Online Supplementary Materials for "Modelling right-skewed financial data stream with outliers: a Likelihood-based inference via finite mixture of the generalized Birnbaum-Saunders distributions"

Mehrdad Naderi[a,*], Farzaneh Hashemi[a,b], Andriette Bekker[a], Ahad Jamalizadeh[b,c]

[a] *Department of Statistics, Faculty of Natural & Agricultural Sciences, University of Pretoria, Pretoria, South Africa*
[b] *Department of Statistics, Faculty of Mathematics & Computer, Shahid Bahonar University of Kerman, Kerman, Iran*
[c] *Mahani Mathematical Research Center, Shahid Bahonar University of Kerman, Kerman, Iran*

Abstract

The current supplementary material contains details of performing the Kolmogorov-Smirnov (KS) test, and proof of propositions and theorems.

Appendix A. Details of performing KS test

To study the validity of the hypothetical models, we obtain the KS goodness of fit test. The procedure for calculating the KS test statistic D_n, which is defined as the maximum value of the absolute difference between the empirical and estimated cumulative distributions, and the corresponding P-values are described below:

Step 1: Order the n data values and called them $t_{(1)}, t_{(2)}, \ldots, t_{(n)}$.

Step 2: Compute the KS test statistic

$$D_n = \max_{j=1,\ldots,n} \left\{ \frac{j}{n} - \hat{F}(t_{(j)}) \ , \ \hat{F}(t_{(j)}) - \frac{j-1}{n} \right\},$$

where $\hat{F}(\cdot)$ is the fitted cdf of a hypothetical distribution.

Step 3: For $i = 1, 2, \ldots, N$, generate n random numbers from $U(0, 1)$ and order them to obtain $u^{(i)}_{(1)} \leq u^{(i)}_{(2)} \leq \ldots \leq u^{(i)}_{(n)}$.

Step 4: Compute

$$d^{(i)} = \max_{j=1,\ldots,n} \left\{ \frac{j}{n} - u^{(i)}_{(j)} \ , \ u^{(i)}_{(j)} - \frac{j-1}{n} \right\}.$$

Step 5: Let $I_i = 1$ if $d^{(i)} \geq D_n$ and 0 otherwise. The P-value is estimated by $\frac{1}{N}\sum_{i=1}^{N} I_i$.

Appendix B. Detail of proofs

Proof of Proposition 1. Using the pdf of $T \sim \mathcal{NMV}-\mathcal{GBS}(\alpha, \beta, \lambda, \boldsymbol{\theta})$ in (7) and method of change-of-variable, the pdf of $Y = aT$ is obtained as

$$\begin{aligned} f(y; \alpha, \beta, \lambda, \boldsymbol{\theta}) &= \frac{1}{a} f_{\text{NMV-GBS}}\Big(\frac{y}{a}; \alpha, \beta, \lambda, \boldsymbol{\theta}\Big) \\ &= \frac{1}{a} C\Big(\frac{y}{a}, \alpha, \beta\Big) f_{\text{NMV}}\Big(c\Big(\frac{y}{a}, \alpha, \beta\Big); \lambda, \boldsymbol{\theta}\Big) \\ &= C(y, \alpha, a\beta) f_{\text{NMV}}(c(y, \alpha, a\beta); \lambda, \boldsymbol{\theta}), \end{aligned}$$

*Corresponding author

which is a pdf of $\mathcal{NMV}-\mathcal{GBS}(\alpha, a\beta, \lambda, \boldsymbol{\theta})$. Similarly for $Y = T^{-1}$, we have

$$
\begin{aligned}
f(y;\alpha,\beta,\lambda,\boldsymbol{\theta}) &= \frac{1}{y^2} f_{\text{NMV-GBS}}(y^{-1};\alpha,\beta,\lambda,\boldsymbol{\theta}) \\
&= \frac{1}{y^2} C(y^{-1},\alpha,\beta) f_{\text{NMV}}(c(y^{-1},\alpha,\beta);\lambda,\boldsymbol{\theta}) \\
&= C(y,\alpha,\beta^{-1}) f_{\text{NMV}}(c(y,\alpha,\beta^{-1});\lambda,\boldsymbol{\theta}).
\end{aligned}
$$

Proof of Proposition 2. The proof can be completed by using the presented representations in Definition 1. Through the stochastic representation of NMV-GBS distribution (6), we have

$$
W\lambda + W^{1/2}Z = \alpha^{-1}\left(\sqrt{T/\beta} - \sqrt{\beta/T}\right) \;\Rightarrow\; Z = -W^{1/2}\lambda + \left(\sqrt{W}\alpha\right)^{-1}\left(\sqrt{T/\beta} - \sqrt{\beta/T}\right) \sim N(0,1). \tag{B.1}
$$

Now, since the skew-normal distribution (Azzalini, 1985) contains the normal model as λ approaches zero, applying theorem 3.2 and corollaries 3.2 and 3.3 of Leiva et al. (2010) gives the results.

Proof of Proposition 3. To prove this proposition, we first note that the special cases of the GH distribution with pdf (2) are:

i. The GHST distribution is obtained when ψ approaches zero and $\kappa = -\nu/2, \chi = \nu$.

ii. The hyperbolic and VG distributions are obtained by setting $\kappa = 1$ and $\psi = 0$, respectively.

iii. When $\kappa = -0.5$, the GH distribution reduce to the normal inverse Gaussian model.

iv. Setting $\kappa = 1, \psi = 1, \chi = 0$, the skew-Laplace model is obtained.

v. The GH distribution includes the scale-mixture of normal distribution as λ tends to zero.

Using the pdf of NMV-GBS distribution (7), one can obtain the pdf of GH-BS distribution as

$$
f_{GH-BS}(t;\alpha,\beta,\lambda,\kappa,\chi,\psi) = C(t,\alpha,\beta) f_{\text{GH}}(c(t,\alpha,\beta);\lambda,\kappa,\chi,\psi). \tag{B.2}
$$

Therefore, the proof is completed by considering the special cases of GH distribution.

Proof of Proposition 4.

i. The proof of part *i* is trivial.

ii. Using (6), (7) and the pdf of NMVBS distribution in Pourmousa et al. (2015), the pdf of the NMVBS-BS distribution can be obtained.

iii. Similarly, by (6) and (7) and through the pdf of NMVL distribution in Naderi et al. (2017), the pdf of the NMVL-BS distribution can be obtained.

Proof of Theorem 1. From Proposition 2, we have $T|W = w \sim \mathcal{EBS}(\alpha\sqrt{w}, \beta, 2, -\sqrt{w}\lambda, 0)$. Using B.2 and some algebraic factorization, the conditional pdf can be obtained by applying Bayes' rule as

$$
\begin{aligned}
f(w|t) &= \frac{f(t,w)}{f(t)} = \frac{f(t|w)f(w)}{f(t)} \\
&= \frac{\dfrac{C(t,\alpha,\beta)}{\sqrt{2\pi w}}\exp\Big\{-\dfrac{1}{2w}(c(t,\alpha,\beta)-\lambda w)^2\Big\}\Big(\dfrac{\psi}{\chi}\Big)^{\kappa/2}\dfrac{w^{\kappa-1}}{2K_\kappa(\sqrt{\psi\chi})}\exp\left\{-\frac{1}{2}\left(w^{-1}\chi + w\psi\right)\right\}}{C(t,\alpha,\beta)\dfrac{\sqrt{(\psi/\chi)^\kappa}(\psi+\lambda^2)^{0.5-\kappa}}{\sqrt{2\pi}K_\kappa(\sqrt{\psi\chi})}\dfrac{K_{0.5-\kappa}(\sqrt{(\psi+\lambda^2)(\chi+c^2(t,\alpha,\beta))})}{(\sqrt{(\psi+\lambda^2)(\chi+c^2(t,\alpha,\beta))})^{0.5-\kappa}}\exp\left\{\lambda c(t,\alpha,\beta)\right\}} \\
&= \left(\frac{\psi+\lambda^2}{\chi+c^2(t,\alpha,\beta)}\right)^{\frac{\kappa-0.5}{2}}\frac{w^{\kappa-0.5-1}\exp\left\{-\frac{1}{2}\left(w^{-1}(\chi+c^2(t,\alpha,\beta))+w(\psi+\lambda^2)\right)\right\}}{2K_{0.5-\kappa}(\sqrt{(\psi+\lambda^2)(\chi+c^2(t,\alpha,\beta))})},
\end{aligned} \tag{B.3}
$$

which is a pdf of $\mathcal{GIG}(\kappa-0.5, \chi+c^2(t,\alpha,\beta), \psi+\lambda^2)$. Moreover, using the moments of the GIG distribution (Jørgensen, 1982), we have

$$
\begin{aligned}
\mathrm{E}[W^r|T=t] &= \Big(\frac{\chi+\rho}{\psi+\lambda^2}\Big)^{r/2} R_{(\kappa,r)}\left(\sqrt{(\chi+\rho)(\psi+\lambda^2)}\right), \qquad \text{for} \quad r=\pm1,\pm2,\ldots,\\
\mathrm{E}[\log W|T=t] &= \frac{\partial E[W^{\vartheta}|T=t]}{\partial\vartheta}\Big|_{\vartheta=0} = \log\left(\sqrt{\frac{\chi+\rho}{\psi+\lambda^2}}\right) + \frac{1}{K_{\kappa-0.5}(\sqrt{(\chi+\rho)(\psi+\lambda^2)})}\frac{\partial}{\partial\kappa}K_{\kappa-0.5}(\sqrt{(\chi+\rho)(\psi+\lambda^2)}).
\end{aligned}
$$

References

Azzalini, A., 1985. A class of distributions which includes the normal ones. Scandinavian Journal of Statistics 12, 171–178.

Jørgensen, B., 1982. Statistical properties of the generalized inverse Gaussian distribution. volume 9. Springer Science & Business Media.

Leiva, V., Vilca, F., Balakrishnan, N., Sanhueza, A., 2010. A skewed sinh-normal distribution and its properties and application to air pollution. Communications in Statistics - Theory and Methods 39, 426–443.

Naderi, M., Arabpour, A., Jamalizadeh, A., 2017a. Multivariate normal mean-variance mixture distribution based on Lindley distribution. Communications in Statistics - Simulation and Computation 47, 1179–1192.

Pourmousa, R., Jamalizadeh, A., Rezapour, M., 2015. Multivariate normal mean-variance mixture distribution based on Birnbaum-Saunders distribution. Journal of Statistical Computation and Simulation 85, 2736–2749.