Le pire des cas dans le choix de la copule

Comment éviter le pire

Brice Franke

Département de Mathématique
Université de Bretagne Occidentale
29200 Brest

Solvency 2 impose aux assureurs une analyse des risques accumulés sur plusieurs produits d'assurances.

- le risque est quantifié par la valeur à risque au niveau α:

$$\mathrm{VaR}_\alpha := \inf\left\{z \in \mathbb{R} : F_p(z) \geq \alpha\right\};$$

- F_p est la fonction de répartition des pertes annuelles;
- la VaR est difficile à calculer car la fonction de répartition des pertes F_p est difficile à estimer;
- un des problèmes vient du fait qu'on a souvent à faire à des risques composés qui viennent de différents produits;
- dans ce cas il est nécessaire d'avoir des outils empriques qui permettent d'estimer les risques accumulés.

- X la perte annuelle due au contrat de type A;
- Y la perte annuelle due au contrat de type B;
- $X + Y$ pertes annuelles accumulées;
- $\gamma > 0$ capital en réserve pour couvrir les pertes annuelles;
- $X + Y > \gamma$ situation de ruine.

Modélisation de la probabilité de ruine:

- (X, Y) un couple de variables aléatoires avec loi jointe μ;
- la probabilité de ne pas être confronté à la ruine est:

$$\mathbb{P}\Big(X + Y \leq \gamma\Big) = \iint \mathbf{1}_{(-\infty,\gamma)}(x + y)\mu(dx, dy).$$

Problème: La loi jointe n'est pas connue en général!

Supposons que $\mu(dx, dy) = \rho(x, y)dxdy$, alors la loi jointe μ est paramétrisé par

- les fonctions de répartitions F_X et F_Y de X et Y;
- on a $F_X(X) \sim \mathcal{U}[0, 1]$ et $F_Y(Y) \sim \mathcal{U}[0, 1]$;
- la fonction de répartition bivariée $C_{(X,Y)}$ du couple de variables aléatoires $(F_X(X), F_Y(Y))$; i.e.:

$$C_{(X,Y)}(z_1, z_2) := \mathbb{P}\Big(F_X(X) \leq z_1, F_Y(Y) \leq z_2\Big).$$

Définition

On appelle copule toute fonction de répartition bivariée d'un couple de variables aléatoires (U_1, U_2) *avec* $\mathcal{L}(U_i) = \mathcal{U}[0, 1]$ *pour* $i = 1, 2$.

Théorème (Sklar)

Pour tout couple de variables aléatoires (X, Y) il existe une copule $C(z_1, z_2)$ de sorte que

$$\mathbb{P}\Big(X \leq z_1, Y \leq z_2\Big) = C(F_X(z_1), F_Y(z_2)).$$

La copule est unique si la loi jointe de (X, Y) a une densité.

Souvent la copule associée au couple (X, Y) et la fonction de répartition bivariée du couple $(F_X(X), F_Y(Y))$. La probabilité de ne pas être ruiné en termes de la copule est alors:

$$\begin{aligned}
\mathbb{P}(X + Y \leq \gamma) &= \mathbb{P}\Big(F_X^{-1}(F_X(X)) + F_Y^{-1}(F_Y(Y)) \leq \gamma\Big) \\
&= \iint \mathbf{1}_{(-\infty,\gamma)}(F_X^{-1}(x) + F_Y^{-1}(y))C(dx, dy)
\end{aligned}$$

Les différentes composantes d'un modèle multivarié sont souvent connues avec une fiabilité de différente qualité:

Les lois marginales sont plutôt bien connues:

- longue période de données enregistrées;
- bonne résolution statistique du problème unidimensionnel;
- multitude de méthodes statistiques pour justifier le choix du modèle paramétrique et ajuster les paramètres;
- résultats théoriques pour justifier le choix des modèles.

La copule est plutôt mal connue

- peu de données;
- problème de haute dimensionnalité;
- peu de méthodes statistiques;
- peu de résultats théoriques.

- le choix de la copule est souvent moins bien justifié que le choix des lois marginales;
- une analyse du pire des cas peut donner des informations complémentaires sur les risques;
- si les lois marginales P et Q de X resp. Y sont connues le pire des cas pour la probabilité de non-ruine est:

$$W(P, Q, \gamma) := \inf_{\mu \in \mathcal{M}(P,Q)} \iint \mathbf{1}_{(-\infty,\gamma)}(x+y)\mu(dx, dy);$$

- $\mathcal{M}(P, Q) := \left\{\mu \in \mathcal{M}(\mathbb{R}^2);\ \mu \circ \mathrm{pr}_1^{-1} = P,\ \mu \circ \mathrm{pr}_2^{-1} = Q\right\};$
- $\mathcal{M}(\mathbb{R}^2)$ mesures de Borel sur $\mathbb{R}^2$;
- $\mathrm{pr}_i : \mathbb{R}^2 \to \mathbb{R}; (z_1, z_2) \mapsto z_i$ projections cartésiennes.

- évaluer $W(P, Q, \gamma)$ est un problème posé par Kolmogorov;

Théorème (Makarov (1981), Frank, Nelson, Schweizer (1987))

$$W(P, Q, \gamma) = \sup_{y \in \mathbb{R}} \Big(F_X(y) + F_Y(\gamma - y) - 1 \Big) \vee 0;$$

- la démonstration peut être basée sur un théorème de dualité

$$W(P, Q, \gamma) = \sup_{(f,g) \in \mathcal{F}(\gamma)} \left(\int f(x) P(dx) + \int g(y) Q(dy) \right);$$

- $\mathcal{F}(\gamma) := \Big\{ (f, g) \in C(\mathbb{R}^2);\ f(x) + g(y) \leq \mathbf{1}_{(-\infty,\gamma)}(x, y) \Big\};$

Comment utiliser le nuage de données bivariées $(x_i, y_i)_{i=1,...N}$?

- le choix d'une famille paramétrique de copules est dans beaucoup de situations difficile à justifier;
- l'estimation de moments mixtes (covariance etc.) donne de l'information fiable sur les lois bivariées;
- un grand échantillon de données bivariées permet d'estimer des moments mixtes de grand ordre;
- si tous les moments mixtes sont connus, alors la loi bivariée peut être reconstituée

Il est donc envisageable d'utiliser une estimation des moments mixtes pour réduire l'ensemble des lois bivariées dans l'analyse du pire des cas.

- on peut par exemple estimer le premier moment mixte:

$$m_{11}(\mu) := \iint xy\mu(dx,dy) \quad \text{par} \quad \hat{m}_{11} := \frac{1}{n}\sum_j x_j y_j;$$

- le pire des cas avec covariance prescrite est alors

$$W(P,Q,\hat{m}_{11},\gamma) := \inf_{\mu\in\mathcal{M}(P,Q,\hat{m})} \iint \mathbf{1}_{(-\infty,\gamma)}(x+y)\mu(dx,dy)$$

- avec

$$\mathcal{M}(P,Q,\hat{m}) = \left\{\mu \in \mathcal{M}(P,Q); \iint xyd\mu = \hat{m}_{11}\right\}$$

- on peut aussi rajouter d'autres moments mixtes.

Théorème (Franke, Stolz)

On a pour tous les $m \in \mathbb{R}$ qui satisfont

$$\inf_{\mu \in \mathcal{M}(P,Q)} \iint xy d\mu \leq k \leq \sup_{\mu \in \mathcal{M}(P,Q)} \iint xy d\mu$$

la dualité suivante

$$W(P,Q,k,\gamma) = \sup_{(f,g,\alpha) \in \mathcal{U}(\gamma)} \left(\int f dP + \int g dQ + \alpha k \right)$$

avec

$$\mathcal{U}(\gamma) := \left\{ (f,g,\alpha) : f(x) + g(y) + \alpha xy \leq \mathbf{1}_{(-\infty,\gamma)}(x+y) \right\}.$$

Comment faire le choix sur les moments à utilser:

- déterminer les moments qui peuvent être estimés de façon robuste

$$(k,l) \in \mathcal{K} \iff \sup_{1 \le i \le N} \left| \frac{1}{N} \sum_{j=1}^{N} x_j^k y_j^l - \frac{1}{N-1} \sum_{i \neq j} x_j^k y_j^l \right| < \epsilon;$$

- utiliser l'estimateur de moment $\hat{m}_{kl} := \frac{1}{N} \sum_{j=1}^{N} x_j^k y_j^l$ pour réduire l'ensemble des lois

$$\mathcal{M}(P,Q,\mathcal{K}) := \left\{ \mu \in \mathcal{M}(P,Q);\ \iint x^k y^l \mu(dx,dy) = \hat{m}_{kl} \right\};$$

- le pire des cas basé sur les moments mixtes est alors

$$W_N(P,Q,\mathcal{K},\gamma) := \inf_{\mu \in \mathcal{M}(P,Q,\mathcal{K})} \iint \mathbf{1}_{(-\infty,\gamma)}(x+y)\mu(dx,dy).$$

Une autre façon de faire le choix des moments: :

- utiliser les intervalles de confiance pour μ fixe

$$\hat{\beta}_{kl}^{(\delta)}(\mu) := \inf\left\{\beta > 0; \mu^{\otimes N}\left(\left|\hat{m}_{kl} - m_{kl}(\mu)\right| \leq \beta\right) \geq 1 - \delta\right\}$$

avec $\hat{m}_{kl} := \frac{1}{N}\sum_{j=1}^{N} x_j^k y_j^l$ et $m_{kl}(\mu) := \iint x^k y^l d\mu$;

- déterminer les intervalles de confiance pire des cas

$$\hat{\beta}_{kl}^{(\delta)} := \inf_{\mu \in \mathcal{M}(P,Q)} \hat{\beta}_{kl}^{(\delta)}(\mu);$$

- le pire des cas basé sur les intervalles de confiances est

$$W_N(P, Q, \delta, \gamma) := \inf_{\mu \in \mathcal{M}(P,Q,\delta)} \iint \mathbf{1}_{(-\infty,\gamma)}(x + y)\mu(dx, dy);$$

- $\mathcal{M}(P, Q, \delta) := \left\{\mu \in \mathcal{M}(P, Q); \left|m_{kl}(\mu) - \hat{m}_{kl}\right| \leq \hat{\beta}_{kl}^{(\delta)}\right\}$

On peut obtenir une expression asymptotique pour $\hat{\beta}_{kl}^{(\delta)}$

- soit $\sigma_{kl}(\mu)^2 := \iint (x^k y^l)^2 d\mu - \left(\int x^k y^l d\mu \right)^2$;
- le théorème limite central implique pour $N \to \infty$

$$
\begin{aligned}
1 - \delta &= \mu^{\otimes N}\left(\left| \hat{m}_{kl} - m_{kl}(\mu) \right| \leq \beta_N \right) \\
&= \mu^{\otimes N}\left(\frac{\sqrt{N}}{\sigma_{kl}(\mu)} \left| \hat{m}_{kl} - m_{kl}(\mu) \right| \leq \frac{\sqrt{N}\beta_N}{\sigma_{kl}(\mu)} \right) \\
&\longrightarrow 2\Phi\left(\frac{\sqrt{N}\beta_N}{\sigma_{kl}(\mu)} \right) - 1;
\end{aligned}
$$

- donc asymptotiquement

$$\beta_N = \frac{\sigma_{kl}(\mu)}{\sqrt{N}} \Phi^{-1}(1 - \delta/2)$$

- Cauchy-Schwarz implique

$$\sigma_{kl}(\mu) \leq \int (x^k)^2 dP \int (y^l)^2 dQ$$

- donc on a quand N est grand:

$$\hat{\beta}_{kl}^{(\delta)} \leq \frac{\sigma_{kl}(\mu)}{\sqrt{N}} \Phi^{-1}(1 - \frac{\delta}{2});$$

- l'ensemble des lois bivariées admissibles est alors

$$\mathcal{M}_o(P, Q, \delta) := \left\{ \mu \in \mathcal{M}(P, Q); \left| m_{kl}(\mu) - \hat{m}_{kl} \right| \leq ... \right\}$$

- et le pire des cas peut être borné par

$$W_N^o(P, Q, \delta, \gamma) := \inf_{\mu \in \mathcal{M}_o(P,Q,\delta)} \iint \mathbf{1}_{(-\infty,\gamma)}(x + y)\mu(dx, dy);$$

Théorème (Consistance de l'estimation du pire des cas)

Soit $(X_i, Y_i)_{i\in\mathbb{N}}$ une suite de variables aléatoires bivariées indépendantes et de même loi jointe μ_o avec densité ρ_o. Alors on a quand $N \to \infty$

$$W_N(P, Q, \delta, \gamma) \xrightarrow{p.s.} \iint \mathbf{1}_{(-\infty,\gamma)}(x+y)\mu_o(dx, dy).$$

Démonstration:

- La loi du grand nombre implique pour $n \to \infty$:

$$\hat{m}_{kl} = \frac{1}{N}\sum_{j=1}^{N} x_j^k y_j^l \xrightarrow{\text{p.s.}} \iint x^k y^l \mu(dx, dy) = m_{kl}(\mu);$$

- de plus $\hat{\beta}_{kl}^{(\delta)} \to 0$ quand $N \to \infty$;

- soit $\mu_N \in \mathcal{M}(P, Q, \delta)$ la mesure bivariée qui minimise

$$W_N(P, Q, \delta, \gamma) = \inf_{\mu \in \mathcal{M}(P,Q,\delta)} \iint \mathbf{1}_{(-\infty,\gamma)}(x+y)\mu(dx, dy);$$

- donc μ_N satisfait $\left|m_{kl}(\mu_N) - \hat{m}_{kl}\right| \leq \hat{\beta}_{kl}^{(\delta)} \to 0$
- ceci implique $\left|m_{kl}(\mu_N) - m_{kl}(\mu_o)\right| \xrightarrow{\text{p.s.}} 0;$
- la suite $(\mu_N)_{N\in\mathbb{N}}$ est tendue et il existe donc une sous-suite qui converge en lois vers une mesure μ_∞;
- on a $m_{kl}(\mu_o) = m_{kl}(\mu_\infty)$ pour tout couple (k, l);
- ici les moments déterminent la mesure; i.e.: $\mu_o = \mu_\infty$;
- donc on a

$$W_N(P, Q, \delta, \gamma) \longrightarrow \iint \mathbf{1}_{(-\infty,\gamma)}(x+y)\mu_o(dx, dy);$$

$\square$

- Le choix d'une famille de copules $(C_\theta(x,y))_{\theta\in\Theta}$ nous expose au risque de sous-estimer la probabilité d'être ruiné;
- ce risque peut être détecté en comparant les pires des cas;

$$W(P,Q,C,\gamma) := \sup_{\theta\in\Theta} \iint \mathbf{1}_{(-\infty,\gamma)}(x+y) C_\theta(F_X(dx), F_Y(dy))$$

$$W(P,Q,\mathcal{K},\gamma) := \sup_{\mu\in\mathcal{M}(P,Q,\mathcal{K})} \iint \mathbf{1}_{(-\infty,\gamma)}(x+y)\mu(dx,dy);$$

- si $W(P,Q,C,\gamma)$ est beaucoup plus petit que $W(P,Q,\mathcal{K},\gamma)$ alors il existe un risque non-negligeable de sous-estimer la probabilité de la ruine;

- Souvent les moments d'ordre suppérieur n'existent pas;
- dans ce cas on peut transformer le nuage avec les marges:

$$(\tilde{x}_i, \tilde{y}_i) := (F_X(x_i), F_Y(y_i)); \quad i = 1, ..., N;$$

- les marges suivent des lois uniformes et tous les moments purs et mixtes existent;
- le pire des cas devient alors

$$\tilde{W}(\gamma) = \inf_{C \in \mathcal{C}(\delta)} \iint \mathbf{1}_{(-\infty,\gamma)}(F_X^{-1}(x) + F_Y^{-1}(y))C(dx, dy);$$

- avec

$$\mathcal{C}(\delta) = \left\{ C \text{ copule; } \iint x^k y^l C(dx, dy) = \hat{m}_{kl} \, \forall (k, l) \in \tilde{\mathcal{K}}(\delta) \right\}.$$

Programme et questions ouvertes:

- la fonction de répartition pire des cas pour $X + Y$ est

$$\gamma \mapsto W(P, Q, \gamma) = \inf_{\mu \in \mathcal{M}(P,Q)} \iint \mathbf{1}_{(-\infty,\gamma)}(x + y)\mu(dx, dy);$$

- peut on dire quelque chose sur les variations de la fonction $\gamma \mapsto W(P, Q, \gamma)$ pour obtenir de l'information sur les pires extrêmes du couple $X + Y$?
- peut on calculer une probabilité du ruine pire des cas à partir de cette fonction de répartition?

Litérature:

- Embrechts P., Frey R., McNeil A.: Quantitative Risk Management, Princeton Series in Finance, 2005.
- Frank M., Nelse, R., Schweizer B.: Best possible bounds on the distribution of a sum, Probability and Related Fields, Vol. 74, p.199-211
- Franke B., Stolz M.: A duality approach to the worst case value at risk for a sum of dependent random variables with known covariances, arXiv:0912.1841v1
- Makarov G.: Estimates for the distribution function of a sum of two random variables when the marginal distributions are fixed, Theory of Probability and its Applications, Vol. 26, p. 803-806
- Mikosch T.: Copulas: tales and facts, Extremes, Vol. 9, p. 3-20, (2006)

Merci pour votre attention!