple 100% du nominal) est fixé dès le départ et ne dépend donc pas de  $R$ .
- Taux de recouvrement implicite peut ainsi être estimé en confrontant swaps digitaux et CDS pour un même sous-jacent.

## 1 Risque de marché

- Risque actions
- Risque de taux
- Risque de change

## 2 Risque de crédit

- Valorisation sous risque de crédit
- Modèle structurel de crédit
- Dérivés de crédit

## 3 Risque de liquidité

- Notion de liquidité
- Mesures de liquidité
- Modèles d'impact

## 4 Risque opérationnel

- Aperçu général
- Risque de modèle
  - Surapprentissage et sélection de modèle
  - Méthodes de lissage
  - LASSO

# Liquidity : confronting many viewpoints

L'écart entre le cours vendeur et le coût acheteur peut être divisé en deux du point de vue du **faiseur de marché** :

- **Le coût de la liquidité** : l'écart réel qui rémunère le teneur de marché au titre de son rôle de garant de la liquidité du marché et qui correspond aux coûts de la technologie et du stock encourus par le faiseur de marché.
- **Le risque encouru par le faiseur de marché** : l'asymétrie d'informations qui rémunère les faiseurs de marché pour les pertes qu'ils pourraient subir lorsqu'ils font un pari sans en connaître les tenants et les aboutissants : comme ils ne peuvent pas distinguer les paris informés des non-informés, ils peuvent être la "partie opposée" d'un ordre informé à cause de leur rôle de garant de la liquidité de marché.

# Liquidity : confronting many viewpoints

Selon Amihud et Mendelson (2006), il y aurait trois composantes dans les coûts de liquidité :

- **Coûts de négociation directs** : des coûts de transaction déterministes qui englobent les commissions de courtage, les taxes sur les transactions et les frais de bourse.
- **L'impact des ordres sur le prix du marché** : d'autant plus important que l'ordre est conséquent, car un gros ordre va impliquer plusieurs contreparties et donc aller loin dans le carnet d'ordre.
- **Le coût induit par un retard d'exécution** : pour limiter l'impact d'un gros ordre, on peut passer des ordres en bloc, ce qui fait que certains blocs sont échangés avec un certain retard, ce qui expose d'avantage à l'éventualité de mouvements de marché.

# Liquidity : confronting many viewpoints

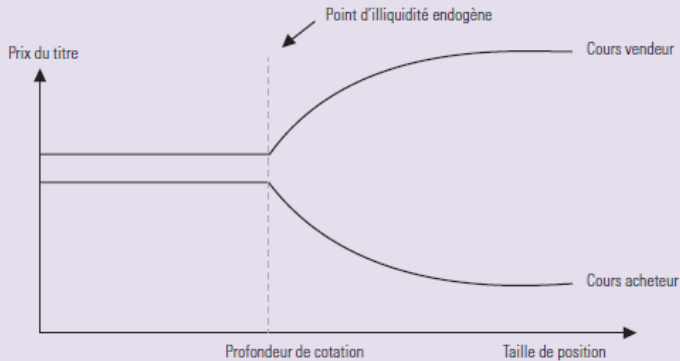
Selon Harris (1990), la notion de marché liquide est définie comme un marché sur lequel on peut acheter et vendre à tout instant une grande quantité, rapidement et à bas coût. Elle est mesurée par quatre piliers :

- **La largeur** : coût encouru par un achat/vente simultané, ce qui correspond à la somme des coûts directs et de l'impact chez Amihud et Mendelson.
- **La profondeur** : la quantité de titres pouvant être négociés sans impact (seul coût : spread bid/ask) car absorbée sans difficulté par le faiseur de marché (cf graphique suivant).
- **L'immédiateté**, qui détermine la rapidité avec laquelle les positions peuvent être négociées et qui correspond au délai entre le passage de l'ordre et son règlement.
- **La résistance**, qui désigne la capacité du marché à absorber les chocs aléatoires, comme des ordres non informatifs.

# Liquidity : confronting many viewpoints

*Au-dessus d'une taille limite, l'illiquidité devient endogène et son poids supérieur*

*Source: Bangia et al. (1999)*

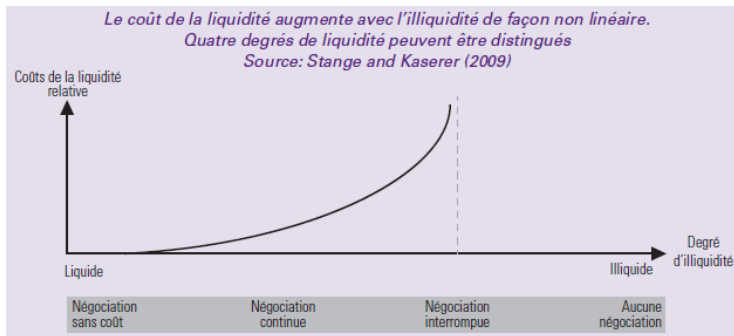




# Liquidity : confronting many viewpoints

Selon Stange et Kaserer (2009), les coûts de la liquidité peuvent être définis de manière continue et ainsi former quatre degrés de liquidité :

- **La négociation sans coûts** (ex : cash).
- **La négociation continue** : ordres exécutés à un coût donné.
- **La négociation interrompue** : ordres exécutés ponctuellement.
- **Aucune négociation** : le marché est entièrement illiquide, les prix ne sont pas disponibles et sont récupérés par des techniques adaptées



# Quelques remarques

- Le risque de liquidité n'est pas diversifiable : pas de couverture possible, pas de produits dérivés de la liquidité, etc.
- Quel est le prix d'un actif lorsqu'il n'est pas échangé ? Peut-on se utiliser une comparaison avec un univers liquide ? (exemple de courbes de taux pour les obligations d'entreprise)
- Prime de liquidité : lien positif entre rendement et illiquidité (Amihud et Mendelson) mais concave (les rendements augmentent moins pour les actifs très illiquides).
- Pour les obligations d'entreprise, l'illiquidité augmente quand la notation de crédit se dégrade (prime de risque de 0.6% en IG, 1.5% en HY).

## 1 Risque de marché

- Risque actions
- Risque de taux
- Risque de change

## 2 Risque de crédit

- Valorisation sous risque de crédit
- Modèle structurel de crédit
- Dérivés de crédit

## 3 Risque de liquidité

- Notion de liquidité
- Mesures de liquidité
- Modèles d'impact

## 4 Risque opérationnel

- Aperçu général
- Risque de modèle
  - Surapprentissage et sélection de modèle
  - Méthodes de lissage
  - LASSO

# Mesure de liquidité sur les marchés actions

- La liquidité sur ce marché est fournie essentiellement par le trading haute-fréquence. La comparaison entre trading volume et quantité à échanger donne une bonne idée de la liquidité.
- Dans le modèle de Kyle (1985), l'illiquidité est mesurée comme le ratio moyen du rendement quotidien en valeur absolue sur le volume échangé ce jour-là (sorte d'impact moyen).
- Autre mesure, le ratio de rotation, qui est le rapport du volume échangé pour un titre sur le nombre de titres restants pour la même entreprise. Ratio négativement corrélé avec coûts d'illiquidité selon Amihud et Mendelson.
- Autres facteurs que l'on peut considérer : la capitalisation, le volume échangé, etc.

# Mesure de liquidité sur les marchés obligataires

- Les obligations ont une maturité : la liquidité pour un titre de maturité résiduelle 10 ans n'est pas la même que pour 1 an.
- D'autres caractéristiques à prendre en compte : taille de l'émission (500 M€ pour une entreprise vs 20 G€ pour une émission souveraine, par exemple), rendement à l'échéance, duration, spread, subordination...
- Une mesure de la liquidité basée sur le volume échangé peut induire en erreur car une obligation négociée n'est pas nécessairement liquide (ex. : vente forcée et titres IG devenus HY). Inversement, les obligations non négociées ne sont pas nécessairement illiquides.
- Mesurer la liquidité si les écarts entre cours vendeur et cours acheteurs sont observés, c'est facile, sinon c'est délicat... Modèle factoriel en fonction des facteurs cités ci-dessus. Fin 2008 début 2009, les perpétuelles étaient valorisées vers 25%, il n'était pas rare de voir 5% d'écart entre prix acheteur et prix vendeur.

# VaR ajustée de la liquidité

La L-VaR peut être obtenue de plusieurs façons :

- Si la liquidité entraîne des délais dans la liquidation d'un portefeuille, on peut considérer que la L-VaR à un jour est égale à la VaR classique à deux jours, par exemple.
- Si la liquidité est comprise comme un impact, il faut l'intégrer au rendement qui sert au calcul de la VaR : cette L-VaR sera donc adaptée au montant détenu dans le portefeuille pour le titre étudié.
- Approximation à la Bangia, Diebold, Schuermann, Stroughair : intégration de la moitié du spread bid/ask,  $S_t$ , ce qui correspond à l'hypothèse de liquidation. Le prix en  $t + 1$  est  $P_{mid,t+1} = P_{mid,t} \exp(r_{t+1}) - S_{t+1}/2$ , où  $r_{t+1}$ , le rendement, et  $S_{t+1}$  sont aléatoires. Si elles sont gaussiennes, parfaitement corrélées et que le rendement est centré, la L-VaR de niveau  $\alpha$  (e.g. 90%) s'écrit comme somme d'une composante standard et d'une composante liquidité :

$$P_{mid,t}(1 - \exp(Q_{1-\alpha}\sigma_r)) + (\mu_S + Q_\alpha\sigma_S)/2,$$

avec  $Q_\alpha$  le quantile de probabilité  $\alpha$  d'une loi gaussienne standard.

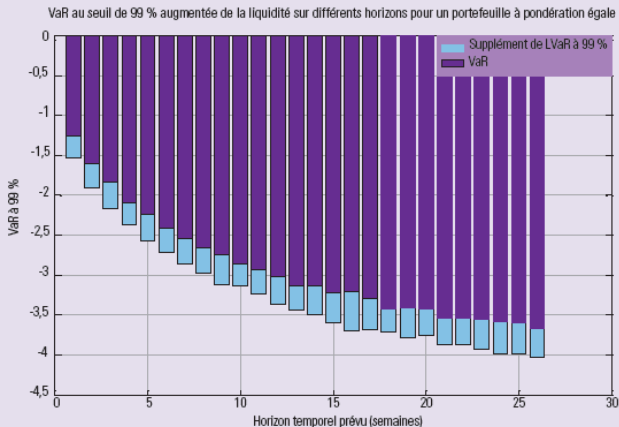
- Si la stratégie est de maintenir les titres en portefeuille, cet ajustement n'a pas de raison d'être.

# L-VaR de Bangia *et alii*

## Graphique 14 : la LVaR à 99 % pour différents horizons

Les bâtons rouges correspondent à l'écart entre la LVaR et la VaR traditionnelle (en bleu)

Source: Natixis Asset Management



## 1 Risque de marché

- Risque actions
- Risque de taux
- Risque de change

## 2 Risque de crédit

- Valorisation sous risque de crédit
- Modèle structurel de crédit
- Dérivés de crédit

## 3 Risque de liquidité

- Notion de liquidité
- Mesures de liquidité
- Modèles d'impact

## 4 Risque opérationnel

- Aperçu général
- Risque de modèle
  - Surapprentissage et sélection de modèle
  - Méthodes de lissage
  - LASSO



# Modèle d'impact permanent : Almgren et Chriss (1/5)

**Almgren et Chriss** : cadre de liquidation optimale (métaordre : prendre du temps pour limiter impact vs se dépêcher pour limiter le risque de marché).

- Quantité  $X$  à liquider par tranches  $n_1, n_2, \dots$ , avec  $x_k = X - \sum_{j=1}^k n_j$ .
- Trading à temps discret, pas de temps  $\tau$ .
- Dynamique du prix selon marche aléatoire arithmétique :  
 $S_k = S_{k-1} + \sigma\sqrt{\tau}\varepsilon_k - \tau g(n_k/\tau)$ , où  $g$  est la fonction d'impact permanent, dépendant de la vitesse de liquidation  $n_k/\tau$  entre les instants  $t_{k-1}$  et  $t_k$ .
- Font hypothèse que prix observé est différent de  $S$  (dynamique sous-jacente) à cause d'impact transitoire (trader épuise le carnet d'ordre entre  $t_{k-1}$  et  $t_k$ , mais il se reconstitue instantanément en  $t_k$ ) :  $\bar{S}_k = S_{k-1} - h(n_k/\tau)$ .
- Prix obtenu pour la liquidation (prix initial sans friction + pertes d'illiquidité) :

$$\sum_{k=1}^N n_k \bar{S}_k = XS_0 + \sum_{k=1}^N \left[ \left( \sigma\sqrt{\tau}\varepsilon_k - \tau g\left(\frac{n_k}{\tau}\right) \right) x_k - n_k h\left(\frac{n_k}{\tau}\right) \right].$$

# Modèle d'impact permanent : Almgren et Chriss (2/5)

- Problème minimiser pertes attendues et variance perte avec  $\varepsilon_k$  les seules variables aléatoires :  $E(x) + \lambda V(x)$ , avec :

$$\begin{cases} E(x) &= \sum_{k=1}^N \left[ \tau x_k g\left(\frac{n_k}{\tau}\right) + n_k h\left(\frac{n_k}{\tau}\right) \right] \\ V(x) &= \sigma^2 \sum_{k=1}^N \tau x_k^2. \end{cases}$$

- Hypothèse d'Almgren et Chriss : impact linéaire :

- $g(v) = \gamma v$ , qui conduit à un coût total

$$\begin{aligned} \sum_{k=1}^N \tau x_k g\left(\frac{n_k}{\tau}\right) &= \gamma \sum_{k=1}^N x_k (x_{k-1} - x_k) \\ &= \frac{1}{2} \gamma \sum_{k=1}^N x_{k-1}^2 - x_k^2 - (x_k - x_{k-1})^2 \\ &= \frac{1}{2} \gamma X^2 - \frac{1}{2} \gamma \sum_{k=1}^N n_k^2. \end{aligned}$$

- $h(n_k/\tau) = \xi \operatorname{sgn}(n_k) + \eta n_k/\tau$ , avec  $\xi$  la moitié du spread bid-ask. Ce modèle pour  $h$  est souvent appelé modèle de coût quadratique car le coût associé est  $n_k h(n_k/\tau) = \xi |n_k| + \eta n_k^2/\tau$ .

# Modèle d'impact permanent : Almgren et Chriss (3/5)

- Impact minimum en prenant vitesse constante :  $n_k = X/N$ , mais variance peut être élevée.
- Variance minimale en liquidant tout au premier instant :  $n_1 = X$  et  $n_2 = \dots = n_N = 0$ . Alors,  $V = 0$  et  $E = \xi X + \eta X^2/\tau$ .
- Entre les deux : une frontière efficiente pour plusieurs niveaux d'aversion au risque correspondant au lagrangien  $\lambda$ . Résolution par Euler-Lagrange :

$$x_k = \frac{\sinh(K(T - (k - \frac{1}{2}\tau)))}{\sinh(KT)} X,$$

où  $K \xrightarrow{\tau \rightarrow 0} \sqrt{\frac{\lambda \sigma^2}{\eta}} + \mathcal{O}(\tau)$ .

# Modèle d'impact permanent : Almgren et Chriss (4/5)

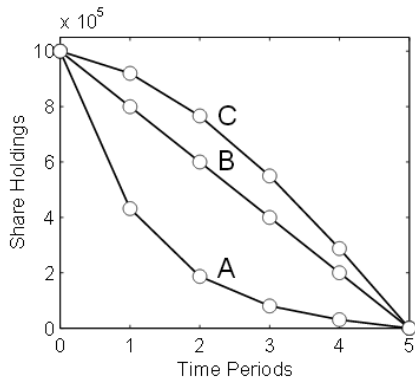


Figure 2: *Optimal trajectories.* The trajectories corresponding to the points shown in Figure 1. (A)  $\lambda = 2 \times 10^{-6}$ , (B)  $\lambda = 0$ , (C)  $\lambda = -2 \times 10^{-7}$ .

FIGURE –  $x$  pour diverses aversions au risque. Source : Almgren et Chriss.

# Modèle d'impact permanent : Almgren et Chriss (5/5)

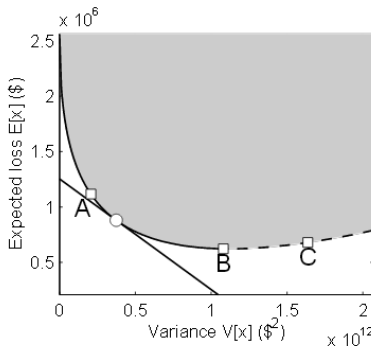


Figure 1: *The efficient frontier.* The parameters are as in Table 1. The shaded region is the set of variances and expectations attainable by some time-dependent strategy. The solid curve is the efficient frontier; the dashed curve is strategies that have higher variance for the same expected costs. Point B is the “naïve” strategy, minimizing expected cost without regard to variance. The straight line illustrates selection of a specific optimal strategy for  $\lambda = 10^{-6}$ . Points A,B,C are strategies illustrated in Figure 2.

FIGURE – Source : Almgren et Chriss.

# Modèle d'impact transitoire

Le prix est la somme des impacts passés (Bouchaud) :

$$p_t = p_{-\infty} + \sum_{s=-\infty}^{t-1} G(t-s) \varepsilon_s S_s V_s^r,$$

où :

- $S$  est le *spread bid-ask* ;
- $\varepsilon$  vaut -1 ou 1 selon que la transaction soit un achat (prix = ask) ou une vente (prix = bid) sur le marché ;
- $V$  est le volume de la transaction, avec  $r$  proche de zéro pour avoir une fonction concave du volume ;
- $G$  est une fonction valant 0 sur  $\mathbb{R}^-$ , que l'on peut interpréter comme l'impact d'un unique ordre. Une étude statistique sur la corrélation de long terme, par exemple, peut permettre de fixer cette fonction, sous la forme, par exemple, d'une fonction puissance décroissante.

# Liquidité et crise

- Dans une bulle, liquidité de marché et prix croissent concomitamment.
- Ce mécanisme est instable, avec forte demande et faible offre.
- Quand la liquidité de marché cesse de croître, on peut entrer dans une spirale de liquidité, impliquant les deux types de ce que l'on appelle *liquidité* :
  - liquidité de marché ;
  - financement.
- La contagion à tout un marché peut se faire par le risque de financement.

# Spirale de liquidité

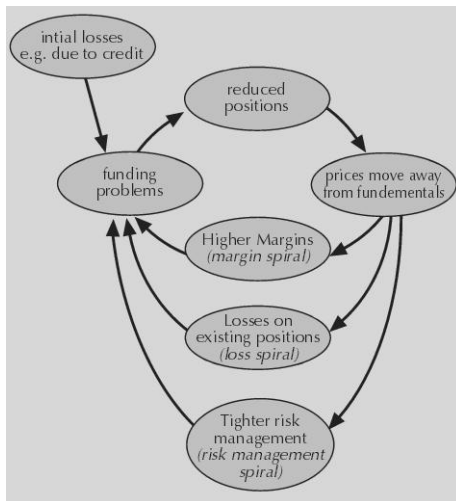


FIGURE – Spirale de liquidité. Source : Pedersen.



# Spirale de liquidité

vol initiale 1%, un choc en date 100 la fait monter à 1.3%. Puis spirale liquidité induite par 10 agents qui ont intervalle admissible de vol et une cible, dans cet intervalle, qu'ils visent lorsqu'ils sont contraints de réallouer. Impact prix : modèle de Bouchaud.

Paramètres :

- $V$ , facteur d'échelle dans modèle d'impact du prix ;
- $\gamma$ , dans la fonction temporelle du modèle d'impact du prix  $G : t \geq 0 \mapsto 1/t^\gamma$  ;
- $r$ , exposant du volume dans modèle d'impact de prix ;
- $\lambda$ , paramètre de pondération exponentielle pour la vol (plus  $\lambda$  est grand, moins les valeurs ancienne ont de poids).

# Spirale de liquidité

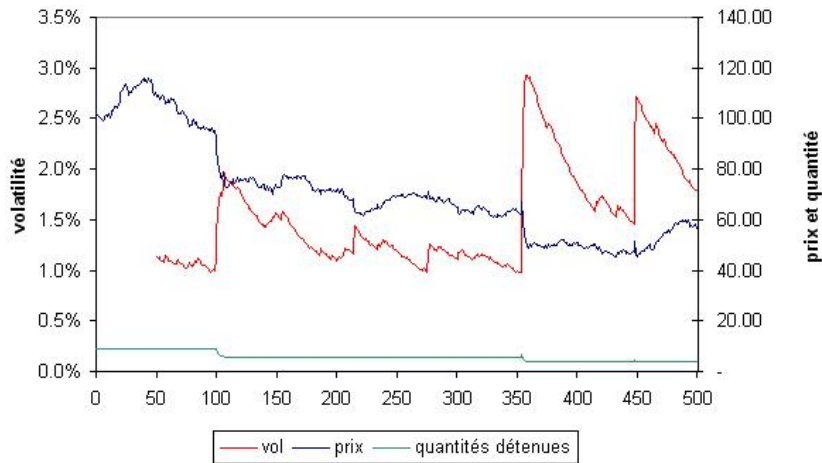


FIGURE –  $V = 5$ ,  $\gamma = 0.2$ ,  $r = 0.25$ ,  $\lambda = 0.03$  (demie vie en 24 jours).

# Spirale de liquidité

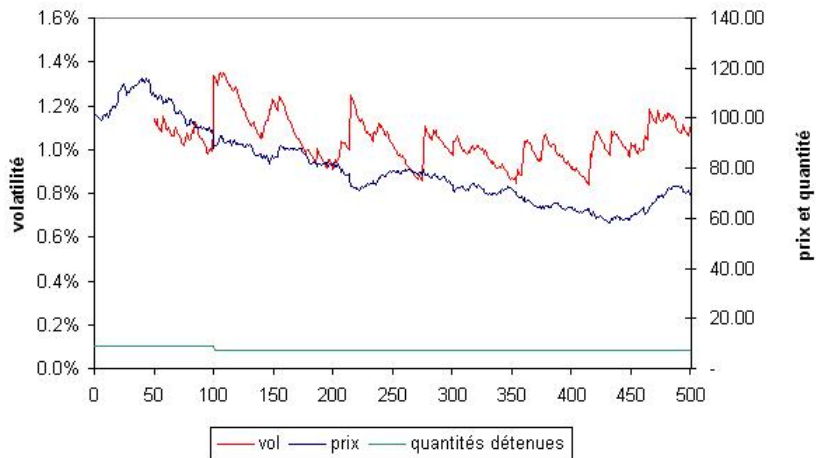


FIGURE –  $V = 0$ ,  $\gamma = 0.2$ ,  $r = 0.25$ ,  $\lambda = 0.03$  (demie vie en 24 jours).

# Spirale de liquidité

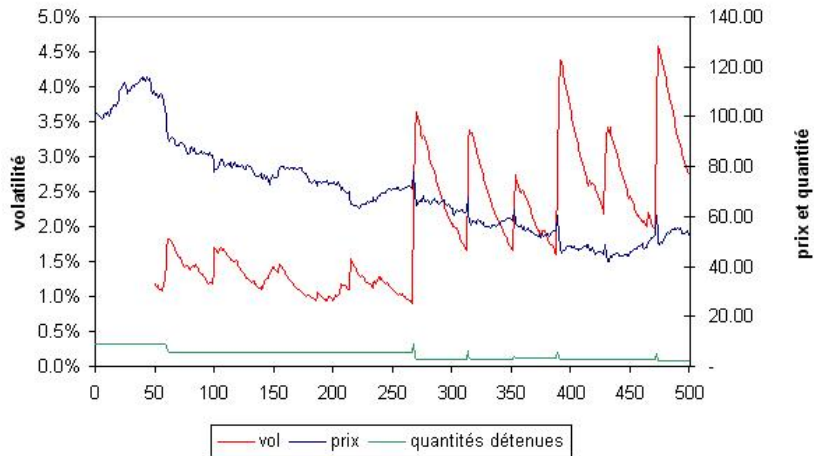


FIGURE –  $V = 5$ ,  $\gamma = 0.2$ ,  $r = 0.25$ ,  $\lambda = 0.05$  (demie vie en 15 jours).

# Spirale de liquidité

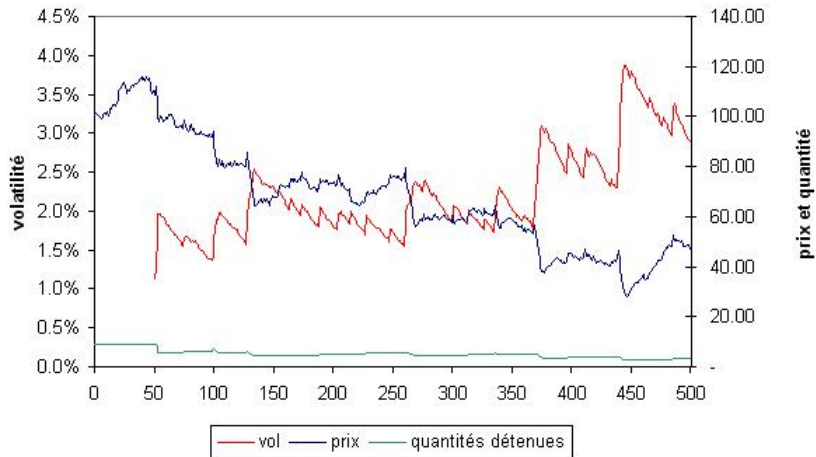


FIGURE –  $V = 5$ ,  $\gamma = 0.2$ ,  $r = 0.25$ ,  $\lambda = 0.03$  (demie vie en 24 jours). Vol min de 0.8% au lieu de 0.5%.

# Spirale de liquidité

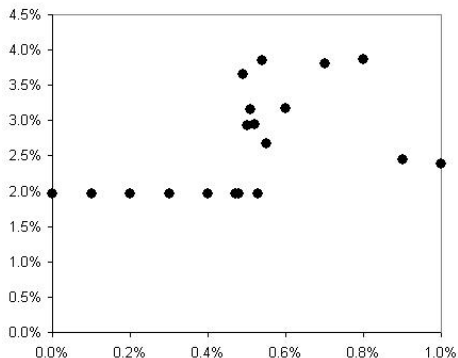


FIGURE –  $V = 5$ ,  $\gamma = 0.2$ ,  $r = 0.25$ ,  $\lambda = 0.03$  (demie vie en 24 jours). Vol max atteinte pour une borne inférieure d'intervalle de vol admissible allant de 0 à 1%.

# Spirale de liquidité

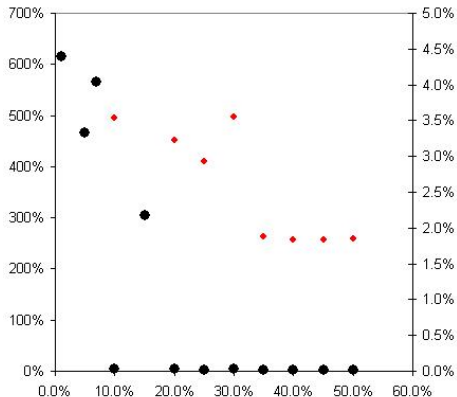


FIGURE –  $V = 5$ ,  $\gamma = 0.2$ ,  $\lambda = 0.03$  (demie vie en 24 jours). Vol max atteinte pour plusieurs valeurs de  $r$ . En rouge, axe de droite, zoom.

## 1 Risque de marché

- Risque actions
- Risque de taux
- Risque de change

## 2 Risque de crédit

- Valorisation sous risque de crédit
- Modèle structurel de crédit
- Dérivés de crédit

## 3 Risque de liquidité

- Notion de liquidité
- Mesures de liquidité
- Modèles d'impact

## 4 Risque opérationnel

- Aperçu général
- Risque de modèle
  - Surapprentissage et sélection de modèle
  - Méthodes de lissage
  - LASSO



# Définition du risque opérationnel

Dans le document de Bâle II : *"Operational risk is defined as the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events. This definition includes legal risk, but excludes strategic and reputational risk."*

**Exemples** : fraudes internes ou externes, pertes à cause d'erreurs ou de défaillances informatiques, erreurs dans le règlement des transactions, litiges et pertes dues à évènements externes (inondation, incendie, tremblement de terre, braquage).

Les mauvais choix de management (de portefeuille ou de l'institution) n'en font pas partie.

Un cas frappant, la Barings, cumule insuffisance des processus de contrôle, fraude interne et évènement externe (tremblement de terre de Kobe, lequel provoque un décalage du marché).

# Mesure du risque opérationnel

Deux approches sont possibles :

- Une mesure historique avec les techniques propres aux séries temporelles, ce qui est préconisé pour tout ce qui est exogène aux portefeuilles : fraudes (*a priori* impossible à prévoir), risques climatiques (risque plus prévisible mais demande des compétences non-financières).
- Une description du mécanisme (plus spécifiquement pour le risque de modèle) avec une mesure qui n'est pas nécessairement historique et dépend par exemple des la natures des actifs gérés.

# Mesure du risque opérationnel

Dans les accords de Bâle II, il y a trois options pour calculer le risque opérationnel selon l'exposition de chaque banque à ce type de risque (en backtest, la VaR de risque opérationnel à 99.9% doit être vérifiée à un horizon un an) :

- **Basic Indicator Approach** : proportion des revenus bruts positifs moyens des trois dernières années.
- **Standardized Approach** : proche de l'indicateur de base mais avec disjonction en *business lines*.
- **Advanced Measurement Approaches** : pour les banques ayant une activité internationale. Le comité exige une historique de pertes internes de 5 ans minimum (3 ans si méthode nouvellement mise en place).

# Basic Indicator and Standardized Approaches

- **Basic Indicator Approach** : Capital requis en année  $t$  est proportion  $\alpha = 15\%$  des revenus bruts positifs moyens des trois dernières années :

$$\frac{1}{Z_t} \sum_{i=1}^3 \alpha \max(R^{t-i}, 0),$$

où  $R^{t-i}$  est le revenu brut de l'institution en année  $t - i$  et  $Z_t = \sum_{i=1}^3 \mathbf{1}_{\{R^{t-i} > 0\}}$ .

- **Standardized Approach** : 8 *business lines* avec une proportion de capital requis  $\beta_j$  propre à chacune :

$$\frac{1}{3} \sum_{i=1}^3 \max \left( \sum_{j=1}^8 \beta_j R_j^{t-i}, 0 \right).$$

La division par 3 plutôt que  $Z_t$ , ainsi que le fait que des performances négatives sur une ligne de  $\beta_j$  élevé peut compenser les gains d'autres lignes, est une incitation à passer à cette approche.

# Standardized Approache

*Business lines :*

$j = 1$  *corporate finance* :  $\beta_1 = 18\%$  ;

$j = 2$  *trading and sales* :  $\beta_2 = 18\%$  ;

$j = 3$  *retail banking* :  $\beta_3 = 12\%$  ;

$j = 4$  *commercial banking* :  $\beta_4 = 15\%$  ;

$j = 5$  *payment and settlement* :  $\beta_5 = 18\%$  ;

$j = 6$  *agency services* :  $\beta_6 = 15\%$  ;

$j = 7$  *asset management* :  $\beta_7 = 12\%$  ;

$j = 8$  *retail brokerage* :  $\beta_8 = 12\%$ .

# Advanced Measurement Approaches

Dans le document de Bâle II : *"Given the continuing evolution of analytical approaches for operational risk, the Committee is not specifying the approach or distributional assumptions used to generate the operational risk measure for regulatory capital purposes. However, a bank must be able to demonstrate that its approach captures potentially severe tail loss events. Whatever approach is used, a bank must demonstrate that its operational risk measure meets a soundness standard comparable to that of the internal ratings-based approach for credit risk (comparable to a one year holding period and the 99.9 percent confidence interval)."*

De plus, les pertes doivent être réparties en 8 lignes, comme dans l'approche standard, et pour chaque ligne en **7 types de pertes** : internal fraud ; external fraud ; employment practices & workplace safety ; clients, products & business practices ; damage to physical assets ; business disruption & system failures ; execution, delivery & process management.

Comité impose historique haute fréquence pour événements internes et plus basse fréquence pour événements externes, avec corrélation entre types de pertes (en l'absence de corrélation, simple somme, ce qui correspond à borne haute de Fréchet-Hoeffding, cas comonotonique : hypothèse la plus conservatrice).

# Advanced Measurement Approaches

- Distribution jointe en général inconnue.
- On filtre les pertes inférieures à un seuil.
- La perte suit une variable aléatoire et la fréquence aussi.

Pour chaque business line, on est donc ramené au calcul de quelque chose comme

$$VaR_{0.999} \left( \sum_{k=1}^N X_k^j \right),$$

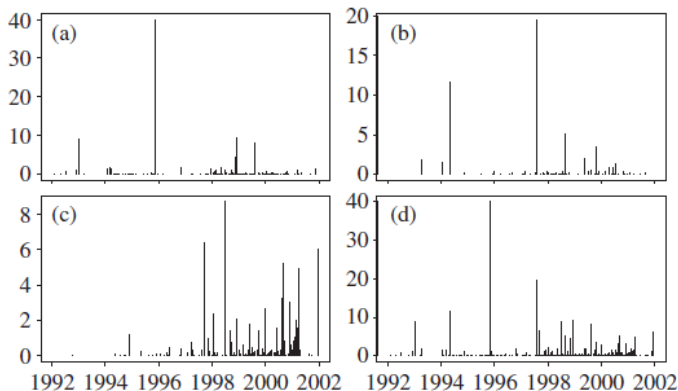
où le nombre d'occurrences  $N$  est aléatoire et la  $k$ -ème perte dans la business line  $j$ ,  $X_k^j$  (la sévérité), aussi.

Sur historiques 1992-2001, McNeil, Frey et Embrechts citent les faits stylisés suivants :

- sévérité des pertes distribuée selon loi à queue épaisse (EVT, paramètre de queue  $\xi$  de 0.85 à 1.39 selon *business lines*) ;
- les pertes arrivent à des temps aléatoire ;
- la fréquence des pertes peut varier substantiellement dans le temps (même si cela est partiellement dû au fait que le monitoring est plus complet ces dernières années)

Par ailleurs, effets cycliques (fraude liée à conjoncture économique ?) et corrélation à variables économiques (erreur de back-office liée aux volumes échangés), mais on suppose que toutes les sévérités sont iid.

# Advanced Measurement Approaches



**Figure 10.1.** Operational risk losses: (a) type 1,  $n = 162$ ; (b) type 2,  $n = 80$ ; (c) type 3,  $n = 175$ ; and (d) pooled losses  $n = 417$ .

**FIGURE – 3** types de pertes et dernier graphique somme. Source : McNeil, Frey et Embrechts.



# Advanced Measurement Approaches – Somme composée

On cherche la distribution de la perte agrégée.

**Definition 10.1 (total loss amount and distribution).** Denote by  $N(t)$  the (random) number of losses over a fixed time period  $[0, t]$  and write  $X_1, X_2, \dots$  for the individual losses. The *total loss amount* (or *aggregate loss*) is defined as

$$S_{N(t)} = \sum_{k=1}^{N(t)} X_k, \quad (10.9)$$

with df  $F_{S_{N(t)}}(x) = P(S_{N(t)} \leq x)$ , the *total* (or *aggregate*) *loss df*. Whenever  $t$  is fixed,  $t = 1$  say, we may drop the time index from the notation and simply write  $S_N$  and  $F_{S_N}$ .

# Advanced Measurement Approaches – Somme composée

**Assumption 10.3 (independence, compound sums).** We assume that the rvs  $(X_k)$  are iid with common df  $G$ ,  $G(0) = 0$ . We further assume that the rvs  $N$  and  $(X_k)$  are independent; in that case we refer to (10.9) as a compound sum. The probability mass function of  $N$  is denoted by  $p_N(k) = P(N = k)$ ,  $k = 0, 1, 2, \dots$ . The rv  $N$  is referred to as the compounding rv.

**Proposition 10.4 (compound distribution).** Let  $S_N$  be a compound sum and suppose that Assumption 10.3 holds. Then, for all  $x \geq 0$ ,

$$F_{S_N}(x) = P(S_N \leq x) = \sum_{k=0}^{\infty} p_N(k) G^{(k)}(x), \quad (10.10)$$

where  $G^{(k)}(x) = P(S_k \leq x)$ , the  $k$ th convolution of  $G$ . Note that  $G^{(0)}(x) = 1$  for  $x \geq 0$ , and  $G^{(0)}(x) = 0$  for  $x < 0$ .

Preuve en notant que  $F_{S_N}(x) = \sum_{k=0}^{\infty} \mathbb{P}(S_N \leq x | N = k) \mathbb{P}(N = k)$ .

La formule est explicite mais le calcul exact de la convolution est en général impossible. Solution possible en passant que transformée de Laplace, car la transformée d'une convolution est un produit des transformées. La difficulté sera alors l'inversion de la transformée de Laplace pour obtenir la distribution.

# AMA – Somme composée – Transformée de Laplace

## Définition

La **transformée de Laplace** de la variable aléatoire réelle  $X$  est la fonction  $L_X$ , définie de  $\mathbb{C}$  dans  $\mathbb{C}$  par la relation :

$$L_X(a) = \mathbb{E} \left[ e^{aX} \right].$$

Dans notre cas de figure, on note la transformée  $\hat{F}(s) = \int_0^\infty e^{-sx} dF(x)$ . Alors :

$$\widehat{G^{(k)}}(s) = \int_0^\infty \dots \left( \int_0^\infty e^{-s(x_1 + \dots + x_k)} dG(x_1) \right) \dots dG(x_k) = (\hat{G}(s))^k.$$

Alors, de la proposition (10.4), on obtient une formule très simple pour la transformée de Laplace (erratum : pas fonction génératrice des moments,  $\mathbb{E}(e^{tX})$ , mais fonction génératrice,  $\mathbb{E}(t^X)$ ) :

$$\hat{F}_{S_N}(s) = \sum_{k=0}^{\infty} p_N(k) \hat{G}^k(s) = M_N(\hat{G}(s)), \quad s \geq 0, \quad (10.11)$$

where  $M_N$  denotes the *moment-generating function* of  $N$ .

# AMA – Somme composée – Transformée inverse

Plusieurs possibilités pour faire la transformée inverse de Laplace et ainsi récupérer la fonction de répartition :

- **Méthode analytique** : intégrale sur droite complexe (**formule de Bromwich-Mellin**), avec  $\gamma \in \mathbb{R}$  choisi pour respecter certaines conditions de convergence assez précises (ainsi que la condition d'avoir  $\gamma$  supérieur à la partie réelle de toute singularité de  $F$  : si fonction régulière, on peut prendre  $\gamma = 0$  et on retrouve la **transformée inverse de Fourier**) :

$$F(t) = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} e^{st} \hat{F}(s) ds.$$

- **Méthode de Stehfest** : Méthode numérique (choix de  $n$  tronqué)

$$F(t) = \lim_{n \rightarrow +\infty} \frac{\ln(2)}{t} \sum_{j=1}^n V_j \hat{F}\left(\frac{j \ln(2)}{t}\right) \text{ avec } n \text{ pair et}$$

$$V_i = (-1)^{\frac{n}{2}+i} \sum_{k=\lfloor \frac{i+1}{2} \rfloor}^{\min(i, \frac{n}{2})} \frac{k^{\frac{n}{2}} (2k)!}{(\frac{n}{2}-k)! k! (k-1)! (i-k)! (2k-i)!} \text{ où } \lfloor \cdot \rfloor \text{ partie entière.}$$

- **Inversion de Post-Widder** : Si  $F$  est continue et bornée, on obtient en utilisant la loi forte des grands nombres :  $\lim_{n \rightarrow +\infty} \frac{(-1)^{n-1} n^n \hat{F}^{(n-1)}(\frac{n}{x})}{x^n (n-1)!} = F(x).$

# AMA – Somme composée – Exemple

**Example 10.5 (the compound Poisson df).** Suppose that  $N$  has a Poisson df with intensity parameter  $\lambda > 0$ , denoted  $N \sim \text{Poi}(\lambda)$ . In that case,  $p_N(k) = e^{-\lambda} \lambda^k / k!$ ,  $k \geq 0$ , and, for  $s \in \mathbb{R}$ ,

$$M_N(s) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} s^k = \exp(-\lambda(1-s)).$$

Hence from (10.11) it follows that, for  $s \geq 0$ ,

$$\hat{F}_{S_N}(s) = \exp(-\lambda(1 - \hat{G}(s))).$$

In this case, the df of  $S_N$  is referred to as the *compound Poisson df* and we write  $S_N \sim \text{CPoi}(\lambda, G)$ . Formula (10.11) facilitates the calculation of moments of  $S_N$

**Example 10.6 (continuation of Example 10.5).** In the case of the compound Poisson df, one obtains

$$\begin{aligned} E(S_N) &= (-1) \frac{d}{ds} \hat{F}_{S_N}(s) \Big|_{s=0} = \exp(-\lambda(1 - \hat{G}(0))) \lambda (-\hat{G}'(0)) \\ &= \lambda \mu_1 = E(N) E(X_1). \end{aligned}$$

Similar calculations yield  $\text{var}(S_N) = E(S_N^2) - (E(S_N))^2 = \lambda \mu_2$ .

avec  $\mu_k = \mathbb{E}(X_1^k)$ .

# AMA – Somme composée

Dans le cas général ( $N$  n'est pas forcément distribué selon loi de Poisson), on démontre, en notant que  $\mathbb{E}(S_N) = \mathbb{E}(\mathbb{E}(S_N|N))$  :

**Proposition 10.7 (moments of compound dfs).** *Under Assumption 10.3 and assuming that  $E(N) < \infty$ ,  $\mu_2 < \infty$ , we have that*

$$E(S_N) = E(N)E(X_1) \quad \text{and} \quad \text{var}(S_N) = \text{var}(N)(E(X_1))^2 + E(N) \text{var}(X_1). \quad (10.12)$$

Retour à l'exemple avec une loi de Poisson composée. Si on calcule ainsi le risque pour une ligne de métier (on a la distribution de la perte agrégée en une année), on peut calculer le risque (on a aussi la distribution de la perte agrégée) pour plusieurs lignes de métier agrégées, grâce à cette propriété (prouvée en passant par la transformée de Laplace, unique) :

**Proposition 10.9 (sums of compound Poisson rvs).** *Suppose that the compound sums  $S_{N_i} \sim \text{CPoi}(\lambda_i, G_i)$ ,  $i = 1, \dots, d$ , and that these rvs are independent, then  $S_N = \sum_{i=1}^d S_{N_i} \sim \text{CPoi}(\lambda, G)$ , where  $\lambda = \sum_{i=1}^d \lambda_i$  and  $G = \sum_{i=1}^d (\lambda_i / \lambda) G_i$ .*

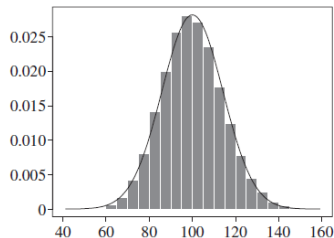
L'intensité des événements est la somme des intensités et la distribution de la sévérité est le mélange discret des deux distributions pondérées par leur intensité relative.

# AMA – Somme composée – Approximation

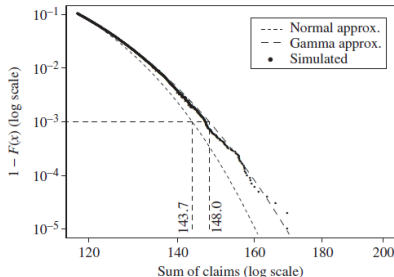
De nombreuses approximations numériques de  $F_{S_N}$  :

- Approximation par des lois paramétriques ayant les mêmes premiers moments, mais ces approches s'intéressent plus aux moments et au centre de la distribution qu'à la queue :
  - 2 premiers moments : loi normale ;
  - 3 premiers moments : loi gamma traduite, de densité
$$f(x) = (x - k)^{\alpha-1} \frac{\beta^\alpha e^{-\beta(x-k)}}{\Gamma(\alpha)} \mathbf{1}_{\{x > k\}},$$
avec une espérance  $k + \alpha/\beta$ , une variance  $\alpha/\beta^2$  et une skewness  $2/\sqrt{\alpha}$  ;
  - davantage de cumulants : décomposition d'Edgeworth.
- Une simulation Monte Carlo de la perte agrégée : on peut tirer les pertes individuelles et  $N$ , donc on peut observer empiriquement les quantiles élevés.
- Une récurrence de Panjer.

# AMA – Somme composée – Approximation moments



**Figure 10.3.** Histogram of simulated compound loss data ( $n = 100\,000$ ) for  $S_N \sim \text{CPoi}(100, \text{Exp}(1))$  together with normal approximation (10.14).



**Figure 10.4.** Simulated  $\text{CPoi}(100, \text{Exp}(1))$  data together with normal- and translated-gamma approximations (log-log scale). The 99.9% quantile estimates are also given.



# AMA – Somme composée – Approximation Edgeworth

**Cumulants** : caractérisent variable aléatoire, comme moments. Générés par fonction génératrice des cumulants :

$$K(t) = \log \mathbb{E}[e^{tX}] = \sum_{r=1}^{\infty} C_r(X) \frac{t^r}{r!},$$

à comparer avec la fonction génératrice des moments  $M(t) = \mathbb{E}[e^{tX}] = \sum_{r=1}^{\infty} M_r(X) \frac{t^r}{r!}$ .

**Exemples** :

- Pour une loi gaussienne de moyenne  $\mu$  et de variance  $\sigma^2$ ,  $C_1(X) = \mu$ ,  $C_2(X) = \sigma^2$ ,  $C_3(X) = C_4(X) = \dots = 0$ . Néanmoins, il n'existe pas de distribution pour laquelle  $K$  est un polynôme d'ordre fini strictement supérieur à 2.
- Pour une loi exponentielle de paramètre  $\lambda$ ,  $C_r(X) = \lambda^{-r}(r-1)!$ .
- Pour une loi de Poisson d'intensité  $\mu$ ,  $K(t) = \mu(e^t - 1)$  et  $C_r(X) = \mu$  pour tout  $r \geq 1$ .

# AMA – Somme composée – Approximation Edgeworth

It is easy to get moments from cumulants and cumulants from moments thanks to the fact that  $C_1(X) = M_1(X)$  and to the recursive formula, for  $r \in \{1, 2, \dots\}$  :

$$C_r(X) = M_r(X) - \sum_{s=1}^{r-1} \binom{r}{s} C_s(X) M_{r-s}(X),$$

where  $C$  denotes any finite cumulant,  $M$  any finite moment and  $X$  a random variable for which cumulants and moments are well defined. Moreover, two properties concerning cumulants are very interesting :

- The cumulant of any sum of independent random variables is the sum of the cumulants of each random variable : ideal for sums (like  $S_N$ ) of random variables whose cumulants are known !
- The cumulants are homogeneous : the  $r$ -th cumulant of  $aX$ , for any  $a \in \mathbb{R}$ , is  $a^r C_r(X)$ .

# AMA – Somme composée – Approximation Edgeworth

Développement d'Edgeworth : déformation d'une densité gaussienne  $\phi$  en considérant l'effet des cumulants de la densité  $\delta$  d'une variable aléatoire  $Z$ , ayant ses moments  $M_r(X)$  (et donc ses cumulants  $C_r(X)$ ) bien définis. En particulier, moyenne  $\mu = M_1(X)$  et variance  $\sigma^2 = M_2(X) - M_1(X)^2$ . En effet,  $\forall x \in \mathbb{R}$  :

$$\delta(x) = \phi(x) \left[ 1 + \sum_{m=1}^{+\infty} \mathbb{E} \left[ \mathcal{H}_m \left( \frac{X - \mu}{\sqrt{2}\sigma} \right) \right] \mathcal{H}_m \left( \frac{x - \mu}{\sqrt{2}\sigma} \right) \frac{\sigma^m}{(\sqrt{2})^m m!} \right],$$

où les  $(\mathcal{H}_m)_{m \in \mathbb{N}}$  sont les polynômes d'Hermite :

$\forall (m, x) \in \mathbb{N} \times \mathbb{R}$ ,  $\mathcal{H}_m(x) = (-1)^m e^{x^2/2} \frac{d^m}{dx^m} e^{-x^2/2}$ . On a aussi la fonction de répartition de  $X$  développée autour de la fonction de répartition gaussienne  $\Phi$  :

$$\Delta(x) = \Phi(x) + \sum_{m=1}^{+\infty} \mathbb{E} \left[ \mathcal{H}_m \left( \frac{X - \mu}{\sqrt{2}\sigma} \right) \right] \left[ \frac{\mathcal{H}_{m-1}(0)}{\sqrt{2}\sigma} - \phi(x) \mathcal{H}_{m-1} \left( \frac{x - \mu}{\sqrt{2}\sigma} \right) \right] \frac{\sigma^{m+1}}{(\sqrt{2})^{m-1} m!}.$$

# AMA – Somme composée – Approximation Edgeworth

## Algorithme (Espérance des polynômes d'Hermite (1/2))

Let  $M \in \mathbb{N}$  and  $\mu, \sigma \in \mathbb{R}$ . Then, the following algorithm allows  $\mathbb{E} \left[ \mathcal{H}_m \left( \frac{X-\mu}{\sqrt{2}\sigma} \right) \right]$  to be calculated for every  $m \in \{1, \dots, M\}$  :

- 1 We begin with  $r = 1$ . At each iteration, we replace  $r$  by  $r + 1$ .
- 2 We calculate  $C_r \left( \frac{X-\mu}{\sqrt{2}\sigma} \right)$  as :

$$C_r \left( \frac{X-\mu}{\sqrt{2}\sigma} \right) = \frac{C_r(X)}{\left[ \sqrt{2}\sigma \right]^r} - \frac{C_r(\mu)}{\left[ \sqrt{2}\sigma \right]^r}.$$

- 3 For each  $j$  and  $k$  indexing the considered empirical wavelet coefficients, we calculate  $M_r \left( \frac{X-\mu}{\sqrt{2}\sigma} \right)$  as :

$$\begin{cases} M_r \left( \frac{X-\mu}{\sqrt{2}\sigma} \right) = C_r \left( \frac{X-\mu}{\sqrt{2}\sigma} \right) & \text{if } r = 1 \\ M_r \left( \frac{X-\mu}{\sqrt{2}\sigma} \right) = C_r \left( \frac{X-\mu}{\sqrt{2}\sigma} \right) + \sum_{s=1}^{r-1} \binom{r}{s} C_s \left( \frac{X-\mu}{\sqrt{2}\sigma} \right) M_{r-s} \left( \frac{X-\mu}{\sqrt{2}\sigma} \right) & \text{else.} \end{cases}$$

# AMA – Somme composée – Approximation Edgeworth

## Algorithme (Espérance des polynômes d'Hermite (2/2))

- 4 We go back to the step 1 of the algorithm until  $r = M$  is reached.
- 5 For each  $m \in \{0, \dots, M\}$ , we deduct  $\mathbb{E} \left[ \mathcal{H}_m \left( \frac{X-\mu}{\sqrt{2}\sigma} \right) \right]$  from the previously calculated moments, say :

$$\begin{cases} \mathbb{E} \left[ \mathcal{H}_0 \left( \frac{X-\mu}{\sqrt{2}\sigma} \right) \right] &= 1 \\ \mathbb{E} \left[ \mathcal{H}_1 \left( \frac{X-\mu}{\sqrt{2}\sigma} \right) \right] &= M_1 \left( \frac{X-\mu}{\sqrt{2}\sigma} \right) \\ \mathbb{E} \left[ \mathcal{H}_2 \left( \frac{X-\mu}{\sqrt{2}\sigma} \right) \right] &= M_2 \left( \frac{X-\mu}{\sqrt{2}\sigma} \right) - 1 \\ \mathbb{E} \left[ \mathcal{H}_3 \left( \frac{X-\mu}{\sqrt{2}\sigma} \right) \right] &= M_3 \left( \frac{X-\mu}{\sqrt{2}\sigma} \right) - 3M_1 \left( \frac{X-\mu}{\sqrt{2}\sigma} \right) \end{cases}$$

and so on until the  $M$ -th polynomial.

# AMA – Somme composée – Approximation Edgeworth

Quelques limites du développement d'Edgeworth :

- le résultat n'est pas à proprement parler une densité ou une fonction de répartition (intégrale de la densité différente de 1, probabilités négatives possibles) ;
- imprécision, notamment dans les queues, car :
  - développement de Taylor autour d'une moyenne (pas des queues) ;
  - minimise l'erreur absolue et non pas relative, de sorte que pour faibles probabilités (queues de la distribution), l'erreur relative peut être très importante alors que l'erreur absolue est maîtrisée.

Par ailleurs, on préfère parfois le développement de Cornish-Fisher qui est identique mais qui ordonne différemment les termes, rendant le résultat d'un développement tronqué différent du cas Edgeworth.

# AMA – Somme composée – Approximation Panjer

$X_1$  a maintenant distribution discrète, avec  $g_k = \mathbb{P}(X_1 = k)$ ,  $p_k = \mathbb{P}(N = k)$  et  $s_k = \mathbb{P}(S_N = k)$ . On suppose  $g_0 = 0$  et on note la convolution discrète  $g_k^{(n)} = \mathbb{P}(X_1 + \dots + X_n = k)$ .

$$\left. \begin{aligned} s_0 &= P(S_N = 0) = P(N = 0) = p_0, \\ s_n &= P(S_N = n) = \sum_{k=1}^{\infty} p_k g_n^{(k)}, \quad n \geq 1, \end{aligned} \right\} \quad (10.16)$$

Simplifié pour classe particulière pour  $N$ , classe de Panjer, à laquelle appartiennent les lois binomiales ( $p_r = (n!/(r!(n-r)!))p^r(1-p)^{n-r}$  pour  $r \leq n$ ) et Poisson ( $p_r = e^{-\lambda}\lambda^r/r!$ ).

**Definition 10.10 (Panjer class).** The probability mass function ( $p_k$ ) of  $N$  belongs to the Panjer( $a, b$ ) class for some  $a, b \in \mathbb{R}$  if the following relationship holds for  $r \geq 1$ :  $p_r = (a + (b/r))p_{r-1}$ .

# AMA – Somme composée – Approximation Panjer

## Théorème et preuve (1/3)

**Theorem 10.15 (Panjer recursion).** *Suppose that  $N$  satisfies the Panjer( $a, b$ ) class condition and  $g_0 = P(X_1 = 0) = 0$ , then  $s_0 = p_0$  and, for  $r \geq 1$ ,  $s_r = \sum_{i=1}^r (a + (bi/r))g_i s_{r-i}$ .*

$$\begin{aligned} E\left(X_1 \mid \sum_{i=1}^n X_i = r\right) &= \frac{1}{n} \sum_{j=1}^n E\left(X_j \mid \sum_{i=1}^n X_i = r\right) \\ &= \frac{1}{n} E\left(\sum_{j=1}^n X_j \mid \sum_{i=1}^n X_i = r\right) = \frac{r}{n}. \end{aligned}$$



# AMA – Somme composée – Approximation Panjer

## Preuve du théorème (2/3)

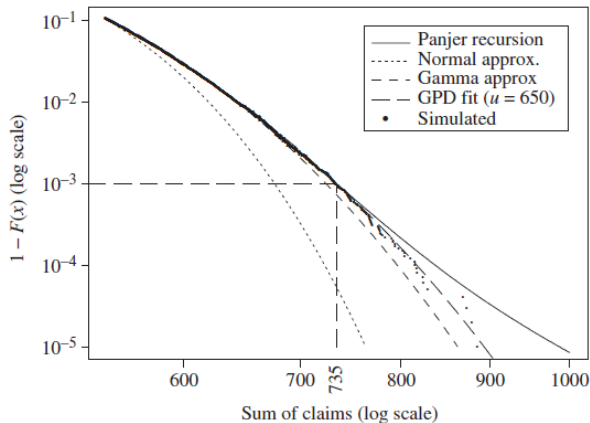
$$\begin{aligned}
 p_{n-1} \sum_{i=1}^{r-1} \left( a + \frac{bi}{r} \right) g_i g_{r-i}^{(n-1)} &= p_{n-1} \sum_{i=1}^r \left( a + \frac{bi}{r} \right) g_i g_{r-i}^{(n-1)} \\
 &= p_{n-1} \sum_{i=1}^r \left( a + \frac{bi}{r} \right) P \left( X_1 = i, \sum_{j=2}^n X_j = r - i \right) \\
 &= p_{n-1} \sum_{i=1}^r \left( a + \frac{bi}{r} \right) P \left( X_1 = i, \sum_{j=1}^n X_j = r \right) \\
 &= p_{n-1} \sum_{i=1}^r \left( a + \frac{bi}{r} \right) P \left( X_1 = i \mid \sum_{j=1}^n X_j = r \right) g_r^{(n)} \\
 &= p_{n-1} E \left( a + \frac{bX_1}{r} \mid \sum_{j=1}^n X_j = r \right) g_r^{(n)} \\
 &= p_{n-1} \left( a + \frac{b}{n} \right) g_r^{(n)} = p_n g_r^{(n)},
 \end{aligned}$$

# AMA – Somme composée – Approximation Panjer

## Preuve du théorème (3/3)

$$\begin{aligned}
 s_r &= \sum_{n=1}^{\infty} p_n g_r^{(n)} = p_1 g_r + \sum_{n=2}^{\infty} p_n g_r^{(n)} \\
 &= (a+b)p_0 g_r + \sum_{n=2}^{\infty} \sum_{i=1}^{r-1} \left(a + \frac{bi}{r}\right) g_i p_{n-1} g_{r-i}^{(n-1)} \\
 &= (a+b)s_0 g_r + \sum_{i=1}^{r-1} \left(a + \frac{bi}{r}\right) g_i \sum_{n=2}^{\infty} p_{n-1} g_{r-i}^{(n-1)} \\
 &= (a+b)g_r s_0 + \sum_{i=1}^{r-1} \left(a + \frac{bi}{r}\right) g_i s_{r-i} \\
 &= \sum_{i=1}^r \left(a + \frac{bi}{r}\right) g_i s_{r-i}.
 \end{aligned}$$

# AMA – Somme composée – Approximation Panjer



**Figure 10.5.** Simulated CPoi(100,  $LN(1, 1)$ ) data ( $n = 100\,000$ ) with normal-, translated-gamma, GPD and Panjer recursion (see Example 10.17) approximations (on log-log scale).

FIGURE – 735 est quantile à 99.9%. Source : McNeil, Frey et Embrechts.

## 1 Risque de marché

- Risque actions
- Risque de taux
- Risque de change

## 2 Risque de crédit

- Valorisation sous risque de crédit
- Modèle structurel de crédit
- Dérivés de crédit

## 3 Risque de liquidité

- Notion de liquidité
- Mesures de liquidité
- Modèles d'impact

## 4 Risque opérationnel

- Aperçu général
- Risque de modèle
  - Surapprentissage et sélection de modèle
  - Méthodes de lissage
  - LASSO

# Plusieurs approches

Le risque de modèle peut provenir de plusieurs sources :

- le modèle utilisé est **mal spécifié**<sup>1</sup> : les résidus sont distribués différemment de ce qui est supposé (utiliser modèles non-paramétriques, mais risque d'overfitting), un paramètre supplémentaire est nécessaire (mouvement brownien vs mouvement brownien fractionnaire), un facteur de risque n'est pas identifié ;
- il y a une **incertitude dans l'estimation** : présence de bruit (débruitage partiel), variation des paramètres dans le temps et impossibilité de prévoir leur valeur future ;
- même bien estimé, la sortie du modèle peut être incertaine : intervalles de confiance d'un Monte-Carlo, approximations **numériques** d'un schéma aux différences finies.

---

1. Détaillé dans les sous-sections qui suivent : suaprentissage et sélection de modèle, méthodes de lissage, LASSO.

# Risque d'estimation dans le modèle Black-Scholes

Stratégie de couverture du risque action d'une option dans le modèle de Black-Scholes avec une volatilité *fausse*  $\bar{\sigma}$  au lieu de  $\sigma$ . On achète en 0 une option de payoff  $h(S_T)$  et de prix  $C_t$  et on vend un portefeuille répliquant  $\bar{V}_t$  investi dans  $\bar{\delta}$  actions, où les grecques surmontées d'une barre sont les sensibilité du *faux* prix, i.e. du prix de notre réplication imparfaite, aux vrais facteurs de risque. On dénoue en  $T$ . Profit est

$$e_T = \bar{V}_T - h(S_T) = \int_0^T e^{-r_t(T-t)} \frac{1}{2} (\bar{\sigma}_t^2 - \sigma_t^2) S_t^2 \bar{\Gamma}_t dt.$$

Si le  $\Gamma$  est positif (c'est le cas pour les calls), cette erreur de couverture conduit à une perte si la volatilité est sous-estimée.

- ◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ▶ ↺ 🔍 ↻

# Overfitting

## Définition

**Surapprentissage** : *Ajout de variables explicatives (ou paramètres) peu pertinentes : améliore précision in-sample mais la détériore out-of-sample.*

C'est le danger en statistiques non-paramétriques ou paramétriques avec grand nombre de coefficients. Corrigé avec :

- régularisation,
- diminution du nombre de paramètres (ex : LASSO).

Exemple : estimation de densité avec histogramme.



# Sélection de modèle

Ainsi, la sélection de modèle ne peut pas se faire uniquement en prenant en compte un écart de vraisemblance. On utilise donc l'un des critères suivants, qu'il faut minimiser :

- AIC (Akaike information criterion) :

$$AIC = 2k - 2 \ln(\hat{L}),$$

où  $k$  est le nombre de paramètres du modèle,  $\hat{L}$  la vraisemblance la plus forte pour le modèle estimé.

- BIC (Bayesian information criterion) :

$$BIC = \ln(n)k - 2 \ln(\hat{L}),$$

où  $n$  est le nombre d'observations.

Ces deux critères ne permettent pas de dire qu'un modèle est bon mais simplement de faire le tri entre plusieurs modèles. Critère à utiliser alors pour quantifier cet écart : **vraisemblance relative**, définie par

$$\rho = \exp \left( \frac{AIC_1 - AIC_2}{2} \right).$$

Le modèle 2 est  $\rho$  fois autant vraisemblable que le modèle 1 de minimiser la perte d'information induite par le recours à un modèle.

1 Risque de marché

2 Risque de crédit

3 Risque de liquidité

4 Risque opérationnel

- Aperçu général
- Risque de modèle
  - Surapprentissage et sélection de modèle
  - Méthodes de lissage
  - LASSO

# Smoothing nonparametric statistics

Smoothing time-dependent (or depending on another parameter, such as a strike, a residual maturity in the framework of model calibration) raw statistics :

- (exponentially weighted) moving average : rule of thumb, thus inaccurate description, lagged predictions, **but** easy to compute.
- decomposition in a series of orthogonal functions (Fourier, wavelets) : many methodological choices (truncation, linear or non-linear filtering of coefficients, mother wavelet), high computational cost (nested integrals), tricky predictions (border effects), **but** accurate description.

The problem, when smoothing the raw statistic  $\hat{H}$  :

$$\mathcal{H} = \operatorname{argmin}_{h \in C^2([0,1], \mathbb{R})} \int_0^1 \left[ (h(t) - \hat{H}(t))^2 + \lambda h'(t)^2 \right] dt.$$

Fidelity term and smoothness penalty : not that new... e.g. Whittaker-Henderson problem in actuarial literature, or in variational image processing. But always solved (sometimes for more general frameworks) by numerical schemes.

# Variational problem

The problem, when smoothing the raw statistic  $\hat{H}$  :

$$\mathcal{H} = \operatorname{argmin}_{h \in \mathcal{C}^2([0,1], \mathbb{R})} \int_0^1 \left[ (h(t) - \hat{H}(t))^2 + \lambda h'(t)^2 \right] dt. \quad (1)$$

It can be written

$$\mathcal{H} = \operatorname{argmin}_{h \in \mathcal{C}^2([0,1], \mathbb{R})} \int_0^1 \mathcal{L}(t, h(t), h'(t)) dt,$$

so that we can use the Euler-Lagrange equation (explained therefater) :

$$\frac{\partial}{\partial h} \mathcal{L}(t, \mathcal{H}(t), \mathcal{H}'(t)) = \frac{d}{dt} \frac{\partial}{\partial h'} \mathcal{L}(t, \mathcal{H}(t), \mathcal{H}'(t)). \quad (2)$$

In our framework, we have a simple non-homogeneous linear differential equation :

$$\mathcal{H}''(t) - \frac{\mathcal{H}(t)}{\lambda} = -\frac{\hat{H}(t)}{\lambda}. \quad (3)$$

# Euler-Lagrange equation

We are looking for  $h$  minimizing

$$J(h) = \int_0^1 \mathcal{L}(t, h(t), \dot{h}(t)) dt.$$

Differential calculus provides us with :

$$J(h + \varepsilon) = J(h) + DJ(h, \varepsilon) + o(\varepsilon),$$

for a given function  $\varepsilon$  such that  $\varepsilon(0) = \varepsilon(1) = 0$ , and where  $DJ(h, \varepsilon)$  is linear in  $\varepsilon$ . More exactly,  $DJ(h, \varepsilon)$  is equal to :

$$DJ(h, \varepsilon) = \int_0^1 \left[ \frac{\partial \mathcal{L}}{\partial h}(t, h(t), \dot{h}(t)) \varepsilon + \frac{\partial \mathcal{L}}{\partial \dot{h}}(t, h(t), \dot{h}(t)) \dot{\varepsilon} \right] dt.$$

At the optimum,  $DJ(h, \varepsilon) = 0$  and we can integrate by parts the second part of the integral :

$$\int_0^1 \frac{\partial \mathcal{L}}{\partial \dot{h}} \times \frac{d\varepsilon}{dt} dt = \left[ \frac{\partial \mathcal{L}}{\partial \dot{h}} \varepsilon \right]_{t=0}^1 - \int_0^1 \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{h}} \times \varepsilon dt,$$

in which the  $[\cdot]_{t=0}^1$  is equal to zero because  $\varepsilon(0) = \varepsilon(1) = 0$ .

Finally :

$$\frac{\partial}{\partial h} \mathcal{L}(t, h(t), \dot{h}(t)) = \frac{d}{dt} \frac{\partial}{\partial \dot{h}} \mathcal{L}(t, h(t), \dot{h}(t)).$$

# The solution with two parameters

## Théorème

Let  $\hat{H} : [0, 1] \rightarrow \mathbb{R}$  be two times differentiable and  $\lambda > 0$ . Then, a solution  $\mathcal{H} : [0, 1] \rightarrow \mathbb{R}$  of equation (1) is necessarily of the form :

$$\mathcal{H} : t \mapsto e^{t/\sqrt{\lambda}} \left( A - \frac{1}{2\sqrt{\lambda}} \int_0^t e^{-s/\sqrt{\lambda}} \hat{H}(s) ds \right) + e^{-t/\sqrt{\lambda}} \left( B + \frac{1}{2\sqrt{\lambda}} \int_0^t e^{s/\sqrt{\lambda}} \hat{H}(s) ds \right),$$

where  $A, B \in \mathbb{R}$ .

- 1 First (general solution), compute  $\Phi \hat{H}$ , where  $\Phi$  is the operator such that, for  $h \in \mathcal{C}^2([0, 1], \mathbb{R})$  and  $t \in [0, 1]$  :

$$\Phi h(t) = -\frac{e^{t/\sqrt{\lambda}}}{2\sqrt{\lambda}} \int_0^t e^{-s/\sqrt{\lambda}} h(s) ds + \frac{e^{-t/\sqrt{\lambda}}}{2\sqrt{\lambda}} \int_0^t e^{s/\sqrt{\lambda}} h(s) ds. \quad (4)$$

- 2 Then (fit to *initial conditions*), project the residual in the functional space generated by  $t \mapsto e^{t/\sqrt{\lambda}}$  and  $t \mapsto e^{-t/\sqrt{\lambda}}$ .

# A closed formula in the time series approach

## Théorème

Let  $\hat{H}(t_1), \dots, \hat{H}(t_n)$  be  $n$  statistics associated to times  $t_1 = 0 < t_2 < \dots < t_n = 1$ , with  $n > 1$ . Then, the unique function  $\mathcal{H} : \{t_0, \dots, t_n\} \rightarrow \mathbb{R}$  such that  $\sum_{i=1}^n (\mathcal{H}(t_i) - \hat{H}(t_i))^2 dt$  is minimal and of the form

$$\mathcal{H} : t \mapsto \Phi^d \hat{H}(t) + Ae^{t/\sqrt{\lambda}} + Be^{-t/\sqrt{\lambda}},$$

where  $\Phi^d$  is the discretized version of the operator introduced in equation (4) and  $\lambda > 0$ , is characterized by :

$$\begin{pmatrix} A \\ B \end{pmatrix} = \frac{1}{\alpha} \begin{pmatrix} \left( \sum_{i=1}^n e^{-2t_i/\sqrt{\lambda}} \right) \left( \sum_{i=1}^n e^{t_i/\sqrt{\lambda}} [1 - \Phi^d] \hat{H}(t_i) \right) - n \left( \sum_{i=1}^n e^{-t_i/\sqrt{\lambda}} [1 - \Phi^d] \hat{H}(t_i) \right) \\ \left( \sum_{i=1}^n e^{2t_i/\sqrt{\lambda}} \right) \left( \sum_{i=1}^n e^{-t_i/\sqrt{\lambda}} [1 - \Phi^d] \hat{H}(t_i) \right) - n \left( \sum_{i=1}^n e^{t_i/\sqrt{\lambda}} [1 - \Phi^d] \hat{H}(t_i) \right) \end{pmatrix},$$

with  $\alpha = \left( \sum_{i=1}^n e^{2t_i/\sqrt{\lambda}} \right) \left( \sum_{i=1}^n e^{-2t_i/\sqrt{\lambda}} \right) - n^2$ .

# Accuracy of the variational smoother

## Théorème

Let  $\Phi$  be the operator introduced in equation (4) and  $H$  an analytic function  $[0, 1] \rightarrow \mathbb{R}$ . Then, for  $\lambda \rightarrow 0$ , there exists  $\xi \in [0, 1]$  such that :

$$\min_{(A,B) \in \mathbb{R}^2} \int_0^1 \left( \Phi H(t) + Ae^{t/\sqrt{\lambda}} + Be^{-t/\sqrt{\lambda}} - H(t) \right)^2 dt \leq 4\lambda^2 H''(\xi)^2 + o(\lambda^2).$$

In particular, if  $H$  is an affine function, the variational smoothing reproduces it without any error. Indeed, as proved by integration by parts in the computation of  $\Phi$ , if  $H$  is an affine function such that  $H(t) = a_0 + a_1 t$ , then  $\Phi H(t) = H(t) - Ae^{t/\sqrt{\lambda}} - Be^{-t/\sqrt{\lambda}}$ , with  $A = (a_0 + a_1 \sqrt{\lambda})/2$  and  $B = (a_0 - a_1 \sqrt{\lambda})/2$ .



# A simulation study - Metrics for a comparison

- The mean integrated squared error (MISE), average accuracy of the method :

$$R_0 = \frac{1}{N} \sum_{i=1}^N \left[ \mathcal{H} \left( \frac{i}{N} \right) - H \left( \frac{i}{N} \right) \right]^2,$$

for  $N$  observations in  $[0, 1]$ ,  $\mathcal{H}$  being the smooth estimate of  $H$ .

- The average bias in  $t = 1$ , that is on the right of the sample :

$$R_1 = \mathcal{H}(1) - H(1).$$

Ability of the technique not to create artificial gaps at the border of the sample.  
Important feature in a predictive framework.

- The average quadratic slope, which determines to which extent the erratic estimate has been smoothed :

$$R_2 = \frac{1}{N-1} \sum_{i=2}^N \left[ \frac{\mathcal{H} \left( \frac{i}{N} \right) - \mathcal{H} \left( \frac{i-1}{N} \right)}{1/N} \right]^2.$$

# An example : exponential function

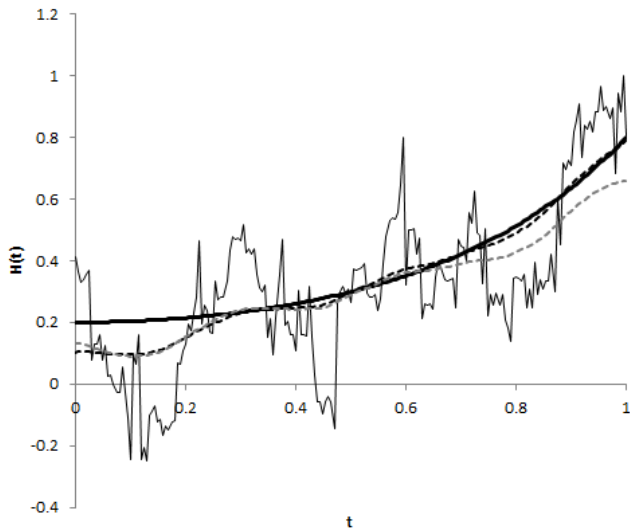
We first simulate a mBm on the time interval  $[0, 1]$  with an exponential function in  $[0, 1]$  for the Hurst exponent ranging from 0.2 to 0.8 :

$$t \mapsto 0.2 + 0.6 \frac{\exp(t^2) - 1}{\exp(1) - 1}.$$

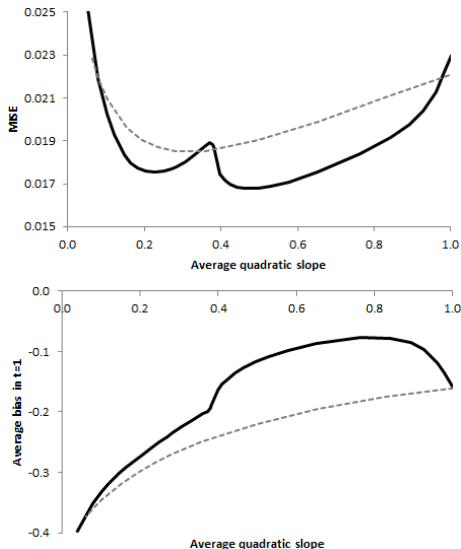
- The time interval is discretized in 200 points.
- Each raw estimate is computed on a window of 25 observations.
- We then obtain the smoothed Hurst exponents by two methods :
  - 1 a variational smoothing of parameter 0.02
  - 2 an exponentially weighted moving average of parameter 0.04.

Next figures : a trajectory, the average metrics for 100 trajectories (black dots : variational smoothing ; grey dots : moving average).

# Example : a simulation



# Example : average accuracy



1 Risque de marché

2 Risque de crédit

3 Risque de liquidité

4 Risque opérationnel

- Aperçu général
- Risque de modèle
  - Surapprentissage et sélection de modèle
  - Méthodes de lissage
  - LASSO

# Lasso

LASSO : *Least Absolute Shrinkage and Selection Operator*, introduit par Robert Tibshirani en 1996.

- Alternative à la régression *ridge*, ie régularisation de Tikhonov  

$$\underset{x}{\operatorname{arginf}} \|Tx - y\|_2^2 + \lambda \|x\|_2^2$$
, comme solution du problème inverse mal posé  $Tx = y$ .
- Ridge* ne réduit pas le nombre de paramètres et n'améliore donc par BIC ou AIC (le détériore même).

Recall our example:  $n = 50$ ,  $p = 30$ ; true coefficients: 10 are nonzero and pretty big, 20 are zero

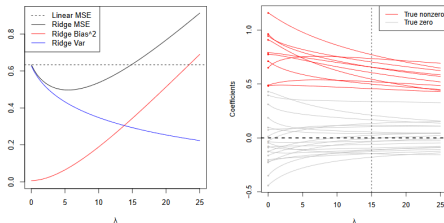


FIGURE – Ridge regression. Source : Tibshirani.

# Lasso

- Dans le lasso, la pénalité est une norme 1 et non 2 :

$$\hat{x} = \underset{x \in \mathbb{R}^p}{\operatorname{arginf}} \|Tx - y\|_2^2 + \lambda \|x\|_1,$$

où  $\|x\|_1 = \sum_{i=1}^p |x_i|$ .

- Simplement grâce à la norme 1, quand  $\lambda$  augmente, de plus en plus de paramètres sont réduits exactement à la valeur 0 : le lasso opère automatiquement une sélection de variables.
- On peut ne pas inclure certains paramètres dans la pénalité si on veut les garder.
- Tous les paramètres ne sont pas forcément à la même échelle (par exemple des prix et des dates). Il faut donc d'abord les convertir pour qu'ils soient comparables.
- Cette sélection des paramètres pertinents permet l'**interprétation** d'un modèle compliqué.

# Lasso

Our running example from last time:  $n = 50$ ,  $p = 30$ ,  $\sigma^2 = 1$ , 10 large true coefficients, 20 small. Here is a visual representation of lasso vs. ridge coefficients (with the same degrees of freedom):

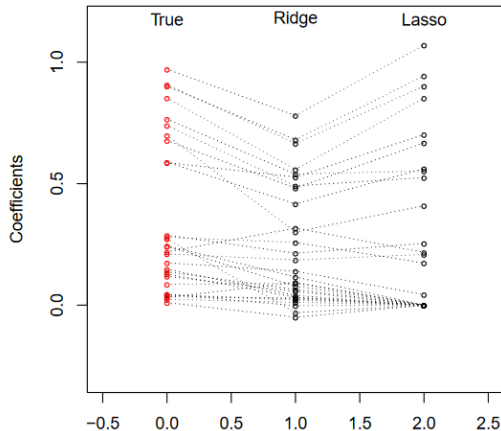


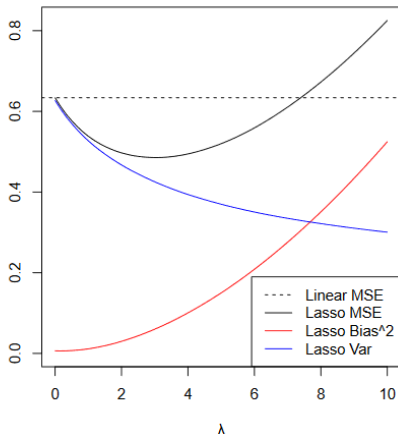
FIGURE – Source : Tibshirani



# Lasso

Interprétation en termes de lien entre  $\lambda$  et biais/variance similaire à Tikhonov : quand  $\lambda$  augmente, le biais augmente et la variance diminue.

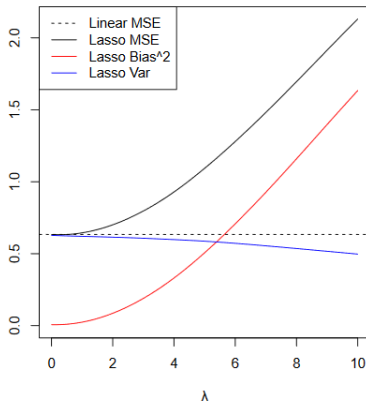
Example:  $n = 50$ ,  $p = 30$ ; true coefficients: 10 large, 20 small



# Lasso

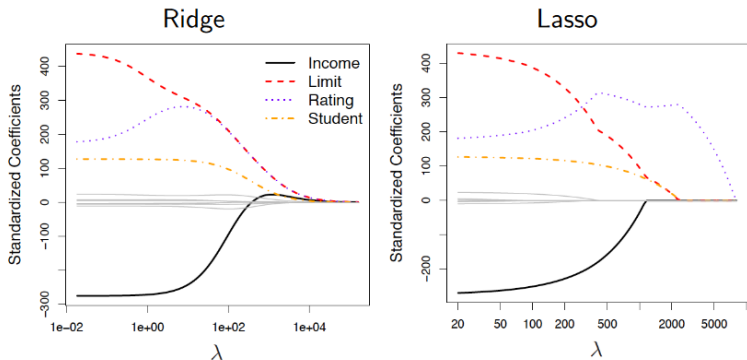
Ne fonctionne pas dans tous les cas, notamment dans ceux où parcimonie pas de mise.

Example:  $n = 50$ ,  $p = 30$ ; true coefficients: 30 moderately large



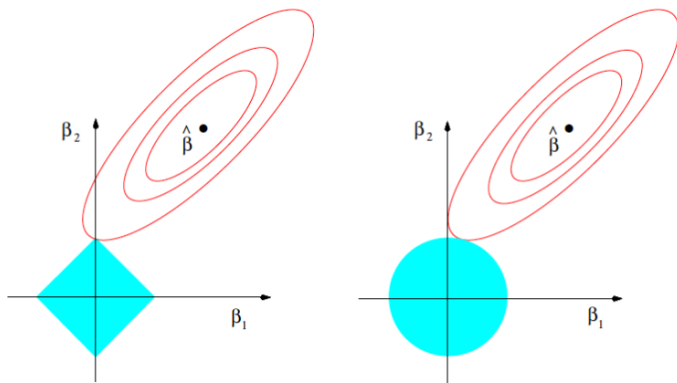
Note that here, as opposed to ridge regression the variance doesn't decrease fast enough to make the lasso favorable for small  $\lambda$

# Lasso



**FIGURE** – Variation des paramètres standardisés dans cas ridge/lasso. Source : Tibshirani.

# Lasso



**FIGURE** – Pourquoi le lasso réduit-il des paramètres à zéro ? C'est une question de normes (1 à gauche, 2 à droite). Le problème peut être vu comme optimisation sous contrainte :  $\hat{x} = \underset{x \in \mathbb{R}^p}{\operatorname{arginf}} \|Tx - y\|_2^2$  sous contrainte que  $\|x\|_1 \leq \theta$  (ou  $\|x\|_2^2 \leq \theta$  pour ridge). Quand la contrainte est saturée, à cause du pic sur la norme, l'un des deux paramètres vaut zéro. Source : Tibshirani.

# Discussion sur le risque de modèle

A partir des articles :

- Cont (2004), *Model uncertainty and its impact on the pricing of derivative instruments*
- Danielsson, James, Valenzuela, Zer (2014), *Model risk of risk models*
- Roustant, Laurent, Bay, Carraro (2003), *Model risk in the pricing of weather derivatives*
- Sibbertsen, Stahl, Luedtke (2008), *Measuring model risk*

Questions :

- En quoi consiste le risque de modèle ? Quelle sont ses sources ?
- Comment le mesurer ?

Les réponses sont autant dans les méthodes proposées par les auteurs des articles que dans la revue de littérature qu'ils font.

# Discussion sur la régulation, les crises et l'industrie financière

A partir des articles :

- Beker (2011), *On the economic crisis and the crisis of economics*
- Danielsson, Shin et Zigrand (2004), *The impact of risk regulation on price dynamics*
- Lee et Chih (2013), *Does financial regulation affect the profit efficiency and risk of banks? Evidence from China's commercial banks*
- Llewellyn (2013), *Fifty years in the evolution of bank business models*

Questions :

- Quels liens entre régulation, crises et industrie financière (par exemple : impact régulation sur activité sur un certain type de titre, ou régulation source de bulles)
- Comment les économistes peuvent-ils prévoir les crises et intégrer l'impact de la régulation dans leurs prévisions ?

Les réponses sont autant dans les méthodes proposées par les auteurs des articles que dans la revue de littérature qu'ils font.

# Discussion sur le risque de liquidité et le risque systémique

A partir des articles :

- Benoit, Colliard, Hurlin, and Pérignon (2016), *Where the risks lie : a survey on systemic risk* (cet article en groupes de 2)
- Ericsson et Renault (2006), *Liquidity and credit risk*
- Papavassiliou (2013), *A new method for estimating liquidity risk : insights from a liquidity-adjusted CAPM framework*

Questions :

- En quoi consiste le risque de liquidité (resp. systémique) ? Quelle sont ses sources ?
- Comment le mesurer ?

Les réponses sont autant dans les méthodes proposées par les auteurs des articles que dans la revue de littérature qu'ils font.

# Bibliographie

- Almgren, Chriss (2001), *Optimal execution of portfolio transactions*
- El Karoui, Gobet (2004), *Modèles stochastiques en finance*
- Embrechts, Furrer, Kaufmann (2003), *Quantifying regulatory capital for operational risk*
- Hull (2007 pour la 6ème édition), *Options, futures et autres actifs dérivés*
- McNeil, Frey, Embrechts (2005), *Quantitative risk management - concepts, techniques and tools*
- Natixis AM (2014), *Le risque de liquidité sur les marchés obligataires*