

CONSTRUCTION ET SIMULATION DES LOIS STABLE-TWEEDIE MULTIVARIÉES

Johann Cuenin ¹, Bent Jørgensen ² & Célestin C. Kokonendji ¹

¹ *Université de Franche-Comté, UFR Sciences et Techniques,
Laboratoire de Mathématiques de Besançon - UMR 6623 CNRS-UFC,*

16 route de Gray F-25030 Besançon Cedex, France

*Emails : johann.cuenin@univ-fcomte.fr (Johann Cuenin),
celestin.kokonendji@univ-fcomte.fr (Célestin C. Kokonendji)*

² *University of Southern Denmark,*

Department of Mathematics and Computer Science,

Campusvej 55 DK-5230 Odense M, Denmark

Email : bentj@stat.sdu.dk

Résumé. De manière comparable à la loi normale multivariée, une construction et une méthode de simulation pour les lois stable-Tweedie multivariées sont introduites. Cette famille contient en particulier les lois normale, Poisson, gamma décentrée, gamma et gaussienne inverse. La méthode de construction est basée sur une combinaison linéaire de variables aléatoires indépendantes de lois stable-Tweedie univariées. La corrélation entre les marginales est établie en utilisant les fonctions génératrices des cumulants. Cette méthode est simplifiée par les propriétés de convolution et d'échelle des lois stable-Tweedie univariées. En outre, elle permet de simuler un vecteur aléatoire stable-Tweedie, lequel est caractérisé par un vecteur position et une matrice de dispersion, à l'instar du vecteur gaussien. Les résultats numériques obtenus sont très satisfaisants tant pour des corrélations positives que négatives.

Mots-clés. Fonction génératrice des cumulants, corrélation, modèles exponentiels de dispersion, lois α -stables.

Abstract. Comparably to the well known Gaussian vector, a construction and a way of simulation are introduced for multivariate stable-Tweedie distributions. This family contains the following distributions: Gaussian, Poisson, non-central gamma, gamma and inverse Gaussian. The construction method is based on a linear combination of independent univariate stable-Tweedie variables. The correlation between the margins is established by using cumulant generating functions. The method is facilitated by the convolution and scale properties of univariate stable-Tweedie distributions. Moreover, it allows simulations of a random stable-Tweedie vector, which is characterized by a position vector and a dispersion matrix, like the Gaussian vector. The results of simulations obtained are very satisfactory for both positive and negative correlations.

Keywords. Cumulant generating function, correlation parameter, multivariate exponential dispersion models, α -stable distributions.

1 Introduction

On s'intéresse dans cette communication à la construction et à une méthode de simulation des lois stable-Tweedie multivariées que l'on va nommer *vecteurs aléatoires stable-Tweedie*. On peut se référer à Furman et Landsman (2010) ainsi qu'à Alai et al (2013) pour une approche pratique des modèles multivariés. Introduites par M. C. K. Tweedie (1984) comme extension des lois α -stables, les lois stable-Tweedie univariées ont été étudiées par Jørgensen (1997) dans le cadre des modèles exponentiels de dispersion. Ces lois ont des applications en écologie, finance ou encore en physique (e.g. Kendal, 2014). On rappelle qu'une loi α -stable est une loi telle que si X est une variable aléatoire suivant ladite loi, alors pour n réplifications X_1, \dots, X_n de cette variable, on a $X_1 + \dots + X_n \stackrel{\mathcal{D}}{=} n^{1/\alpha}X + b_n$, avec $\alpha \in]0, 2]$ et b_n une constante déterministe et où $\stackrel{\mathcal{D}}{=}$ représente l'égalité en loi. Les lois stable-Tweedie ont leur paramètre α étendu à $[-\infty, 2]$. Du point de vu des modèles de dispersion, les lois stable-Tweedie ont une fonction de variance définie par $V(\mu) = \mu^p$, où μ est le paramètre de position et $p \in \mathbb{R} \setminus]0, 1[$ un paramètre de puissance lié à α par la relation $(1-p)(\alpha-1) = 1$. En variant p on peut retrouver plusieurs lois connues telles que la loi normale ($p = 0$), la loi de Poisson ($p = 1$), les lois gamma ($p = 2$) et gamma décentrée ($p = 3/2$) ainsi que la loi gaussienne inverse ($p = 3$). On introduit maintenant la fonction génératrice des cumulants des lois stable-Tweedie :

$$\log \mathbb{E}e^{\theta X} = \kappa_\alpha(\theta) = \begin{cases} \exp(\theta) & \text{pour } p = 1 \\ -\log(-\theta) & \text{pour } p = 2 \\ [(\alpha-1)/\alpha] [\theta/(\alpha-1)]^\alpha & \text{pour } p \neq 1, 2, \end{cases}$$

telle que $\theta \in \Theta_p$, avec

$$\Theta_p = \begin{cases} \mathbb{R} & \text{pour } p = 0, 1 \\ [0, +\infty[& \text{pour } p < 0 \\]-\infty, 0[& \text{pour } 1 < p \leq 2 \\]-\infty, 0[& \text{pour } 2 < p. \end{cases}$$

Nous pouvons mettre en relation cette fonction, qui est analytique à l'intérieur de $\overset{\circ}{\Theta}_p$, avec les moyenne et variance de la distribution. En effet la dérivée première à l'intérieur de $\overset{\circ}{\Theta}_p$ donnera la moyenne théorique du modèle, et la dérivée seconde, la variance.

À l'instar des modèles exponentiels, il existe deux formes de lois stable-Tweedie : la forme additive, que l'on notera $X \sim \mathcal{T}w^{*p}(\mu, \lambda)$ telle que $\mathbb{E}X = \lambda\mu$ et $\text{Var}(X) = \lambda\mu^p$, et la forme reproductrice, notée $Y \sim \mathcal{T}w^p(\mu, \sigma^2)$ et telle que $\mathbb{E}Y = \mu$ et $\text{Var}(Y) = \sigma^2\mu^p$. Ces formes admettent respectivement les propriétés suivantes :

Propriété 1. Si $X_1 \sim \mathcal{T}w^{*p}(\mu, \lambda_1)$ et $X_2 \sim \mathcal{T}w^{*p}(\mu, \lambda_2)$, alors $X_1 + X_2 \sim \mathcal{T}w^{*p}(\mu, \lambda_1 + \lambda_2)$.

Propriété 2. Soit $Y \sim \mathcal{T}w^p(\mu, \sigma^2)$ et soit $c \in \mathbb{R}_+ \setminus \{0\}$, alors $cY \sim \mathcal{T}w^p(c\mu, c^{2-p}\sigma^2)$.

Ces propriétés impliquent que les deux formes sont en fait les mêmes après une reparamétrisation. Nous parlerons donc uniquement de la forme additive dans la suite. Nous pouvons alors introduire la fonction génératrice des cumulants pour un modèle stable-Tweedie additif :

$$K(\cdot; \mu(\theta), \lambda) : s \mapsto \begin{cases} \lambda \exp(\theta) [\exp(s) - 1] & \text{pour } p = 1 \\ -\lambda \log(1 + s/\theta) & \text{pour } p = 2 \\ \lambda [(1 + s/\theta)^\alpha - 1] \kappa_\alpha(\theta) & \text{pour } p \neq 1, 2, \end{cases}$$

avec $s \in \mathcal{S}_p$ et $\mu(\theta) \in M_p$ tels que

$$\mathcal{S}_p = \begin{cases} \mathbb{R} & \text{pour } p = 0, 1, +\infty \\]-\theta, +\infty[& \text{pour } 1 < p < +\infty \\ [-\theta, +\infty[& \text{pour } p < 0, \end{cases} \quad M_p = \begin{cases} \mathbb{R} & \text{pour } p = 0, +\infty \\]0, +\infty] & \text{pour } p \neq 0, +\infty. \end{cases}$$

2 Résultats

Nous présentons ici la construction théorique du vecteur aléatoire stable-Tweedie avec une corrélation positive entre les marginales. Cette construction est basée sur les résultats se trouvant dans Jørgensen et Martínez (2013). Nous donnons ensuite une écriture matricielle permettant la simulation d'un tel modèle et permettant d'introduire une corrélation négative. Les détails et illustrations des résultats se trouvent dans Cuenin et al. (2014) La construction dans le cas positivement corrélé est présentée ci-dessous :

Théorème 1. *Soit $k > 1$ et $p \in \{0\} \cup [1, +\infty[$. Soit $\boldsymbol{\mu} = (\mu_1, \dots, \mu_k)^\top \in M_p^k \subseteq \mathbb{R}^k$ et $\boldsymbol{\Lambda} = (\lambda_{ij})_{i,j=1,\dots,k}$ une matrice symétrique définie positive telle que $\lambda_{ij} > 0$ pour tout $i, j = 1, \dots, k$ et soit $\mathbf{X}^{(+)} = (X_1^{(+)}, \dots, X_k^{(+)})^\top$ un vecteur aléatoire de dimension k . alors les assertions suivantes sont équivalentes :*

(i) $X_j^{(+)} \sim \mathcal{T}w^{*p}(\mu_j, \lambda_{jj})$ et $\text{Cov}(X_i^{(+)}, X_j^{(+)}) = \lambda_{ij}(\mu_i \mu_j)^{p/2}$, pour tout $i, j = 1, \dots, k$;

(ii) Il existe une loi en dimension k pour $\mathbf{X}^{(+)}$, que l'on note $\mathcal{T}w_k^{*p(+)}(\boldsymbol{\mu}, \boldsymbol{\Lambda})$ et qui a pour fonction génératrice de cumulants la fonction

$$K_p(\mathbf{s}; \boldsymbol{\mu}, \boldsymbol{\Lambda}) = \sum_{i < j} K_{p;ij}(\mathbf{s}; \boldsymbol{\mu}, \boldsymbol{\Lambda} \boxtimes \boldsymbol{\mu}) + \sum_{j=1}^k K_{p;j}(\mathbf{s}; \boldsymbol{\mu}, \boldsymbol{\Lambda} \boxtimes \boldsymbol{\mu}), \quad (1)$$

avec $\mathbf{s} = (s_1, \dots, s_k)^\top \in \mathcal{S}_p^k$, $\boldsymbol{\mu} = \boldsymbol{\mu}(\boldsymbol{\theta})$ tel que $\boldsymbol{\theta} = (\theta_j)_{j=1,\dots,k}$ avec θ_j vu comme inverse de μ_j et $\boldsymbol{\Lambda} \boxtimes \boldsymbol{\mu} = (\gamma_{ij})_{i,j=1,\dots,k}$ une matrice symétrique dépendant de $\boldsymbol{\mu}$ et $\boldsymbol{\Lambda}$ avec $\gamma_{ij} = \gamma_{ij}(\boldsymbol{\mu}, \boldsymbol{\Lambda}) \in \mathbb{R} \setminus \{0\}$ donné par

$$\gamma_{ij} = \gamma_{ij}(\boldsymbol{\mu}, \boldsymbol{\Lambda}) = \begin{cases} \lambda_{ij} & \text{pour } p = 1, 2 \\ \lambda_{ij} [(\alpha - 1)\alpha] [\theta_i/(\alpha - 1)]^{\alpha/2} [\theta_j/(\alpha - 1)]^{\alpha/2} & \text{pour } p \neq 1, 2. \end{cases} \quad (2)$$

Les fonctions $K_{p;ij}$ et $K_{p;j}$ sont définies respectivement comme suit :

$$K_{p;ij}(\mathbf{s}; \boldsymbol{\mu}, \boldsymbol{\Lambda} \boxtimes \boldsymbol{\mu}) = \begin{cases} \gamma_{ij} [\exp(s_i + s_j) - 1] & \text{pour } p = 1 \\ -\gamma_{ij} \log(1 + s_i/\theta_i + s_j/\theta_j) & \text{pour } p = 2 \\ \gamma_{ij} [(1 + s_i/\theta_i + s_j/\theta_j)^\alpha - 1] & \text{pour } p \neq 1, 2, \end{cases}$$

$$K_{p;j}(\mathbf{s}; \boldsymbol{\mu}, \boldsymbol{\Lambda} \boxtimes \boldsymbol{\mu}) = \begin{cases} \left(\gamma_{jj} - \sum_{\ell \neq j} \gamma_{\ell j} \right) [\exp(s_j) - 1] & \text{pour } p = 1 \\ - \left(\gamma_{jj} - \sum_{\ell \neq j} \gamma_{\ell j} \right) \log(1 + s_j/\theta_j) & \text{pour } p = 2 \\ \left(\gamma_{jj} - \sum_{\ell \neq j} \gamma_{\ell j} \right) [(1 + s_j/\theta_j)^\alpha - 1] & \text{pour } p \neq 1, 2. \end{cases}$$

On obtient alors une loi stable-Tweedie multivariée, notée $\mathcal{T}w_k^{*p(+)}(\boldsymbol{\mu}, \boldsymbol{\Lambda})$. On remarque que pour $k = 1$, nous obtenons les lois stable-Tweedie, univariées connues. Les restrictions de cette construction pour des corrélations positives ou nulles sont exposées ci-dessous.

Corollaire 1. Soit $\mathbf{X}^{(+)} = (X_1^{(+)}, \dots, X_k^{(+)})^\top \sim \mathcal{T}w_k^{*p(+)}(\boldsymbol{\mu}, \boldsymbol{\Lambda})$ un vecteur aléatoire stable-Tweedie. Alors la corrélation entre $X_i^{(+)}$ et $X_j^{(+)}$, que l'on note $\rho_{ij} = \rho(X_i^{(+)}, X_j^{(+)})$, est donnée par

$$\rho_{ij} = \frac{|\gamma_{ij}(\boldsymbol{\mu}, \boldsymbol{\Lambda})|}{\sqrt{\gamma_{ii}(\boldsymbol{\mu}, \boldsymbol{\Lambda})\gamma_{jj}(\boldsymbol{\mu}, \boldsymbol{\Lambda})}} = \frac{\lambda_{ij}}{\sqrt{\lambda_{ii}\lambda_{jj}}} = \rho_{ij}(\boldsymbol{\Lambda}),$$

avec $\gamma_{ij} = \gamma_{ij}(\boldsymbol{\mu}, \boldsymbol{\Lambda})$. De plus, $\rho_{ij} \in (0, \varepsilon_{p;ij})$ tel que $\varepsilon_{p;ij} = \min\{R(i, j), R(j, i)\}$ et $\varepsilon_{p;ij} \in (0, 1)$, où R est défini par

$$R(i, j) = \begin{cases} \sqrt{\frac{\lambda_{ii}}{\lambda_{jj}}} \left(1 - \frac{1}{\lambda_{ii}} \sum_{\ell \neq i, j} \lambda_{i\ell} \right) & \text{pour } p = 1, 2; \\ \sqrt{\frac{\lambda_{ii}\mu_i^{\alpha/(\alpha-1)}}{\lambda_{jj}\mu_j^{\alpha/(\alpha-1)}}} \left(1 - \frac{1}{\lambda_{ii}\mu_i^{\alpha/(\alpha-1)}} \sum_{\ell \neq i, j} \lambda_{i\ell} \frac{\alpha}{\alpha-1} (\mu_i\mu_j)^{\frac{\alpha/2}{\alpha-1}} \right) & \text{pour } p \neq 1, 2. \end{cases}$$

En outre, $\rho(X_i^{(+)}, X_j^{(+)}) = 0$ est équivalent à ce que $X_i^{(+)}$ soit indépendant de $X_j^{(+)}$.

Cette construction théorique ne permettant pas une implémentation aisée et ne permettant pas d'introduire une corrélation négative, nous proposons une écriture matricielle.

Théorème 2. Soit $\mathbf{X}^{(+)}$ un vecteur aléatoire stable-Tweedie tel que les marginales soit positivement corrélées. Alors

$$\mathbf{X}^{(+)} = [\text{diag}_k(\boldsymbol{\mu})]^{p-1} \mathbf{U}_p^{(+)} \mathbf{1}, \quad (3)$$

avec $\mathbf{1} = (1, \dots, 1)^\top \in \mathbb{R}^k$ et où $\mathbf{U}_p^{(+)} = \left(U_{p;i,j}^{(+)} \right)_{i,j=1,\dots,k}$ est une matrice aléatoire symétrique telle que

$$U_{p;i,j}^{(+)} \sim \begin{cases} \mathcal{T}w^{*p}(1, \gamma_{ij}) & \text{pour } p = 1, 2 \\ \mathcal{T}w^{*p}\left(1, \gamma_{ij} \frac{\alpha}{\alpha-1}\right) & \text{pour } p \neq 1, 2, \end{cases} \quad (4)$$

pour $i \neq j$ et

$$U_{p;ii}^{(+)} \sim \begin{cases} \mathcal{T}w^{*p}\left(1, \gamma_{ii} - \sum_{\ell \neq i} \gamma_{i\ell}\right) & \text{pour } p = 1, 2 \\ \mathcal{T}w^{*p}\left(1, \left(\gamma_{ii} - \sum_{\ell \neq i} \gamma_{i\ell}\right) \frac{\alpha}{\alpha-1}\right) & \text{pour } p \neq 1, 2. \end{cases}$$

Les résultats des simulations sont ceux attendus pour une corrélation positive. En effet la distribution s'agglutine autour d'une ligne droite croissante en cas de forte corrélation.

Nous nous servons ensuite de l'écriture du Théorème 2 pour obtenir une corrélation négative entre les lois marginales de la manière suivante :

Corollaire 2. *On considère l'écriture du Théorème 2 dans laquelle on remplace la matrice $\mathbf{U}_p^{(+)}$ par une autre matrice aléatoire, non nécessairement symétrique, notée $\mathbf{U}_p^{(-)}$, telle que $U_{p;ii}^{(-)} = U_{p;ii}^{(+)}$ et $U_{p;i,j}^{(-)}$ soient définis par*

$$U_{p;i,j}^{(-)} = \begin{cases} F_{p;i,j}^{-1}(Z) & \text{pour } i < j \\ F_{p;i,j}^{-1}(1 - Z) & \text{pour } i > j; \end{cases} \quad (5)$$

où Z est une variable aléatoire uniforme sur $]0, 1[$ et où la fonction $F_{p;i,j}^{-1}$ est la fonction inverse généralisée (i.e. $F_{p;i,j}^{-1}(y) = \inf \{x \in \mathbb{R}, F_{p;i,j}(x) = y, y \in (0, 1)\}$) de la fonction de répartition $F_{p;i,j}$ de la loi Tweedie univariée:

$$\begin{cases} \mathcal{T}w^{*p}(1, \gamma_{ij}) & \text{pour } p = 1, 2 \\ \mathcal{T}w^{*p}\left(1, \gamma_{ij} \frac{\alpha}{\alpha-1}\right) & \text{pour } p \neq 1, 2. \end{cases}$$

Les fonctions inverses généralisées nous garantissent l'unicité de cette écriture pour $p \in \{0\} \cup [1, +\infty[$. Quant à l'existence, elle est garantie par le lemme suivant.

Lemme 3. *Soit $\varphi :]0, 1[\rightarrow \mathbb{R}$ une fonction strictement croissante et $Z \sim \mathcal{U}]0, 1[$. Alors $\text{Cov}[\varphi(Z), \varphi(1 - Z)] \leq 0$.*

L'écriture matricielle permet également une implémentation correcte. Les résultats de simulations sont très cohérents. En effet, nous obtenons des distributions qui s'agglutinent autour d'une fonction décroissante convexe en cas d'une forte corrélation négative et d'une loi à support le cadran positif. Nous supposons que ce comportement vient du fait que les supports des lois soient majoritairement positifs. Il faut noter que pour la loi normale ($p = 0$), la décroissance se fait en ligne droite.

3 Conclusions et perspectives

Une construction théorique a été présentée pour les modèles stable-Tweedie multivariés tels que les marginales soient positivement corrélées. Cette construction étant pour le moment inconnue pour le cas négativement corrélé, les investigations doivent se poursuivre dans ce sens. Cependant les résultats de simulations sont acceptables mais quelques points méritent également d’approfondir les recherches, en particulier en ce qui concerne les comportements des distributions. Nous pouvons donc proposer l’écriture

$$\mathbf{X} = \mathbf{X}^{(+)} \oplus \mathbf{X}^{(-)} \sim Tw_k^{*p}(\boldsymbol{\mu}, \boldsymbol{\Lambda}),$$

où \oplus représente le *OU exclusif*. Après avoir réalisé d’autres tests de l’algorithme sous-jacent, il sera intéressant de pouvoir le confronter à des données réelles. Une autre perspective intéressante pourra être de s’appuyer sur notre méthode pour envisager la construction des modèles Tweedie discrets multivariés, aussi appelés *modèles Poisson-Tweedie multivariés*. Enfin, l’étude des modèles *stable-Tweedie multiples*, dont un récent cas particulier a été étudié par Boubacar Maïnassara et Kokonendji (2014) pourra s’inspirer de notre méthode.

Bibliographie

- [1] Alai, D. H., Landsman, Z. et Sherris, M. (2013), Multivariate Tweedie lifetime: the impact of dependence, *ARC CEPAR, working paper 2013/11*.
- [2] Boubacar Maïnassara, Y., Kokonendji, C. C. (2014) On normal stable Tweedie models and power-generalized variance functions of only one component, *TEST*, in press.
- [3] Cuenin, J., Jørgensen, B. et Kokonendji, C. C. (2014), Tweedie random vectors: construction and simulation, *Preprint LMB numéro 2014/03*.
- [4] Furman, E. et Landsman, Z. (2010), Multivariate Tweedie distributions and some related capital-at-risk analyses, *Insurance: Mathematics and Economics*, **46**, 351–361.
- [5] Jørgensen, B. (1997), *The Theory of Dispersion Models*, Chapman & Hall, London.
- [6] Jørgensen, B. et Martínez, J. R. (2013), Multivariate exponential dispersion models, *Multivariate Statistics: Theory and Applications. Proceedings of the IX Tartu Conference on Multivariate Statistics & XX International Workshop on Matrices and Statistics (Ed. Tõnu Kollo)*, 73–98, World Scientific, Singapore.
- [7] Kendal, W. S. (2014), Multifractability attributed to dual central limit-like convergence effects, *Physica A*, **401**, 22–23.
- [8] Tweedie, M. C. K. (1984), An index which distinguishes between some important exponential families, In: Ghosh, J.K., Roy, J. (Eds.), *Statistics: Applications and New Directions. Proceedings of Indian Statistical Golden Jubilee International Conference*, 579–604.