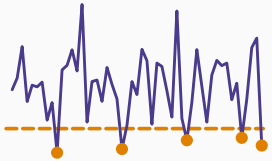
Mesure et gestion du risque de marché

Chapitre 4 : Le contrôle a posteriori de la VaR



Jaouad Madkour

Faculté d'Économie et de Gestion de Tanger · 30 août 2026

Plan du chapitre

Avant de commencer

Le principe

Le test de défaillance

Les règles de Bâle

La couverture conditionnelle

Choix et constats pratiques

Synthèse

Avant de commencer

Pourquoi ce chapitre

Pourquoi ce chapitre ?

Une VaR n'a de valeur que si elle est **bien calibrée**. Le contrôle a posteriori — le backtesting — compare systématiquement les VaR annoncées aux pertes effectivement observées. C'est sur cette discipline que repose l'autorisation, accordée par les régulateurs, d'utiliser les modèles internes des banques pour le calcul du capital réglementaire.

Le principe

Compter les dépassements

L'exception

Une **exception** (ou dépassement) survient lorsque la perte réalisée excède la VaR annoncée la veille. Pour une VaR à 99 %, on en attend environ 1 % des jours. Trop d'exceptions signale un modèle qui **sous-estime** le risque ; trop peu, un modèle inutilement **conservateur**, qui immobilise du capital en excès.

Un exemple bien calibré

Sur un graphique opposant le P&L quotidien absolu à la VaR à 99 %, environ 2 % des observations devraient dépasser la diagonale — soit environ 5 points sur une année. En observer 4 signale un modèle globalement bien calibré, même si le nombre exact varie par pur hasard d'une année à l'autre.

Rendement réel ou hypothétique ?

Deux notions de rendement

Le rendement **hypothétique** correspond à un portefeuille figé sur l'horizon — c'est lui que la VaR prédit réellement. Le rendement **réel** intègre les transactions intrajournalières, commissions et revenus d'intérêts nets — c'est lui que surveillent les régulateurs. Idéalement, on utilise les deux.

Ce que révèle un écart entre les deux

Si le modèle passe le contrôle avec le rendement hypothétique mais pas avec le rendement réel, le problème vient de l'activité **intrajournalière**, pas du modèle de risque lui-même. L'inverse indique une méthodologie de modélisation à revoir en profondeur.

Le test de défaillance

Le nombre d'exceptions suit une loi binomiale

Sous l'hypothèse que le modèle est correctement calibré au niveau p, le nombre d'exceptions N sur T jours suit une loi binomiale d'espérance pT et de variance $p(1-p)T$. Pour T grand, l'approximation normale

$$z = \frac{N - pT}{\sqrt{p(1-p)T}}$$

fournit un test simple, qui ne fait **aucune hypothèse** sur la forme de la distribution des rendements — seul le comptage des exceptions importe.

La divulgation de J.P. Morgan, 1998

En 1998, J.P. Morgan a observé 20 jours où le résultat est sorti de la bande de VaR à 95 %, sur 252 jours — alors que $0{,}05 \times 252 = 12{,}6$ étaient attendus. Le test donne $z = (20 - 12{,}6)/\sqrt{0{,}05 \times 0{,}95 \times 252} = 2{,}14$, supérieur au seuil critique de 1,96 : le modèle est rejeté au niveau de confiance de test de 95 %. La banque a d'ailleurs adopté par la suite un modèle de simulation historique, mieux à même de capter les queues épaisses.

Les règles de Bâle

Les zones de pénalité

Vert, jaune, rouge

Sur une VaR à 99 % et $T = 250$ jours, le comité de Bâle tolère jusqu'à **quatre** exceptions sans pénalité — la zone verte. À partir de **cinq**, la banque entre en zone jaune, où un facteur multiplicatif croissant s'applique au capital requis ; à partir d'un seuil plus élevé, en zone rouge, avec pénalité automatique.

Pourquoi ce seuil de quatre

Avec $p = 0{,}01$ et $T = 250$, on attend en moyenne 2,5 exceptions par an. Fixer le seuil à quatre — plutôt qu'à trois, ce que suggérerait naïvement l'espérance — revient à accepter un **taux d'erreur de type 1** de 10,8 % : la probabilité de pénaliser à tort une banque dont le modèle est en réalité correct.

L'arbitrage entre deux erreurs

Erreur de type 1 et de type 2

L'**erreur de type 1** rejette un modèle correct — la banque subit une pénalité injustifiée. L'**erreur de type 2** ne rejette pas un modèle incorrect — un régulateur laisse passer une sous-estimation du risque. Concevoir un test, c'est choisir le point d'équilibre entre ces deux risques.

La faible puissance du cadre bâlois

Au niveau de confiance de 99 % retenu par Bâle, le test manque de puissance : avec cinq exceptions ou plus comme seuil de rejet, l'erreur de type 1 atteint déjà 10,8 %, et l'erreur de type 2 — la probabilité de ne pas rejeter un modèle dont la vraie couverture n'est que de 97 % — culmine à 12,8 %. Ce cadre a du mal à distinguer un modèle correct d'un modèle légèrement défaillant.

Un remède simple : abaisser le niveau de confiance

En passant d'une VaR à 99 % à une VaR à 95 %, le nombre d'exceptions attendu monte à environ 13 par an — largement assez pour qu'un test statistique retrouve de la puissance. Avec un seuil de rejet à 17 exceptions, l'erreur de type 1 reste comparable à celle du cadre bâlois, mais l'erreur de type 2 chute à 7,4 % contre 12,8 % à 99 %. Le prix : il faut alors un multiplicateur k plus élevé pour retrouver le même niveau de capital.

La couverture conditionnelle

Au-delà du simple comptage

Pourquoi compter ne suffit pas

Un modèle correctement calibré doit produire des exceptions **indépendantes** dans le temps. Si dix exceptions sur vingt surviennent en l'espace de deux semaines, le nombre total peut sembler acceptable alors que leur **regroupement** trahit un problème — une hausse de volatilité non captée, ou des positions déplacées vers des zones de risque non couvertes.

Le test de Christoffersen

Ce test étend le cadre binomial en comparant la probabilité d'une exception sachant qu'il y en a eu une la veille à la probabilité d'une exception sachant qu'il n'y en a pas eu. Le test combiné de **couverture conditionnelle**, $LR_{cc} = LR_{uc} + LR_{ind}$, suit asymptotiquement une loi du χ^2 à deux degrés de liberté ; on rejette au niveau de test de 95 % si $LR_{cc} > 5{,}991$.

Regroupement chez J.P. Morgan

Sur les 20 exceptions de 1998, 6 sont survenues au lendemain d'une autre exception. Cela donne une probabilité conditionnelle d'exception après exception de 30 %, contre 6 % après un jour sans exception — un écart considérable. Le test d'indépendance donne $LR_{ind} = 9{,}53$, largement au-delà du seuil de 3,84 : l'hypothèse d'indépendance est rejetée, les exceptions se regroupent anormalement.

Choix et constats pratiques

Les paramètres du test

Deux recommandations

L'horizon de backtesting doit rester **court** — typiquement un jour — afin de maximiser le nombre d'observations et de limiter les effets d'un changement de composition du portefeuille. Le niveau de confiance ne doit pas être trop **élevé**, sous peine de sacrifier la puissance du test, comme le montre l'écart entre 95 % et 99 % ci-dessus.

Les banques sont souvent trop conservatrices

Des études empiriques sur des banques américaines montrent une VaR déclarée souvent supérieure de 60 % ou plus au 99$^{?}$ centile réellement observé des P&L, et un nombre d'exceptions anormalement faible même durant des périodes de marché agitées. Les distributions de P&L des banques sont généralement symétriques mais présentent des queues plus épaisses que la normale — un rappel que le modèle sous-jacent, même prudent, reste rarement parfaitement calibré.

Synthèse

Ce qu'il faut retenir

Synthèse du chapitre

Le backtesting compare systématiquement les VaR annoncées aux pertes réalisées, idéalement sur des rendements à la fois **réels** et **hypothétiques**. Le **test par taux de défaillance**, fondé sur la loi binomiale, ne fait aucune hypothèse de distribution mais souffre d'un arbitrage difficile entre erreurs de **type 1** et de **type 2** — arbitrage que les règles de **Bâle**, avec leurs zones verte, jaune et rouge à 99 %, résolvent au prix d'une puissance statistique limitée, améliorable en abaissant le niveau de confiance testé. Le simple comptage ne suffit d'ailleurs pas : la **couverture conditionnelle** vérifie aussi l'indépendance des exceptions dans le temps, un regroupement trahissant un risque non capté par le modèle.