

## **Nouvelle réglementation internationale du risque de marché : rôles de la VaR et de la CVaR dans la validation des modèles**

**Samir Saissi Hassani  
Georges Dionne**

**Novembre 2020**

**Bureau de Montréal**  
Université de Montréal  
C.P. 6128, succ. Centre-Ville  
Montréal (Québec) H3C 3J7  
Tél : 15 M-343-7575  
Télécopie : 15 M-343-7121

**Bureau de Québec**  
Université Laval  
2325, rue de la Terrasse  
Pavillon Palasis-Prince, local 2415  
Québec (Québec) G1V 0A6  
Tél : 14 B-656-2073  
Télécopie : 14 B-656-2624

# Nouvelle réglementation internationale du risque de marché : rôles de la VaR et de la CVaR dans la validation des modèles

Samir Saissi Hassani<sup>1</sup>, Georges Dionne<sup>1,2,\*</sup>

<sup>1</sup> Chaire de recherche du Canada en gestion des risques, HEC Montréal

<sup>2</sup> Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT)

**Résumé.** Dans cette note, nous modélisons les nouveaux aspects quantitatifs de la gestion du risque de marché des banques que Bâle a décidés en 2016 et mis en vigueur en janvier 2019. Le risque de marché est mesuré par la valeur à risque conditionnelle, ou CVaR, à un degré de confiance de 97,5 %. Le backtest réglementaire reste, en grande partie, basé sur la VaR à 99 %. De plus, à titre de procédures statistiques supplémentaires et tel que suggéré par Bâle, des backtests complémentaires sur la VaR et la CVaR doivent être effectués. Nous appliquons ces tests sur différentes distributions paramétriques et utilisons des mesures non paramétriques de la CVaR, dont la CVaR<sup>-</sup> et la CVaR<sup>+</sup> comme compléments de validation des distributions utilisées. Nos données sont relatives à une période de turbulences extrêmes des marchés. Avec huit distributions paramétriques mises à l'épreuve par ces données, nos résultats montrent que l'information obtenue sur leurs performances empiriques est très liée aux conclusions des backtests des modèles.

**Mots-clés :** Bâle, VaR, CVaR, backtest, modèle paramétrique, modèle non paramétrique, mélange de distributions, distribution à queue épaisse.

**Remerciements.** Recherche financée par la Chaire de recherche du Canada en gestion des risques. Nous remercions Claire Boisvert pour son aide dans la préparation du manuscrit.

Results and views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect those of CIRRELT.

Les résultats et opinions contenus dans cette publication ne reflètent pas nécessairement la position du CIRRELT et n'engagent pas sa responsabilité.

---

\* Auteur correspondant : [georges.dionne@hec.ca](mailto:georges.dionne@hec.ca)

## Introduction

En 2016, le Comité de Bâle a décidé que le capital de couverture du risque de marché des banques devait être calculé avec la CVaR au degré de confiance 97,5 % en maintenant le *backtest* des modèles, comme auparavant, sur la VaR à 99 % (BCBS, 2016, 2019). Ce changement vers la CVaR serait motivé par des problèmes de cohérence de la VaR et l'insuffisance de la couverture du risque par cette mesure, constatée au cours du temps. Désormais, le risque de marché est conjointement géré par la CVaR et la VaR, à deux probabilités différentes :  $p = 2,5 \%$  et  $p = 1 \%$  respectivement.<sup>1</sup>

D'autre part, Bâle suggère d'ajouter des procédures statistiques utiles pour s'assurer de l'adéquation des modèles en ex-post (BCBS, 2016, page 82 ; BCBS, 2019, paragraphe 32.13). Nous déployons donc quatre *backtests* complémentaires au *backtest* de la VaR à 1 %, dont deux sur la VaR à  $p = 2,5 \%$  et  $p = 5 \%$ , et deux autres sur la CVaR à  $p = 2,5 \%$  et  $p = 5 \%$ . Nous utilisons des mesures non paramétriques, dont la CVaR<sup>-</sup> et la CVaR<sup>+</sup>, comme compléments de validation des distributions utilisées. L'objectif de cette note est d'orchestrer tous ces aspects dans une démarche de validation compatible avec la réglementation en vigueur.

Nous travaillons avec des données réelles obtenues de trois titres risqués : IBM, General Electric et Wal-Mart, dont les fluctuations de prix confirment l'invalidité des hypothèses habituelles de normalité des rendements durant la période utilisée. La période choisie englobe les fluctuations extrêmes des prix lors de la dernière récession économique aux États-Unis (NBER<sup>2</sup>, de décembre 2007 à juin 2009) et de la crise financière de 2007-2009. Nous évaluons le comportement de la VaR et de la CVaR en faisant appel à plusieurs distributions paramétriques pour modéliser les rendements<sup>3</sup> : la loi normale, la *t* de Student, la loi EGB2 (GB2 exponentielle), la SN2 (*Skewed Normal* de type 2), et enfin la SEP3 (*Skewed Exponential Power* de type 3)<sup>4</sup>. Nous construisons également des mélanges (*mixtures*) de densités paramétriques homogènes ou

---

<sup>1</sup> Dans l'article, nous utilisons la lettre *p* pour désigner la probabilité que la VaR soit dépassée et  $1-p$  pour le degré de confiance correspondant. Nous utilisons la notation *p*-value pour les tests statistiques.

<sup>2</sup> <https://www.nber.org/cycles.html>.

<sup>3</sup> La CVaR est aussi appelée *Expected Shortfall* dans la littérature. Les deux mesures sont équivalentes avec des distributions continues sans sauts (Rockafellar et Uryasev, 2002). Voir aussi Dionne (2019).

<sup>4</sup> Pour une description de SN2 et SEP3, voir Fernandez et al, (1995) et Rgiby et al, (2014).

hétérogènes. Huit modèles sont analysés afin d'identifier la ou les distributions qui représentent le mieux les données pour gérer le risque de marché qu'elles contiennent.

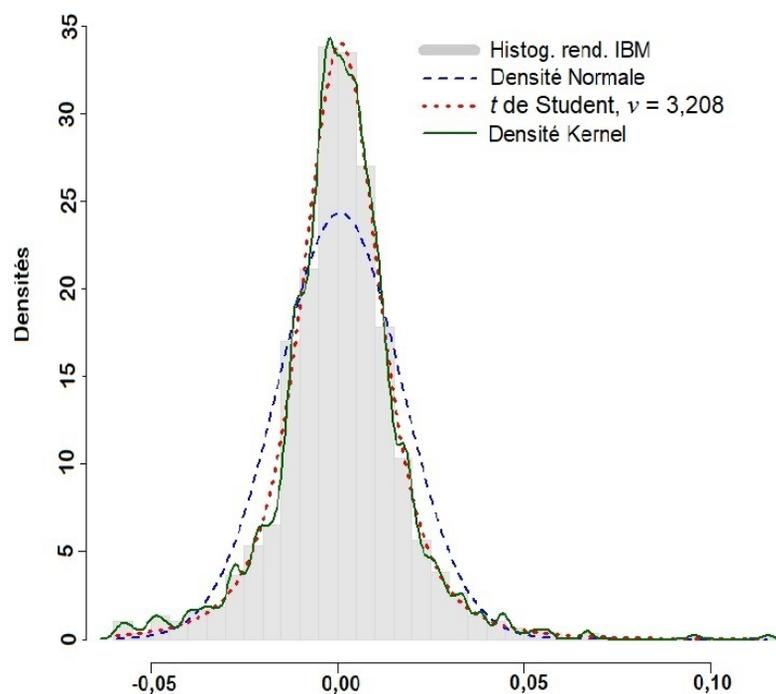
L'analyse des modèles se fait en trois étapes. D'abord, l'estimation des paramètres des modèles est validée par les mesures standards telles que les critères AIC, BIC et le test de Kolmogorov-Smirnov. La deuxième validation consiste en la comparaison du kurtosis et de l'asymétrie obtenus des modèles paramétriques en comparaison aux mêmes moments déterminés par une approche non paramétrique des données. Le point le plus important dans cette étape est l'encadrement des CVaR paramétriques de chaque modèle par l'intervalle  $[CVaR_{-np}, CVaR_{+np}]$ , où  $np$  désigne non paramétrique. Comme la  $CVaR_{np}$  obéit aux inégalités fondamentales  $CVaR_{-np} \leq CVaR_{np} \leq CVaR_{+np}$  (Rockafellar et Uryasev, 2002), un bon modèle paramétrique devrait aussi assurer le même encadrement. La troisième validation est le *backtesting* des mesures de risque que nous effectuons en respectant la réglementation de Bâle en vigueur pour le risque de marché. De plus, nous ajoutons des *backtests* complémentaires pour les VaR et les CVaR pour s'assurer de la qualité de l'adéquation des modèles aux données. Nous constatons que les résultats des deux dernières étapes sont liés ; ne pas valider qu'une distribution statistique représente adéquatement les données affecte significativement les résultats du *backtesting*.

Les calculs de la CVaR paramétrique étant nettement plus complexes que ceux de la VaR, nous devons être précis dans les définitions et les dérivations mathématiques concernant chacune des lois ou mélanges de lois que nous employons dans les huit modèles. Les développements mathématiques sont exposés dans les annexes. L'annexe A1 décrit les symboles des différents modèles. L'annexe A2 présente l'expression générale de la CVaR, alors que l'annexe A3 présente l'expression générale d'une CVaR provenant d'un mélange de distributions. Les détails des propriétés des modèles statistiques et de la procédure du *backtesting* sont présentés dans les annexes qui suivent.

La section suivante présente les données utilisées. La section 2 offre une analyse préliminaire des données. La section 3 est vouée à l'estimation des paramètres des huit modèles concurrents et à la vérification empirique de leurs performances respectives. La section 4 procède aux *backtests* des modèles et la dernière section conclut.

## 1. Données

Commençons par la présentation des données. Les trois titres risqués choisis sont IBM, General Electric (GE) et Wal-Mart (WM). La période est constituée de 1 200 jours, du 18 juin 2007 au 20 mars 2012<sup>5</sup>. Les rendements sont calculés en tenant compte des paiements de dividendes. La figure 1 montre la distribution des rendements des titres de cette période, qui sont loin d'être normaux. En ajustant les rendements à la loi  $t$  de Student, les degrés de liberté ( $\nu$ ) sont égaux à 3,2, 2,4 et 3,2 respectivement, ce qui indique la présence d'une épaisseur de queue très importante.



<sup>5</sup> En fait, la période d'extraction des prix quotidiens est du 15 juin 2007 au 20 mars 2012, ce qui représente 1 201 jours, permettant d'avoir 1 200 rendements quotidiens du 18 juin 2007 au 20 mars 2012.

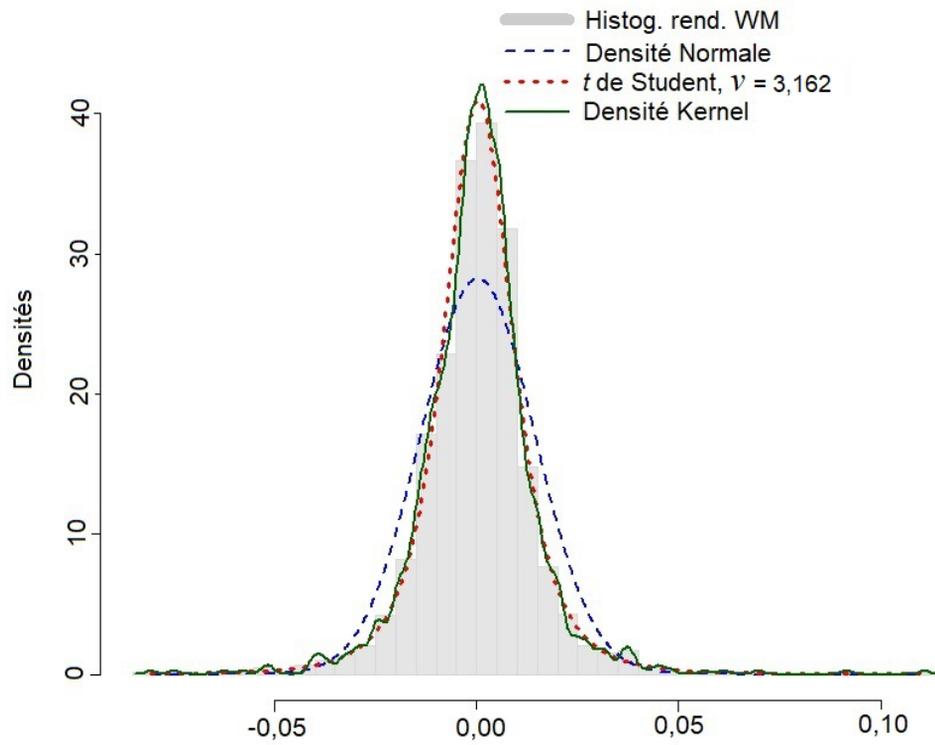
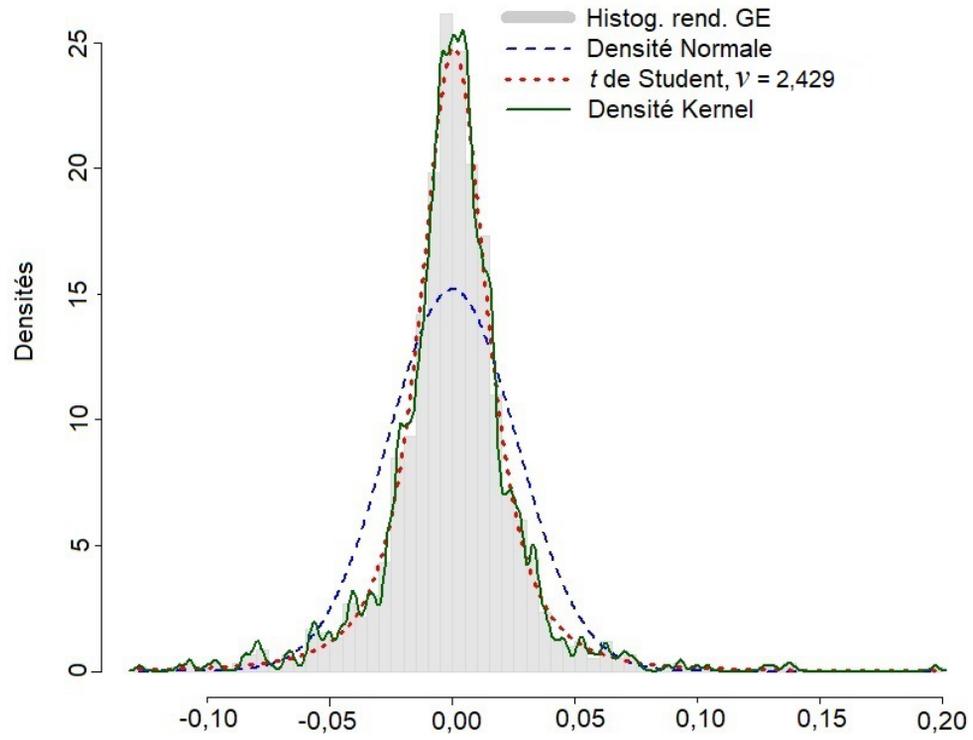


Figure 1: Histogrammes et densités des titres IBM, GE et WM

Le tableau 1 présente des statistiques descriptives dont les matrices de corrélations, de variances-covariances et les quatre premiers moments non paramétriques. Les corrélations positives sont très fortes durant cette période de crise financière, étant d'environ 50 %.

Tableau 1  
Matrices des corrélations, variances-covariances et moments non paramétriques

	IBM	General Electric	Wal Mart
IBM	1		
General Electric	0,567592	1	
Wal Mart	0,491463	0,430835	1
IBM	0,02688 %	0,02443 %	0,01138 %
General Electric	0,02443 %	0,06894 %	0,01598 %
Wal Mart	0,01138 %	0,01598 %	0,01995 %
Moyenne	0,07580 %	-0,00286 %	0,03472 %
Variance	0,02688 %	0,06894 %	0,01995 %
Asymétrie	0,27190	0,35375	0,35429
Kurtosis	7,43415	9,95718	10,68244

Les moyennes des rendements quotidiens sont pratiquement nulles. Les coefficients d'asymétrie sont positifs, montrant que les trois distributions sont tirées vers des rendements positifs. Ceci est surprenant, car on aurait pu s'attendre à ce que ce soit plutôt vers la queue gauche des pertes durant cette période de pertes extrêmes. Le point le plus important est que les coefficients de kurtosis sont très grands, de l'ordre de trois fois plus grands que le kurtosis d'une distribution normale. Cela confirme l'épaisseur de queue très importante signalée dans les paragraphes précédents.

## 2. Analyse préliminaire des données

Nous commençons par calculer les poids optimaux d'un portefeuille A qui minimise la VaR relative à un degré de confiance de 95% ( $p = 5\%$ ) sous contrainte que les poids somment à 1. Les rendements sont supposés normaux pour le moment.

La VaR relative du portefeuille s'écrit :

$$\text{VaR}_r = -\sigma_{\text{portefeuille}} \times \Phi_0^{-1}(p) = -\sqrt{\beta^T \Sigma \beta} \times q_0 > 0 \quad (1)$$

où  $\Phi_0^{-1}(\cdot)$  est l'inverse de la fonction cumulative de la loi normale  $N(0,1)$  évaluée à  $p$ ,  $\beta$  est le vecteur des poids des titres,  $\beta^T$  est la transposée de  $\beta$ ,  $\Sigma$  est la matrice variance-covariance des rendements des titres et  $q_0 (q_0 < 0)$  est le quantile de la densité  $N(0,1)$  relatif à  $p$ . L'expression de la VaR est  $> 0$  puisque  $q_0 < 0$ . Le fichier Excel donne les résultats du tableau 2 en pourcentages pour  $p = 5\%$ . On remarque que les VaR des portefeuilles sont inférieures à celles des actifs pondérés, ce qui respecte le principe de diversification.<sup>6</sup>

Nous calculons maintenant le portefeuille optimal qui minimise la CVaR en supposant, pour l'instant, la normalité des rendements avec une moyenne des rendements nulle.

Tableau 2  
Portefeuille A : minimisation de la VaR<sup>1</sup> (modèle normal)

		IBM	General Electric	Wal Mart	Total	
Poids A		0,3889444	-0,0465131	0,6575686	1,00000	
Portefeuille	$p = 5\%$	$q_0 = -1,64485$				
Moyenne	0,05244 %	0,07580 %	-0,00286 %	0,03472 %		
Variance	0,01680 %					
Écart type	1,29631 %	1,63961 %	2,62559 %	1,41254 %		
Asymétrie	0,3588781					
Kurtosis	9,8157828					
	Portefeuille				Somme pondérée	Différence
VaRa 5 %	2,07980 %	2,62112 %	4,32157 %	2,28871 %	2,72546 %	-0,64566 %
VaRr 5 %	2,13224 %	2,69692 %	4,31871 %	2,32343 %	2,77764 %	-0,64540 %

<sup>1</sup> VaRa et VaRr désignent respectivement la VaR absolue et la VaR relative.

Dans les mêmes conditions, la CVaR s'écrit selon l'équation (2) (A6 en annexe) :

$$\text{CVaR} = \frac{1}{p} \left\{ \sigma_{\text{portefeuille}} \times \phi_0(q_0) \right\} = \sqrt{\beta^T \Sigma \beta} \times \frac{\phi_0(q_0)}{p} > 0 \quad (2)$$

<sup>6</sup> Le fichier Excel est disponible sur le site de la Chaire de recherche du Canada à l'adresse suivante : <https://chairegestiondesrisques.hec.ca/seminaires/livre-economica>.

où  $\phi_0(\cdot)$  est la fonction de densité  $N(0,1)$ . Nous cherchons cette fois le portefeuille B dont les poids optimaux minimisent la CVaR. Nous obtenons les mêmes coefficients qu'avec la minimisation de la VaR en négligeant les moyennes des rendements. La raison est que les deux formules (1) et (2) reviennent à minimiser  $(\beta^T \Sigma \beta)^{1/2}$  à une constante multiplicative près qui ne change rien au résultat. Le tableau 3 donne les résultats du portefeuille B. La CVaR du portefeuille est naturellement inférieure à celle des actifs pondérés, ce qui signifie que le principe de base de diversification est respecté.

Nous calculons maintenant la VaR non paramétrique, la CVaR, la CVaR- et la CVaR+ non paramétriques des données pour  $p = 5\%$  en utilisant la méthode de la simulation historique (Rockafellar et Uryasev, 2002). Les résultats sont présentés au tableau 4. La CVaR<sub>np</sub> est égale à la CVaR<sub>+np</sub>. Ce résultat est également vérifié à  $p = 2,5\%$  et  $p = 1\%$ , car il n'y a pas d'atomes coupés. L'écart relatif entre la CVaR<sub>+np</sub> et la CVaR<sub>-np</sub> est extrêmement serré :  $(2,97795\% - 2,96269\%) / 2,96269\% = 0,51\%$ . Par conséquent, exiger que la CVaR paramétrique soit encadrée par CVaR- et CVaR+ est un critère qui va ajouter une contrainte très sélective sur les modèles.<sup>7</sup>

Tableau 3  
Portefeuille B : minimisation de la CVaR (modèle normal)

	IBM	General Electric	Wal Mart	Total		
Poids B	0,3889444	-0,0465131	0,6575686	1,00000		
$p = 5\%$	$q_0 = -1,64485$					
Portefeuille				Somme pondérée	Différence	
CVaR 5 %	2,67392 %	2,69692 %	4,31871 %	2,32343 %	2,77764 %	-0,10372 %

La figure 2 montre clairement qu'une densité normale ne conviendrait pas pour représenter les données du portefeuille B. D'autre part, la  $t$  de Student ne serait pas assez "pointue" et ne garde

<sup>7</sup> En supposant qu'il y a une certaine discontinuité dans la distribution des rendements autour de la VaR.

pas assez de masse autour du mode, comme l'exigerait le tracé de la densité kernel des données. Ces dernières remarques suggèreraient plutôt une loi de Laplace.

Tableau 4  
Mesures non paramétriques des VaR, CVaR- et CVaR+

	IBM	General Electric	Wal Mart			
Poids B	0,3889444	-0,0465131	0,6575686	1,00000		
$p = 5 \%$	$q_0 = -1,64485$					
		Ordre <sup>1</sup> : 61		Somme pondérée		
VaR <sub>np</sub> <sup>2</sup>	2,04736 %	2,69619 %	4,12926 %	2,02620 %	2,57310 %	-0,52573 %
CVaR <sub>-np</sub>	2,96269 %	3,80857 %	6,39598 %	3,19753 %	3,88142 %	-0,91873 %
CVaR <sub>+np</sub>	2,97795 %	3,82711 %	6,43375 %	3,21706 %	3,90322 %	-0,92528 %

<sup>1</sup> La VaR non paramétrique est égale à la 61<sup>ème</sup> plus petite valeur de l'échantillon.

<sup>2</sup> np : non paramétrique.

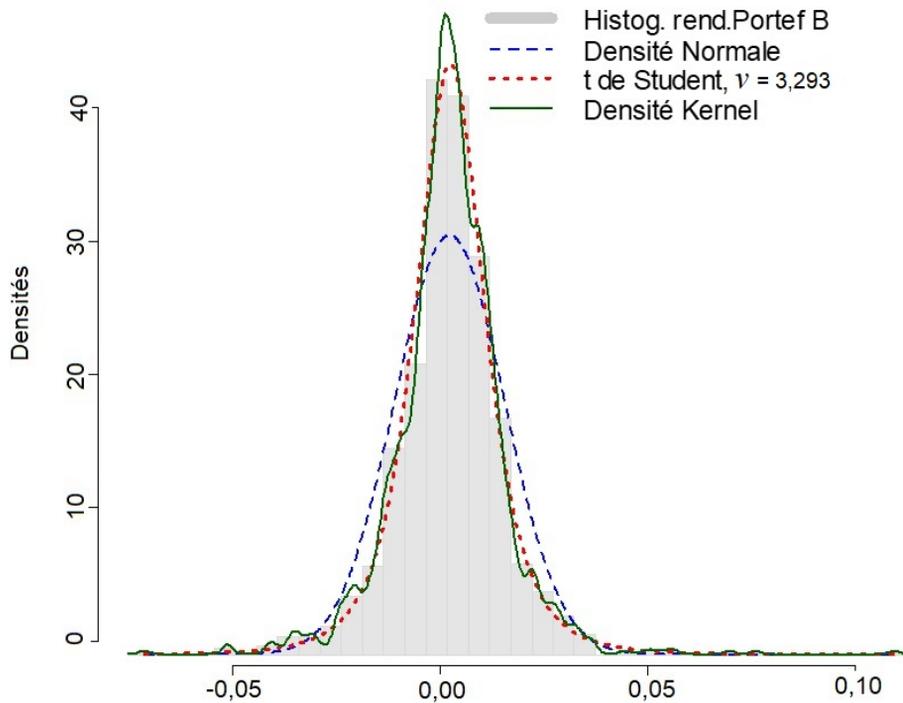


Figure 2: Histogramme et densités du portefeuille B

### 3. Estimation des paramètres de différentes distributions paramétriques

À partir de maintenant, nous prenons le portefeuille B comme portefeuille de référence. Les VaR calculées des modèles paramétriques sont des VaR absolues déterminées par rapport à 0, comme les CVaR. Les modèles sont notés de M1 à M8. L'estimation complète des coefficients des modèles est exposée dans les tableaux qui suivent et dans les tableaux A2 et A3 en annexe.

Nous débutons avec le modèle M1, en supposant que les données suivent une loi normale (voir définitions et expressions dans l'annexe A4). Nous ne supposons plus que  $\mu = 0$ . Le modèle normal M1 est noté 1:NO, dans le sens qu'il est constitué d'une seule loi normale. Les résultats sont présentés au tableau 5. La VaR de ce modèle est supérieure à la  $VaR_{np}$  (VaR non paramétrique), alors que la CVaR est très inférieure aux deux CVaR non paramétriques. Sans surprise, le test de Kolmogorov-Smirnov (KS) rejette ce modèle ( $p$ -value = 0,0015 < 10 %). Les valeurs d'asymétrie et de kurtosis rapportées, 0,3589 et 9,8158, correspondent aux moments empiriques des rendements du portefeuille (tableau 2). Elles sont très différentes de celles relatives à la distribution normale.

Tableau 5  
Modèle M1 = 1:NO

Paramètres M1	Loi normale (1:NO)	p = 5 %				Ordre <sup>1</sup> : 61
$\mu$	0,0005254	Quantiles, coefficients et probabilités				
$\sigma$	0,0129559	$q_0$	$\mu$	$\Phi[q_0]$	$-\sigma$	$\phi[q_0]$
		-1,644854	0,000525	0,050000	-0,012956	0,10313564
					q	$F_{\text{Modèle}}[q]$
VaR	2,07852 %	VaR <sub>np</sub> : 2,04736 %			-2,07852 %	0,050000
CVaR	2,61989 %	CVaR <sub>-/+np</sub> : [2,96269 %, 2,97795 %]				
Asymétrie	0,0000	0,3589	AIC -7021,1012	KS stat.		0,0775
Kurtosis	3,0000	9,8158	BIC -7010,9210	KS p-value		0,0015

<sup>1</sup> La VaR non paramétrique est égale à la 61<sup>ième</sup> plus petite valeur de l'échantillon.

Nous passons maintenant à la loi  $t$  de Student (M2 = 1:T ; voir définitions et expressions à l'annexe A5). Le paramètre de degré de liberté estimé est  $\nu = 3,28871$ . Le kurtosis est ainsi

indéfini (car  $\nu < 4$ ). Cette fois, la  $VaR < VaR_{np}$  et la  $CVaR > [CVaR_{-np}, CVaR_{+np}]$ , ce qui est la situation inverse par rapport à 1:NO. D'autre part, la  $p$ -value = 0,1285 > 10 % fait que le test KS ne rejette pas ce modèle. Les critères AIC et BIC s'améliorent grandement par rapport au modèle normal M1. Plus encore, ces valeurs AIC et BIC sont les plus petites de tous les modèles. Malgré tout, le fait que la  $CVaR > [CVaR_{-np}, CVaR_{+np}]$  est problématique. La raison est probablement liée au fait que la  $t$  de Student permet de tenir compte de l'épaisseur des queues, mais que son kurtosis est indéfini. De plus, la  $t$  de Student ne capture pas l'asymétrie des données.

Tableau 6  
Modèle M2 = 1:T

Paramètres M2		Loi $t$ de Student (1:T)		p = 5 %		Ordre <sup>1</sup> : 61
$\mu$	0,0006974	Quantiles, coefficients et probabilités				
$\sigma$	0,0085310	$q_0$	$\mu$	$F_{T,\nu}^0 [q_0]$	$-\sigma$	$Tail_{T,\nu}^0 [q_0]$
$\nu$	3,2887197	-2,271479	0,000697	0,050000	-0,008531	0,180675
					$q$	$F_{Modèle} [q]$
VaR	1,86806 %	VaR <sub>np</sub> : 2,04736 %			-1,86806 %	0,050000
CVaR	3,01294 %	CVaR <sub>-/+np</sub> : [2,96269 %, 2,97795 %]				
Asymétrie	0,0000	0,3589	AIC -7249,1447	KS stat.		0,0483
Kurtosis	indéfini	9,8158	BIC -7233,8745	KS $p$ -value		0,1285

<sup>1</sup> La VaR non paramétrique est égale à la 61<sup>ième</sup> plus petite valeur de l'échantillon.

Nous passons maintenant au modèle M3, en utilisant la loi EGB2 (les calculs de la CVaR sont faits par intégrales numériques, l'expression analytique n'étant pas disponible ; voir définitions et expressions à l'annexe A6). Avec ce modèle, nous avons un paramètre de plus. En effet, les paramètres  $\nu$  et  $\tau$  caractérisent à la fois l'épaisseur de la queue et l'asymétrie de la distribution. La distribution présente une asymétrie négative, positive ou symétrique quand  $\nu < \tau$ ,  $\nu > \tau$  ou  $\nu = \tau$ , respectivement. Quant à l'épaisseur de la queue, plus  $\nu$  est petit, plus la queue est épaisse (tous les autres paramètres maintenus égaux par ailleurs). L'estimation du modèle M3 donne  $\nu = 0,1587$  et  $\tau = 0,1652$ . Vu que  $\nu$  est très près de  $\tau$ , nous avons une légère asymétrie

négative =  $-0,081$  (tableau 7), ce qui n'est pas compatible avec le coefficient d'asymétrie non paramétrique des données =  $0,359$ .

Tableau 7  
Modèle M3 = 1:EGB2

Paramètres M3		Loi GB exponentielle type 2 (1:EGB2)		p = 5 %	Ordre <sup>1</sup> :61
$\mu$	0,0008884				
$\sigma$	0,0014108				
$\nu$	0,1587161			q	$F_{\text{Modèle}}[q]$
$\tau$	0,1652522			-2,00674 %	0,050000
VaR	2,00674 %	VaR <sub>np</sub> : 2,04736 %			
CVaR	2,89562 %	CVaR <sub>-/+np</sub> : [2,96269 %, 2,97795 %]			
Asymétrie	-0,0813	0,3589	AIC -7243,1659	KS stat.	0,0375
Kurtosis	5,8076	9,8158	BIC -7222,8056	KS p-value	0,3490

<sup>1</sup> La VaR non paramétrique est égale à la 61<sup>ième</sup> plus petite valeur de l'échantillon.

La VaR de M3 est la plus proche de la VaR<sub>np</sub> jusqu'à présent. Cependant, la CVaR =  $2,89562\% < [CVaR_{-np}, CVaR_{+np}]$ . Le kurtosis de 5,8 est encore insuffisant par rapport à 9,8. Malgré la grande flexibilité connue dans la littérature concernant la EGB2 à quatre paramètres, ces résultats semblent indiquer qu'une seule distribution paramétrique ne pourrait pas suffire pour bien cerner les risques inhérents à nos données, malgré le fait que le test KS ne rejette pas ce modèle ( $p$ -value =  $0,3490$ ).

Un dernier mot concernant les valeurs  $\nu = 0,1587$  et  $\tau = 0,1652$ . Vu que les paramètres  $\nu$  et  $\tau$  sont très petits et pratiquement nuls, nous connaissons le lemme 2 de Caivano et Harvey (2014), qui permet de dire que la EGB2 tend vers une densité de Laplace lorsque  $\nu \approx \tau \approx 0$ . Ceci rejoint directement la remarque du mode pointu de la densité kernel tracée dans la figure 2. Nous reverrons ce fait de convergence vers une Laplace dans la suite. Nous passons maintenant à des mélanges de distributions.

Nous estimons la VaR et la CVaR du modèle M4 construit avec un mélange de deux lois normales (2:NO) en utilisant les expressions exposées dans les annexes A2, A3 et A4. Le quantile

$q_m$  à un degré de confiance  $(1-p)$  d'un mélange de densités s'obtient par méthode numérique. La VaR est égale à  $-q_m$ . La CVaR est calculée en utilisant la valeur de  $q_m$ . Les résultats du modèle 2:NO, présentés au tableau 8, montrent que nous sommes peut-être sur la bonne voie en utilisant un mélange de densités. La VaR = 1,95397 % et la CVaR = 3,11363 % s'approchent clairement des mesures non paramétriques comparativement aux mesures obtenues pour la distribution 1:NO. Le kurtosis = 6,7 s'améliore également. Le test KS donne une  $p$ -value = 0,2181, ce qui est confortablement supérieur à 10 %. Cela étant, la CVaR du modèle 2:NO est encore loin de l'intervalle  $[CVaR_{-np}, CVaR_{+np}]$ .

Tableau 8  
Modèle M4 = 2:NO

Paramètres M4	Mélange de 2 lois normales (2:NO)		p = 5 %		Ordre <sup>1</sup> : 61			
Loi 1	$\mu_1$	-0,0004845						
	$\sigma_1$	0,0226636	Quantiles, coefficients et probabilités					
Loi 2	$\mu_2$	0,0008151	densité	$q_0$	$\mu$	$\Phi^0[q_0]$	$-\sigma$	$\phi^0[q_0]$
	$\sigma_2$	0,0082545	1	-0,840784	-0,000484	0,200234	-0,022664	0,280159
	$c_1$	0,2231962	2	-2,465888	0,000815	0,006834	-0,008255	0,019078
							$q_m$	$F_{\text{Modèle}}[q_m]$
							-1,95397 %	0,050000
VaR <sup>2</sup>	1,95397 %	VaR <sub>np</sub> : 2,04736 %						
CVaR	3,11363 %	CVaR <sub>-/+np</sub> : [2,96269 %, 2,97795 %]						
Asymétrie	-0,1386	0,3589	AIC	-7228,0889	KS stat.	0,0433		
Kurtosis	6,6789	9,8158	BIC	-7202,6385	KS $p$ -value	0,2180		

<sup>1</sup> La VaR non paramétrique est égale à la 61<sup>ième</sup> plus petite valeur de l'échantillon.

<sup>2</sup> Obtenue numériquement par le solveur d'Excel en minimisant  $(F_m(q_m) - p)^2$ .

L'estimation du mélange de deux lois  $t$  de Student (2:T, voir définitions et expressions aux annexes A2, A3 et A5) présentée au tableau 9 documente un très grand paramètre de degré de liberté de la première  $t$  de Student  $v_1 = 23642,3$ , clairement  $> 30$ , indiquant que la première  $t$  de Student est pratiquement une normale. La deuxième distribution ayant un degré de liberté  $v_2 = 6,4162 > 4$  permet au mélange de posséder maintenant un kurtosis = 8,4 bien défini et proche

du kurtosis des données 9,8 (non paramétrique). Nous avons malgré tout une  $p$ -value égale à 0,1100, étant à la limite du rejet à 10 %. Le BIC = -7207,65 se dégrade par rapport à celui de 1:T et 1:EGB2. La VaR = 2,02945 % est très proche de la non paramétrique, mais la CVaR = 3,04198 % > [CVaR<sub>-np</sub>, CVaR<sub>+np</sub>]. Remarquons également que le coefficient d'asymétrie de la 2:T = -0,15 est négatif alors que celui non paramétrique = 0,36 > 0. Ceci semble indiquer que l'asymétrie présente dans les données devrait être mieux intégrée à la modélisation. En effet, ce mélange 2:T s'avère une amélioration, mais reste insuffisant car il ne semble pas permettre de modéliser directement l'asymétrie.

Tableau 9  
Modèle M5 = 2:T

Paramètres M5		Mélange de 2 lois $t$ de Student (2:T)			p = 5 %			Ordre <sup>1</sup> : 61
Loi 1	$\mu_1$	-0,0012920		Quantiles, coefficients et probabilités				
	$\sigma_1$	0,0066854						
	$\nu_1$	23642,3157236	densité					
Loi 2	$\mu_2$	-0,0004740	1	-3,228904	0,001292	0,000622	-0,006685	-0,002176
	$\sigma_2$	0,0140598	2	-1,409730	-0,000474	0,102601	-0,014060	0,218922
	$\nu_2$	6,4162601						
	$c_1$	0,5158049						
	VaR <sup>2</sup>	2,02945 %		VaR <sub>np</sub> : 2,04736 %				
	CVaR	3,04198 %		CVaR <sub>-/+np</sub> : [2,96269 %, 2,97795 %]				
	Asymétrie	-0,1544	0,3589	AIC	-7243,2842	KS stat.	0,0500	
	Kurtosis	8,3993	9,8158	BIC	-7207,6537	KS $p$ -value	0,1100	

<sup>1</sup> La VaR non paramétrique est égale à la 61<sup>ème</sup> plus petite valeur de l'échantillon.

<sup>2</sup> Obtenue numériquement par le solveur d'Excel en minimisant  $(F_m(q_m) - p)^2$ .

Avant d'explorer l'ajout d'un paramètre capturant l'asymétrie, nous voulons examiner ce qui se passe pour un mélange de trois densités normales. Le modèle M6 est construit avec un mélange 3:NO. Dans le tableau 10, la  $p$ -value du test KS est 0,2280 > 10 %. De plus, c'est le premier modèle dont le kurtosis = 9,4 est quasiment identique au non paramétrique. Cette fois, le coefficient d'asymétrie est positif, comme celui non paramétrique.

La VaR (3:NO) = 2,03847 % est pratiquement identique à la non paramétrique. Quant à la CVaR (3:NO) = 3,00452 %, elle est la plus proche de l'intervalle  $[CVaR_{-np}, CVaR_{+np}]$  jusqu'à maintenant. Ce modèle semble mieux convenir aux données. Nous reviendrons à ce constat lorsque nous effectuerons les *backtests* des modèles.

Tableau 10  
Modèle M6 = 3:NO

Paramètres M6		Mélange de 3 lois normales (3:NO)			p = 5 %			Ordre <sup>1</sup> : 61
Loi 1	$\mu_1$	-0,0004753						
	$\sigma_1$	0,0150441	Quantiles, coefficients et probabilités					
Loi 2	$\mu_2$	0,0043390	densité	$q_0$	$\mu$	$\Phi^0[q_0]$	$-\sigma$	$\phi^0[q_0]$
	$\sigma_2$	0,0376531	1	-1,323399	-0,000475	0,092851	-0,015044	0,166189
Loi 3	$\mu_3$	0,0011752	2	-0,656617	0,004339	0,255714	-0,037653	0,321579
	$\sigma_3$	0,0065771	3	-3,278018	0,001175	0,000523	-0,006577	0,001852
	$c_1$	0,4433715						
	$c_2$	0,0334707						
						$q_m$		$F_{\text{Modèle}}[q_m]$
						-2,03847 %		0,050000
	VaR <sup>2</sup>	2,03847 %	VaR <sub>np</sub> : 2,04736 %					
	CVaR	3,00452 %	CVaR <sub>-/+np</sub> : [2,96269 %, 2,97795 %]					
Asymétrie	0,1224	0,3589	AIC	-7244,1307	K-S stat.			0,0425
Kurtosis	9,43212	9,8158	BIC	-7203,4101	K-S p-value			0,2280

<sup>1</sup> La VaR non paramétrique est égale à la 61<sup>ème</sup> plus petite valeur de l'échantillon.

<sup>2</sup> Obtenue numériquement par le solveur d'Excel en minimisant  $(F_m(q_m) - p)^2$ .

Pour avancer dans la modélisation, nous explorons maintenant l'effet d'ajouter un paramètre d'asymétrie comme amélioration du modèle précédent 3:NO. La densité SN2 (*Skewed Normal type 2*, Fernandez et al, 1995, annexe A7) permet cela. Nous injectons un paramètre d'asymétrie dans deux densités normales et gardons la troisième telle quelle. Le mélange de ce modèle M7 devient 2:SN2 + 1:NO.

Le modèle 2:SN2+1:NO comprend 10 paramètres. L'effet de capturer l'asymétrie est clair dans tous les résultats présentés au tableau 11. Le coefficient d'asymétrie est le plus proche de celui

non paramétrique, et surtout le kurtosis de M7 est même légèrement supérieur au non paramétrique. La  $p$ -value du test KS est  $0,1980 > 10\%$ . La  $VaR = 2,05018\% \approx$  la  $VaR_{np}$ . La  $CVaR = 2,98338\%$  est presque collée à la borne supérieure de l'intervalle des  $CVaR$ . Nous avons probablement un candidat sérieux pour représenter les risques des données, même si la  $CVaR$  n'est pas tout à fait encadrée par la  $[CVaR_{-np}, CVaR_{+np}]$ . En dernière remarque, le paramètre  $v_1 = 0,8831 < 1$  confirme la capture d'un certain degré d'asymétrie pour la première densité SN2. La deuxième ayant  $v_2 = 0,9939 \approx 1$  dégénère en une simple loi normale. Un mélange 1:SN2 + 2:NO aurait probablement suffi, tout en économisant un paramètre pour modéliser.

Tableau 11  
Modèle M7 = 2:SN2 + 1:NO

Paramètres M7	Mélange de 2 SN2 + 1 normale (2:SN2+1:NO)		$p = 5\%$	Ordre <sup>1</sup> : 61
Loi 1	$\mu_1$	0,00259297		
	$\sigma_1$	0,01467882		
	$v_1$	0,88306710	Quantiles, coefficients et probabilités	
Loi 2	$\mu_2$	0,00092266	densité	$q_0$
	$\sigma_2$	0,00638969	1	-1,573342
	$v_2$	0,99395520	2	-3,352978
Loi 3	$\mu_3$	0,00918332	3	-0,763277
	$\sigma_3$	0,03889170		
	$c_1$	0,47293330		
	$c_2$	0,50005730		
			$q_m$	$F_{\text{Modèle}}[q_m]$
			-2,05018 %	0,050000
	$VaR^2$	2,05018 %	$VaR_{np}$ :	2,04736 %
	$CVaR$	2,98338 %	$CVaR_{-/+np}$ :	[2,96269 %, 2,97795 %]
Asymétrie	0,2433	0,3589	AIC	-7241,4466
Kurtosis	9,8409	9,8158	BIC	-7190,5458
			KS stat.	0,0442
			KS $p$ -value	0,1980

<sup>1</sup> La  $VaR$  non paramétrique est égale à la 61<sup>ème</sup> plus petite valeur de l'échantillon.

<sup>2</sup> Obtenue numériquement par le solveur d'Excel en minimisant  $(F_m(q_m) - p)^2$ .

En plus de la paramétrisation directe de l'asymétrie, nous voulons également capturer l'épaisseur de la queue. Fernandez et al (1995) proposent la densité SEP3 (*Skewed Exponential Power* de type 3 ; voir aussi Rigby et al, 2014). Nous souhaitons réduire le nombre de paramètres

du même coup. Le modèle M8 est construit avec le mélange 2:SEP3 (mélange de 2 lois SEP3, voir définitions et calculs en annexes A3, A8 et A9).

D'abord, la  $p$ -value du test KS du tableau 12 est la plus grande. Le coefficient d'asymétrie est très petit et positif. Le kurtosis est grand, mais moins grand que pour le modèle précédent. Vu les valeurs de AIC et BIC, le modèle s'ajuste mieux aux données que le précédent. La VaR(2:SEP3) = 1,99295 % s'éloigne un peu de la non paramétrique, mais le plus important est de noter que finalement, la CVaR = 2,97397 % tombe dans l'intervalle  $[CVaR_{-np}, CVaR_{+np}]$  pour ce modèle 2:SEP3 et ce, malgré l'étroitesse de cet intervalle.

Remarquons les paramètres d'asymétrie  $\nu_1 = 1,0315 \approx 1$  et  $\nu_2 = 0,6137 < 1$ . Nous nous retrouvons dans la même configuration que le modèle précédent, avec une densité qui prend en charge de capturer l'asymétrie. Quant aux paramètres d'épaisseur de queue, ils valent  $\tau_1 = 0,9599 \approx 1$  et  $\tau_2 = 2,1084 \approx 2$ . Nous avons donc une première distribution SEP3 qui est pratiquement une Laplace ( $\nu = 1$ ,  $\tau = 1$ ). La deuxième SEP3 dégénère en une normale ( $\tau = 2$ ) asymétrique ( $\nu \neq 1$ ), qui est finalement une SN2. Dans ce cas, un mélange Laplace ajouté à une SN2 aurait probablement convenu aux données. Ce résultat rejoint directement des conclusions similaires de la littérature récente du risque de marché mettant en exergue les qualités du mélange d'une Laplace et une Gaussienne (voir Haas et al, 2006 ; Haas, 2009 ; Broda et Paoletta, 2011 ; Miao et al, 2016 ; Taylor, 2019).

Tableau 12  
Modèle M8 = 2:SEP3

Paramètres M8		Mélange de 2 lois SEP3 (2:SEP3)		p = 5 %		Ordre <sup>1</sup> : 61		
Loi 1	$\mu_1$	-0,0007520						
	$\sigma_1$	0,0004529						
	$\nu_1$	1,0315089						
	$\tau_1$	0,9598700	Quantiles, coefficients et probabilités					
Loi 2	$\mu_2$	0,0075456	densité	$q_0$	$\mu$	$F_{T,v}^0 [q_0]$	$-\sigma$	$Tail_{T,v}^0 [q_0]$
	$\sigma_2$	0,0065018	1	-4,234276	-0,000752	0,066185	-0,004529	0,425876
	$\nu_2$	0,6137048	2	-4,225762	0,007546	0,004189	-0,006502	0,019551
	$\tau_2$	2,1083901						
	$c_1$	0,7389303						
						$q_m$	$F_{Modèle} [q_m]$	
						-1,99295 %	0,050000	
	VaR <sup>2</sup>	1,99295 %	VaR <sub>np</sub> : 2,04736 %					
	CVaR	2,97397 %	CVaR <sub>-/+np</sub> : [2,96269 %, 2,97795 %]					
Asymétrie	0,0051	0,3589	AIC	-7245,9268	KS stat.	0,0392		
Kurtosis	7,1752	9,8158	BIC	-7200,1161	KS p-value	0,3040		

<sup>1</sup> La VaR non paramétrique est égale à la 61<sup>ème</sup> plus petite valeur de l'échantillon.

<sup>2</sup> Obtenue numériquement par le solveur d'Excel en minimisant  $(F_m(q_m) - p)^2$ .

Avant de passer à l'étape des *backtests* des VaR et des CVaR, il faut signaler que les huit modèles gardent les mêmes comportements aux probabilités  $p = 2,5 \%$  et  $1 \%$ . Cependant, à  $p = 1 \%$ , la CVaR des modèles M7 et M8 va plutôt quasiment coller à la borne supérieure de  $[CVaR_{-np}, CVaR_{+np}]$  plutôt que d'être dans cet intervalle (voir le tableau A4 de l'annexe A11). Ce centile est en fait trop avancé dans la queue des pertes pour que la CVaR soit suffisamment précise. Cela ne devrait pas poser de problème par rapport aux exigences actuelles de la réglementation, vu que Bâle requiert des *backtests* sur la VaR plutôt que sur la CVaR à ce centile de  $1 \%$ .

Pour finir cette section, la figure 3 récapitule graphiquement le comportement des VaR et des CVaR des modèles 2:SEP3 et 2:SN2 + 1:NO, ainsi que 1:NO et 1:T dans la queue gauche des rendements du portefeuille.

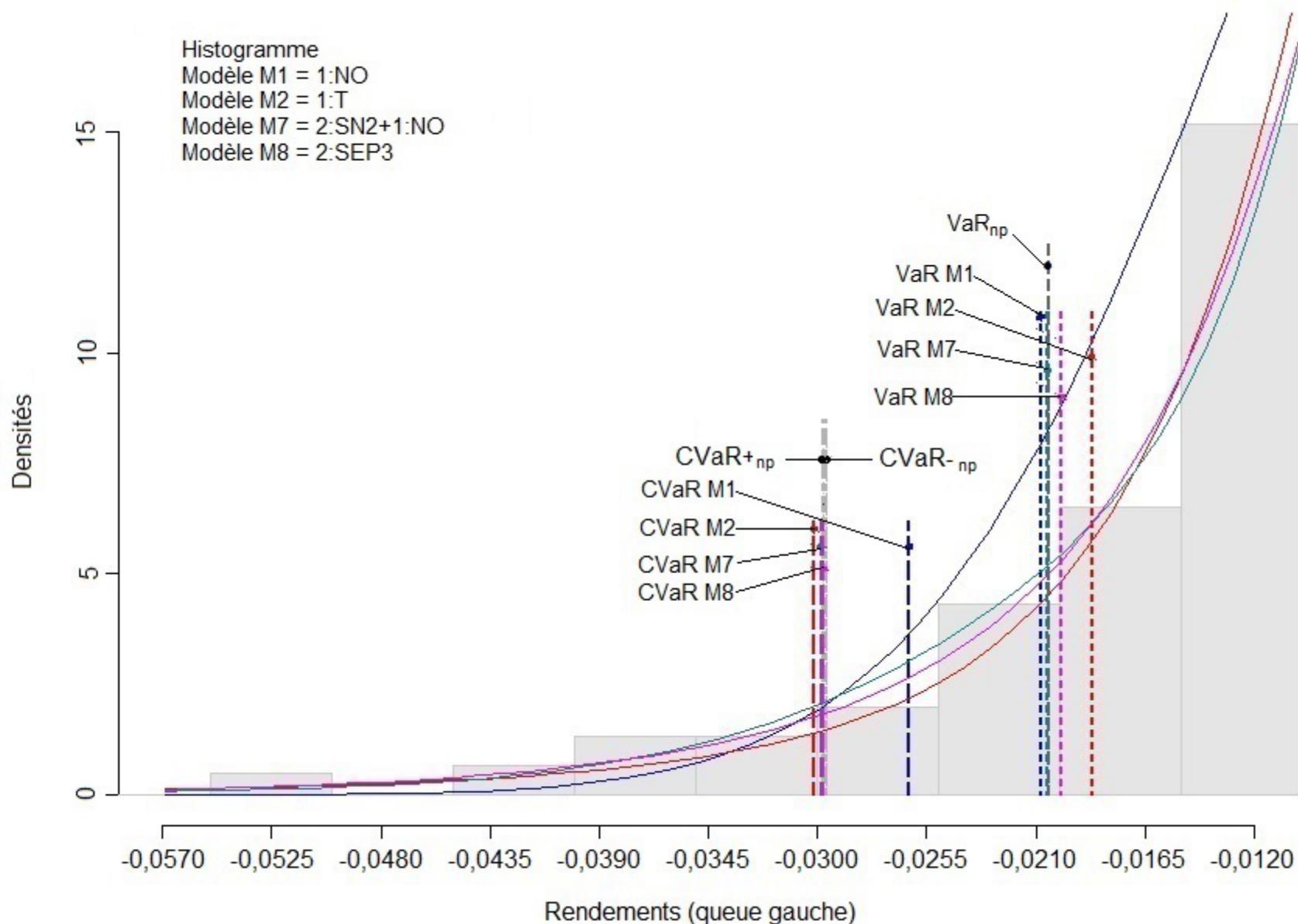


Figure 3 – Tracés des VaR et CVaR de certains modèles dans la queue gauche des rendements

## 4. *Backtests* des VaR et CVaR en conformité avec la réglementation de Bâle en vigueur

### 4.1 Méthodologie de validation des modèles de VaR et de CVaR

La validation des VaR des différents modèles sera effectuée par trois *backtests*. Le *backtest* cc valide la fréquence des hits en inconditionnel (Kupiec 1995). Ensuite, le *backtest* cc de Christoffersen (1998) est conditionnel à l'indépendance inter-hits. Le dernier test est le *backtest* DQ de Engle et Manganelli (2004). DQ est employé en parallèle avec cc pour détecter aussi bien les exceptions qui se suivent que celles espacées avec un lag pouvant aller jusqu'à une semaine environ avec des données quotidiennes. Le test cc de Christofferson détecte les dépassements qui se suivent avec un lag égal à un jour seulement. S'il y a des regroupements de dépassements plus ou moins proches avec des lags supérieurs à un jour, cc ne les détecte pas alors que DQ le fait.

Quant aux CVaR des modèles, nous nous appuyons sur le *backtest*  $Z_{ES}$  d'Acerbi et Szekely (2017) et sur le *backtest* de Righi et Ceretta (2015), qui sera désormais nommé RC. Nous présentons également les résultats des *backtests*  $Z_1$  et  $Z_2$  d'Acerbi et Szekely (2014), à titre informatif uniquement.<sup>8</sup>

Nous déployons ici cinq *backtests* pour valider les VaR et les CVaR des modèles candidats afin de : (i) satisfaire l'exigence réglementaire d'effectuer le backtest de la VaR à 1 % (BCBS, 2016, page 77 ; BCBS, 2019, paragraphe 32.5) ; (ii) à titre complémentaire, valider la CVaR à 2,5 % et la VaR à  $p = 2,5$  % ; (iii) à titre complémentaire également, valider la CVaR à 5 % et la VaR à 5 %.

Pour le moment, aucun des quatre *backtests* des points (ii) et (iii) n'est explicitement exigé pour valider la couverture globale du risque de marché des banques. Cela étant, nous les proposons dans le cadre de la suggestion de Bâle de prévoir des procédures statistiques supplémentaires à différents degrés de confiance pour étayer l'adéquation des modèles (BCBS, 2016, page 82 ; BCBS, 2019, paragraphe 32.13).

---

<sup>8</sup> En effet, bien que pour le moment assez populaires dans la littérature, ils comportent des défauts importants qui font que nous ne nous baserons pas sur leurs résultats pour tirer des conclusions.

Dans ce sens, il est naturel de penser à ajouter la validation à 2,5 % des mesures de risque de (ii), vue que la VaR à 2,5 % détermine la couverture. Les *backtests* à 5 % de (iii) seraient d'une importance moindre, mais devraient participer à la confirmation de la robustesse des modèles. Il reste à mentionner que les cinq *backtests* sont effectués en hors échantillon, dont la démarche est exposée à l'annexe A10.

#### 4.2 Résultats des *backtests* des modèles VaR et CVaR

Les résultats des *backtests* des huit modèles sont présentés au tableau 13. Sans surprise, le modèle normal 1:NO est rejeté à cause de sa VaR à 1 %, sa VaR à 2,5 % et sa CVaR à 2,5 %. Ce qui est étonnant est que le modèle *t* de Student 1:T soit rejeté pour des raisons similaires. Les *p*-values sont plus grandes, mais restent  $< 10\%$ , qui est le seuil critique de rejet.

Les *backtests* de la VaR à 1 % de EGB2 ont des *p*-values plus grandes, mais toujours inférieures à 10 %. On pourrait être tentés de ne pas rejeter la VaR à 1 %, surtout que la VaR à 2,5 % se comporte plutôt bien, avec des *p*-values des *backtests* uc, cc et DQ toutes  $> 10\%$  (0,1552, 0,2875 et 0,1131 respectivement). Par contre, la CVaR à 2,5 % est rejetée par  $Z_{ES}$  (*p*-value = 0,00  $< 10\%$ ) et par RC (*p*-value = 0,0124  $< 10\%$ ). Ce modèle est un exemple concret où la CVaR à 2,5 % ne passe pas le *backtest*, alors que la VaR performe relativement bien pour le *backtest* réglementaire à 1 %. Elle le fait également à 2,5 %, ce qui est une validation complémentaire.

Le cas suivant est la situation inverse. En effet, avec le modèle 2:NO, la VaR à 2,5 % est rejetée par les *backtests* uc et DQ (0,0783  $< 10\%$  et 0,0898  $< 10\%$ ). Par contre, la CVaR à 2,5% est bien validée par  $Z_{ES}$  et RC (*p*-value = 0,3516  $> 10\%$  et *p*-value = 0,3396  $> 10\%$ ). La 2:T réplique quasiment ce même comportement. La 2:T serait une amélioration du modèle 2:NO, certes, mais demeure insuffisante pour les données (comme c'est le cas de 1:T versus 1:NO).

Le mélange 3:NO, malgré ses huit paramètres, serait plutôt inférieur aux modèles précédents, dont la 2:NO, qui n'a que cinq paramètres. Pourtant, la 3:NO semblait bien se comporter lors de la discussion autour du tableau 10 de la section précédente (voir section 3). Cela confirme le bien-fondé d'injecter des paramètres supplémentaires pour capter l'asymétrie des données.

Justement, la 2:SN2+1:NO, incluant deux paramètres pour l'asymétrie, semble mieux s'ajuster à la VaR à 1 %. Les  $p$ -values respectives uc, cc et DQ sont 0,2695, 0,4378 et 0,3496. La VaR à 2,5 % semble également bien se comporter pour uc ( $p$ -value = 0,2819) et cc ( $p$ -value = 0,4004), sauf pour l'indépendance des hits DQ, dont la  $p$ -value est 0,0685. La CVaR à 2,5 % n'est pas rejetée selon  $Z_{ES}$  ( $p$ -value = 0,1516 > 10 %), mais l'est pour RC ( $p$ -value = 0,0516 < 10 %).

Un mot concernant le critère (iii). Le *backtest* uc valide la VaR à 5% ( $p$ -value=0.4338) et le *backtest*  $Z_{ES}$  à son tour valide la CVaR à 5 % ( $p$ -value = 0,4020), mais RC la rejette ( $p$ -value = 0,0596). En somme, le mélange 2:SN2+1:NO présente bien une amélioration nette par rapport aux mélanges précédents avec l'injection des deux paramètres d'asymétrie. Cependant, l'amélioration n'est pas encore suffisante pour bien modéliser les risques des données.

On arrive aux *backtests* concernant la 2:SEP3. Clairement, la VaR à 1 % est validée vu les  $p$ -values respectives de uc, cc et DQ (0,2695 > 10 %, 0,4378 > 10 % et 0,0994  $\approx$  10 %). La VaR à 2,5 % est aussi validée selon uc ( $p$ -value = 0,2819), cc ( $p$ -value = 0,4004) et DQ ( $p$ -value = 0,1200). La CVaR à 2,5 % est confortablement validée, aussi bien par  $Z_{ES}$  ( $p$ -value = 0,5000 > 10 %) que par RC 0,3168 > 10 %). Quant au critère (iii), on peut considérer que le *backtest* uc valide la VaR à 5 % ( $p$ -value = 0,0954  $\approx$  10 %). Les *backtests*  $Z_{ES}$  et RC valident la CVaR à 5 % avec des  $p$ -values confortables ( $p$ -value = 0,8412 et 0,3776 respectivement). En conclusion, le mélange 2:SEP3, capturant à la fois l'asymétrie et l'épaisseur des queues, semble posséder des aptitudes supérieures pour modéliser les risques incorporés dans nos données. Ces résultats confirment directement des conclusions de travaux récents sur la supériorité d'un mélange d'une loi normale et une loi de Laplace, comme vu précédemment.

## Conclusion

Cette note présente une méthodologie de validation des modèles concernant le risque de marché. La démarche déploie conjointement les *backtests* de la CVaR et la VaR, et ce en conformité avec la réglementation internationale en vigueur (couverture avec la CVaR à 2,5 % et *backtest* requis sur la VaR à 1 %). D'autre part, avec des données réelles qui incluent celles relevées durant une période de turbulences extrêmes des marchés, l'hypothèse de normalité des rendements

est définitivement révolue. L'identification d'un modèle paramétrique est orientée en comparant les grandeurs issues des calculs utilisant les paramètres des modèles avec les grandeurs équivalentes estimées en non paramétrique. La clé de voûte de cette note est l'encadrement de la CVaR du modèle à évaluer par rapport à l'intervalle  $[CVaR_{-np}, CVaR_{+np}]$  qui fonctionne comme un critère sélectif important d'évaluation des modèles et s'avère être très lié aux conclusions des *backtests*. Comme vu dans les différentes estimations, le kurtosis et l'asymétrie non paramétriques permettent également de guider l'approche de recherche pour savoir dans quel sens avancer.

D'autre part, cette recherche est un exercice de mise en oeuvre concrète des *backtests* des VaR et des CVaR pour choisir un modèle paramétrique permettant de gérer le risque de marché incorporé dans les données. L'identification du mélange 2:SEP3 qui semble bien fonctionner avec nos données n'est pas un hasard. En effet, le mélange d'une loi normale avec une loi de Laplace vient directement joindre les conclusions de la littérature récente, qui positionne ce mélange comme un remplacement naturel de la loi normale pour le risque de marché (voir Haas et al, 2006 ; Haas, 2009 ; Broda et Paoletta, 2011 ; Miao et al, 2016 ; Taylor, 2019).

Tableau 13  
*Backtest* hors échantillon des VaR et CVaR

$p$	#Hits	Modèle	— uc —		cc	DQ	— $Z_{ES}$ —		— RC —		— $Z_1$ —		— $Z_2$ —	
			Stat	$p$ -value	$p$ -value	$p$ -value	Stat	$p$ -value	Stat	$p$ -value	Stat	$p$ -value	Stat	$p$ -value
0,050	71	1 :NO	2,010	0,1563	0,0326	0,0012	-0,257	0,0000	-0,915	0,0000	-0,184	0,0000	-0,401	0,0024
0,025	47	1 :NO	8,450	0,0037	0,0105	0,0035	-0,379	0,0000	-0,611	0,0000	-0,182	0,0000	-0,852	0,0000
0,010	26	1 :NO	12,372	0,0004	0,0012	0,0000	-0,649	0,0000	-0,535	0,0000	-0,229	0,0000	-1,663	0,0000
0,050	78	1 :T	5,215	0,0224	0,0033	0,0001	-0,152	0,0388	-0,193	0,3012	-0,051	0,2876	-0,366	0,0160
0,025	45	1 :T	6,685	0,0097	0,0343	0,0163	-0,168	0,0724	0,010	0,9056	-0,028	0,5724	-0,542	0,0148
0,010	20	1 :T	4,487	0,0342	0,0756	0,0303	-0,229	0,0860	-0,058	0,4304	-0,062	0,4144	-0,771	0,0236
0,050	70	1 :EGB2	1,669	0,1964	0,0037	0,0000	-0,144	0,0008	-0,464	0,0056	-0,087	0,0220	-0,268	0,0028
0,025	38	1 :EGB2	2,020	0,1552	0,2875	0,1131	-0,185	0,0000	-0,294	0,0124	-0,102	0,0340	-0,396	0,0040
0,010	19	1 :EGB2	3,504	0,0612	0,1277	0,0791	-0,292	0,0008	-0,206	0,0072	-0,122	0,0724	-0,777	0,0024
0,050	74	2 :NO	3,211	0,0731	0,0004	0,0000	-0,143	0,4956	-0,213	0,4496	-0,058	0,1892	-0,305	0,5700
0,025	40	2 :NO	3,100	0,0783	0,1817	0,0898	-0,193	0,3516	-0,139	0,3396	-0,076	0,0492	-0,435	0,5140
0,010	17	2 :NO	1,864	0,1722	0,3084	0,0880	-0,306	0,1024	-0,293	0,0064	-0,170	0,0173	-0,657	0,4044
0,050	74	2 :T	3,211	0,0731	0,0004	0,0000	-0,147	0,2604	-0,212	0,2796	-0,063	0,2544	-0,311	0,2528
0,025	40	2 :T	3,100	0,0783	0,1817	0,0670	-0,204	0,2508	-0,146	0,2540	-0,081	0,1184	-0,441	0,2500
0,010	18	2 :T	2,627	0,1051	0,2044	0,0532	-0,312	0,2108	-0,174	0,0836	-0,140	0,0984	-0,709	0,2436
0,050	73	3 :NO	2,782	0,0954	0,0013	0,0000	-0,135	0,1836	-0,384	0,1432	-0,060	0,1356	-0,254	0,2996
0,025	42	3 :NO	4,387	0,0362	0,1017	0,0267	-0,175	0,1204	-0,220	0,1428	-0,089	0,0336	-0,343	0,2676
0,010	19	3 :NO	3,504	0,0612	0,1023	0,0126	-0,283	0,0544	-0,306	0,0052	-0,152	0,0096	-0,633	0,1688
0,050	66	2 :SN2 + 1 :NO	0,613	0,4338	0,0086	0,0000	-0,105	0,4020	-0,410	0,0596	-0,068	0,1316	-0,175	0,6520
0,025	36	2 :SN2 + 1 :NO	1,158	0,2819	0,4004	0,0685	-0,137	0,1516	-0,277	0,0516	-0,079	0,1140	-0,295	0,4116
0,010	16	2 :SN2 + 1 :NO	1,219	0,2695	0,4378	0,3496	-0,219	0,0296	-0,236	0,0144	-0,129	0,0441	-0,505	0,2004
0,050	73	2 :SEP3	2,782	0,0954	0,0301	0,0006	-0,111	0,8412	-0,153	0,3776	-0,037	0,3308	-0,262	0,8640
0,025	36	2 :SEP3	1,158	0,2819	0,4004	0,1200	-0,127	0,5000	-0,131	0,3168	-0,064	0,1699	-0,277	0,9912
0,010	16	2 :SEP3	1,219	0,2695	0,4378	0,0994	-0,198	0,1516	-0,206	0,0388	-0,101	0,1097	-0,468	0,6016

## Références

- Acerbi, C. et Szekely, B. (2014). Back-testing expected shortfall. *Risk*, 27(11):76–81.
- Acerbi, C. et Szekely, B. (2017). General properties of backtestable statistics. Available at SSRN: <https://ssrn.com/abstract=2905109> or <http://dx.doi.org/10.2139/ssrn.2905109>.
- Basel Committee on Banking Supervision (BCBS) (2016). Minimum capital requirements for market risk, publication no 352. *Bank For International Settlements (BIS)*, (Jan-2016):1–92.
- Basel Committee on Banking Supervision (BCBS) (2019). Minimum capital requirements for market risk, publication no 457. *Bank For International Settlements (BIS)*, (Jan-2019):1–136.
- Broda, S.A. et Paoletta, M.S. (2011). Expected shortfall for distributions in finance. *In Statistical Tools for Finance and Insurance*, pages 57–99. Springer.
- Caivano, M. et Harvey, A. (2014). Time-series models with an egb2 conditional distribution. *Journal of Time Series Analysis*, 35(6): 558–571.
- Christoffersen, P.F. (1998). Evaluating interval forecasts. *International Economic Review*, 39(4): 841–862.
- Cummins, J.D., Dionne, G., McDonald, J.B. et Pritchett, B.M. (1990). Applications of the GB2 family of distributions in modeling insurance loss processes. *Insurance: Mathematics and Economics*, 9(4):257–272.
- Dionne, G. (2019). Corporate risk management: Theories and applications. John Wiley, 384 pages.
- Dionne, G. et Saissi Hassani, S. (2017). Hidden Markov regimes in operational loss data: Application to the recent financial crisis. *Journal of Operational Risk*, 12(1): 23–51.
- Efron, B. et Tibshirani, R.J. (1994). *An introduction to the bootstrap*. Chapman and Hall/CRC press, 456 p.
- Engle, R.F. et Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics*, 22(4):367–381.
- Fernandez, C., Osiewalski, J. et Steel, M.F. (1995). Modeling and inference with  $\uparrow$ -spherical distributions. *Journal of the American Statistical Association*, 90(432):1331–1340.
- Haas, M. (2009). Modelling skewness and kurtosis with the skewed Gauss-Laplace sum distribution. *Applied Economics Letters*, 16(12): 1277-1283.
- Haas, M., Mittnik, S. et Paoletta, M.S. (2006). Modelling and predicting market risk with laplace–gaussian mixture distributions. *Applied Financial Economics*, 16(15): 1145–1162.

- Kerman, S.C. et McDonald, J.B. (2015). Skewness-kurtosis bounds for egb1, egb2, and special cases. *Communications in Statistics-Theory and Methods*, 44(18):3857–3864.
- Kupiec, P.H. (1995). Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*, 3(2):73–84.
- McDonald, J.B. (2008). Some generalized functions for the size distribution of income. Dans : *Modeling Income Distributions and Lorenz Curves*, Springer, p. 37–55.
- McDonald, J.B. (1984). Some generalized functions for the size distribution of income. *Econometrica*, 52(3):647–663.
- McDonald, J.B. et Michelfelder, R.A. (2016). Partially adaptive and robust estimation of asset models: accommodating skewness and kurtosis in returns. *Journal of Mathematical Finance*, 7(1): 219.
- McDonald, J.B. et Xu, Y.J. (1995). A generalization of the beta distribution with applications. *Journal of Econometrics*, 66(1-2): 133–152.
- Miao, D.W.C., Lee, H.C. et Chen, H. (2016). A standardized normal-Laplace mixture distribution fitted to symmetric implied volatility smiles. *Communications in Statistics-Simulation and Computation* 45(4): 1249-1267.
- Rigby, B., Stasinopoulos, M., Heller, G. et Voudouris, V. (2014). The distribution toolbox of GAMLSS. *gamlss.org*.
- Righi, M. et Ceretta, P.S. (2015). A comparison of expected shortfall estimation models. *Journal of Economics and Business*, 78: 14–47.
- Rockafellar, R.T. et Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26(7): 1443–1471.
- Taylor, J.W. (2019). Forecasting value at risk and expected shortfall using a semiparametric approach based on the asymmetric laplace distribution. *Journal of Business and Economic Statistics* 37 (1): 121-133.
- Theodossiou, P. (2018). Risk measures for investment values and returns based on skewed-heavy tailed distributions: Analytical derivations and comparison. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3194196](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3194196).

## Annexes

### A1. Modèles estimés

Les annexes présentent les développements mathématiques des équations retenues et des tableaux de résultats d'estimation des paramètres de différentes distributions statistiques des rendements. Ces développements étant algébriques, il convient d'inverser les signes des expressions finales des VaR et CVaR pour avoir des mesures positives. Le tableau A1 présente les symboles des modèles estimés.

Tableau A.1  
Définitions des symboles des modèles

Modèle	Symbole	Description des modèles
M1	1:NO	Loi normale
M2	1:T	Loi $t$ de Student
M3	1:EGB2	Loi GB2 Exponentielle
M4	2:NO	Mélange de 2 lois normales
M5	2:T	Mélange de 2 lois $t$ de Student
M6	3:NO	Mélange de 3 lois normales
M7	2:SN2+1:NO	Mélange de 2 SN2 + 1 normale
M8	2:SEP3	Mélange de 2 lois SEP3

Commençons par des démonstrations des formules générales de la CVaR pour une distribution statistique (annexe A2) et pour un mélange de distributions statistiques (annexe A3).

### A2. Expression de la CVaR

#### Expression de la densité et de la fonction cumulative d'une loi réduite

On s'intéresse à la famille  $\bar{F}$  de lois paramétriques de type *location-scale* ayant un paramètre de location  $\mu$  et un paramètre de dispersion  $\sigma$ . Si  $F \in \bar{F}$  et  $y \sim F$ , la variable réduite  $z = (y - \mu)/\sigma$  suit alors la distribution  $F^0$  définie avec l'égalité :

$$F(y) = F^0(z).$$

$F^0$  est dite fonction cumulative réduite de  $F$ . La densité réduite  $f^0$  est liée à la densité  $f(y)$  en écrivant :

$$f(y) = \frac{f^0(z)}{\sigma} \quad (\text{A0})$$

Toutes les densités de ce document appartiennent à  $\bar{F}$ , dont la normale et la  $t$  de Student. Les paramètres de location et de dispersion coïncident avec la moyenne et l'écart type de la loi normale. Ce n'est pas toujours le cas pour les autres lois  $\in \bar{F}$ .

### Expression générale de la CVaR

On note  $q < 0$  le quantile de la VaR correspondant au degré de confiance  $(1 - p)$ . Comme dans Broda et Paoletta (2011), on définit la quantité *tail* d'une densité  $f(\cdot)$  au point  $x$  par :

$\text{tail}_f(x) \stackrel{\text{def}}{=} \int_{-\infty}^x t f(t) dt$ . On développe l'expression de la CVaR à partir de cette définition :

$$\text{CVaR}_f = E[y|y \leq q] = \frac{1}{F(q)} \int_{-\infty}^q y f(y) dy \quad (\text{A1})$$

$$= \frac{1}{p} \left\{ \int_{-\infty}^{\frac{q-\mu}{\sigma}} (\mu + \sigma z) \frac{f^0(z)}{\sigma} \sigma dz \right\} \quad (\text{A2})$$

$$= \frac{1}{p} \left\{ \int_{-\infty}^{\frac{q-\mu}{\sigma}} \mu f^0(z) dz + \sigma \int_{-\infty}^{\frac{q-\mu}{\sigma}} z f^0(z) dz \right\} \quad (\text{A3})$$

$$= \frac{1}{p} \left\{ \mu F^0\left(\frac{q-\mu}{\sigma}\right) + \sigma \text{Tail}_{f^0}\left(\frac{q-\mu}{\sigma}\right) \right\} \quad (\text{A4})$$

On obtient l'équation (A2) en utilisant (A0) après un changement de variable  $z = (y - \mu)/\sigma$ , ou  $y = \mu + \sigma z$ , d'où  $dy = \sigma dz$ . Les égalités (A3) et (A4) proviennent juste de calculs algébriques sur la précédente. À noter qu'il y a deux parties dans la formule (A4) : la première est  $\mu/p$  fois la cumulative centrée réduite  $F^0(\cdot)$  évaluée à la variable centrée réduite  $(q - \mu)/\sigma$ . La deuxième est  $\sigma/p$  fois le *tail* de  $f^0$ , évalué également à  $(q - \mu)/\sigma$ .

Une dernière remarque est qu'on a bien sûr  $F^0((q-\mu)/\sigma) = p$ , ce qui permet de simplifier l'expression (A4). Malgré cela, on va la laisser telle quelle pour qu'elle soit de la même forme que pour les mélanges de distributions où, justement, il y aura plusieurs cumulatives  $F_i^0(\cdot)$  pour lesquelles  $F_i^0((q-\mu)/\sigma) \neq p$ .

### A3. CVaR d'un mélange de distributions

Soit un mélange  $m(\cdot)$  de  $n$  densités  $f_i(\cdot)$ ,  $i=1, \dots, n$ . Les densités  $f_i \in \bar{F}$  ont chacune un paramètre de location  $\mu_i$  et de dispersion  $\sigma_i$ . La densité de mélange  $f_m(y)$  et sa répartition  $F_m(y)$  s'écrivent :

$$f_m(y) = \sum_1^n c_i f_i(y), \quad F_m(y) = \sum_1^n c_i F_i(y)$$

où  $c_i$  est la probabilité, à estimer, de la densité  $f_i(\cdot)$ . La somme des  $c_i$  est égale à 1. Soit  $q_m$  le quantile correspondant à la VaR du mélange au degré de confiance  $(1-p)$ . On note  $f_i^0(\cdot)$  et  $F_i^0(\cdot)$  la densité et la cumulative réduites de la  $i^{\text{ième}}$  densité. L'expression de la  $CVaR_m$  se développe en sortant la somme ( $\Sigma$ ) de l'intégrale :

$$\begin{aligned} CVaR_m &= E[y|y \leq q_m] = \frac{1}{p} \int_{-\infty}^{q_m} y f_m(y) dy \\ &= \frac{1}{p} \sum_1^n c_i \int_{-\infty}^{q_m} y f_i(y) d(y) \\ &= \frac{1}{p} \sum_1^n c_i \int_{-\infty}^{\frac{q_m - \mu_i}{\sigma_i}} (\mu_i + \sigma_i z_i) \frac{f_i^0(z_i)}{\sigma_i} \sigma_i d(z_i) \\ &= \frac{1}{p} \sum_1^n c_i \left\{ \mu_i F_i^0\left(\frac{q_m - \mu_i}{\sigma_i}\right) + \sigma_i \text{Tail}_{f_i^0}\left(\frac{q_m - \mu_i}{\sigma_i}\right) \right\} \end{aligned}$$

ou, en forme vectorielle pratique pour les calculs :

$$\text{CVaR}_m = \frac{1}{p} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}^T \times \left[ \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} \times \begin{pmatrix} F_1^0 \left( \frac{q_m - \mu_1}{\sigma_1} \right) \\ \vdots \\ F_n^0 \left( \frac{q_m - \mu_n}{\sigma_n} \right) \end{pmatrix} + \begin{pmatrix} \sigma_1 \\ \vdots \\ \sigma_n \end{pmatrix} \times \begin{pmatrix} \text{Tail}_{f_1}^0 \left( \frac{q_m - \mu_1}{\sigma_1} \right) \\ \vdots \\ \text{Tail}_{f_n}^0 \left( \frac{q_m - \mu_n}{\sigma_n} \right) \end{pmatrix} \right].$$

Dans le cas général,  $q_m$  est trouvé numériquement comme solution à l'équation  $F_m(q_m) - p = 0$ . Le fichier Excel permet cela. On remarquera que pour une distribution  $i$ ,  $F_i^0((q_m - \mu_i)/\sigma_i) \neq p$ . Le seul cas où on a égalité est lorsque la distribution  $F_i^0$  est unique (pas de mélange).

#### A4. Expression de la CVaR d'une loi normale

La densité  $\phi_{\mu,\sigma}(\cdot)$  d'une  $N(\mu,\sigma)$  est

$$\phi_{\mu,\sigma}(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right).$$

On note  $\phi_0(\cdot)$  et  $\Phi_0(\cdot)$  la densité et la cumulative de la loi normale standard  $N(0,1)$ . Il est aisé de montrer que :

$$\frac{\partial}{\partial x} \phi_0(x) = -x\phi_0(x). \quad (\text{A5a})$$

Pour  $y \sim N(\mu,\sigma)$ , le quantile  $q$  de la VaR au degré de confiance  $(1-p)$  est trouvé par  $P(y = \mu + \sigma z \leq q) = p$ , d'où la  $\text{VaR}_{p,\mu,\sigma}$  :

$$q = \mu + \sigma\Phi_0^{-1}(p).$$

D'autre part, avec la définition du *tail* et à l'aide de l'égalité A5a, on trouve :

$$\text{Tail}_{\phi_0}(x) = \int_{-\infty}^x z\phi_0(z)dz = -\phi_0(x). \quad (\text{A5})$$

On applique (A4) et (A5) pour trouver :

$$CVaR_{\phi, \mu, \sigma} = \frac{1}{p} \left\{ \mu \Phi_0 \left( \frac{q - \mu}{\sigma} \right) - \sigma \phi_0 \left( \frac{q - \mu}{\sigma} \right) \right\}. \quad (A6)$$

## A5. Expression de la CVaR de la loi $t$ de Student

La densité  $f_{T, \mu, \sigma, \nu}(\cdot)$  de la  $t$  de Student de paramètres  $\mu$  (location),  $\sigma$  (dispersion) et  $\nu$  (degrés de liberté) s'écrit :

$$f_{T, \mu, \sigma, \nu}(y) = \frac{A}{\sigma} \left( 1 + \left( \frac{y - \mu}{\sigma} \right)^2 \frac{1}{\nu} \right)^{-b}$$

où  $A = \left[ \sqrt{\nu} \times B(1/2, \nu/2) \right]^{-1}$  et  $b = (\nu + 1)/2$ .  $B(\cdot)$  est la fonction beta et  $b$  est un paramètre défini en fonction de  $\nu$ . Il existe une autre façon d'écrire la constante  $A$  avec la fonction gamma  $\Gamma(\cdot)$  au lieu de la fonction beta. Les fonctions réduites sont notées  $f_{T, \nu}^0(\cdot)$  et  $F_{T, \nu}^0(\cdot)$ . On détermine  $q$  de la VaR de  $y \sim t(\mu, \sigma, \nu)$  au degré de confiance  $(1 - p)$  :

$$\begin{aligned} P(y \leq q) &= P(\mu + \sigma z \leq q) \Rightarrow F_{T, \nu}^0 \left( \frac{q - \mu}{\sigma} \right) = p \\ q &= \mu + \sigma F_{T, \nu}^{0-1}(p) \end{aligned}$$

où  $F_{T, \nu}^{0-1}(\cdot)$  est la fonction quantile (ou inverse) de  $F_{T, \nu}^0(\cdot)$ . Le *tail* à un point  $x$  est par définition :

$$\text{Tail}_{f_{T, \nu}^0}(x) = \int_{-\infty}^x z f_{T, \nu}^0(z) dz = A \int_{-\infty}^x z \left( 1 + z^2/\nu \right)^{-b} dz. \quad (A7)$$

On opère un changement de variable  $u = z^2/\nu$ , d'où  $z dz = \nu du/2$ . L'intégrale de l'équation (A7) devient :

$$\begin{aligned} \text{Tail}_{f_{T, \nu}^0}(x) &= A \frac{\nu}{2} \int_{-\infty}^{\frac{x^2}{\nu}} (1 + u)^{-b} du \\ &= \frac{A}{2} \frac{\nu}{-b+1} \left[ (1 + u)^{-b+1} \right]_{-\infty}^{\frac{x^2}{\nu}} = \frac{\nu}{2(1-b)} \left( 1 + \frac{x^2}{\nu} \right) \times A \left( 1 + \frac{x^2}{\nu} \right)^{-b} \end{aligned} \quad (A8)$$

$$= -\frac{\nu + x^2}{\nu - 1} \times f_{T, \nu}^0(x). \quad (A9)$$

Dans la ligne (A8), on remplace  $b$  par sa valeur  $(v+1)/2$ . L'expression finale du *tail* se simplifie en (A9). On doit contraindre  $v > 1$  pour qu'elle soit valide. On applique maintenant (A9) dans (A4) pour trouver :

$$\text{CVaR}_{f_{T,\mu,\sigma,v}} = \frac{1}{p} \left\{ \mu F_{T,v}^0 \left( \frac{q-\mu}{\sigma} \right) - \sigma \frac{v + \left( \frac{q-\mu}{\sigma} \right)^2}{v-1} f_{T,v}^0 \left( \frac{q-\mu}{\sigma} \right) \right\}.$$

**Attention :** Dans Excel, les fonctions relatives à la loi  $t$  de Student considèrent que le degré de liberté  $v$  est un nombre entier. De ce fait, les calculs ne peuvent pas être faits sous forme standard et un module complémentaire est nécessaire. On fait appel au module XRealStats.xlam qui doit être téléchargé<sup>9</sup> du répertoire C:/TP5 et activé pour utiliser les fonctions permettant les calculs avec  $v \in \mathbb{R}$ . La fonction cumulative et de densité sont appelées par T\_DIST. La fonction inverse de la cumulative est T\_INV.

## A6. La loi EGB2 : GB2 exponentielle

La densité EGB2 (*Exponential Generalized Beta type 2*) a quatre paramètres et s'écrit, selon Kerman et McDonald (2015), pour  $y \in \mathbb{R}$  :

$$f(y|\mu, \sigma, \nu, \tau) = \frac{e^{y\nu}}{|\sigma| \times B(\nu, \tau) (1 + e^z)^{\nu+\tau}}$$

où  $z = (y - \mu)/\sigma$ ,  $\mu, \sigma \in \mathbb{R}$ ,  $\nu, \tau > 0$ .  $B(\cdot)$  est la fonction beta standard. La densité exponentielle de la GB2 a été proposée à l'origine par McDonald (1984).

Les paramètres  $\nu$  et  $\tau$  caractérisent à la fois l'épaisseur de la queue et l'asymétrie de la distribution. La distribution présente une asymétrie négative, positive ou symétrique quand  $\nu < \tau$ ,  $\nu > \tau$  ou  $\nu = \tau$  respectivement. Quant à l'épaisseur de la queue, plus  $\nu$  est petit, plus la queue de distribution est épaisse (tous les autres paramètres maintenus égaux par ailleurs).

<sup>9</sup> <http://www.real-statistics.com/free-download/>

La EGB2 englobe beaucoup de lois paramétriques comme cas particuliers. Précisément, quand  $v \approx \tau \rightarrow +\infty$ , la distribution converge vers une loi normale. Dans la pratique, cette convergence peut être considérée atteinte quand  $v \approx \tau > 15$ . Quand  $v = \tau = 1$ , la EGB2 devient une loi logistique. D'autre part, le lemme 2 de Caivano et Harvey (2014) démontre que la EGB2 tend vers une densité de Laplace lorsque  $v \approx \tau \approx 0$ . D'autres cas particuliers intéressants de la EGB2 et de la GB2 sont présentés dans Kerman et McDonald (2015), McDonald (2008) et McDonald et Xu (1995).

Pour l'utilisation de la GB2 en assurance, voir Cummins, Dionne, McDonald et Pritchett (1990). La EGB2 en finance est de plus en plus présente comme dans Caivano et Harvey (2014), McDonald et Michelfelder (2016) et Theodossiou (2018).

### A7. La loi SN2 : *Skewed Normal de type 2*

La définition de la densité de la Skewed Normal de type 2 (SN2) de Fernandez et al (1995) pour  $y \in \mathbb{R}$  s'écrit :

$$f_{SN2,\mu,\sigma,v}(y) = \frac{2v}{\sigma\sqrt{2\pi}(1+v^2)} \left\{ \exp\left(-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2 v^2\right) \mathbf{I}_{(y<\mu)} + \exp\left(-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2 \frac{1}{v^2}\right) \mathbf{I}_{(y\geq\mu)} \right\} \quad (\text{A10})$$

où  $\mu \in \mathbb{R}$ ,  $\sigma > 0$ ,  $v > 0$ . Si  $v < 1$ , l'asymétrie est vers la gauche (rendements négatifs) ; si  $v > 1$ , l'asymétrie est positive. Quand  $v = 1$ , on revient à une loi normale (symétrique). Cette densité est utilisée en gestion des risques opérationnels, comme dans Dionne et Saissi Hassani (2017). Une variable aléatoire  $y \sim F_{SN2,\mu,\sigma,v} \Rightarrow z = (y - \mu)/\sigma \sim F_{SN2,v}^0$ . Pour  $z < 0$ , seul le membre de gauche de l'équation (A10) est non nul. La densité réduite s'écrit alors :

$$f_{SN2,v}^0(z) = \frac{2v}{1+v^2} \phi_0(z \times v).$$

La cumulative au point  $z < 0$  s'écrit :

$$F_{SN2,v}^0(z) = \int_{-\infty}^z f_{SN2,v}^0(t) dt = \frac{2}{1+v^2} \Phi_0(z \times v).$$

Les fonctions  $\Phi_0(\cdot)$  et  $\phi_0(\cdot)$  désignent la cumulative et la densité normale centrée réduite  $N(0,1)$ . L'équation précédente permet de trouver l'expression de la VaR au degré de confiance  $(1-p)$ :

$$\begin{aligned} P(y \leq q) &= P\left(z \leq \frac{q-\mu}{\sigma}\right) \Rightarrow F_{SN2,v}^0\left(\frac{q-\mu}{\sigma}\right) = p \\ \frac{2}{1+v^2} \Phi_0\left(\frac{q-\mu}{\sigma} \times v\right) &= p \\ q &= \mu + \sigma \frac{1}{v} \Phi_0^{-1}\left(p \frac{1+v^2}{2}\right). \end{aligned} \quad (A11)$$

L'expression (A11) n'est valable que si  $p(1+v^2)/2 \leq 1$ , sinon  $\Phi^{-1}(\cdot)$  ne serait pas définie. Cela impose que  $v \leq \sqrt{2/p-1}$ .

L'expression du *tail* se développe de cette façon :

$$\begin{aligned} \text{Tail}_{SN2,v}(x) &= \int_{-\infty}^x z f^0(z) dz = \int_{-\infty}^x \frac{2v}{1+v^2} \phi_0(z \times v) dz \\ &= \frac{2v}{1+v^2} \int_{-\infty}^{x \times v} \frac{u}{v} \phi_0(u) \frac{du}{v} = \frac{2}{v(1+v^2)} \left[ -\phi_0(u) \right]_{-\infty}^{x \times v} \\ &= -\frac{2}{v(1+v^2)} \phi_0(x \times v). \end{aligned} \quad (A12)$$

L'égalité (A12) est obtenue avec un changement de variable  $u = z \times v$  et à l'aide de l'égalité A5a. Les équations (A11) et (A12) dans (A4) donnent l'expression de la CVaR :

$$\text{CVaR}_{SN2,\mu,\sigma,v} = \frac{1}{p} \left\{ \frac{2}{1+v^2} \left[ \mu \Phi_0\left(\frac{q-\mu}{\sigma} v\right) - \sigma \frac{1}{v} \phi_0\left(\frac{q-\mu}{\sigma} v\right) \right] \right\}.$$

Encore une fois, quand  $v = 1$  on retrouve la CVaR de la  $N(\mu, \sigma)$ .

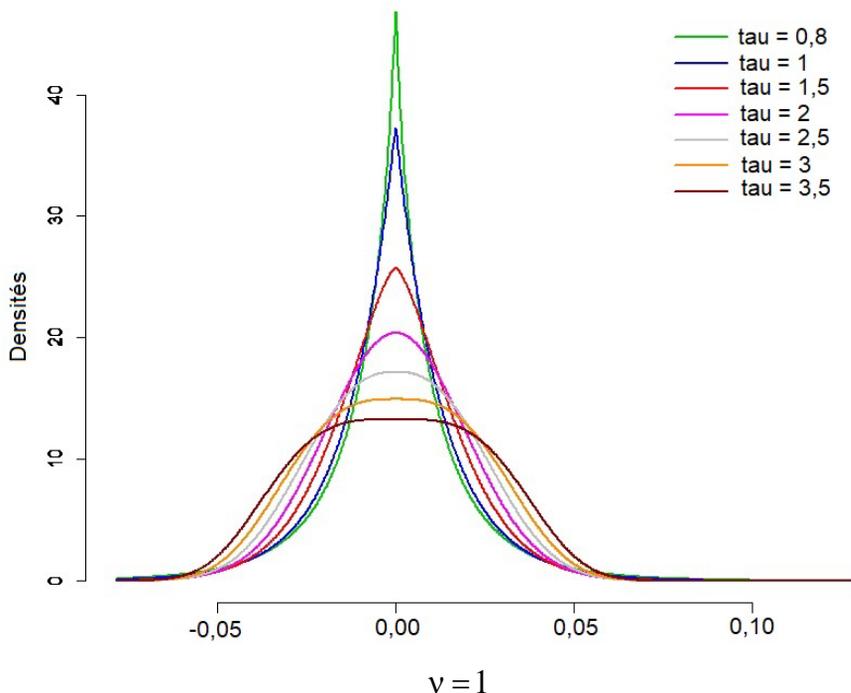
### A8. La loi SEP3 : *Skewed Exponential Power type 3*

La définition de cette loi revient à Fernandez et al (1995). L'appellation SEP est due à ces auteurs, l'ajout du numéro 3 dans SEP3 est selon le classement de Rigby et al (2014). La densité de la SEP3 s'écrit :

$$f_{SEP3,\mu,\sigma,v,\tau}(y) = \frac{c}{\sigma} \left\{ \exp\left(-\frac{1}{2} \left| \frac{y-\mu}{\sigma} v \right|^\tau\right) I_{(y<\mu)} + \exp\left(-\frac{1}{2} \left| \frac{y-\mu}{\sigma} \frac{1}{v} \right|^\tau\right) I_{(y\geq\mu)} \right\}$$

où  $c = v \times \tau \times \left[ (1+v^2) 2^{1/\tau} \Gamma(1/\tau) \right]^{-1}$  et où  $\mu \in \mathbb{R}, \sigma > 0, v \in \mathbb{R}, \tau > 0$ . Ils sont respectivement les paramètres de location, de dispersion, d'asymétrie, d'épaisseur de queue. La SEP3 a comme cas particuliers la SN2 quand le quatrième paramètre  $\tau=2$  ou une loi de Laplace (version asymétrique) quand  $\tau=1$ . À noter que d'autres appellations coexistent dans la littérature pour désigner des distributions comparables à la SEP3, comme la classe AP (*Asymmetric Power*) et AEP (*Asymmetric Exponential Power*).

La SEP3 peut être leptokurtique lorsque  $\tau < 2$  ou platykurtique quand  $\tau > 2$  (voir figure A1). Les calculs de la VaR et la CVaR font appel aux fonctions gamma et à la distribution gamma, comme présenté dans la section suivante.



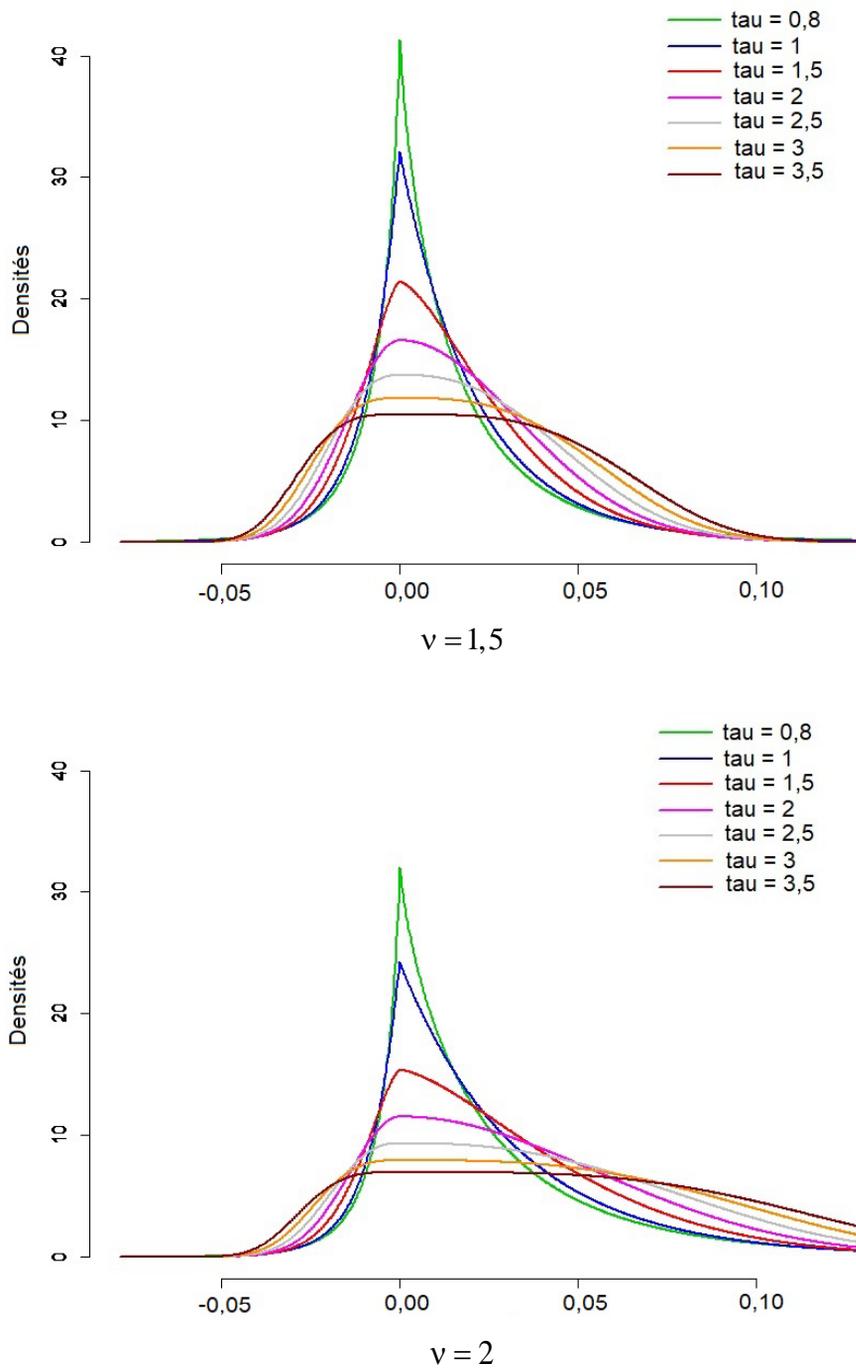


Figure A1: Tracés de la SEP3 avec différentes valeurs de  $\tau$  et de  $\nu$

## A9. Expression de la VaR et la CVaR d'une SEP3

Comme fait pour la SN2, on développe l'expression de la cumulative réduite de la SEP3 pour  $z < 0$  (queue gauche) en écrivant :

$$\begin{aligned} F_{\text{SEP3},v,\tau}^0(z) &= \int_{-\infty}^z \frac{\tau v}{(1+v^2)^{1/\tau} \Gamma(1/\tau)} \times \exp\left(-\frac{1}{2}|wv|^\tau\right) dw \\ &= \frac{2^{1/\tau}}{v\tau} \frac{\tau v}{(1+v^2)^{1/\tau} \Gamma(1/\tau)} \int_{(zv)^\tau/2}^{+\infty} u^{1/\tau-1} e^{-u} du \end{aligned} \quad (\text{A13})$$

$$= \frac{1}{(1+v^2)\Gamma(1/\tau)} \int_{(zv)^\tau/2}^{+\infty} u^{1/\tau-1} e^{-u} du. \quad (\text{A14})$$

L'égalité (A13) est immédiate après le changement de variable  $u = (-wv)^\tau/2$  et en posant  $s = -z > 0$ . À remarquer que l'intérieur de l'intégrale  $u^{1/\tau-1} e^{-u} du$  rappelle la fonction gamma. On aura besoin de la fonction gamma complète  $\Gamma(\cdot)$  et de sa version incomplète  $\gamma(\cdot, \cdot)$  qui sont définies par :

$$\gamma(a, r) = \int_0^r t^{a-1} e^{-t} dt \quad a > 0, r > 0$$

$$\Gamma(a) = \int_0^{+\infty} t^{a-1} e^{-t} dt \quad a > 0.$$

Le paramètre  $a$  est la *shape* de ces fonctions. Il est facile de voir que  $\Gamma(a) = \gamma(a, +\infty)$ . D'autre part, on a aussi la distribution qui porte le même nom gamma<sup>10</sup> dont la cumulative de paramètre *shape* =  $a$  (et *scale* = 1 car normée) évaluée au point  $x > 0$  s'écrit  $G_a(x) = [\Gamma(a)]^{-1} \gamma(a, x)$ . Le calcul de  $F_{\text{SEP3},v,\tau}^0(z)$  peut être obtenu à partir de l'égalité (A14) :

<sup>10</sup> Sous la même désignation « gamma », trois entités sont à distinguer : la fonction  $\Gamma(\cdot)$  (complète de 0 jusqu'à  $+\infty$ ), la fonction incomplète (son intégrale arrête à un point  $r < +\infty$ ). La troisième entité est la distribution gamma ayant deux paramètres : le *shape* et le *scale*.

$$\begin{aligned}
 F_{\text{SEP3},v,\tau}^0(z) &= \frac{1}{(1+v^2)\Gamma(1/\tau)} \int_{(zv)^\tau/2}^{+\infty} u^{1/\tau-1} e^{-u} du \\
 &= \frac{1}{(1+v^2)\Gamma(1/\tau)} \left\{ \int_0^{+\infty} - \int_0^{(zv)^\tau/2} \right\} = \frac{1}{(1+v^2)} \frac{\Gamma(1/\tau) - \gamma\left(1/\tau, (zv)^\tau/2\right)}{\Gamma(1/\tau)} \quad (\text{A15})
 \end{aligned}$$

$$= \frac{1}{1+v^2} \left( 1 - G_{1/\tau} \left( \frac{|zv|^\tau}{2} \right) \right) \quad (\text{A16})$$

L'égalité (A15) est un découpage des bornes de l'intégrale qui permet de retrouver les fonctions gamma. Pour sauver de l'espace, nous n'avons pas inséré les intégrants des deux intégrales dans (A15), qui sont les mêmes que dans l'équation précédente. L'expression se simplifie en utilisant la cumulative  $G_{1/\tau}$  (*shape* =  $1/\tau$  et *scale* = 1). En inversant (A16), le quantile de la VaR à un degré de confiance  $(1-p)$  est immédiat :

$$\begin{aligned}
 F_{\text{SEP3},v,\tau}^0\left(\frac{q-\mu}{\sigma}\right) &= p \\
 q &= \mu + \sigma \times \frac{\left[ 2 \times G_{1/\tau}^{-1}(1-p(1+v^2)) \right]^{1/\tau}}{v}
 \end{aligned}$$

Le calcul du *tail* de la SEP3 est similaire à celui fait pour la cumulative, mais avec une *shape*  $2/\tau$ , pour  $x < 0$ :

$$\begin{aligned}
 \text{Tail}_{\text{SEP3},v,\tau}(x) &= \int_{-\infty}^x z f^0(z) dz = \int_{-\infty}^x c \times z \times \exp\left(-\frac{1}{2}|zv|^\tau\right) dz \\
 &= \frac{-2^{1/\tau}}{v(1+v^2)\Gamma(1/\tau)} \int_{(-xv)^\tau/2}^{+\infty} u^{2/\tau-1} e^{-u} du \\
 &= \frac{-2^{1/\tau}}{v(1+v^2)\Gamma(1/\tau)} \Gamma(2/\tau) \left( 1 - G_{2/\tau} \left( \frac{|xv|^\tau}{2} \right) \right). \quad (\text{A17})
 \end{aligned}$$

Enfin, en mettant (A16) et (A17) dans (A4) on trouve :

$$\text{CVaR}_{\text{SEP3},\mu,\sigma,v,\tau} = \frac{1}{p} \left\{ \frac{1}{1+v^2} \left[ \mu \times \left( 1 - G_{1/\tau} \left( \frac{|q-\mu|v^\tau}{2} \right) \right) - \sigma \times \frac{2^{1/\tau}}{v} \frac{\Gamma(2/\tau)}{\Gamma(1/\tau)} \left( 1 - G_{2/\tau} \left( \frac{|q-\mu|v^\tau}{2} \right) \right) \right] \right\}.$$

On rappelle que  $G_{n/\tau}(x)$  est la cumulative de la loi Gamma de *shape* =  $n/\tau$  et *scale* = 1 évaluée au point  $x$ . Quand  $\tau=2$ , on revient à la SN2. Si  $\tau=2$ , et  $v=1$ , on obtient une loi normale. La loi gamma et la fonction gamma complète existent dans Excel.

### A10. Notations et expressions des *backtests* de la CVaR

Les *backtests* effectués dans cette note sont en hors échantillon. Notons la série des 1 200 rendements quotidiens  $\{X_t\}_{t=1}^{t=1200}$ . Nous avons huit modèles  $M_i, i=1\dots 8$ . Pour le modèle  $M$  et le jour  $t$ , nous prenons les 250 rendements précédant ce jour pour estimer le vecteur de paramètres du modèle  $M$  qu'on notera  $\theta_t$ . À partir de ce vecteur  $\theta_t$ , on calcule les mesures  $\text{VaR}_{p,t}$  et  $\text{CVaR}_{p,t}$  relatives au degré de confiance  $(1-p)$ . On rappelle ici que  $\text{VaR}_{p,t} > 0$  et  $\text{CVaR}_{p,t} > 0$  pour tout  $t$  par convention. Pour chaque modèle, on aura construit deux séries de taille 1 200 chacune :

$$\left\{ \text{VaR}_{p,t} \right\}_{t=1}^{t=1200} \text{ et } \left\{ \text{CVaR}_{p,t} \right\}_{t=1}^{t=1200}.$$

Le premier *backtest* utilisé est  $Z_{\text{ES}}$  d'Acerbi et Szekely (2017), dont l'expression de la statistique du test est :

$$Z_{\text{ES}}(X_t) = \frac{1}{T} \sum_{t=1}^T \frac{p \times (\text{CVaR}_{p,t} - \text{VaR}_{p,t}) + (X_t + \text{VaR}_{p,t})(X_t + \text{VaR}_{p,t}) < 0}{p \times \text{CVaR}_{p,t}}. \quad (\text{A18})$$

L'hypothèse nulle  $H_0$  du test  $Z_{\text{ES}}$  est que la CVaR convient et, dans ce cas, la statistique  $Z_{\text{ES}}(X_t)$  doit être statistiquement nulle. L'hypothèse alternative  $H_1$  est que la CVaR est sous-estimée ou surestimée :  $H_0 : Z_{\text{ES}}(X_t) = 0; H_1 : Z_{\text{ES}}(X_t) \neq 0$ . Voici la procédure pour calculer la distribution de l'hypothèse nulle de  $Z_{\text{ES}}(X_t)$ . Pour chaque jour  $t$ , on effectue un tirage de  $N$

valeurs aléatoires utilisant le modèle  $M$  avec les paramètres  $\theta_t$ . En prenant par exemple  $N = 5\,000$ , les tirages génèrent une matrice  $\{Y_t^n\}$  de 1 200 colonnes et 5 000 lignes. En application de l'équation (A18) et en remplaçant  $X_t$  par  $Y_t^n$  on calcule la série de 5 000 valeurs  $\{Z_{ES}(Y_t^n)\}_{n=1}^{n=5000}$ . La  $p$ -value du test  $Z_{ES}$  est alors égale à  $\min[\Pr(Z_{ES}(Y) < Z_{ES}(X)), \Pr(Z_{ES}(Y) > Z_{ES}(X))]$ .

Le deuxième *backtest* est noté RC et est proposé par Righi et Ceretta (2015). Sa statistique est définie par l'expression :

$$RC(X_t) = \frac{1}{T} \sum_{t=1}^T \frac{(X_t + CVaR_{p,t}) \times (X_t + VaR_{p,t}) < 0}{SD_{p,t}} \quad (A19)$$

où  $SD_{p,t} = \sqrt{\text{variance}(X_t \times (X_t + VaR_{p,t}) < 0)}$  est l'écart type des  $X_t$  qui dépassent la VaR. En version standard de Righi et Ceretta (2015), la  $p$ -value est obtenue par *bootstrapping*, selon Efron et Tibshirani (1994). Ici, on va plutôt l'obtenir en suivant exactement la même construction que pour  $Z_{ES}$ .

Pour finir, et pour information uniquement, les statistiques des *backtests*  $Z_1$  et  $Z_2$  sont définies par :

$$Z_1(X_t) = \frac{\sum_{t=1}^T \frac{X_t}{CVaR_{p,t}} \times (X_t + VaR_{p,t}) < 0}{\sum_{t=1}^T 1 \times (X_t + VaR_{p,t}) < 0} + 1$$

$$Z_2(X_t) = \frac{1}{T \times p} \sum_{t=1}^T \frac{X_t \times (X_t + VaR_{p,t}) < 0}{CVaR_{p,t}} + 1.$$

## A11. Estimation des modèles et calculs des VaR et des CVaR paramétriques et non paramétriques

Les paramètres estimés des distributions sont donnés dans les tableaux suivants.

Tableau A.2  
Estimation des modèles - Panel A

	1 :NO	1 :T	1 :EGB2	2 :NO	2 :T
$\mu_1$	0,0005254 (0,0003740)	0,0006974** (0,0002977)	0,0008884* (0,0004982)	-0,0004845 (0,0015691)	0,0012920** (0,0005553)
$\sigma_1$	0,0129559*** (0,0002645)	0,0085310*** (0,0003410)	0,0014108** (0,0006812)	0,0226636*** (0,0018632)	0,0066854*** (0,0009171)
$\nu_1$		3,2887197*** (0,3809600)	0,1587161** (0,0796200)		23 642,3157236*** (0,0000001)
$\tau_1$			0,1652522* (0,0851634)		
$\mu_2$				0,0008151** (0,0003448)	-0,0004740 (0,0008931)
$\sigma_2$				0,0082545*** (0,0005136)	0,0140598*** (0,0025828)
$\nu_2$					6,4162601** (2,5707612)
$\tau_2$					
$c_1$				0,2231962*** (0,0497856)	0,5158049*** (0,1538992)
Nombre param.	2	3	4	5	7
LogLik	3 512,5505	3 627,5723	3 625,5829	3 619,0444	3 628,6421
AIC	-7 021,1011	-7 249,1446	-7 243,1659	-7 228,0889	-7 243,2842
BIC	-7 010,9210	-7 233,8744	-7 222,8056	-7 202,6385	-7 207,6536
KS ( <i>p</i> -value)	0,0015	0,1285	0,3490	0,2180	0,1100
Nombre obs.	1 200	1 200	1 200	1 200	1 200

\*\*\* $p < 0,01$  \*\* $p < 0,05$  \* $p < 0,1$

Tableau A.3  
Estimation des modèles - Panel B

	3 :NO	2 :SN2 + 1 :NO	2 :SEP3
$\mu_1$	-0,0004753 (0,0009649)	0,0025930 (0,0065819)	-0,0007520*** (0,0001560)
$\sigma_1$	0,0150441*** (0,0022041)	0,0146788*** (0,0020169)	0,0045291*** (0,0014071)
$\nu_1$		0,8830671*** (0,2038416)	1,0315089*** (0,0376383)
$\tau_1$			0,9598700*** (0,1180946)
$\mu_2$	0,0043390 (0,0098212)	0,0009227 (0,0013866)	0,0075456** (0,0032033)
$\sigma_2$	0,0376531*** (0,0101797)	0,0063897*** (0,0012556)	0,0065018** (0,0025539)
$\nu_2$		0,9939552*** (0,2581567)	0,6137048*** (0,2182171)
$\tau_2$			2,1083901* (1,1436395)
$\mu_3$	0,0011752*** (0,0004491)	0,0091833 (0,0149751)	
$\sigma_3$	0,0065771*** (0,0008483)	0,0388917*** (0,0081479)	
$c_1$	0,4433715*** (0,1089861)	0,4729333*** (0,1594929)	0,7389303*** (0,1181466)
$c_2$	0,0334707 (0,0303812)	0,5000573*** (0,1734363)	
Nombre param.	8	10	9
LogLik	3 630,0653	3 630,7232	3 631,9633
AIC	-7 244,1307	-7 241,4465	-7 245,9267
BIC	-7 203,4101	-7 190,5457	-7 200,1160
KS ( <i>p</i> -value)	0,2280	0,1980	0,3040
Nombre obs.	1 200	1 200	1 200

\*\*\**p* < 0,01 \*\**p* < 0,05 \**p* < 0,1

Tableau A.4  
Calcul et comparaison des CVaR

p	Densités /VaR Mélanges (en %)	CVaR-/CVaR+ CVaR (en %)	Moyenne (en %)	Variance (en %)	Asymétrie	Kurtosis
0,050	nparam 2,04736	2,96269/2,97795	0,0524	0,0168	0,3589	9,8158
0,050	1:NO 2,07852	2,61989	0,0525	0,0168	0,0000	3,0000
0,050	1 :T 1,86805	3,01294	0,0697	0,0186	0,0000	
0,050	1 :EGB2 2,00674	2,89562	0,0525	0,0157	-0,0813	5,8076
0,050	2:NO 1,95397	3,11363	0,0525	0,0168	-0,1386	6,6789
0,050	2 :T 2,02945	3,04197	0,0437	0,0163	-0,1544	8,3993
0,050	3:NO 2,03846	3,00451	0,0549	0,0172	0,1224	9,4321
0,050	2 :SN2 + 1 :NO 2,05018	2,98338	0,0524	0,0168	0,2433	9,8409
0,050	2 :SEP3 1,99293	2,97395	0,0544	0,0163	0,0051	7,1752
0,025	nparam 2,54290	3,63040/3,66665				
0,025	1:NO 2,48677	2,97630				
0,025	1 :T 2,51522	3,87890				
0,025	1 :EGB2 2,62287	3,51175				
0,025	2:NO 2,81354	3,90424				
0,025	2 :T 2,71654	3,74976				
0,025	3:NO 2,66598	3,68928				
0,025	2 :SN2 + 1 :NO 2,68920	3,62898				
0,025	2 :SEP3 2,66110	3,66159				
0,010	nparam 3,59575	4,44800/4,51902				
0,010	1 :NO 2,96146	3,40049				
0,010	1 :T 3,54473	5,29712				
0,010	1 :EGB2 3,43734	4,32622				
0,010	2 :NO 3,89559	4,82632				
0,010	2 :T 3,62577	4,72258				
0,010	3 :NO 3,47885	4,71115				
0,010	2 :SN2 + 1 :NO 3,47913	4,53241				
0,010	2 :SEP3 3,57259	4,58396				