WORKING PAPER 6/2024 (ECONOMICS)

The Method of Moments for Multivariate Random Sums

Farrukh Javed, Nicola Loperfido and Stepan Mazur

ISSN 1403-0586

Örebro University School of Business
SE-701 82 Örebro, Sweden

The Method of Moments for Multivariate Random Sums

Farrukh Javed
Department of Statistics, Lund University
Tycho Brahes väg 1, 223 63 Lund, Sweden
e-mail: farrukh.javed@stat.lu.se

Nicola Loperfido
Dipartimento di Economia, Società e Politica,
Università degli Studi di Urbino "Carlo Bo", 61033 ITALY
e-mail: nicola.loperfido@uniurb.it

Stepan Mazur
Örebro University School of Business
Fakultetsgatan 1, SE-70182 Örebro, Sweden
e-mail: stepan.mazur@oru.se

Abstract

Multivariate random sums appear in many scientific fields, most notably in actuarial science, where they model both the number of claims and their sizes. Unfortunately, they pose severe inferential problems. For example, their density function is analytically intractable, in the general case, thus preventing likelihood inference. In this paper, we address the problem by the method of moments, under the assumption that the claim size and the claim number have a multivariate skew-normal and a Poisson distribution, respectively. In doing so, we also derive closed-form expressions for some fundamental measures of multivariate kurtosis and highlight some limitations of both projection pursuit and invariant coordinate selection.

Keywords. Fourth cumulant, Kurtosis, Poisson distribution, Skew-normal distribution.

1 Introduction

Let $\mathbf{s} = \mathbf{x}_1 + \ldots + \mathbf{x}_N$ be the sum of the first N components appearing in the sequence of independent and identically distributed p-dimensional random vectors $\{\mathbf{x}_i,\, i \in \mathbb{N}\}$, where N is a random variable independent of the sequence and whose support is the set of nonnegative integers. The random vector $\mathbf{s}$ is a random sum, with the convention that $\mathbf{s}$ coincides with the p-dimensional null vector when $N = 0$. The distributions of random sums are known as compound distributions and occur in many research areas (Gnedenko and Korolev, 1996; Kalashnikov, 1997; Klebanov *et al*, 2006), but have been mainly studied in actuarial science, where the random vector $\mathbf{s}$, $\mathbf{x}_i$ and N are the aggregated claim, the i-th claim size and the claim counts, respectively (Ambagaspitiya, 1999).

Compound distributions pose severe inferential problems: analytic expressions for their densities are not available, in the general case, thus preventing the use of likelihood methods. The problem has often been addressed by means of approximations (Lin, 2006), either asymptotic or parametric. Asymptotic methods are often used to approximate the compound distributions, when the aggregating variable is expected to be high enough (Klebanov *et al*, 2006). However, the asymptotic approach based on stochastically increasing sequences does not necessarily lead to valid asymptotic normal approximations (Javed *et al*, 2021). Alternatively, compound distributions can be approximated by properly chosen parametric distributions, as for example skew-normal distributions and their generalizations (Eling, 2012) or finite mixture distributions (Bernardi *et al*, 2012). In both cases, the parameters lack a clear interpretation.

In this paper, we address the above mentioned problems by assuming well-known parametric families for both the claim size and the claim number. The latter is assumed to be Poisson with mean $\lambda > 0$, so that $\mathrm{P}\left(N = n\right) = e^{-\lambda}\lambda^{n}/n!$, $n \in \mathbb{N}$. The i-th claim size $\mathbf{x}_i$ is assumed to be multivariate skew-normal with location parameter $\boldsymbol{\xi} \in \mathbb{R}^p$, scale parameter $\boldsymbol{\Omega} \in \mathbb{R}^p$, shape parameter $\boldsymbol{\alpha} \in \mathbb{R}^p$, denoted by $\mathbf{x}_i \sim SN\left(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}\right)$. Its probability density function is $f\left(\mathbf{x}; \boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}\right) = 2\phi_p\left(\mathbf{x}; \boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}\right) \cdot \Phi\left\{\boldsymbol{\alpha}^{\top}\left(\mathbf{x} - \boldsymbol{\xi}\right)\right\}$,where $\phi_p\left(\cdot; \boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}\right)$ is the probability density function of a p-dimensional normal distribution with mean $\boldsymbol{\xi}$ and covariance $\boldsymbol{\Omega}$, while $\Phi\left(\cdot\right)$ denotes the cumulative distribution function of a univariate, standard normal distribution. The compound distribution is known as the Poisson-Skew-Normal distribution and has been motivated and investigated by Loperfido *et al* (2018). We shall estimate the parameters λ, $\boldsymbol{\xi}$, $\boldsymbol{\Omega}$ and $\boldsymbol{\alpha}$ by the method of moments. The rest of the paper is structured as follows. Sections 2 and 3 derive the fourth cumulant and the Mardia's kurtosis of the Poisson-Skew-Normal model, respectively. Section 4 uses previous section's results and those in Loperfido *et al* (2018) to obtain method-of-moments estimates. Section 5 highlights some limitations of projection pursuit and invariant coordinate selection, when applied to the Poisson-Skew-Normal model. Section 6 contains some final remarks and hints for future research.

2 Fourth cumulant

The $\left(i, j, h, k\right)$ $-$th cumulant of a p-dimensional random vector $\mathbf{x} = \left(x_1, \ldots, x_p\right)^{\top}$ with mean $\boldsymbol{\mu} = \left(\mu_1, \ldots, \mu_p\right)^{\top}$, covariance $\boldsymbol{\Sigma} = \left\{\sigma_{ij}\right\}$ and finite fourth-order moments is

$$\kappa_{ijhk}=\frac{\log \mathrm{E}\left\{\exp\left(\iota\mathbf{t}^{\top}\mathbf{x}\right)\right\}}{\partial t_i\partial t_j\partial t_h\partial t_k}, \text{ where } \iota=\sqrt{-1} \text{ and } \mathbf{t}=\begin{pmatrix} t_1 \\ ... \\ t_p \end{pmatrix}\in\mathbb{R}^p,\ i,j,h,k\in\{1,...,p\}\,.$$

The scalars κ_{ijhk} might be conveniently arranged into the $p^2\times p^2$ block matrix

$$\mathbf{K}_{4,\mathbf{x}}=\mathrm{E}\left\{\mathbf{y}\otimes\mathbf{y}^{\top}\otimes\mathbf{y}\otimes\mathbf{y}^{\top}\right\}-\left(\mathbf{I}_{p^2}+\mathbf{C}_{p,p}\right)(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma})-vec\,(\boldsymbol{\Sigma})\,vec^{\top}(\boldsymbol{\Sigma})\,,$$

where $\mathbf{y}=\mathbf{x}-\boldsymbol{\mu}$, "$\otimes$" is the Kronecker product, $vec\,(\mathbf{A})$ is the vectorization of the matrix $\mathbf{A}$ and $\mathbf{C}_{p,p}$ is the $p^2\times p^2$ commutation matrix. The matrix $\mathbf{K}_{4,\mathbf{x}}$ is commonly referred to as to the fourth cumulant of $\mathbf{x}$ and possesss many and interesting properties (Loperfido, 2015, 2017). In order to remove linear dependencies, the fourth cumulant of the standardized random vector $\mathbf{z}=\boldsymbol{\Sigma}^{-1/2}(\mathbf{x}-\boldsymbol{\mu})$, where $\boldsymbol{\Sigma}^{-1/2}$ is the unique positive definite matrix such that $\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}^{-1/2}=\boldsymbol{\Sigma}^{-1}$. The fourth cumulant of $\mathbf{z}$ is known as the fourth standardized cumulant of $\mathbf{x}$ and is denoted by $\mathbf{K}_{4,\mathbf{z}}$.

The following theorem shows that the fourth cumulant of a multivariate Poisson-Skew-Normal aggregate claim model as a remarkably simple analytical form.

Theorem 1 *Let N be a Poisson random variable with mean λ. Also, let $\mathbf{s}=\mathbf{x}_1+...+\mathbf{x}_N$, where $\mathbf{x}_1$, ..., $\mathbf{x}_N$ are p-dimensional skew-normal random vectors with location parameter $\boldsymbol{\varsigma}$, scale parameter $\boldsymbol{\Omega}$ and shape parameter $\boldsymbol{\alpha}$. If N, $\mathbf{x}_1$, ..., $\mathbf{x}_N$ are mutually independent, the fourth cumulant of $\mathbf{s}$ is*

$$\mathbf{K}_{4,\mathbf{s}}=\lambda\boldsymbol{\Omega}\otimes\boldsymbol{\Omega}+\lambda\mathbf{C}_{p,p}(\boldsymbol{\Omega}\otimes\boldsymbol{\Omega})+\lambda vec\,(\boldsymbol{\Omega})\,vec^{\top}(\boldsymbol{\Omega})\,.$$

Proof. Let the first, second, third and fourth cumulants of $\mathbf{x}_i$ be $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$, $\boldsymbol{\xi}_3$ and $\boldsymbol{\xi}_4$. Similarly, let the first, second, third and fourth cumulants of N be ν_1, ν_2, ν_3 and ν_4. The fourth cumulant of $\mathbf{s}$, denoted by $\mathbf{K}_{4,\mathbf{s}}$, is

$$\nu_1\boldsymbol{\xi}_4+\nu_2\left\{\left(\mathbf{I}_{p^2}+\mathbf{C}_{p,p}\right)(\boldsymbol{\xi}_2\otimes\boldsymbol{\xi}_2)+vec\,(\boldsymbol{\xi}_2)\,vec^{\top}(\boldsymbol{\xi}_2)\right\}+\nu_2\left(\boldsymbol{\xi}_3\otimes\boldsymbol{\xi}_1^{\top}+\boldsymbol{\xi}_3^{\top}\otimes\boldsymbol{\xi}_1\right.$$

$$\left.+\boldsymbol{\xi}_1^{\top}\otimes\boldsymbol{\xi}_3+\boldsymbol{\xi}_1\otimes\boldsymbol{\xi}_3^{\top}\right)+\nu_3\left\{\boldsymbol{\xi}_2\otimes\boldsymbol{\xi}_1\boldsymbol{\xi}_1^{\top}+vec\,(\boldsymbol{\xi}_2)\,\boldsymbol{\xi}_1^{\top}\otimes\boldsymbol{\xi}_1^{\top}+\boldsymbol{\xi}_1^{\top}\otimes\boldsymbol{\xi}_2\otimes\boldsymbol{\xi}_1\right.$$

$$\left.+\boldsymbol{\xi}_1\otimes\boldsymbol{\xi}_2\otimes\boldsymbol{\xi}_1^{\top}+\boldsymbol{\xi}_1\otimes vec^{\top}(\boldsymbol{\xi}_2)\otimes\boldsymbol{\xi}_1+\boldsymbol{\xi}_1\otimes\boldsymbol{\xi}_1^{\top}\otimes\boldsymbol{\xi}_2\right\}+\nu_4\boldsymbol{\xi}_1\boldsymbol{\xi}_1^{\top}\otimes\boldsymbol{\xi}_1\boldsymbol{\xi}_1^{\top}$$

(Javed *et al*, 2021). By assumption, N is a Poisson random variable with mean

λ, so that $\nu_1 = \nu_2 = \nu_3 = \nu_4 = \lambda$ and

$$\begin{aligned}\frac{\mathbf{K}_{4,\mathbf{s}}}{\lambda} &= \boldsymbol{\xi}_4 + \left(\mathbf{I}_{p^2} + \mathbf{C}_{p,p}\right)\left(\boldsymbol{\xi}_2 \otimes \boldsymbol{\xi}_2\right) + vec\left(\boldsymbol{\xi}_2\right) vec^{\top}\left(\boldsymbol{\xi}_2\right) + \boldsymbol{\xi}_3 \otimes \boldsymbol{\xi}_1^{\top} + \boldsymbol{\xi}_3^{\top} \otimes \boldsymbol{\xi}_1 \\ &\quad + \boldsymbol{\xi}_1^{\top} \otimes \boldsymbol{\xi}_3 + \boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_3^{\top} + \boldsymbol{\xi}_2 \otimes \boldsymbol{\xi}_1 \boldsymbol{\xi}_1^{\top} + vec\left(\boldsymbol{\xi}_2\right) \boldsymbol{\xi}_1^{\top} \otimes \boldsymbol{\xi}_1^{\top} + \boldsymbol{\xi}_1^{\top} \otimes \boldsymbol{\xi}_2 \otimes \boldsymbol{\xi}_1 \\ &\quad + \boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_2 \otimes \boldsymbol{\xi}_1^{\top} + \boldsymbol{\xi}_1 \otimes vec^{\top}\left(\boldsymbol{\xi}_2\right) \otimes \boldsymbol{\xi}_1 + \boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_1^{\top} \otimes \boldsymbol{\xi}_2 + \boldsymbol{\xi}_1 \boldsymbol{\xi}_1^{\top} \otimes \boldsymbol{\xi}_1 \boldsymbol{\xi}_1^{\top}.\end{aligned}$$

By assumption, the distribution of $\mathbf{x}_i$ is $SN\left(\boldsymbol{\varsigma}, \boldsymbol{\Omega}, \boldsymbol{\alpha}\right)$. The fourth cumulant is location invariant, so that we can assume without loss of generality that the location parameter of the skew-normal distribution is a null vector and its first four cumulants are

$$\boldsymbol{\xi}_1 = \sqrt{\frac{2}{\pi}}\boldsymbol{\delta},\ \boldsymbol{\xi}_2 = \boldsymbol{\Omega} - \frac{2}{\pi}\boldsymbol{\delta}\boldsymbol{\delta}^{\top},\ \boldsymbol{\xi}_3 = \sqrt{\frac{2}{\pi}}\left(\frac{4}{\pi} - 1\right)\boldsymbol{\delta}\boldsymbol{\delta}^{\top} \otimes \boldsymbol{\delta}\ ,$$

$$\boldsymbol{\xi}_4 = 2\left(\pi - 3\right)\left(\frac{2}{\pi}\right)^2 \boldsymbol{\delta}\boldsymbol{\delta}^{\top} \otimes \boldsymbol{\delta}\boldsymbol{\delta}^{\top},\ \text{where } \boldsymbol{\delta} = \frac{\boldsymbol{\Omega}\boldsymbol{\alpha}}{\sqrt{1 + \boldsymbol{\alpha}^{\top}\boldsymbol{\Omega}\boldsymbol{\alpha}}}.$$

By expressing $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$, $\boldsymbol{\xi}_3$ and $\boldsymbol{\xi}_4$ by means of $\boldsymbol{\delta}$ and $\boldsymbol{\Omega}$ we obtain

$$\frac{\mathbf{K}_{4,\mathbf{s}}}{\lambda} = \left\{2\left(\pi - 3\right)\left(\frac{2}{\pi}\right)^2 + \frac{8}{\pi}\left(\frac{4}{\pi} - 1\right) + \left(\frac{2}{\pi}\right)^2\right\}\boldsymbol{\delta} \otimes \boldsymbol{\delta}^{\top} \otimes \boldsymbol{\delta} \otimes \boldsymbol{\delta}^{\top} +$$

$$\left(\mathbf{I}_{p^2} + \mathbf{C}_{p,p}\right)\left\{\left(\boldsymbol{\Omega} - \frac{2}{\pi}\boldsymbol{\delta}\boldsymbol{\delta}^{\top}\right) \otimes \left(\boldsymbol{\Omega} - \frac{2}{\pi}\boldsymbol{\delta}\boldsymbol{\delta}^{\top}\right)\right\} + vec\left(\boldsymbol{\Omega} - \frac{2}{\pi}\boldsymbol{\delta}\boldsymbol{\delta}^{\top}\right) vec^{\top}\left(\boldsymbol{\Omega} - \frac{2}{\pi}\boldsymbol{\delta}\boldsymbol{\delta}^{\top}\right) +$$

$$\frac{2}{\pi}\left(\boldsymbol{\Omega} - \frac{2}{\pi}\boldsymbol{\delta}\boldsymbol{\delta}^{\top}\right) \otimes \boldsymbol{\delta}\boldsymbol{\delta}^{\top} + \frac{2}{\pi} vec\left(\boldsymbol{\Omega} - \frac{2}{\pi}\boldsymbol{\delta}\boldsymbol{\delta}^{\top}\right)\boldsymbol{\delta}^{\top} \otimes \boldsymbol{\delta}^{\top} + \frac{2}{\pi}\boldsymbol{\delta}^{\top} \otimes \left(\boldsymbol{\Omega} - \frac{2}{\pi}\boldsymbol{\delta}\boldsymbol{\delta}^{\top}\right) \otimes \boldsymbol{\delta} +$$

$$\frac{2}{\pi}\boldsymbol{\delta} \otimes \left(\boldsymbol{\Omega} - \frac{2}{\pi}\boldsymbol{\delta}\boldsymbol{\delta}^{\top}\right) \otimes \boldsymbol{\delta}^{\top} + \frac{2}{\pi}\boldsymbol{\delta} \otimes vec^{\top}\left(\boldsymbol{\Omega} - \frac{2}{\pi}\boldsymbol{\delta}\boldsymbol{\delta}^{\top}\right) \otimes \boldsymbol{\delta} + \frac{2}{\pi}\boldsymbol{\delta} \otimes \boldsymbol{\delta}^{\top} \otimes \left(\boldsymbol{\Omega} - \frac{2}{\pi}\boldsymbol{\delta}\boldsymbol{\delta}^{\top}\right).$$

Some algebra leads to

$$\frac{\mathbf{K}_{4,\mathbf{s}}}{\lambda}=\frac{12}{\pi^2}\boldsymbol{\delta}\otimes\boldsymbol{\delta}^\top\otimes\boldsymbol{\delta}\otimes\boldsymbol{\delta}^\top+\boldsymbol{\Omega}\otimes\boldsymbol{\Omega}-\frac{2}{\pi}\boldsymbol{\Omega}\otimes\boldsymbol{\delta}\otimes\boldsymbol{\delta}^\top-\frac{2}{\pi}\boldsymbol{\delta}\otimes\boldsymbol{\delta}^\top\otimes\boldsymbol{\Omega}+\frac{4}{\pi^2}\boldsymbol{\delta}\otimes\boldsymbol{\delta}^\top\otimes\boldsymbol{\delta}\otimes\boldsymbol{\delta}^\top+$$

$$\mathbf{C}_{p,p}\left(\boldsymbol{\Omega}\otimes\boldsymbol{\Omega}-\frac{2}{\pi}\boldsymbol{\Omega}\otimes\boldsymbol{\delta}\boldsymbol{\delta}^\top-\frac{2}{\pi}\boldsymbol{\delta}\boldsymbol{\delta}^\top\otimes\boldsymbol{\Omega}+\frac{4}{\pi^2}\boldsymbol{\delta}\boldsymbol{\delta}^\top\otimes\boldsymbol{\delta}\boldsymbol{\delta}^\top\right)+vec\left(\boldsymbol{\Omega}\right)vec^\top\left(\boldsymbol{\Omega}\right)-\frac{2}{\pi}vec\left(\boldsymbol{\Omega}\right)\left(\boldsymbol{\delta}^\top\otimes\boldsymbol{\delta}^\top\right)-$$

$$\frac{2}{\pi}\left(\boldsymbol{\delta}\otimes\boldsymbol{\delta}\right)vec^\top\left(\boldsymbol{\Omega}\right)+\left(\frac{2}{\pi}\right)^2\boldsymbol{\delta}\otimes\boldsymbol{\delta}^\top\otimes\boldsymbol{\delta}\otimes\boldsymbol{\delta}^\top+\frac{2}{\pi}\boldsymbol{\Omega}\otimes\boldsymbol{\delta}\boldsymbol{\delta}^\top-\frac{4}{\pi^2}\boldsymbol{\delta}\boldsymbol{\delta}^\top\otimes\boldsymbol{\delta}\boldsymbol{\delta}^\top+\frac{2}{\pi}vec\left(\boldsymbol{\Omega}\right)\left(\boldsymbol{\delta}^\top\otimes\boldsymbol{\delta}^\top\right)-$$

$$\frac{4}{\pi^2}\boldsymbol{\delta}\otimes\boldsymbol{\delta}^\top\otimes\boldsymbol{\delta}\otimes\boldsymbol{\delta}^\top+\frac{2}{\pi}\boldsymbol{\delta}^\top\otimes\boldsymbol{\Omega}\otimes\boldsymbol{\delta}-\frac{4}{\pi^2}\boldsymbol{\delta}^\top\otimes\boldsymbol{\delta}\boldsymbol{\delta}^\top\otimes\boldsymbol{\delta}+\frac{2}{\pi}\boldsymbol{\delta}\otimes\boldsymbol{\Omega}\otimes\boldsymbol{\delta}^\top-\frac{4}{\pi^2}\boldsymbol{\delta}\otimes\boldsymbol{\delta}\otimes\boldsymbol{\delta}^\top\otimes\boldsymbol{\delta}^\top$$

$$+\frac{2}{\pi}\boldsymbol{\delta}\otimes vec^\top\left(\boldsymbol{\Omega}\right)\otimes\boldsymbol{\delta}-\frac{4}{\pi^2}\boldsymbol{\delta}\otimes vec^\top\left(\boldsymbol{\delta}\otimes\boldsymbol{\delta}^\top\right)\otimes\boldsymbol{\delta}+\frac{2}{\pi}\boldsymbol{\delta}\otimes\boldsymbol{\delta}^\top\otimes\boldsymbol{\Omega}-\frac{4}{\pi^2}\boldsymbol{\delta}\otimes\boldsymbol{\delta}^\top\otimes\boldsymbol{\delta}\otimes\boldsymbol{\delta}^\top.$$

The above identity might be simplified by removing redundant terms:

$$\frac{\mathbf{K}_{4,\mathbf{s}}}{\lambda}=\boldsymbol{\Omega}\otimes\boldsymbol{\Omega}+\mathbf{C}_{p,p}\left(\boldsymbol{\Omega}\otimes\boldsymbol{\Omega}-\frac{2}{\pi}\boldsymbol{\Omega}\otimes\boldsymbol{\delta}\boldsymbol{\delta}^\top-\frac{2}{\pi}\boldsymbol{\delta}\boldsymbol{\delta}^\top\otimes\boldsymbol{\Omega}\right)+vec\left(\boldsymbol{\Omega}\right)vec^\top\left(\boldsymbol{\Omega}\right)$$

$$-\frac{2}{\pi}\left(\boldsymbol{\delta}\otimes\boldsymbol{\delta}\right)vec^\top\left(\boldsymbol{\Omega}\right)+\frac{2}{\pi}\boldsymbol{\delta}^\top\otimes\boldsymbol{\Omega}\otimes\boldsymbol{\delta}+\frac{2}{\pi}\boldsymbol{\delta}\otimes\boldsymbol{\Omega}\otimes\boldsymbol{\delta}^\top+\frac{2}{\pi}\boldsymbol{\delta}\otimes vec^\top\left(\boldsymbol{\Omega}\right)\otimes\boldsymbol{\delta}.$$

Basic properties of the commutation matrix (Kollo and von Rosen, 2005, p. 80) and the Kronecker product yield the identities

$$\mathbf{C}_{p,p}\left(\boldsymbol{\Omega}\otimes\boldsymbol{\delta}\boldsymbol{\delta}^\top\right)=\boldsymbol{\delta}\otimes\boldsymbol{\Omega}\otimes\boldsymbol{\delta}^\top,\ \mathbf{C}_{p,p}\left(\boldsymbol{\delta}\boldsymbol{\delta}^\top\otimes\boldsymbol{\Omega}\right)=\boldsymbol{\delta}^\top\otimes\boldsymbol{\Omega}\otimes\boldsymbol{\delta}\ \text{ and }\ \boldsymbol{\delta}\otimes vec^\top\left(\boldsymbol{\Omega}\right)\otimes\boldsymbol{\delta}=\boldsymbol{\delta}\otimes\boldsymbol{\delta}\otimes vec^\top\left(\boldsymbol{\Omega}\right).$$

In turn, the above identities lead to the following one:

$$\frac{\mathbf{K}_{4,\mathbf{s}}}{\lambda}=\boldsymbol{\Omega}\otimes\boldsymbol{\Omega}+\mathbf{C}_{p,p}\left(\boldsymbol{\Omega}\otimes\boldsymbol{\Omega}\right)+vec\left(\boldsymbol{\Omega}\right)vec^\top\left(\boldsymbol{\Omega}\right).$$

3 Mardia's kurtosis

Mardia (1970, 1974) introduced the following measures of multivariate kurtosis:

$$\beta_{2,M}\left(\mathbf{x}\right)=\mathrm{E}\left\{\left[\left(\mathbf{x}-\boldsymbol{\mu}\right)^\top\boldsymbol{\Sigma}^{-1}\left(\mathbf{x}-\boldsymbol{\mu}\right)\right]^2\right\},\ \gamma_{2,M}\left(\mathbf{x}\right)=\beta_{2,M}\left(\mathbf{x}\right)-p\left(p+2\right)=\sum_{i,j}\kappa_{iijj}.$$

The former and the latter are known as Mardia's kurtosis and Mardia's excess kurtosis. They are the best known and most used measure of multivariate kurtosis (Kollo, 2008). The following theorem shows that they have simple analytical forms of the Poisson-Skew-Normal model have remarkably simple analytical forms.

Theorem 2 *Let N be a Poisson random variable with mean λ. Also, let $\mathbf{s} = \mathbf{x}_1 + ... + \mathbf{x}_N$, where $\mathbf{x}_1$, ..., $\mathbf{x}_N$ are p-dimensional skew-normal random vectors. If N, $\mathbf{x}_1$, ..., $\mathbf{x}_N$ are mutually independent, the Mardia's kurtosis of $\mathbf{s}$ is*

$$\beta_{2M}(\mathbf{s}) = \frac{(\lambda+1)(p+2)p}{\lambda}.$$

Proof. We use the same notation as in Theorem 1 and its proof. The fourth cumulant of $\mathbf{s}$ is

$$\mathbf{K}_{4,\mathbf{s}} = \lambda \boldsymbol{\Omega} \otimes \boldsymbol{\Omega} + \lambda \mathbf{C}_{p,p} (\boldsymbol{\Omega} \otimes \boldsymbol{\Omega}) + \lambda vec(\boldsymbol{\Omega}) vec^{\top}(\boldsymbol{\Omega}).$$

Let $\mathbf{z}$ be the standardization of $\mathbf{s}$:

$$\mathbf{z} = \boldsymbol{\Sigma}^{-1/2}(\mathbf{s} - \boldsymbol{\mu}), \text{ where } \mathrm{E}(\mathbf{s}) = \boldsymbol{\mu} \text{ and } cov(\mathbf{s}) = \boldsymbol{\Sigma}.$$

The covariance matrix of $\mathbf{s}$ is $\boldsymbol{\Sigma} = \lambda \boldsymbol{\Omega}$ (Loperfido *et al*, 2018), so that $\mathbf{z} = \boldsymbol{\Omega}^{-1/2}(\mathbf{s} - \boldsymbol{\mu}) / \lambda^{-1/2}$. By ordinary properties of fourth cumulants the fourth cumulant of $\mathbf{z}$, that is the fourth standardized cumulant of $\mathbf{s}$, is

$$\mathbf{K}_{4,\mathbf{z}} = \frac{1}{\lambda^2}\left(\boldsymbol{\Omega}^{-1/2} \otimes \boldsymbol{\Omega}^{-1/2}\right) \mathbf{K}_{4,\mathbf{s}} \left(\boldsymbol{\Omega}^{-1/2} \otimes \boldsymbol{\Omega}^{-1/2}\right) = \frac{1}{\lambda}\left(\boldsymbol{\Omega}^{-1/2} \otimes \boldsymbol{\Omega}^{-1/2}\right) \boldsymbol{\Omega} \otimes \boldsymbol{\Omega} \left(\boldsymbol{\Omega}^{-1/2} \otimes \boldsymbol{\Omega}^{-1/2}\right) +$$

$$\frac{1}{\lambda}\left(\boldsymbol{\Omega}^{-1/2} \otimes \boldsymbol{\Omega}^{-1/2}\right) \mathbf{C}_{p,p} (\boldsymbol{\Omega} \otimes \boldsymbol{\Omega}) \left(\boldsymbol{\Omega}^{-1/2} \otimes \boldsymbol{\Omega}^{-1/2}\right) + \frac{1}{\lambda}\left(\boldsymbol{\Omega}^{-1/2} \otimes \boldsymbol{\Omega}^{-1/2}\right) vec(\boldsymbol{\Omega}) vec^{\top}(\boldsymbol{\Omega}) \left(\boldsymbol{\Omega}^{-1/2} \otimes \boldsymbol{\Omega}^{-1/2}\right)$$

Recall the distributive property of the Kronecker product with respect to the ordinary matrix product (see, e.g., Kollo and von Rosen, 2005, page 81):

$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}$, where $\mathbf{A} \in \mathbb{R}^p \times \mathbb{R}^q$, $\mathbf{B} \in \mathbb{R}^r \times \mathbb{R}^s$, $\mathbf{C} \in \mathbb{R}^q \times \mathbb{R}^u$, $\mathbf{D} \in \mathbb{R}^s \times \mathbb{R}^v$.

We apply this property to obtain a simplified expression for the fourth cumulant of $\mathbf{z}$:

$$\mathbf{K}_{4,\mathbf{z}} = \frac{1}{\lambda}\mathbf{I}_{p^2} + \frac{1}{\lambda}\left(\boldsymbol{\Omega}^{-1/2} \otimes \boldsymbol{\Omega}^{-1/2}\right) \mathbf{C}_{p,p} \left(\boldsymbol{\Omega}^{1/2} \otimes \boldsymbol{\Omega}^{1/2}\right) + \frac{1}{\lambda}\left(\boldsymbol{\Omega}^{-1/2} \otimes \boldsymbol{\Omega}^{-1/2}\right) vec(\boldsymbol{\Omega}) vec^{\top}(\boldsymbol{\Omega}) \left(\boldsymbol{\Omega}^{-1/2} \otimes \boldsymbol{\Omega}^{-1/2}\right)$$

The following identity connects the vectorization operator, the Kronecker product and the ordinary matrix product (see, e.g., Kollo and von Rosen, 2005, page 89):

$vec(\mathbf{ABC}) = \left(\mathbf{C}^{\top} \otimes \mathbf{A}\right) vec(\mathbf{B})$, where $\mathbf{A} \in \mathbb{R}^p \times \mathbb{R}^q$, $\mathbf{B} \in \mathbb{R}^q \times \mathbb{R}^r$, $\mathbf{C} \in \mathbb{R}^r \times \mathbb{R}^s$.

We apply this identity to achieve further simplification of the analytical expression of the fourth cumulant of $\mathbf{z}$:

$$\mathbf{K}_{4,\mathbf{z}} = \frac{1}{\lambda}\mathbf{I}_{p^2} + \frac{1}{\lambda}\left(\boldsymbol{\Omega}^{-1/2} \otimes \boldsymbol{\Omega}^{-1/2}\right) \mathbf{C}_{p,p} \left(\boldsymbol{\Omega}^{1/2} \otimes \boldsymbol{\Omega}^{1/2}\right) + \frac{1}{\lambda} vec(\mathbf{I}_p) vec^{\top}(\mathbf{I}_p).$$

The commutation matrix $\mathbf{C}_{p,p}$ coincides with its inverse, so that $\mathbf{C}_{p,p}\mathbf{C}_{p,p} = \mathbf{I}_{p^2}$. Moreover, the commutation matrix appears when exchanging the order of matrices in a Kronecker product:

$$\mathbf{A} \otimes \mathbf{B} = \mathbf{C}_{p,r}\left(\mathbf{B} \otimes \mathbf{A}\right)\mathbf{C}_{s,q}, \text{ where } \mathbf{A} \in \mathbb{R}^p \times \mathbb{R}^q, \mathbf{B} \in \mathbb{R}^r \times \mathbb{R}^s.$$

These properties of the commutation matrix imply the identity

$$\left(\mathbf{\Omega}^{-1/2} \otimes \mathbf{\Omega}^{-1/2}\right)\mathbf{C}_{p,p} = \mathbf{C}_{p,p}\left(\mathbf{\Omega}^{-1/2} \otimes \mathbf{\Omega}^{-1/2}\right).$$

This identity, together with further application of the above mentioned distributive property, leads to the following representation of the fourth standardized cumulant of $\mathbf{s}$:

$$\mathbf{K}_{4,\mathbf{z}} = \frac{1}{\lambda}\left\{\mathbf{I}_{p^2} + \mathbf{C}_{p,p} + vec\left(\mathbf{I}_p\right)vec^{\top}\left(\mathbf{I}_p\right)\right\}.$$

The identity $tr\left(\mathbf{C}_{p,p}\right) = p$ and ordinary properties of the trace imply that

$$tr\left(\mathbf{K}_{4,\mathbf{z}}\right) = \frac{p^2 + 2p}{\lambda}.$$

The Mardia's kurtosis of $\mathbf{s}$ (Mardia, 1970) is the fourth moment of the Euclidean norm of $\mathbf{z}$:

$$\beta_{2M}\left(\mathbf{s}\right) = \mathrm{E}\left(\left\|\mathbf{z}\right\|^4\right) = \mathrm{E}\left[\left\{\left(\mathbf{s} - \boldsymbol{\mu}\right)^{\top}\mathbf{\Sigma}^{-1}\left(\mathbf{s} - \boldsymbol{\mu}\right)\right\}^2\right].$$

Mardia's excess kurtosis (Mardia, 1974) is just the trace of its fourth standardized cumulant (see, e.g., Loperfido, 2020):

$$\gamma_{2M}\left(\mathbf{s}\right) = \beta_{2M}\left(\mathbf{s}\right) - p\left(p+2\right) = tr\left(\mathbf{K}_{4,\mathbf{z}}\right).$$

With a little algebra, the Mardia's kurtosis of $\mathbf{s}$ is then found to be

$$\beta_{2M}\left(\mathbf{s}\right) = \frac{\left(\lambda+1\right)\left(p+2\right)p}{\lambda}.$$

4 Method of moments

In this section we derive method-of-moments estimates $\widehat{\boldsymbol{\xi}}$, $\widehat{\mathbf{\Omega}}$, $\widehat{\boldsymbol{\alpha}}$, $\widehat{\lambda}$ of the parameters $\boldsymbol{\xi}$, $\mathbf{\Omega}$, $\boldsymbol{\alpha}$, λ indexing the distribution of the random sum $\mathbf{s} = \mathbf{x}_1 + ... + \mathbf{x}_N$, where $\mathbf{x}_i \sim SN\left(\boldsymbol{\xi}, \mathbf{\Omega}, \boldsymbol{\alpha}\right)$ and $N \sim Po\left(\lambda\right)$. The mean and the variance of $\mathbf{s}$ are

$$\boldsymbol{\mu} = \lambda\left(\boldsymbol{\xi} + \sqrt{\frac{2}{\pi}}\boldsymbol{\delta}\right), \text{ where } \boldsymbol{\delta} = \frac{\mathbf{\Omega}\boldsymbol{\alpha}}{\sqrt{1 + \boldsymbol{\alpha}^{\top}\mathbf{\Omega}\boldsymbol{\alpha}}}, \text{ and } \mathbf{\Sigma} = \lambda\mathbf{\Omega}.$$

The partial and directional skewnesses of $\mathbf{s}$ are

$$\beta_{1,p}^{P}\left(\mathbf{s}\right) = \frac{2\varsigma}{\pi\lambda}\left(p+2+\varsigma\right)^2, \beta_{1,p}^{D}\left(\mathbf{s}\right) = \frac{2\varsigma}{\pi\lambda}\left(3-\varsigma\right)^2, \text{ where } \varsigma = \frac{\boldsymbol{\alpha}^{\top}\mathbf{\Omega}\boldsymbol{\alpha}}{1 + \boldsymbol{\alpha}^{\top}\mathbf{\Omega}\boldsymbol{\alpha}}.$$

The unit norm vector $\boldsymbol{\nu}$ which maximizes the skewness of a projection of $\mathbf{s}$ onto its direction is proportional to the shape parameter $\widehat{\boldsymbol{\alpha}}$:

$$\boldsymbol{\nu} = \frac{\boldsymbol{\alpha}}{\|\boldsymbol{\alpha}\|} = \underset{\mathbf{u}\in\mathbb{S}^{p-1}}{\arg\max}\ \mathrm{E}\left\{\left(\frac{\mathbf{u}^{\top}\mathbf{s}-\mathbf{u}^{\top}\boldsymbol{\mu}}{\sqrt{\mathbf{u}^{\top}\boldsymbol{\Sigma}\mathbf{u}}}\right)^{3}\right\}.$$

The Mardia's kurtosis of $\mathbf{s}$ (Mardia, 1970) is

$$\beta_{2M}(\mathbf{s}) = \frac{(\lambda+1)(p+2)p}{\lambda}.$$

Let η be the square root of the ratio of the partial skewness to the total skewness:

$$\eta = \sqrt{\frac{\beta_{1,p}^{P}(\mathbf{s})}{\beta_{1,p}^{D}(\mathbf{s})}} = \frac{2+p+\varsigma}{3-\varsigma}.$$

A little algebra leads to the identities

$$\boldsymbol{\xi} = \frac{\boldsymbol{\mu}}{\lambda} - \sqrt{\frac{2}{\pi}}\boldsymbol{\delta},\ \lambda = \frac{p(p+2)}{\beta_{2,p}(\mathbf{s}) - p(p+2)},\ \boldsymbol{\alpha}^{\top}\boldsymbol{\Omega}\boldsymbol{\alpha} = \frac{\varsigma}{1-\varsigma},\ \varsigma = \frac{3\eta-2-p}{1+\eta},$$

$$\boldsymbol{\Sigma} = \frac{\boldsymbol{\Omega}}{\lambda},\ \boldsymbol{\nu}^{\top}\boldsymbol{\Omega}\boldsymbol{\nu} = \frac{\boldsymbol{\alpha}^{\top}\boldsymbol{\Omega}\boldsymbol{\alpha}}{\boldsymbol{\alpha}^{\top}\boldsymbol{\alpha}},\ \boldsymbol{\alpha}^{\top}\boldsymbol{\Omega}\boldsymbol{\alpha} = \frac{3\eta-2-p}{3-2\eta+p},\ \boldsymbol{\alpha}^{\top}\boldsymbol{\alpha} = \frac{3-2\eta+p}{(3-2\eta+p)\,\boldsymbol{\nu}^{\top}\boldsymbol{\Omega}\boldsymbol{\nu}}.$$

Let $\mathbf{m}$ and $\mathbf{S}$ be the sample mean and the sample variance of the $n \times p$ data matrix $\mathbf{X}$:

$$\mathbf{m} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i \text{ and } \mathbf{S} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_i-\mathbf{m})(\mathbf{x}_i-\mathbf{m})^{\top},$$

where $\mathbf{x}_i^{\top}$ is the i-th row of $\mathbf{X}$. Also, let the matrix $\mathbf{Z} = \mathbf{H}_n\mathbf{X}\mathbf{S}^{-1/2}$ denote the standardized data, where $\mathbf{H}_n = \mathbf{I}_n - n^{-1}\mathbf{1}_n\mathbf{1}_n^{\top}$ is the $n\times n$ centring matrix, $\mathbf{1}_n$ is the $n-$dimensional vector of ones, $\mathbf{I}_n$ is the $n-$dimensional identity matrix, and $\mathbf{S}^{-1/2}$ is the symmetric, positive definite square root of the sample concentration matrix $\mathbf{S}^{-1}$. The third and fourth standardized moments of $\mathbf{X}$ are

$$\mathbf{M}_{3,\mathbf{Z}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\otimes\mathbf{z}_i^{\top}\otimes\mathbf{z}_i \text{ and } \mathbf{M}_{4,\mathbf{Z}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\otimes\mathbf{z}_i^{\top}\otimes\mathbf{z}_i\otimes{\mathbf{z}_i}^{\top},$$

where $\mathbf{z}_i^{\top}$ is the i-th row of $\mathbf{Z}$. The sample partial and directional skewnesses (Loperfido, 2015) are

$$b_{1,P} = vec^{\top}(\mathbf{I}_p)\,\mathbf{M}_{3,\mathbf{Z}}\mathbf{M}_{3,\mathbf{Z}}^{\top}vec(\mathbf{I}_p) \text{ and } b_{1,D} = \max_{\mathbf{u}\in\mathbb{S}^{p-1}}\left\{\left(\mathbf{z}_i^{\top}\otimes{\mathbf{z}_i}^{\top}\right)\mathbf{M}_{3,\mathbf{Z}}\mathbf{z}_i\right\}^{2}.$$

The sample Mardia's kurtosis coincides with the trace of the fourth sample standardized moment Kollo and Srivastava, 2005):

$$b_{2,M} = \frac{1}{n}\sum_{i=1}^{n}\left[(\mathbf{x}_i-\mathbf{m})'\,\mathbf{S}^{-1}(\mathbf{x}_i-\mathbf{m})\right]^{2} = tr(\mathbf{M}_{4,\mathbf{Z}}).$$

The method-of-moments estimating procedure articulates into the following steps.

1. Use the sample Mardia kurtosis to derive a moment estimate of λ:

$$\widehat{\lambda} = \frac{p\,(p+2)}{b_{2,M} - p\,(p+2)}.$$

2. Use this estimate and the sample variance to estimate $\mathbf{\Omega}$:

$$\widehat{\mathbf{\Omega}} = \frac{\mathbf{S}}{\widehat{\lambda}}.$$

3. Use the sample partial and directional skewnesses to estimate η:

$$\widehat{\eta} = \sqrt{\frac{b_{1,P}}{b_{1,D}}}.$$

4. Estimate $\boldsymbol{\nu}$ by the direction maximizing the sample skewness:

$$\widehat{\boldsymbol{\nu}} = \frac{\widehat{\boldsymbol{\alpha}}}{\|\widehat{\boldsymbol{\alpha}}\|} = \underset{\mathbf{u}\in\mathbb{S}^{p-1}}{\arg\max}\left\{\frac{1}{n}\sum\left(\frac{\mathbf{u}^\top\mathbf{x}_i - \mathbf{u}^\top\mathbf{m}}{\sqrt{\mathbf{u}^\top\mathbf{S}\mathbf{u}}}\right)^3\right\}^2.$$

5. Estimate the norm of the shape parameter $\boldsymbol{\alpha}$ using the estimates of η, $\mathbf{\Omega}$ and $\boldsymbol{\nu}$:

$$\widehat{\|\boldsymbol{\alpha}\|} = \sqrt{\frac{3-2\widehat{\eta}+p}{(3-2\widehat{\eta}+p)\,\widehat{\boldsymbol{\nu}}^\top\widehat{\mathbf{\Omega}}\widehat{\boldsymbol{\nu}}}}.$$

6. Estimate the shape parameter $\boldsymbol{\alpha}$ using the estimates of $\boldsymbol{\nu}$ and $\|\boldsymbol{\alpha}\|$:

$$\widehat{\boldsymbol{\alpha}} = \widehat{\boldsymbol{\nu}}\widehat{\|\boldsymbol{\alpha}\|}.$$

7. This estimate, together with the estimate of the scatter matrix, leads to an estimate of $\boldsymbol{\delta}$:

$$\widehat{\boldsymbol{\delta}} = \frac{\widehat{\mathbf{\Omega}}\widehat{\boldsymbol{\alpha}}}{\sqrt{1+\widehat{\boldsymbol{\alpha}}^\top\widehat{\mathbf{\Omega}}\widehat{\boldsymbol{\alpha}}}}.$$

8. Use the estimate of $\boldsymbol{\delta}$, the estimate of $\widehat{\lambda}$ and the sample mean to estimate the location parameter $\boldsymbol{\xi}$:

$$\widehat{\boldsymbol{\xi}} = \frac{\mathbf{m}}{\widehat{\lambda}} - \sqrt{\frac{2}{\pi}}\widehat{\boldsymbol{\delta}}.$$

5 Dimension reduction

Fourth-order moments often occur within the frameworks of projection pursuit and invariant coordinate selection. In the former case, interesting projections $\boldsymbol{\eta}^{\top}\mathbf{x}$ or $\boldsymbol{\upsilon}^{\top}\mathbf{x}$ are sought by either maximizing or minimizing kurtosis, where

$$\boldsymbol{\eta} = \underset{\mathbf{u}\in\mathbb{S}^{p-1}}{\arg\max}\beta_2\left(\mathbf{u}^{\top}\mathbf{x}\right),\ \boldsymbol{\upsilon} = \underset{\mathbf{u}\in\mathbb{S}^{p-1}}{\arg\min}\beta_2\left(\mathbf{u}^{\top}\mathbf{x}\right),$$

$\mathbf{x}$ is a p-dimensional random vector with positive definite covariance matrix and finite fourth-order moments, $\mathbb{S}^{p-1}$ is the p-dimensional unit circle and $\beta_2(X)$ is the fourth standardized moment of the random variable X. Kurtosis-based projection pursuit has been used for cluster analysis (Peña and Prieto, 2001*b*) and outlier detection (Peña and Prieto, 2001*a*).

In kurtosis-based invariant coordinate selecction, interesting projections are sought by projecting the standardized data onto the eigenvectors of the partial kurtosis matrix $\mathrm{E}\left(\mathbf{z}^{\top}\mathbf{z}\mathbf{z}\mathbf{z}^{\top}\right)$, as proposed in Tyler *et al* (2009). As for the previous method, kurtosis-based invariant coordinate selection has been used in cluster analysis (Peña *et al*, 2010) and outlier detection (Archimbaud *et al*, 2018).

The following theorem highlights a shortcoming of both kurtosis-based projection pursuit and invariant coordinate selection, when applied to the Poisson-Skew-Normal model. All projections have the same kurtosis and any p-dimensional real vector is an eigenvector of the partial kurtosis matrix, since it is proportional to an identity matrix.

Theorem 3 *Let N be a Poisson random variable with mean λ. Also, let $\mathbf{s} = \mathbf{x}_1 + ... + \mathbf{x}_N$, where $\mathbf{x}_1$, ..., $\mathbf{x}_N$ are p-dimensional skew-normal random vectors. If N, $\mathbf{x}_1$, ..., $\mathbf{x}_N$ are mutually independent, the fourth standardized cumulant of a projection of $\mathbf{s}$ is*

$$\gamma_2\left(\mathbf{v}^{\top}\mathbf{s}\right) = \frac{3}{\lambda},\ \textit{where } \mathbf{v}\in\mathbb{R}_0^p,\ \textit{and } \mathrm{E}\left(\mathbf{z}^{\top}\mathbf{z}\mathbf{z}\mathbf{z}^{\top}\right) = \frac{p+2}{\lambda}\mathbf{I}_p.$$

Proof. In this proof we use the notation of the previous one. As seen in Loperfido (2017), the fourth cumulant of $\mathbf{u}^{\top}\mathbf{z}$, where $\mathbf{u}$ is a p-dimensional real vector of unit norm, is

$$\gamma_2\left(\mathbf{u}^{\top}\mathbf{z}\right) = \left(\mathbf{u}^{\top}\otimes\mathbf{u}^{\top}\right)\mathbf{K}_{4,\mathbf{z}}\left(\mathbf{u}\otimes\mathbf{u}\right).$$

The analytical formula of the fourth standardized cumulant derived in the previous proof and simple matrix algebra, lead to

$$\lambda\gamma_2\left(\mathbf{u}^{\top}\mathbf{z}\right) = \lambda\left(\mathbf{u}^{\top}\otimes\mathbf{u}^{\top}\right)\mathbf{K}_{4,\mathbf{z}}\left(\mathbf{u}\otimes\mathbf{u}\right) = \left(\mathbf{u}^{\top}\otimes\mathbf{u}^{\top}\right)\mathbf{I}_{p^2}\left(\mathbf{u}\otimes\mathbf{u}\right)$$

$$+\left(\mathbf{u}^{\top}\otimes\mathbf{u}^{\top}\right)\mathbf{C}_{p,p}\left(\mathbf{u}\otimes\mathbf{u}\right) + \left(\mathbf{u}^{\top}\otimes\mathbf{u}^{\top}\right)vec\left(\mathbf{I}_p\right)vec^{\top}\left(\mathbf{I}_p\right)\left(\mathbf{u}\otimes\mathbf{u}\right).$$

The identity $\left(\mathbf{u}^\top \otimes \mathbf{u}^\top\right) \mathbf{I}_{p^2} (\mathbf{u} \otimes \mathbf{u}) = \left(\mathbf{u}^\top \mathbf{u}\right)^2$ is a direct consequence of the following one: $vec(\mathbf{ABC}) = \left(\mathbf{C}^\top \otimes \mathbf{A}\right) vec(\mathbf{B})$, where $\mathbf{A} \in \mathbb{R}^p \times \mathbb{R}^q$, $\mathbf{B} \in \mathbb{R}^q \times \mathbb{R}^r$, $\mathbf{C} \in \mathbb{R}^r \times \mathbb{R}^s$ (see, e.g., Kollo and von Rosen, 2005, p. 89).

The identity $\left(\mathbf{u}^\top \otimes \mathbf{u}^\top\right) vec(\mathbf{I}_p) = \mathbf{u}^\top \mathbf{u}$ follows from $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}$, where $\mathbf{A} \in \mathbb{R}^p \times \mathbb{R}^q$, $\mathbf{B} \in \mathbb{R}^r \times \mathbb{R}^s$, $\mathbf{C} \in \mathbb{R}^q \times \mathbb{R}^u$, $\mathbf{D} \in \mathbb{R}^s \times \mathbb{R}^v$ (see, e.g., Kollo and von Rosen, 2005, p. 81).

Finally, the identity $\mathbf{C}_{p,p} (\mathbf{u} \otimes \mathbf{u}) = \mathbf{u} \otimes \mathbf{u}$ follows from ordinary properties of the commutation matrix, from the identity $vec\left(\mathbf{uu}^\top\right) = \mathbf{u} \otimes \mathbf{u}$ and from $\mathbf{uu}^\top$ being a symmetric matrix. Therefore, the fourth cumulant of $\mathbf{u}^\top \mathbf{z}$ is $\gamma_2\left(\mathbf{u}^\top \mathbf{z}\right) = 3/\lambda$, since $\mathbf{u}^\top \mathbf{u} = 1$, by assumption.

The fourth cumulant is invariant with respect to affine transformations, so that the fourth cumulant of $\mathbf{u}^\top \mathbf{z}$ coincides with the fourth cumulant of any projection of $\mathbf{x}$: $\gamma_2\left(\mathbf{u}^\top \mathbf{z}\right) = \gamma_2\left(\mathbf{v}^\top \mathbf{x}\right)$, where $\mathbf{v}$ is any nonnull p-dimensional real vector. This concludes the first part of the proof. The second part of the proof is very similar, after recalling the identity $vec\left\{\mathrm{E}\left(\mathbf{z}^\top \mathbf{zzz}^\top\right)\right\} = \mathbf{K}_{4,\mathbf{z}} vec(\mathbf{I}_p)$, as seen in Loperfido (2017).

6 Conclusions

We investigated the fourth-order cumulants of the multivariate Poisson-Skew-Normal aggregate claim model, and derived several measures of multivariate kurtosis. As a main statistical application, we obtained easily computable method-of-moments estimates, which are particularly useful, given the unavalilability of maximum likelihood estimates. As another statistical application, we highlighted some limitations of two established dimension reduction techinques: kurtosis-based projection pursuit and invariant coordinate selection.

A natural question to ask is whether the proposed approach could be extended to other multivariate claim models. We conjecture that the answer is positive, encouraged by the results in Loperfido *et al* (2018). They considered multivariate claim models defined by different claim number and size, including the negative binomial and the Generalized multivariate Laplace model. For any of them, they obtained closed-form expressions of the third cumulant and related measures of multivariate skewness.

As an example, consider the multivariate Laplace distribution, a prominent distribution which models both asymmetry and heavier than Gaussian tails (Kotz *et al*, 2001). A p-dimensional random vector $\mathbf{y}$ has a multivariate generalized asymmetric Laplace (GAL), denoted with $\mathbf{y} \sim \mathcal{GAL}_p(\nu, \boldsymbol{\mu}, \boldsymbol{\Sigma})$, if its characteristic function is

$$\phi_{\mathbf{y}}(\mathbf{t}) = \left(1 + \frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma} \mathbf{t} - \iota \boldsymbol{\mu}^\top \mathbf{t}\right)^{-\nu}, \ \mathbf{t} \in \mathbb{R}^p,$$

where $\nu > 0$, $\boldsymbol{\mu}$ is a p-dimensional vector and $\boldsymbol{\Sigma}$ is a $p \times p$ nonnegative definite symmetric matrix. The random vector $\mathbf{y}$ possesses a stochastic representation given by

$$\mathbf{y} \stackrel{d}{=} \boldsymbol{\mu}\Gamma_\nu + \Gamma_\nu^{1/2} \mathbf{z}_0 \stackrel{d}{=} \mathbf{G}_p(\Gamma_\nu),$$

where the symbol "$\stackrel{d}{=}$ " stands for the equality in distribution, Γ_ν has a standard gamma distribution with shape parameter ν, $\mathbf{z}_0 \sim \mathcal{N}_p(\mathbf{0}, \mathbf{\Sigma})$, and $\mathbf{G}_p(\cdot)$ is a p-dimensional Gaussian process with independent increments, $\mathbf{G}_d(0) = \mathbf{0}$, and $\mathbf{G}_d(1) \sim \mathcal{N}_d(\boldsymbol{\mu}, \mathbf{\Sigma})$. Stochastic representations show that GAL distribution can be represented as a subordinated Gaussian process, or as a location-scale mixture of normal distribution. For our discussion we need the fourth cumulant for GAL distributions. For which one can apply a general result about the fourth cumulant of the process that is obtained by subordination of a Brownian motion with drift to a random time change. An alternative approach has been discussed in Hurlimann (2013). Consider a non-negative random variable Γ_s independent of a Brownian motion $\mathbf{B}_d$ and consider

$$\mathbf{y} = \mathbf{B}_d(\Gamma_s).$$

As a second example of multivariate distribution which is often used to describe asymmetric and leptokurtic data, consider the Normal inverse Gaussian distribution. It can be represented as a multivariate Brownian motion $\mathbf{B}_p$ with drift $\boldsymbol{\mu} \in \mathbb{R}^p$ and covariance matrix $\mathbf{\Sigma} \in \mathbb{R}^{p \times p}$ when subordinated to the random variable z having the generalized inverse Gaussian (GIG) distribution with parameters $\lambda \in \mathbb{R}$, $\chi > 0$, and $\psi > 0$, often denoted with $z \sim GIG(\lambda, \chi, \psi)$. When $\lambda = -1/2$, z has the inverse Gaussian (IG) distribution with parameters χ and ψ, often denoted with $z \sim IG(\chi, \psi)$. As a direct consequence, $\mathbf{B}_p(\Gamma)$ has the p-dimensional normal inverse Gaussian (NIG) distribution.

As a venue of future research, we hope to generalize the results in this paper to claim sizes distributed either as multivariate generalized asymmetric Laplace or multivariate Normal inverse Gaussian.

Acknowledgement

Stepan Mazur acknowledges financial support from the project "Improved Economic Policy and Forecasting with High-Frequency Data" (Dnr: E47/22) funded by the Torsten Söderbergs Foundation and the internal research grants at Örebro University.

References

Ambagaspitiya, R. S. (1999). On the distributions of two classes of correlated aggregate claims. *Insurance: Mathematics and Economics* **24**, 301-308.

Archimbaud A., Nordhausen K. and Ruiz-Gazen A. (2018). ICS for multivariate outlier detection with application to quality control. *Comp. Statist. Data Anal.* **128**, 184-199.

Eling, M. (2012). Fitting insurance claims to skewed distributions: are the skew-normal and skew-student good models? *Insurance: Mathematics and Economics* **51**, 239-248.

Gnedenko, B. and Korolev, V. (1996). *Random Summation: Limit Theorems and Applications*. CRC Press, 1st edition.

Javed F., Loperfido N. and Mazur S. (2021). Edgeworth Expansions for Multivariate Random Sums. *Econometrics & Statistics*, in press.

Hürlimann, W. (2013). A moment method for the multivariate asymmetric Laplace distribution. *Statistics and Probability Letters* **83**, 1247-1253.

V. Kalashnikov, V. (1997). *Geometric Sums: Bounds for Rare Events with Applications*. Kluwer Academic, Dordrecht, 1997.

Klebanov L.B., Kozubowski T.J. and Rachev S.T. (2006). *Ill-Posed Problems in Probability and Stability of Random Sums*. Nova Science Publishers Inc, New York.

Kollo, T. (2008). Multivariate skewness and kurtosis measures with an application in ICA. *Journal of Multivariate Analysis* **99**, 2328-2338.

Kollo, T. and Srivastava, M.S. (2005). Estimation and testing of parameters in multivariate Laplace distribution. *Communications in Statistics* **33**, 2363-2687.

Kollo, T. and von Rosen, D. (2005). *Advanced Multivariate Statistics with Matrices*. Springer, Dordrecht.

Kotz S., Kozubowski T. J. and Podgórski, K. (2001). *The Laplace distribution and generalizations: a revisit with applications to communications, economics, engineering and finance.* Boston: Birkhäuser.

Lin, X. S. (2006). *Encyclopedia of actuarial science.* Wiley, 314–317.

Loperfido, N. (2015). Singular Value Decomposition of the Third Multivariate Moment. *Linear Algebra and its Applications* **473**, 202-216.

Loperfido, N. (2017). A New Kurtosis Matrix, with Statistical Applications. *Linear Algebra and its Applications* **512**, 1-17.

Loperfido N., Mazur S., and Podgorski, K. (2018). Third cumulant for multivariate aggregate claim models. *Scandinavian Actuarial Journal* **2**, 109-128.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis, with applications. *Biometrika* **57**, 519-530.

Mardia, K. V. (1974). Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. *Sankhya: The Indian Journal of Statistics Series B* **36**, 115-128.

Peña D. and Prieto F.J. (2001*a*). Multivariate Outlier Detection and Robust Covariance Estimation (with discussion). *Technometrics* **43**, 286-310.

Peña D. and Prieto F.J. (2001*b*). Cluster identification using projections. *J. Amer. Statist. Assoc.* **96**, 1433-1445.

Peña D., Prieto F.J. and Viladomat J. (2010). Eigenvectors of a kurtosis matrix as interesting directions to reveal cluster structure. *J. Multiv. Anal.* **101**, 1995-2007.

Tyler D.E., Critchley F., Dümbgen L. and Oja, H. (2009). Invariant coordinate selection (with discussion). *J. R. Statist. Soc. B* **71**, 549-592.