Received: 15 August 2018 | Revised: 24 August 2018 | Accepted: 24 August 2018
DOI: 10.1002/asmb.2408

DISCUSSION

WILEY

Discussion of "Birnbaum-Saunders distributions: A review of models, analysis and applications"

1 | INTRODUCTION

We congratulate Balakrishnan and Kundu for this wide review on Birnbaum-Saunders (BS) models, which certainly will be helpful for the researchers in the areas of reliability and survival analysis. We are also grateful to the editor for the invitation to comment this interesting paper.

An important tool for assessing the suitability of some postulated model to a data set is the residual analysis, particularly the normal probability plot that has been largely applied in regression models. Discussions on this subject in BS models may be found, for instance, in the works by Leiva et al[1] and Paula et al.[2] However, the majority of the contributions of residual analysis in BS models do not consider the multivariate case. Then, the aim of this comment is to present the extension of the quantile residual[3] to the multivariate BS distribution under uncensored observations.

2 | UNIVARIATE CASE

The quantile residual was originally developed for the independent case and it is based on the cumulative distribution function (cdf) of the postulated error distribution and the standard normal distribution. To illustrate, suppose that $T_i \overset{\text{iid}}{\sim} \mathrm{BS}(\alpha, \beta)$ whose cdf is denoted by $F_T(t_i; \alpha, \beta)$, for $i = 1, \ldots, n$. Then, from the work by Dunn and Smyth[3] and by using the expression $F_T(t_i; \alpha, \beta) = \Phi\left[\frac{1}{\alpha}\left\{\left(\frac{t_i}{\beta}\right)^{1/2} - \left(\frac{\beta}{t_i}\right)^{1/2}\right\}\right]$, the quantile residual becomes given by

$$\begin{aligned} r_{q_i} &= \Phi^{-1}\left\{F_T(t_i; \hat{\alpha}, \hat{\beta})\right\} \\ &= \frac{1}{\hat{\alpha}}\left\{\left(\frac{t_i}{\hat{\beta}}\right)^{1/2} - \left(\frac{\hat{\beta}}{t_i}\right)^{1/2}\right\}, \end{aligned}$$

where $\Phi(\cdot)$ denotes the cdf of $\mathsf{N}(0, 1)$, for $i = 1, \ldots, n$. For large n and under the postulated model r_{q_i} follows a standard normal distribution. Therefore, the normal probability plot of r_{q_i} may be helpful to assess the suitability of the BS distribution to fit the data set.

3 | BIVARIATE CASE

Let $(T_1, T_2) \sim BS_2(\alpha_1, \beta_1, \alpha_2, \beta_2, \rho)$. The conditional probability density function of T_2, given $T_1 = t_1$, is expressed as follows:

$$f_{T_2|T_1=t_1}(t_2; \alpha_1, \beta_1, \alpha_2, \beta_2, \rho) = \frac{1}{2\sqrt{2\pi}\alpha_2\beta_2\sqrt{1-\rho^2}}\left[\left(\frac{\beta_2}{t_2}\right)^{1/2} + \left(\frac{\beta_2}{t_2}\right)^{3/2}\right]\exp\left\{-\frac{1}{2(1-\rho^2)}[v_2 - \rho v_1]^2\right\},$$

and the cdf of T_2, given $T_1 = t_1$, becomes

$$P(T_2 \le t_2 | T_1 = t_1) = \Phi\left(\frac{v_2 - \rho v_1}{\sqrt{1-\rho^2}}\right),$$

where $v_\ell = \frac{1}{\alpha_\ell}\left[\left(\frac{t_\ell}{\beta_\ell}\right)^{1/2} - \left(\frac{\beta_\ell}{t_\ell}\right)^{1/2}\right]$ and ρ denotes the correlation between V_1 and V_2, for $\ell = 1, 2$.

Suppose now that $(T_{i1}, T_{i2}) \overset{\text{iid}}{\sim} BS_2(\alpha_1, \beta_1, \alpha_2, \beta_2, \rho)$, for $i = 1, \ldots, n$. Then, according to the work by Dunn and Smith,[3] we can define the conditional quantile residuals from the conditional distribution of $T_{i2}|T_{i1} = t_{i1}$ as follows:

$$\begin{aligned} r_{q_{i2}} &= \Phi^{-1}\left\{\widehat{F}_{T_{i2}|T_{i1}}(t_{i2})\right\} \\ &= \frac{\hat{v}_{i2} - \hat{\rho}\hat{v}_{i1}}{\sqrt{1-\hat{\rho}^2}}, \end{aligned}$$

where

$$\hat{v}_{i\ell} = \frac{1}{\hat{\alpha}_\ell}\left\{\left(\frac{t_{i\ell}}{\hat{\beta}_\ell}\right)^{1/2} - \left(\frac{\hat{\beta}_\ell}{t_{i\ell}}\right)^{1/2}\right\},$$

and $\hat{\rho}$ denotes the estimated correlation, for $i = 1, \ldots, n$ and $\ell = 1, 2$. The quantile residuals for the bivariate case are formed by $r_{q_{i1}} = \hat{v}_{i1}$ and $r_{q_{i2}}$, for $i = 1, \ldots, n$.

4 | MULTIVARIATE CASE

To extend the quantile residual to the multivariate case, we will assume that $\mathbf{T}_i \overset{\text{iid}}{\sim} \mathrm{BS}_p(\boldsymbol{\alpha}, \boldsymbol{\beta}; \Gamma)$, where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p)^\top$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$, $\mathbf{T}_i = (T_{i1}, \ldots, T_{ip})^\top$, and Γ is the $p \times p$ correlation matrix, for $i = 1, \ldots, n$.

The idea here, generalizing the bivariate case, is to express the cdf of $\mathbf{T}_i$ as a product of conditional cdfs. Therefore, denoting $F_{T_i}(t_{i1}, \ldots, t_{ip})$ the cdf of $\mathbf{T}_i$, one has that

$$F_{T_i}(t_{i1}, \ldots, t_{ip}) = F_{T_{ip}|\mathbf{T}_{i(p-1)}}\left(t_{ip}|t_{i(p-1)}, t_{i(p-2)}, \ldots, t_{i1}\right) \ldots F_{T_{i2}|T_{i1}}(t_{i2}|t_{i1})F_{T_{i1}}(t_{i1}),$$

for $i = 1, \ldots, n$. Then, from Section 9 of the paper by Balakrishnan and Kundu, the quantile residuals for the p ($p \geq 2$) responses of the ith experimental unit are given by

$$\begin{aligned} r_{q_{i\ell}} &= \Phi^{-1}\left\{\widehat{F}_{T_{i\ell}|\mathbf{T}_{i(\ell-1)}}\left(t_{i\ell}|t_{i(\ell-1)}, \ldots, t_{i1}\right)\right\} \\ &= \left|\widehat{\Gamma}_{(\ell-1)(\ell-1).\ell}\right|^{-1/2}\left(\hat{v}_{i\ell} - \widehat{\Gamma}_{\ell(\ell-1)}\widehat{\Gamma}^{-1}_{(\ell-1)(\ell-1)}\hat{\mathbf{v}}_{i(\ell-1)}\right) \end{aligned}$$

and $r_{q_{i1}} = \hat{v}_{i1}$, where $|\Gamma_{(\ell-1)(\ell-1).\ell}|$ denotes the determinant of the matrix

$$\Gamma_{(\ell-1)(\ell-1).\ell} = \Gamma_{(\ell-1)(\ell-1)} - \Gamma_{(\ell-1)(\ell)}\Gamma^{-1}_{\ell\ell}\Gamma_{\ell(\ell-1)}$$

with

$$v_{i\ell} = \frac{1}{\alpha_\ell}\left\{\left(\frac{t_{i\ell}}{\beta_\ell}\right)^{1/2} - \left(\frac{\beta_\ell}{t_{i\ell}}\right)^{1/2}\right\},$$

$\mathbf{v}_{i(\ell-1)} = (v_{i(\ell-1)}, v_{i(\ell-2)}, \ldots, v_{i1})^\top$, $\Gamma_{\ell(\ell-1)}$ denotes the $1 \times (\ell-1)$ correlation matrix among $V_{i\ell}$ and $\mathbf{V}_{i(\ell-1)}$, and $\Gamma_{(\ell-1)(\ell-1)}$ denotes the $(\ell-1) \times (\ell-1)$ correlation matrix of $\mathbf{V}_{i(\ell-1)}$, for $i = 1, \ldots, n$ and $\ell = 2, \ldots, p$. For large n and under the postulated model, $r_{q_{i\ell}}$ follows $\mathrm{N}(0, 1)$. Therefore, the normal probability plot of $r_{q_{i\ell}}$ may be helpful to assess the suitability of the multivariate BS distribution to fit the data set.

Note that, in the bivariate case, one has $\Gamma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$ with $\Gamma_{11} = 1, \Gamma_{12} = \Gamma_{21} = \rho$, and $\Gamma_{22} = 1$. Then, $\Gamma_{11.2} = \Gamma_{11} - \Gamma_{12}\Gamma^{-1}_{22}\Gamma_{21} = 1 - \rho^2$, $r_{q_{i1}} = \hat{v}_{i1}$, and $r_{q_{i2}} = \frac{\hat{v}_{i2} - \hat{\rho}\hat{v}_{i1}}{\sqrt{1-\hat{\rho}^2}}$, where ρ denotes the correlation between V_{i1} and V_{i2}, for $i = 1, \ldots, n$, which agree with the expressions derived in the previous section.

5 | APPLICATIONS

In this section, we will derive the quantile residuals for the Examples 1, 2, and 5 presented in Section 13 of the paper by Balakrishnan and Kundu. For each example, the normal probability plot will be constructed added by a simulated confidence band.

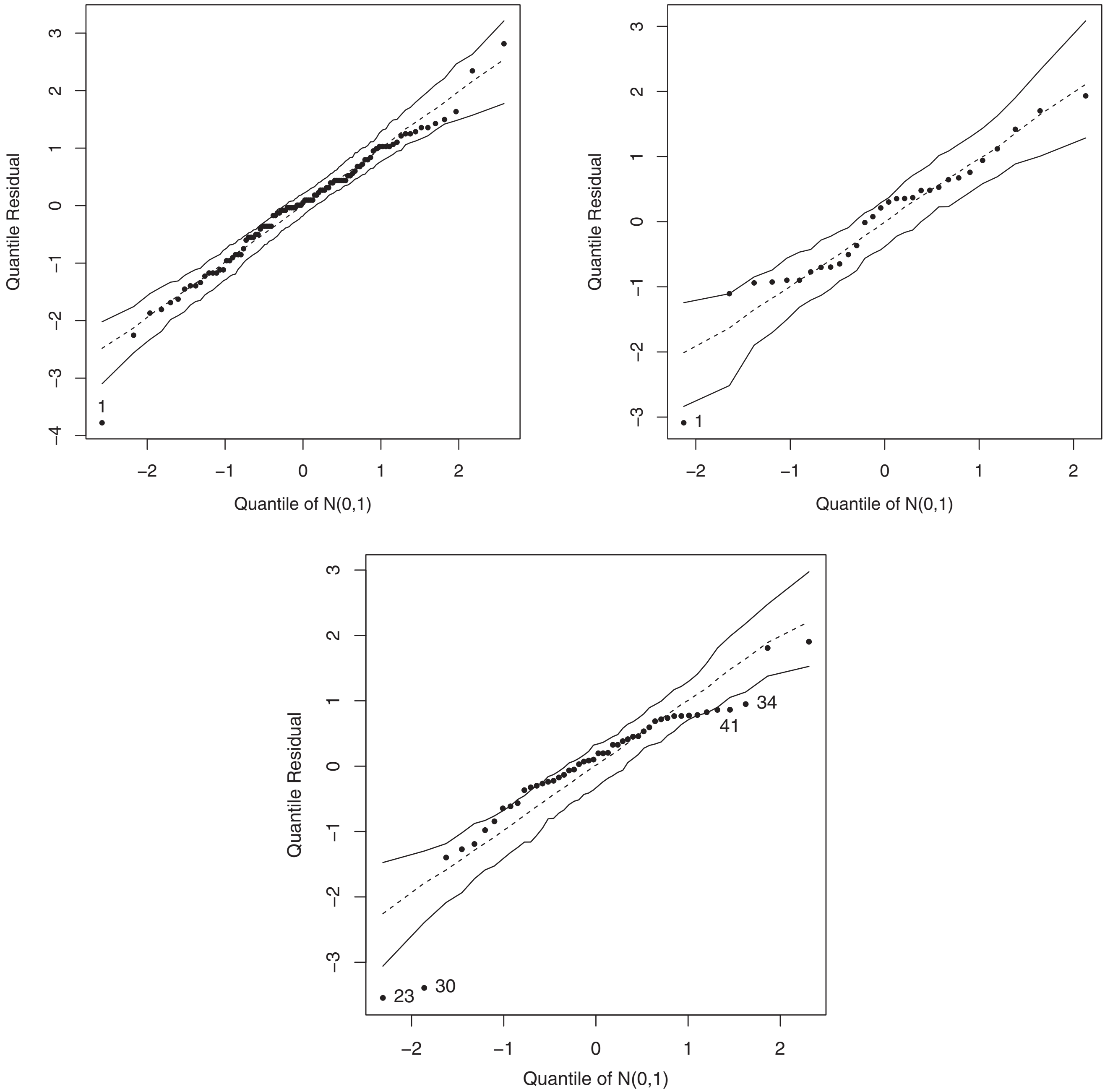

FIGURE 1 Normal probability plots for the quantile residual with a simulated confidence band of 95% for the examples of fatigue data (left), insurance data (right), and bivariate bone mineral data (bottom)

5.1 | Example 1

Figure 1 (left) presents the normal probability plot with a simulated confidence band for the quantile residuals derived from the fit of the univariate BS distribution to the fatigue data. We may notice from this graph that the BS distribution seems to be suitable to fit the maximum stress per cycle 31.000 psi. However, one observation corresponding to the response value #1 appears as a possible outlier.

5.2 | Example 2

Figure 1 (right) describes the normal probability plot with a simulated confidence band for the quantile residuals from the fit of the univariate BS distribution to the insurance data. We may notice from this graph that all points fall within the confidence band, unless one observation corresponding to the response value #1.

5.3 Example 5

Figure 1 (bottom) presents the normal probability plot with a simulated confidence band for the quantile residuals derived from the fit of the bivariate BS distribution to the bivariate bone mineral data. We may notice from this graph that four points fall outside the confidence band, corresponding to the response values #23, #30, #34, and #41. In addition, we also notice a systematic tendency, which may be indication of some departure from the assumption of bivariate BS distribution for the bivariate bone mineral data or even due to the small sample size, ie, 24 observations.

6 CONCLUSION

In this comment, quantile residuals are derived for the multivariate BS distribution for uncensored data. Their calculation is performed and illustrated by normal probability plots for three examples presented in the paper.

Michelli Barros[1]
Gilberto A. Paula[2]

[1]*Departamento de Estatística, Universidade Federal de Campina Grande, Campina Grande, Brazil*

[2]*Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, Brazil*

Correspondence

Michelli Barros, Departamento de Estatística, Universidade Federal de Campina Grande, Campina Grande, CEP: 58429-900, Brazil.
Email: michelli.karinne@ufcg.edu.br

REFERENCES

1. Leiva V, Barros M, Paula GA, Galea M. Influence diagnostics in log-Birnbaum-Saunders regression models with censored data. *Comput Stat Data Anal*. 2007;51(12):5694-5707.
2. Paula GA, Leiva V, Barros M, Liu S. Robust statistical modeling using the Birnbaum-Saunders-t distribution applied to insurance. *Appl Stoch Models Bus Ind*. 2012;28(1):16-34.
3. Dunn PK, Smyth GK. Randomized quantile residuals. *J Comput Graph Stat*. 1996;5(3):236-244.

How to cite this article: Barros M, Paula GA. Discussion of "Birnbaum-Saunders distributions: A review of models, analysis and applications". *Appl Stochastic Models Bus Ind*. 2018;1–4. https://doi.org/10.1002/asmb.2408