Value at Risk

Chapitre 6 : Risque de modèle

Jaouad Madkour

Faculté d'Économie et de Gestion de Tanger · 27 août 2026

Plan du chapitre

Avant de commencer

Pourquoi ce chapitre ?

Les quatre premiers chapitres ont montré, à chaque étape, que le choix d'une méthode ou d'un paramètre — mapping, distribution des facteurs, méthode de résolution, taille d'échantillon — peut faire varier la VaR d'un même portefeuille du simple au double. Ce chapitre systématise cette observation : il identifie les sources de **risque de modèle** et de **risque d'estimation**, puis présente la seule réponse rigoureuse à l'incertitude sur la qualité d'un modèle de VaR : le **backtesting**.

Les trois composantes d’un modèle de risque

Mapping, distribution, résolution

Trois sous-modèles statistiques

Un modèle de risque combine toujours trois choix indépendants : (i) le **mapping** du portefeuille sur ses facteurs de risque ; (ii) la **distribution multivariée** des rendements de ces facteurs ; (iii) la **méthode de résolution** (analytique, historique ou Monte Carlo) qui transforme (i) et (ii) en une mesure de risque. Ces choix sont interdépendants : sous hypothèse i.i.d. normale multivariée avec mapping linéaire, la résolution analytique est optimale — la simulation n'apporterait qu'une erreur d'échantillonnage inutile à un problème qui a une solution fermée.

Risque de modèle vs risque d'estimation

Le **risque de modèle** porte sur le choix des trois sous-modèles eux-mêmes : quelle distribution, quel mapping, quelle méthode de résolution ? Le **risque d'estimation** porte sur la façon d'estimer les paramètres, une fois les choix de modèle fixés — il inclut l'erreur d'échantillonnage mais aussi le choix entre plusieurs méthodes d'estimation compatibles avec les mêmes hypothèses (par exemple, matrice de covariance égalitaire ou EWMA sous hypothèse normale i.i.d.).

Risque de modèle lié au mapping

Taux d’intérêt : un risque de mapping mineur

Exemple : choix des points de la courbe des taux

Pour un flux de trésorerie mappé sur trois points fixes d’une courbe de taux (méthode invariante en valeur actuelle, PV01 et volatilité), le choix des points eux-mêmes — mensuels ou trimestriels — a un effet minime sur la VaR normale linéaire : moins de 0,3 % d’écart relatif dans un exemple typique. En VaR linéaire normale, la VaR est proportionnelle à la volatilité du portefeuille, une quantité invariante au mapping choisi.

Actions : un risque de mapping considérable

Exemple : estimation du bêta d'une action

Pour une position sur une action mappée à un indice de marché, huit méthodes d'estimation du couple (bêta, volatilité de l'indice) — MCO hebdomadaire ou quotidien, sur échantillon long ou court, EWMA avec différentes constantes de lissage — donnent, sur un même exemple réel (HBOS PLC, mai 2008), des VaR à 1 % à 10 jours allant de **4,72 % à 14,68 %** de la valeur du portefeuille, soit un facteur supérieur à 3 entre l'estimation la plus prudente et la plus agressive. Contrairement au cas des taux, le choix de la méthode de mapping actions est une source majeure de risque de modèle.

Pourquoi un tel écart ?

L'EWMA à faible constante de lissage réagit fortement aux données récentes : après un krach, elle peut afficher un bêta **plus faible** (car la corrélation action-indice s'effondre pendant la panique) mais une volatilité de l'indice **plus élevée** — un résultat contre-intuitif où une action à haut risque affiche paradoxalement une VaR systématique modeste. Ce risque de mapping s'ajoute, sans s'y substituer, au risque de modèle propre à la distribution des rendements.

Choix du modèle de rendement des facteurs

Quatre questions clés

Les choix structurants pour les rendements des facteurs

Au-delà du mapping, l'analyste doit trancher quatre questions indépendantes : (1) les rendements sont-ils i.i.d. ou faut-il capturer *clustering* de volatilité et autocorrélation ? (2) la distribution est-elle paramétrique ou empirique (historique) ? (3) si paramétrique, quelle forme — normale, Student t, mélange, copule ? (4) comment estimer les paramètres — quel échantillon, quelle méthode ?

Ce que les quatre premiers chapitres ont montré

L'autocorrélation positive **augmente** la VaR mise à l'échelle, la négative la **diminue**. Le *clustering* de volatilité rend la VaR plus sensible au risque courant. Les distributions à queue épaisse et asymétriques augmentent la VaR aux niveaux de confiance élevés, mais peuvent la réduire aux niveaux de confiance modérés. Enfin, la méthode d'estimation (taille d'échantillon, constante de lissage) peut, à elle seule, doubler ou diviser par deux une estimation de VaR ou d'ETL.

Quatre questions clés (suite)

L'horizon de risque change la donne

Les rendements non normaux et le *clustering* de volatilité importent surtout pour la VaR à **court terme** (quelques jours). Sur un horizon d'un mois ou plus, le théorème central limite rapproche la distribution agrégée de la normalité — même en présence de *clustering* — si bien que l'hypothèse normale i.i.d. redevient raisonnable pour la VaR de long terme. Le principal risque de modèle à long terme n'est alors plus la forme de la distribution, mais la **mise à l'échelle inappropriée** d'une VaR courte à un horizon long via la racine du temps.

Risque d’estimation : intervalles de confiance pour la VaR

Erreur type de la VaR sous normalité

Pour un portefeuille censé rentabiliser au taux sans risque, la VaR normale linéaire est proportionnelle à l'écart-type σ du portefeuille : $\mathrm{VaR}_{\alpha,h} = \Phi^{-1}(1-\alpha)\,\sigma\sqrt{h}$. L'erreur type de l'estimateur de volatilité égalitaire, basé sur un échantillon de taille T, se transmet directement à la VaR :

$$\mathrm{s.e.}(\mathrm{VaR}_{\alpha,h}) \approx \frac{\mathrm{VaR}_{\alpha,h}}{\sqrt{2T}}.$$

Une formule analogue existe pour un estimateur EWMA de constante de lissage λ. Plus l'échantillon est grand (ou λ proche de 1), plus l'intervalle de confiance de la VaR est étroit — mais, comme au chapitre 3, un échantillon plus long est aussi moins réactif aux conditions de marché actuelles.

Exemple

Pour une volatilité de portefeuille de 20 %, une VaR à 1 % à 10 jours de 9,31 % a une erreur type d'environ 0,66 % avec un échantillon égalitaire de 100 observations, mais de 1,16 % avec une EWMA de constante 0,94 — l'estimateur EWMA, bien que plus réactif, est ici **moins précis** (erreur type quasiment double) que l'estimateur égalitaire à cette taille d'échantillon.

Erreur type d'un estimateur de quantile

Pour la VaR historique, l'erreur type du quantile empirique s'obtient sous une hypothèse de densité localement plate autour du quantile visé :

$$\text{s.e.}(q(T,\alpha)) \approx \frac{1}{f(q(T,\alpha))}\sqrt{\frac{\alpha(1-\alpha)}{T}},$$

où f est la densité de la distribution des rendements du portefeuille. Cette approximation est nettement moins précise que celle fondée sur la volatilité : à niveau de confiance et taille d'échantillon égaux, l'erreur type d'un quantile est environ **deux fois plus grande** que celle d'un estimateur de VaR basé sur la volatilité — et cet écart s'accroît aux niveaux de confiance extrêmes.

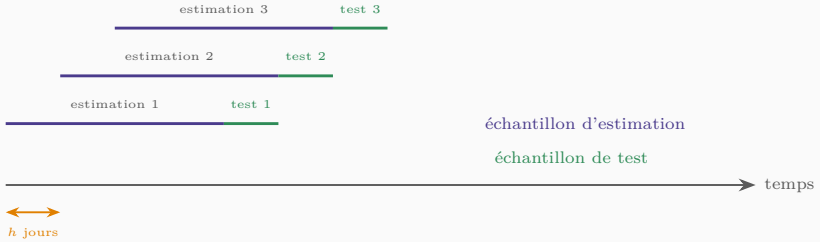
Monte Carlo : deux sources d'erreur

La VaR Monte Carlo cumule une erreur d'échantillonnage (contrôlable en augmentant le nombre de simulations ou via les techniques de réduction de variance du chapitre 4) et le risque d'estimation classique lié aux paramètres du modèle statistique. Contrairement à la simulation historique — où seule la taille de l'échantillon compte, puisqu'il n'y a aucun paramètre à estimer — le risque de modèle et le risque d'estimation des paramètres dominent généralement l'erreur d'échantillonnage en Monte Carlo.

Méthodologie du backtesting

Backtest et fenêtre glissante

Un backtest confronte les prévisions de VaR d'un **portefeuille candidat** (aux poids ou sensibilités fixes) à ses pertes effectivement réalisées, sur une longue période historique. La méthode standard utilise une **fenêtre glissante** : on estime le modèle sur une période d'estimation fixe, on prévoit la VaR sur la période de test qui suit immédiatement, on avance les deux fenêtres de h jours (l'horizon de risque), et on répète jusqu'à épuiser l'échantillon — évitant ainsi le chevauchement des données de test.

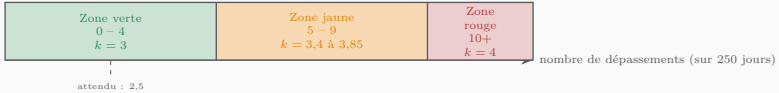


La fonction indicatrice de dépassement

On définit un **dépassement** (*exceedance*) comme un jour où la perte réalisée excède la VaR prévue la

Les zones de Bâle

Le Comité de Bâle recommande un backtest simple : VaR à 1 % à 1 jour, sur 250 jours (nombre attendu de dépassements : 2,5). Les modèles ayant 4 dépassements ou moins sont en **zone verte** (multiplicateur 3) ; entre 5 et 9, en **zone jaune** (multiplicateur croissant de 3,4 à 3,85) ; 10 ou plus, en **zone rouge** (multiplicateur 4, ou modèle interdit).



Un test de faible puissance

Avec seulement 250 observations, la puissance de ce test — sa capacité à rejeter un modèle réellement inexact — est très faible. Un modèle imprécis a donc de bonnes chances de passer le test réglementaire. De plus, la VaR réglementaire utilisée pour le capital est une VaR à 10 jours, mais le backtest porte sur la VaR à 1 jour car un backtest à 10 jours nécessiterait des données sur plus de 10 ans pour être suffisamment précis — un compromis pragmatique qui affaiblit encore la portée du test.

Tests de couverture inconditionnelle et d'indépendance

Test de couverture inconditionnelle (Kupiec, 1995)

Le test de Kupiec confronte la proportion observée de dépassements $\hat{\pi}_1 = n_1/n$ à la proportion attendue α, via un rapport de vraisemblance :

$$-2\ln LR_{uc} = -2\ln\left[\frac{\alpha^{n_1}(1-\alpha)^{n_0}}{\hat{\pi}_1^{n_1}(1-\hat{\pi}_1)^{n_0}}\right] \sim \chi^2_1,$$

où n_1 est le nombre de dépassements et $n_0 = n - n_1$. Une valeur trop élevée conduit à rejeter l'hypothèse que le modèle a la couverture annoncée.

Test d'indépendance (Christoffersen, 1998)

Un modèle peut avoir le bon nombre de dépassements tout en les concentrant dans le temps — signe qu'il ne s'adapte pas assez vite aux conditions de marché changeantes. Le test d'indépendance de Christoffersen compare la probabilité d'un dépassement sachant qu'il y en a eu un la veille, π_1, à la probabilité d'un dépassement sachant qu'il n'y en a pas eu, π_0 : sous l'hypothèse d'indépendance, $\pi_0 = \pi_1 = \alpha$. Le test de **couverture conditionnelle**, combinant les deux, vérifie que $-2\ln LR_{cc} = -2\ln LR_{uc} - 2\ln LR_{ind} \sim \chi^2_2$.

Exemple : S&P 500, 2000-2007

Sur 2000 jours de VaR normale linéaire à 1 % pour une position sur le S&P 500, on observe 33 dépassements (attendu : 20). Le test de couverture inconditionnelle rejette le modèle à 1 % ($-2 \ln LR_{uc} = 7,14$, contre un seuil critique de 6,63). Le test d'indépendance simple, lui, ne détecte pas le regroupement pourtant visible des dépassements (concentrés en 2007, lors de la crise du crédit) : il ne capture que les dépendances de premier ordre, entre jours **consécutifs** — un regroupement espacé de quelques jours lui échappe.

Diagnostiquer la cause d'un échec

Si l'indicatrice de dépassement peut être prédite par une variable retardée x_{t-1} (par exemple, une variation de la volatilité implicite), le modèle n'utilise pas toute l'information disponible. Le backtest régresse l'indicatrice sur x_{t-1} et teste $H_0 : \beta_0 = \alpha,\ \beta_1 = \cdots = \beta_p = 0$ via un test F. Ce type de test ne se contente pas de rejeter le modèle : il aide à **comprendre pourquoi** il échoue, par exemple en révélant un lien entre les dépassements et les sauts de volatilité implicite — suggérant qu'un modèle intégrant le *clustering* de volatilité (EWMA, GARCH) améliorerait les prévisions.

Backtesting de l’ETL et statistiques de biais

Test de McNeil et Frey (2000)

Pour vérifier que l'ETL n'est pas systématiquement sous-estimée, on calcule les **résidus de dépassement standardisés**, définis uniquement sur les jours de dépassement :

$$\xi_t = \frac{-Y_{t+1} - \mathrm{ETL}_{\alpha,t}}{\widehat{\sigma}_t},$$

puis on teste si leur moyenne est nulle (contre l'alternative qu'elle est positive, signe que l'ETL sous-estime la perte moyenne en queue). Ce test perd beaucoup de puissance car il n'utilise que les observations de dépassement — nécessitant des historiques très longs (plus de 10 ans) pour disposer d'un nombre suffisant d'observations.

Statistique de biais

Le biais de Connor (2000)

Sous hypothèse normale i.i.d., le rendement standardisé $Z_t = Y_{t+1}/\hat{\sigma}_t$ devrait avoir un écart-type unitaire si la prévision de volatilité $\hat{\sigma}_t$ est exacte. La statistique de biais b est l'écart-type empirique de $\{Z_t\}$ sur l'échantillon de backtest ; $b < 1$ signale une **sous-estimation** de la VaR, $b > 1$ une **sur-estimation**.

Une mise en garde importante

La statistique de biais n'est valide que si l'hypothèse de normalité i.i.d. sous-jacente est elle-même correcte. Sur des données réelles où cette hypothèse est violée, elle peut conduire à des conclusions **directement contraires** à celles des tests de couverture : un exemple sur le S&P 500 trouve $b > 1$ (VaR prétendument sur-estimée) alors que les tests de couverture, sur les mêmes données, indiquent une sous-estimation flagrante (trop de dépassements). Toujours tester la normalité et l'indépendance des rendements avant d'utiliser une statistique de biais.

Au-delà de la seule queue

Un test de distribution complète évalue la qualité de **toute** la distribution prévue, pas seulement son quantile de VaR. On transforme chaque rendement réalisé en sa probabilité prévue $p_{t,h} = F_t(Y_{t+h})$, censée suivre une loi uniforme $[0, 1]$ si le modèle est exact. Un modèle qui sous-estime systématiquement les queues produit un histogramme des $p_{t,h}$ en forme de W (trop de valeurs près de 0, 1 et — pour une distribution trop pointue — près de 0,5) plutôt que plat.

Résultats empiriques

L'étude d'Alexander et Sheedy (2008)

Huit modèles, trois paires de devises

Une étude de référence backteste huit modèles de VaR (normal, Student t, historique à noyau, mélange de normales — chacun avec et sans *clustering* GARCH) sur trois paires de devises majeures. Les résultats sont sans appel : les modèles combinant **distribution à queue épaisse et clustering de volatilité** (Student t GARCH, simulation historique filtrée) dominent systématiquement les autres, en particulier pour l'ETL. Les modèles normaux — même avec GARCH — sous-estiment presque toujours le risque de queue.

Synthèse

Ce qu'il faut retenir

L'essentiel du chapitre 6

- Un modèle de risque combine trois choix indépendants — mapping, distribution des facteurs, méthode de résolution — chacun source potentielle de **risque de modèle** ; l'estimation des paramètres de ces choix est source de **risque d'estimation**.
- Le risque de mapping est mineur pour les taux, mais considérable pour les actions (choix de la méthode et de l'échantillon d'estimation du bêta).
- L'erreur type d'une VaR fondée sur la volatilité est nettement plus précise que celle d'un estimateur de quantile historique, à échantillon égal.
- Le **backtesting** — confrontation systématique des prévisions aux réalisations sur fenêtre glissante — est le seul moyen rigoureux de juger la qualité d'un modèle de VaR.
- Les tests de Kupiec (couverture), Christoffersen (indépendance, couverture conditionnelle), les backtests de régression, le test ETL de McNeil-Frey et les statistiques de biais offrent des diagnostics complémentaires, chacun avec ses limites.
- L'évidence empirique converge : les modèles avec *clustering* de volatilité et distributions à queue épaisse produisent des VaR et ETL bien plus précises que les modèles normaux statiques.

Vers le chapitre 7