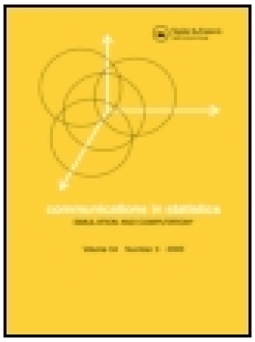

Communications in Statistics - Simulation and Computation

ISSN: 0361-0918 (Print) 1532-4141 (Online) Journal homepage: http://www.tandfonline.com/loi/lssp20

Improved maximum likelihood estimators for the parameters of the Johnson S_B distribution

André Felipe Berdusco Menezes & Josmar Mazucheli

To cite this article: André Felipe Berdusco Menezes & Josmar Mazucheli (2018): Improved maximum likelihood estimators for the parameters of the Johnson S_B distribution, Communications in Statistics - Simulation and Computation, DOI: 10.1080/03610918.2018.1498892

To link to this article: https://doi.org/10.1080/03610918.2018.1498892

Published online: 25 Sep 2018.

COMMUNICATIONS IN STATISTICS - SIMULATION AND COMPUTATION®
https://doi.org/10.1080/03610918.2018.1498892

Taylor & Francis
Taylor & Francis Group

Check for updates

Improved maximum likelihood estimators for the parameters of the Johnson S_B distribution

André Felipe Berdusco Menezes and Josmar Mazucheli

Department of Statistics, Universidade Estadual de Maringá, Maringá, PR, Brazil

ABSTRACT
In this article, considering the two-parameter Johnson S_B distribution, bounded on the unit interval, we derived, for the first time, the analytical expressions for bias-reduction of maximum likelihood estimators applying the Cox and Snell methodology. Although, in general, the analytical expressions are difficult to obtain, for the Johnson distribution they were simple and easy to implement. From Monte Carlo simulations, we estimated and compared the regular biases, the Cox and Snell biases and parametric Bootstrap-based biases. Our numerical results revealed that the biases should not be neglected and the bias reduction approaches based on the analytical expressions and Bootstrap are quite and equally efficient. Finally, a real application is presented and discussed.

ARTICLE HISTORY
Received 31 March 2017

KEYWORDS
Bootstrap bias-correction; Cox–Snell bias-correction; Johnson S_B distribution; maximum likelihood estimators; Monte Carlo simulation

1. Introduction

In reliability, life testing experiments and econometrics, several types of data are modeled by distributions bounded on the unit interval. For example, we see, among others, Johnson, Kotz, and Balakrishnan (1995), Papke and Wooldridge (1996), Bury (1999), Gupta and Nadarajah (2004), Cook, Kieschnick, and McCullough (2008) and Jiang (2013). In this sense, the most used distribution to model random variable in the unit interval is the Beta distribution, also known as Pearson type IV. The reputation of this distribution certainly is due to the flexibility of its probability density function (Johnson, Kotz, and Balakrishnan 1995).

However, several distributions are available as alternatives to the Beta distribution. Some of them, without exhaustion, are the Johnson S_B distribution (Johnson 1949), the Johnson S'_B distribution (Johnson 1955), the Topp–Leone distribution (Topp and Leone 1955), the unit-Gamma distribution (Grassia 1977), the Kumaraswamy distribution (Kumaraswamy 1980), the L_B distribution (Tadikamalla and Johnson 1982), the McDonald's generalized beta type I distribution (McDonald 1984), the Simplex distribution (Barndorff-Nielsen and Jørgensen 1991), the reflected Generalized Topp–Leone distribution (van Drop and Kotz 2006), the McDonald arcsine distribution (Cordeiro and Lemonte 2012), the Log–Lindley distribution (Gómez-Déniz, Sordo, and Calderín-Ojeda 2013), the exponentiated Kumaraswamy distribution (Lemonte, Barreto-Souza, and

CONTACT Josmar Mazucheli jmazucheli@gmail.com Department of Statistics, Universidade Estadual de Maringá, Maringá, PR, Brazil.

Cordeiro 2013), the exponentiated Topp–Leone distribution (Pourdarvish, Mirmostafaee, and Naderi 2015), the Marshall–Olkin extended Kumaraswamy (Castellares and Lemonte 2016), the reflected generalized Topp–Leone power series distribution (Condino and Domma 2016), the transmuted Kumaraswamy distribution (Shuaib, Robert, and Lena 2016), the size biased Kumaraswamy distribution (Sharma and Chakrabarty 2016) and the extended arcsine distribution (Cordeiro, Lemonte, and Campelo 2016). It should be pointed that the majority of these distributions have more than two parameters, which taking into account data limited amount, may produce inaccurate estimates. Moreover, many of them involve special functions in their mathematical expressions.

An interesting system of distributions, whose support can be restricted to the unit interval, was proposed by Johnson (1949) received considerable attention in the second half of the 20th century (Kotz and van Dorp 2004). As pointed out in George (2007), the Johnson system is able to closely approximate many of the standard continuous distributions through one of the three functional forms and is thus highly flexible. The Johnson system accommodates the S_B family of distribution, which has a bounded support and due to its flexibility can be an important alternative to the popular Beta distribution.

According to Kotz and van Dorp (2004), the Johnson S_B distribution was developed as follows. Let X be a standard normal distribution and consider the transformation:

$$Y = g^{-1}\left(\frac{X-\gamma}{\delta}\right), \tag{1}$$

for some suitable function $g(\cdot)$ and parameters $\gamma \in \mathbb{R}$ and $\delta>0$. The choice of $g(\cdot)$ determines the support of the distribution, hence from Johnson (1949), by taking

$$g(Y) = \log\left(\frac{Y}{1-Y}\right), \tag{2}$$

we obtain the Johnson S_B distribution with unity support and probability density function written as:

$$f(y|\gamma,\delta) = \frac{\delta}{\sqrt{2\ \pi}}\ \frac{1}{y\ (1-y)}\ \exp\left\{-\frac{1}{2}\left[\gamma+\delta\ \log\left(\frac{y}{1-y}\right)\right]^2\right\}, \tag{3}$$

where $0<y<1, \gamma \in \mathbb{R}$ and $\delta>0$ are shapes parameters. The corresponding cumulative distribution function and quantile function are written respectively as

$$F(y|\gamma,\delta) = \Phi\left[\gamma+\delta\ \log\left(\frac{y}{1-y}\right)\right] \tag{4}$$

and

$$Q(p|\gamma,\delta) = \frac{\exp\left(\frac{\Phi^{-1}(p)-\gamma}{\delta}\right)}{1+\exp\left(\frac{\Phi^{-1}(p)-\gamma}{\delta}\right)}, \tag{5}$$

where $0<p<1$ and $\Phi(\cdot)$ is the cumulative distribution function of a standard normal distribution.

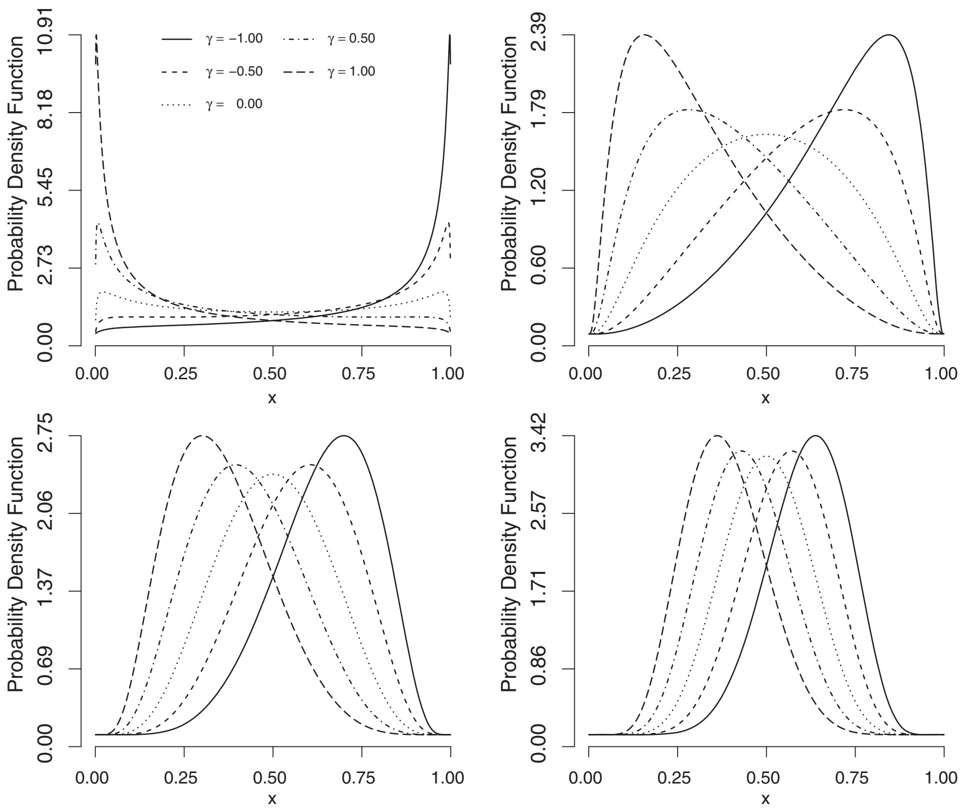

Figure 1. Johnson S_B probability density function considering different values of γ and δ (upper-panel: $\delta = 0.5$ and $\delta = 1.0$, respectively; lower-panel: $\delta = 1.5$ and $\delta = 2.0$, respectively).

Figure 1 illustrates the behavior of the probability density function of the Johnson S_B distribution for different values of γ and δ. It is noteworthy that the densities may display different shapes depending on the values of the two parameters.

Many others Johnson S_B distribution characteristics can be found in Johnson (1949) and Kotz and van Dorp (2004). The parameter estimation of Johnson S_B distribution was studied by several authors. The method of moment was studied by Hill, Hill, and Holder (1976) and Bacon-Shone (1985), while the estimation based on the percentile was considered by Johnson (1949), Bukac (1972), Mage (1980) and Slifker and Shapiro (1980). On the other hand, the maximum likelihood estimation was introduced first by Kottegoda (1987) and latter by Wheeler (1980), Siekierski (1992) and Zhou and McTague (1996). It should be mention that these studies considered the Johnson S_B indexed by four parameters.

Although other estimation methods are useful when the distribution is indexed by at least three parameters, the method of maximum likelihood (Millar 2011; Pawitan 2001) is the most popular method for statistical inference, since it has several attractive properties. For instance, they are asymptotically unbiased, efficient, consistent, functional invariance and asymptotically normally distributed (Edwards 1992; Lehmann 1999). Not all of these properties are shared with other estimation methods. However, it is notable that most of these properties depend on the sample size. Indeed, the maximum likelihood method produces estimates that have biases of order $\mathcal{O}(n^{-1})$, where n is the sample size (Cordeiro and Cribari-Neto 2014). Nevertheless, for small or even moderate sample size it is important to remove the second-order bias in order to obtain estimators with better properties.

In this article, we shall focus on two different approaches that can be employed to obtain modified MLEs that are nearly free of bias, specifically, modified MLEs that are unbiased to second order. First, we derived analytical expressions for the biases through the methodology proposed by Cox and Snell (1968). Lastly, we considered the Bootstrap-based bias-adjusted, which had its pioneered Efron (1982).

In the literature there are many works which introduced bias-corrections for the parameter of others distributions. We may mention: Cordeiro et al. (1997), Cribari-Neto and Vasconcellos (2002), Saha and Paul (2005), Lemonte, Cribari-Neto and Vasconcellos (2007), Giles and Feng (2009), Lagos-Àlvarez, Jiménez-Gamero, and Alba-Fernández (2011), Lemonte (2011), Giles (2012a, 2012b), Schwartz, Godwin, and Giles (2013), Giles, Feng, and Godwin (2013), Teimouri and Nadarajah (2013), Ling and Giles (2014), Zhang and Liu (2015), Singh, Singh, and Murphy (2015), Teimouri and Nadarajah (2016), Schwartz and Giles (2016), Wang and Wang (2017), Mazucheli (2017), Reath, Dong, and Wang (2018), Mazucheli and Dey (2018), Mazucheli, Menezes, and Dey (2018a, 2018b).

The article unfolds as follows. In Sec. 2 we described the maximum likelihood estimators and asymptotic confidence intervals for the parameters of Johnson S_B distribution. Sec. 3 presents the approaches to bias corrections. In Sec. 4, a simulation study is performed to compare the MLEs and bias corrected MLEs. An application using a real data set is presented in Sec. 5. Finally, Sec. 6 closes the article with some concluding remarks.

2. Maximum likelihood estimation

Suppose that $\mathrm{y} = (y_1, \ldots, y_n)$ is a random sample of size n from the Johnson S_B distribution (3) with parameter vector $\theta = (\gamma, \delta)$. The log-likelihood function, dropping constant terms, is written as:

$$l(\theta|y) \propto n \log \delta - \frac{1}{2} \sum_{i=1}^{n} \left[\gamma + \delta \ \log \left(\frac{y_i}{1 - y_i} \right) \right]^2. \tag{6}$$

The maximum likelihood estimates of the γ and δ, $\hat{\gamma}$ and $\hat{\delta}$, respectively, can be obtained solving the nonlinear equations:

$$\frac{\partial}{\partial \gamma} \ l(\theta|\mathrm{y}) = -n \ \gamma - \delta \ \sum_{i=1}^{n} \log \left(\frac{y_i}{1 - y_i} \right) \tag{7}$$

$$\frac{\partial}{\partial \delta} \ l(\theta|y) = \frac{n}{\delta} - \sum_{i=1}^{n} \left[\gamma + \delta \ \log \left(\frac{y_i}{1 - y_i} \right) \right] \log \left(\frac{y_i}{1 - y_i} \right). \tag{8}$$

From (7) we have $\hat{\gamma} = -\frac{\hat{\delta}}{n} \sum_{i=1}^{n} \log \left(\frac{y_i}{1 - y_i} \right)$ while for δ, the maximum likelihood estimate $\hat{\delta}$ must be obtained numerically by solving (8) in δ, replacing γ by $\hat{\gamma}$.

To obtain interval estimation and testing hypothesis for the parameters using the maximum likelihood estimates $\hat{\gamma}$ and $\hat{\delta}$, we can use the expected Fisher information matrix, which is obtained from $n \times E\left(-\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log f(y|\theta)\right)$ for $i, j = 1, 2$ and given by:

$$K(\theta|y) = n \begin{bmatrix} 1 & -\dfrac{\gamma}{\delta} \\ -\dfrac{\gamma}{\delta} & \dfrac{(\gamma^2+2)}{\delta^2} \end{bmatrix}. \tag{9}$$

From (9), we observe that γ and δ are not orthogonal, i.e., the maximum likelihood estimates $\hat{\gamma}$ and $\hat{\delta}$ are not asymptotically independent. The inverse of the expected Fisher information matrix is given by

$$K^{-1}(\theta|y) = \frac{1}{2n} \begin{bmatrix} 2+\gamma^2 & \gamma\ \delta \\ \gamma\ \delta & \delta^2 \end{bmatrix} \tag{10}$$

and evaluated at $\hat{\gamma}$ and $\hat{\delta}$ provides the asymptotic variance-covariance matrix of the maximum likelihood estimates. Since $K(\theta|y)$ is data independent it is equal to the observed information matrix. Naturally, the asymptotic $100 \times (1-\alpha)\%$ confidence intervals of γ and δ, respectively, are then given by

$$\hat{\gamma} \pm z_{\frac{\alpha}{2}} \times \sqrt{\hat{\mathrm{Var}}(\hat{\gamma})} \text{ and } \hat{\delta} \pm z_{\frac{\alpha}{2}} \times \sqrt{\hat{\mathrm{Var}}(\hat{\delta})}$$

where $\hat{\mathrm{Var}}(\hat{\gamma})$ and $\hat{\mathrm{Var}}(\hat{\gamma})$ are elements of the matrix main diagonal defined in (10) and $z_{\frac{\alpha}{2}}$ is the $100 \times (1-\frac{\alpha}{2})\%$ percentile of the standard normal distribution.

3. Bias-corrected MLEs

In what follows, we shall discuss two approaches for bias-reduction the maximum likelihood estimators of the parameters that index the Johnson S_B distribution. First, we shall consider the general formula introduced by Cox and Snell (1968). As reported by the authors, when the sample data are independent, but not necessarily identically distributed, the bias of the s-th element of the MLE of $\theta, \hat{\theta}$, is obtained as:

$$\mathcal{B}(\hat{\theta}_s) = \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{l=1}^{p} \kappa^{si}\ \kappa^{jl}\ [0.5\kappa_{ijl} + \kappa_{ij,l}] + \mathcal{O}(n^{-2}), \tag{11}$$

where $s = 1, \dots, p$, κ^{ij} is the (i, j)-th element of the inverse of the expected Fisher information, $\kappa_{ijl} = \mathbb{E}\left[\frac{\partial^3}{\partial\ \theta_i\ \partial\ \theta_j\ \partial\ \theta_l}\ l(\theta|y)\right]$ and $\kappa_{ij,l} = \mathbb{E}\left[\frac{\partial^2}{\partial\ \theta_i\ \partial\ \theta_j}\ l(\theta|y)\ \frac{\partial}{\partial\ \theta_l}\ l(\theta|y)\right]$. For further details about this methodology interested reader can consult Cordeiro and Cribari-Neto (2014).

In respect to the Johnson S_B distribution, after some algebra, we verified that:

- $\kappa_{111} = \kappa_{112} = \kappa_{121} = \kappa_{122} = \kappa_{211} = \kappa_{212} = \kappa_{221} = 0$,
- $\kappa_{222} = \frac{2n}{\delta^3}$,
- $\kappa_{11,1} = \kappa_{11,2} = 0$,
- $\kappa_{12,1} = \kappa_{21,1} = \frac{n}{\delta}$,
- $\kappa_{12,2} = \kappa_{21,2} = -\frac{n\gamma}{\delta^2}$,
- $\kappa_{22,1} = -\frac{2n\gamma}{\delta^2}$

  and
- $\kappa_{22,2} = \frac{2\ n(\gamma^2+1)}{\delta^3}$.

By replacing these terms in Eq. (11) we achieved the following expressions for the second order biases of $\hat{\gamma}$ and $\hat{\delta}$:

$$\mathcal{B}(\hat{\gamma}) = \frac{5\gamma}{4n} + \mathcal{O}(n^{-2}) \tag{12}$$

and

$$\mathcal{B}(\hat{\delta}) = \frac{5\delta}{4n} + \mathcal{O}(n^{-2}). \tag{13}$$

Employing Eqs (12) and (13), we defined the bias-corrected estimators as:

$$\hat{\gamma}_{BCE} = \hat{\gamma} - \frac{5\hat{\gamma}}{4n} \tag{14}$$

and

$$\hat{\delta}_{BCE} = \hat{\delta} - \frac{5\hat{\delta}}{4n}. \tag{15}$$

We should expect that $\hat{\gamma}_{BCE}$ and $\hat{\delta}_{BCE}$ have better sampling properties than the first-order biased maximum likelihood estimators $\hat{\gamma}$ and $\hat{\delta}$, respectively. Nonetheless, as pointed by Cordeiro and Cribari-Neto (2014), bias-corrections may also increase the mean-squared error.

The other method that we consider to obtain nearly unbiased estimators for the Johnson S_B distribution is based on the Bootstrap scheme (Efron 1982; Efron and Tibshirani 1993; Davison and Hinkley 1997). In particular, the Bootstrap bias-correction handling the data to estimate the bias function. Let $\hat{\theta}_{(\cdot)}$ be the average value of the maximum likelihood estimator from B Bootstrap replications, each of them based on a pseudo-sample of size n generated from (3) using the maximum likelihood estimates $\hat{\theta}$. Thus, its estimated bias is given by

$$\hat{\mathcal{B}}(\hat{\theta}) = \hat{\theta}_{(\cdot)} - \hat{\theta}, \tag{16}$$

hence yielding the Bootstrap bias-corrected estimator as

$$\hat{\theta}_{PBE} = 2\hat{\theta} - \hat{\theta}_{(\cdot)}. \tag{17}$$

As well as the analytically corrected estimators, the Bootstrap corrected estimator is also a method that provides second order bias-correction (Ferrari and Cribari-Neto 1998).

4. Simulation study

In this section, based on Monte Carlo simulations, we shall evaluate the finite-sample behavior of the MLEs of γ and δ and their bias-corrections obtained by Cox–Snell methodology (BCE) and parametric Bootstrap scheme (PBE). We considered random samples of size $n = 10, 20, 30, 40$ and 50 and the parameters values were $\gamma = -1.0, -0.3, 0.0, 0.3$ and 1.0 and $\delta = 0.5, 1.0, 1.5$ and 2.0. To simulate pseudo-random samples from Johnson S_B distribution we used the fact describe in Sec. 1, specifically Eqs (1) and (2). The numbers of Monte Carlo replications in each experiment was set at $M = 10.000$ and the numbers of Bootstrap replications was $B = 1000$, thus totaling 100 millions of replications per experiment. All simulation were carried out in *Ox Console* (Doornik 2007), using the MaxBFGS function to obtain the maximum likelihood estimates for γ and δ. The results are shown in Tables 1–4, where we reported the bias estimates and the root mean-squared errors estimates.

It is observed that the MLEs of γ is extremely biased, while for δ it is moderate, particularly for the small samples size. We may mention, the scenario when $n = 10$, $\gamma = 1.0$ and $\delta = 0.5$ the biases of the MLEs of γ and δ are 0.1553 and 0.0758, respectively. Considering the same scenario above we observed that proposed estimators outperform the MLEs, which the bias of γ_{BCE}, γ_{PBE}, δ_{BCE} and δ_{PBE} are 0.0109, −0.0219, 0.0038 and −0.0126, respectively. Indeed, these estimators achieve substantial bias reduction, mainly in small samples size and therefore they are good alternatives to the uncorrected MLEs. Although the proposed estimators were quite effectively, it should be pointed out that analytical corrections are done immediately, i.e., it is not necessary a computational effort. It is noteworthy that for all scenarios the maximum difference between the bias of MLE and BCE of γ and δ were 28.89 and 14.41%, respectively, for the MLE and PBE the difference were 17.72 and 35.44%, respectively and for the BCE and PBE were 3.28 and 6.56%, respectively. Based on these differences we concluded that the BCE and PBE really provided the bias reduction for both parameters. As expected, the reduction magnitude is generally smaller for larger n.

Table 1. Estimated bias (root mean-squared error), $\delta = 0.5$.

		Estimator of δ			Estimator of γ		
γ	n	MLE	BCE	PBE	MLE	BCE	PBE
−1.0	10	−0.1586 (0.5244)	−0.0137 (0.4376)	0.0192 (0.4241)	0.0776 (0.1751)	0.0054 (0.1375)	−0.0109 (0.1334)
	20	−0.0765 (0.3162)	−0.0093 (0.2877)	−0.0024 (0.2859)	0.0347 (0.0984)	0.0013 (0.0863)	−0.0022 (0.0858)
	30	−0.0482 (0.2421)	−0.0045 (0.2274)	−0.0017 (0.2267)	0.0221 (0.0745)	0.0003 (0.0682)	−0.0011 (0.0680)
	40	−0.0356 (0.2046)	−0.0033 (0.1952)	−0.0016 (0.1950)	0.0164 (0.0627)	0.0002 (0.0586)	−0.0005 (0.0586)
	50	−0.0291 (0.1809)	−0.0034 (0.1741)	−0.0023 (0.1740)	0.0131 (0.0548)	0.0003 (0.0519)	−0.0002 (0.0519)
−0.3	10	−0.0461 (0.3943)	−0.0029 (0.3426)	0.0069 (0.3320)	0.0754 (0.1727)	0.0035 (0.1360)	−0.0128 (0.1325)
	20	−0.0192 (0.2507)	0.0008 (0.2343)	0.0027 (0.2329)	0.0338 (0.0985)	0.0004 (0.0868)	−0.0030 (0.0863)
	30	−0.0119 (0.1979)	0.0011 (0.1894)	0.0020 (0.1889)	0.0208 (0.0741)	−0.0009 (0.0681)	−0.0023 (0.0680)
	40	−0.0088 (0.1688)	0.0008 (0.1633)	0.0013 (0.1631)	0.0152 (0.0620)	−0.0009 (0.0582)	−0.0017 (0.0582)
	50	−0.0082 (0.1491)	−0.0005 (0.1452)	−0.0001 (0.1451)	0.0121 (0.0546)	−0.0007 (0.0519)	−0.0012 (0.0519)
0.0	10	0.0020 (0.3834)	0.0018 (0.3355)	0.0019 (0.3251)	0.0778 (0.1736)	0.0056 (0.1359)	−0.0108 (0.1319)
	20	0.0023 (0.2442)	0.0021 (0.2290)	0.0022 (0.2275)	0.0349 (0.0987)	0.0015 (0.0866)	−0.0019 (0.0861)
	30	0.0029 (0.1945)	0.0027 (0.1864)	0.0027 (0.1860)	0.0231 (0.0764)	0.0013 (0.0698)	−0.0001 (0.0696)
	40	0.0038 (0.1647)	0.0037 (0.1595)	0.0037 (0.1593)	0.0171 (0.0635)	0.0010 (0.0592)	0.0002 (0.0592)
	50	0.0034 (0.1463)	0.0033 (0.1426)	0.0033 (0.1426)	0.0136 (0.0558)	0.0008 (0.0528)	0.0003 (0.0527)
0.3	10	0.0425 (0.3919)	−0.0003 (0.3408)	−0.0099 (0.3301)	0.0754 (0.1729)	0.0035 (0.1362)	−0.0129 (0.1325)
	20	0.0168 (0.2461)	−0.0030 (0.2302)	−0.0049 (0.2286)	0.0330 (0.0979)	−0.0003 (0.0864)	−0.0037 (0.0860)
	30	0.0117 (0.1969)	−0.0013 (0.1884)	−0.0021 (0.1880)	0.0217 (0.0743)	−0.0001 (0.0681)	−0.0015 (0.0680)
	40	0.0090 (0.1685)	−0.0006 (0.1630)	−0.0011 (0.1628)	0.0157 (0.0619)	−0.0004 (0.0580)	−0.0012 (0.0579)
	50	0.0071 (0.1503)	−0.0006 (0.1463)	−0.0008 (0.1462)	0.0126 (0.0540)	−0.0003 (0.0512)	−0.0007 (0.0512)
1.0	10	0.1553 (0.5115)	0.0109 (0.4266)	−0.0219 (0.4134)	0.0758 (0.1717)	0.0038 (0.1349)	−0.0126 (0.1311)
	20	0.0728 (0.3118)	0.0058 (0.2843)	−0.0010 (0.2825)	0.0351 (0.0999)	0.0016 (0.0877)	−0.0017 (0.0871)
	30	0.0473 (0.2436)	0.0036 (0.2290)	0.0008 (0.2284)	0.0227 (0.0755)	0.0010 (0.0691)	−0.0005 (0.0689)
	40	0.0341 (0.2078)	0.0018 (0.1986)	0.0003 (0.1985)	0.0166 (0.0635)	0.0005 (0.0594)	−0.0003 (0.0593)
	50	0.0279 (0.1825)	0.0022 (0.1759)	0.0012 (0.1759)	0.0132 (0.0553)	0.0004 (0.0524)	−0.0001 (0.0524)

Table 2. Estimated bias (root mean-squared error), $\delta = 1.0$.

		Estimator of δ			Estimator of γ		
γ	n	MLE	BCE	PBE	MLE	BCE	PBE
−1.0	10	−0.1586 (0.5244)	−0.0137 (0.4376)	0.0192 (0.4241)	0.0776 (0.1751)	0.0054 (0.1375)	−0.0109 (0.1334)
	20	−0.0765 (0.3162)	−0.0093 (0.2877)	−0.0024 (0.2859)	0.0694 (0.1968)	0.0025 (0.1727)	−0.0043 (0.1716)
	30	−0.0482 (0.2421)	−0.0045 (0.2274)	−0.0017 (0.2267)	0.0441 (0.1489)	0.0006 (0.1363)	−0.0022 (0.1360)
	40	−0.0356 (0.2046)	−0.0033 (0.1952)	−0.0016 (0.1950)	0.0328 (0.1254)	0.0005 (0.1173)	−0.0011 (0.1172)
	50	−0.0291 (0.1809)	−0.0034 (0.1741)	−0.0023 (0.1740)	0.0262 (0.1097)	0.0005 (0.1038)	−0.0004 (0.1038)
−0.3	10	−0.0461 (0.3943)	−0.0029 (0.3426)	0.0069 (0.3320)	0.1509 (0.3455)	0.0070 (0.2720)	−0.0256 (0.2649)
	20	−0.0192 (0.2507)	0.0008 (0.2343)	0.0027 (0.2329)	0.0675 (0.1971)	0.0008 (0.1736)	−0.0060 (0.1726)
	30	−0.0119 (0.1979)	0.0011 (0.1894)	0.0020 (0.1889)	0.0416 (0.1481)	−0.0018 (0.1363)	−0.0046 (0.1360)
	40	−0.0088 (0.1688)	0.0008 (0.1633)	0.0013 (0.1631)	0.0305 (0.1240)	−0.0017 (0.1165)	−0.0033 (0.1164)
	50	−0.0082 (0.1491)	−0.0005 (0.1452)	−0.0001 (0.1451)	0.0242 (0.1092)	−0.0015 (0.1038)	−0.0024 (0.1038)
0.0	10	0.0020 (0.3834)	0.0018 (0.3355)	0.0019 (0.3251)	0.1556 (0.3472)	0.0112 (0.2718)	−0.0216 (0.2639)
	20	0.0023 (0.2442)	0.0021 (0.2290)	0.0022 (0.2275)	0.0698 (0.1975)	0.0030 (0.1732)	−0.0039 (0.1722)
	30	0.0029 (0.1945)	0.0027 (0.1864)	0.0027 (0.1860)	0.0462 (0.1528)	0.0026 (0.1396)	−0.0003 (0.1393)
	40	0.0038 (0.1647)	0.0037 (0.1595)	0.0037 (0.1593)	0.0343 (0.1270)	0.0020 (0.1185)	0.0005 (0.1184)
	50	0.0034 (0.1463)	0.0033 (0.1426)	0.0033 (0.1426)	0.0273 (0.1116)	0.0016 (0.1055)	0.0006 (0.1054)
0.3	10	0.0425 (0.3919)	−0.0003 (0.3408)	−0.0099 (0.3301)	0.1508 (0.3459)	0.0070 (0.2724)	−0.0258 (0.2650)
	20	0.0168 (0.2461)	−0.0030 (0.2302)	−0.0049 (0.2286)	0.0659 (0.1958)	−0.0007 (0.1728)	−0.0074 (0.1719)
	30	0.0117 (0.1969)	−0.0013 (0.1884)	−0.0021 (0.1880)	0.0434 (0.1487)	−0.0001 (0.1363)	−0.0029 (0.1359)
	40	0.0090 (0.1685)	−0.0006 (0.1630)	−0.0011 (0.1628)	0.0314 (0.1238)	−0.0008 (0.1160)	−0.0024 (0.1159)
	50	0.0071 (0.1503)	−0.0006 (0.1463)	−0.0008 (0.1462)	0.0251 (0.1080)	−0.0005 (0.1024)	−0.0015 (0.1024)
1.0	10	0.1553 (0.5115)	0.0109 (0.4266)	−0.0219 (0.4134)	0.1516 (0.3434)	0.0076 (0.2697)	−0.0252 (0.2622)
	20	0.0728 (0.3118)	0.0058 (0.2843)	−0.0010 (0.2825)	0.0701 (0.1997)	0.0033 (0.1753)	−0.0035 (0.1743)
	30	0.0473 (0.2436)	0.0036 (0.2290)	0.0008 (0.2284)	0.0455 (0.1511)	0.0019 (0.1381)	−0.0010 (0.1377)
	40	0.0341 (0.2078)	0.0018 (0.1986)	0.0003 (0.1985)	0.0332 (0.1270)	0.0009 (0.1187)	−0.0006 (0.1186)
	50	0.0279 (0.1825)	0.0022 (0.1759)	0.0012 (0.1759)	0.0264 (0.1107)	0.0007 (0.1048)	−0.0003 (0.1048)

Table 3. Estimated bias (root mean-squared error), $\delta = 1.5$.

		Estimator of δ			Estimator of γ		
γ	n	MLE	BCE	PBE	MLE	BCE	PBE
−1.0	10	−0.1586 (0.5244)	−0.0137 (0.4376)	0.0192 (0.4241)	0.0776 (0.1751)	0.0054 (0.1375)	−0.0109 (0.1334)
	20	−0.0765 (0.3162)	−0.0093 (0.2877)	−0.0024 (0.2859)	0.1040 (0.2952)	0.0038 (0.2590)	−0.0065 (0.2575)
	30	−0.0482 (0.2421)	−0.0045 (0.2274)	−0.0017 (0.2267)	0.0662 (0.2234)	0.0010 (0.2045)	−0.0033 (0.2040)
	40	−0.0356 (0.2046)	−0.0033 (0.1952)	−0.0016 (0.1950)	0.0491 (0.1881)	0.0007 (0.1759)	−0.0016 (0.1758)
	50	−0.0291 (0.1809)	−0.0034 (0.1741)	−0.0023 (0.1740)	0.0393 (0.1645)	0.0008 (0.1557)	−0.0006 (0.1557)
−0.3	10	−0.0461 (0.3943)	−0.0029 (0.3426)	0.0069 (0.3320)	0.2263 (0.5182)	0.0106 (0.4080)	−0.0384 (0.3974)
	20	−0.0192 (0.2507)	0.0008 (0.2343)	0.0027 (0.2329)	0.1013 (0.2956)	0.0012 (0.2604)	−0.0090 (0.2588)
	30	−0.0119 (0.1979)	0.0011 (0.1894)	0.0020 (0.1889)	0.0624 (0.2222)	−0.0027 (0.2044)	−0.0068 (0.2040)
	40	−0.0088 (0.1688)	0.0008 (0.1633)	0.0013 (0.1631)	0.0457 (0.1860)	−0.0026 (0.1747)	−0.0050 (0.1746)
	50	−0.0082 (0.1491)	−0.0005 (0.1452)	−0.0001 (0.1451)	0.0362 (0.1637)	−0.0022 (0.1557)	−0.0036 (0.1557)
0.0	10	0.0020 (0.3834)	0.0018 (0.3355)	0.0019 (0.3251)	0.2335 (0.5208)	0.0168 (0.4077)	−0.0324 (0.3958)
	20	0.0023 (0.2442)	0.0021 (0.2290)	0.0022 (0.2275)	0.1047 (0.2962)	0.0044 (0.2598)	−0.0058 (0.2583)
	30	0.0029 (0.1945)	0.0027 (0.1864)	0.0027 (0.1860)	0.0693 (0.2293)	0.0039 (0.2095)	−0.0004 (0.2089)
	40	0.0038 (0.1647)	0.0037 (0.1595)	0.0037 (0.1593)	0.0514 (0.1905)	0.0030 (0.1777)	0.0007 (0.1776)
	50	0.0034 (0.1463)	0.0033 (0.1426)	0.0033 (0.1426)	0.0409 (0.1674)	0.0024 (0.1583)	0.0009 (0.1581)
0.3	10	0.0425 (0.3919)	−0.0003 (0.3408)	−0.0099 (0.3301)	0.2262 (0.5188)	0.0105 (0.4087)	−0.0387 (0.3974)
	20	0.0168 (0.2461)	−0.0030 (0.2302)	−0.0049 (0.2286)	0.0989 (0.2937)	−0.0010 (0.2592)	−0.0111 (0.2579)
	30	0.0117 (0.1969)	−0.0013 (0.1884)	−0.0021 (0.1880)	0.0651 (0.2230)	−0.0002 (0.2044)	−0.0044 (0.2039)
	40	0.0090 (0.1685)	−0.0006 (0.1630)	−0.0011 (0.1628)	0.0472 (0.1857)	−0.0012 (0.1740)	−0.0035 (0.1738)
	50	0.0071 (0.1503)	−0.0006 (0.1463)	−0.0008 (0.1462)	0.0377 (0.1620)	−0.0008 (0.1536)	−0.0022 (0.1536)
1.0	10	0.1553 (0.5115)	0.0109 (0.4266)	−0.0219 (0.4134)	0.2274 (0.5151)	0.0114 (0.4046)	−0.0378 (0.3933)
	20	0.0728 (0.3118)	0.0058 (0.2843)	−0.0010 (0.2825)	0.1052 (0.2996)	0.0049 (0.2630)	−0.0052 (0.2614)
	30	0.0473 (0.2436)	0.0036 (0.2290)	0.0008 (0.2284)	0.0682 (0.2266)	0.0029 (0.2072)	−0.0015 (0.2066)
	40	0.0341 (0.2078)	0.0018 (0.1986)	0.0003 (0.1985)	0.0498 (0.1904)	0.0014 (0.1781)	−0.0010 (0.1779)
	50	0.0279 (0.1825)	0.0022 (0.1759)	0.0012 (0.1759)	0.0396 (0.1660)	0.0011 (0.1572)	−0.0004 (0.1572)

Table 4. Estimated bias (root mean-squared error), $\delta = 2.0$.

		Estimator of δ			Estimator of γ		
γ	n	MLE	BCE	PBE	MLE	BCE	PBE
−1.0	10	−0.1586 (0.5244)	−0.0137 (0.4376)	0.0192 (0.4241)	0.0776 (0.1751)	0.0054 (0.1375)	−0.0109 (0.1334)
	20	−0.0765 (0.3162)	−0.0093 (0.2877)	−0.0024 (0.2859)	0.1387 (0.3936)	0.0051 (0.3453)	−0.0086 (0.3433)
	30	−0.0482 (0.2421)	−0.0045 (0.2274)	−0.0017 (0.2267)	0.0883 (0.2979)	0.0013 (0.2727)	−0.0044 (0.2720)
	40	−0.0356 (0.2046)	−0.0033 (0.1952)	−0.0016 (0.1950)	0.0655 (0.2508)	0.0010 (0.2346)	−0.0022 (0.2344)
	50	−0.0291 (0.1809)	−0.0034 (0.1741)	−0.0023 (0.1740)	0.0524 (0.2193)	0.0011 (0.2076)	−0.0009 (0.2076)
−0.3	10	−0.0461 (0.3943)	−0.0029 (0.3426)	0.0069 (0.3320)	0.3018 (0.6909)	0.0141 (0.5440)	−0.0513 (0.5298)
	20	−0.0192 (0.2507)	0.0008 (0.2343)	0.0027 (0.2329)	0.1350 (0.3942)	0.0016 (0.3472)	−0.0120 (0.3451)
	30	−0.0119 (0.1979)	0.0011 (0.1894)	0.0020 (0.1889)	0.0832 (0.2963)	−0.0036 (0.2725)	−0.0091 (0.2721)
	40	−0.0088 (0.1688)	0.0008 (0.1633)	0.0013 (0.1631)	0.0609 (0.2480)	−0.0035 (0.2329)	−0.0066 (0.2327)
	50	−0.0082 (0.1491)	−0.0005 (0.1452)	−0.0001 (0.1451)	0.0483 (0.2183)	−0.0029 (0.2076)	−0.0048 (0.2076)
0.0	10	0.0020 (0.3834)	0.0018 (0.3355)	0.0019 (0.3251)	0.3113 (0.6944)	0.0224 (0.5436)	−0.0432 (0.5277)
	20	0.0023 (0.2442)	0.0021 (0.2290)	0.0022 (0.2275)	0.1397 (0.3950)	0.0059 (0.3464)	−0.0077 (0.3444)
	30	0.0029 (0.1945)	0.0027 (0.1864)	0.0027 (0.1860)	0.0924 (0.3057)	0.0052 (0.2793)	−0.0005 (0.2785)
	40	0.0038 (0.1647)	0.0037 (0.1595)	0.0037 (0.1593)	0.0686 (0.2540)	0.0039 (0.2370)	0.0009 (0.2368)
	50	0.0034 (0.1463)	0.0033 (0.1426)	0.0033 (0.1426)	0.0545 (0.2232)	0.0032 (0.2110)	0.0011 (0.2108)
0.3	10	0.0425 (0.3919)	−0.0003 (0.3408)	−0.0099 (0.3301)	0.3016 (0.6917)	0.0139 (0.5449)	−0.0516 (0.5299)
	20	0.0168 (0.2461)	−0.0030 (0.2302)	−0.0049 (0.2286)	0.1319 (0.3915)	−0.0014 (0.3456)	−0.0148 (0.3439)
	30	0.0117 (0.1969)	−0.0013 (0.1884)	−0.0021 (0.1880)	0.0867 (0.2973)	−0.0002 (0.2726)	−0.0059 (0.2719)
	40	0.0090 (0.1685)	−0.0006 (0.1630)	−0.0011 (0.1628)	0.0629 (0.2476)	−0.0016 (0.2320)	−0.0047 (0.2318)
	50	0.0071 (0.1503)	−0.0006 (0.1463)	−0.0008 (0.1462)	0.0502 (0.2159)	−0.0010 (0.2048)	−0.0029 (0.2047)
1.0	10	0.1553 (0.5115)	0.0109 (0.4266)	−0.0219 (0.4134)	0.3032 (0.6868)	0.0153 (0.5394)	−0.0504 (0.5245)
	20	0.0728 (0.3118)	0.0058 (0.2843)	−0.0010 (0.2825)	0.1403 (0.3994)	0.0065 (0.3507)	−0.0069 (0.3485)
	30	0.0473 (0.2436)	0.0036 (0.2290)	0.0008 (0.2284)	0.0909 (0.3022)	0.0038 (0.2762)	−0.0019 (0.2755)
	40	0.0341 (0.2078)	0.0018 (0.1986)	0.0003 (0.1985)	0.0664 (0.2539)	0.0018 (0.2374)	−0.0013 (0.2373)
	50	0.0279 (0.1825)	0.0022 (0.1759)	0.0012 (0.1759)	0.0527 (0.2214)	0.0014 (0.2096)	−0.0005 (0.2096)

Table 5. Integrated bias squared norm.

	Estimator for γ			Estimator for δ		
n	MLE	BCE	PBE	MLE	BCE	PBE
10	0.1032	0.0080	0.0141	0.2093	0.0122	0.0330
20	0.0486	0.0052	0.0029	0.0939	0.0032	0.0071
30	0.0311	0.0030	0.0020	0.0605	0.0023	0.0036
40	0.0228	0.0024	0.0020	0.0444	0.0018	0.0026
50	0.0187	0.0022	0.0019	0.0354	0.0015	0.0018

Furthermore, the root mean squared error of the corrected estimates is smaller than the uncorrected estimates. Hence, it is notorious that the analytical corrections and Bootstrap estimator also accomplish a reduction in mean squared error.

In order to evaluate the overall performance of each of the three different estimators, regarding the bias and root mean squared error, we adopted two measures introduced by Cribari-Neto and Vasconcellos (2002), also considered in Lemonte (2011). These measures, shown in Tables 5 and 6, are the integrated bias squared ($IBSQ$) and the average root mean-squared error ($ARMSE$) are calculated for each value of n as follows:

$$IBSQ_{(n)} = \sqrt{\frac{1}{20}\sum_{h=1}^{20}(r_{h,n})^2} \quad \text{and} \quad ARMSE_{(n)} = \frac{1}{20}\sum_{h=1}^{20} RMSE_{h,n}$$

where $r_{h,n}$ and $RMSE_{h,n}$ are the estimated bias and estimated root mean squared error for the h-th scenario, $h = 1, \ldots, 20$. Overall, the results indicate that both, BCE and PBE estimators, outperforms the MLE. Among the BCE and PBE methods the differences in these measures are negligible.

Table 6. Average root mean-squared error.

	Estimator for γ			Estimator for δ		
n	MLE	BCE	PBE	MLE	BCE	PBE
10	0.4456	0.3794	0.3676	0.4744	0.3727	0.3623
20	0.2758	0.2545	0.2529	0.2703	0.2376	0.2362
30	0.2162	0.2051	0.2045	0.2053	0.1880	0.1876
40	0.1839	0.1768	0.1766	0.1718	0.1607	0.1606
50	0.1626	0.1575	0.1575	0.1504	0.1425	0.1425

Table 7. MLEs and bias-corrected MLEs (Bootstrap standard-error) for Johnson S_B distribution.

Estimators	γ	δ
MLE	−1.4908 (0.3718)	1.4424 (0.2710)
BCE	−1.3976 (0.3613)	1.3522 (0.2533)
PBE	−1.3932 (0.3568)	1.3443 (0.2492)

Finally, based on all the simulation results, it is clear that the second-order bias reduction can be quite efficient in bringing the corrected estimates closer to their true values. Although the correction methods are equally efficient, the BCE is easier to apply.

5. Data analysis

Here, we exemplify the performance of the proposed bias-corrected estimators for the Johnson S_B distribution parameters, analyzing a real data set from the literature. We shall consider the data set corresponding to the monthly water capacity data from the Shasta reservoir in California, USA, and were taken for the month of February from 1991 to 2010, http://cdec.water.ca.gov/reservoir.html (Nadar, Papadopoulos, and Kızılaslan 2013). The data have 20 observations and were used by Nadar, Papadopoulos, and Kızılaslan (2013) and Wang, Wang, and Yu (2017) to fit the Kumaraswamy distribution. It should be mentioned that the authors transformed the data to the interval 0, 1.

The point estimates and the corresponding Bootstrap standard errors (Efron and Tibshirani 1986) in brackets are reported in Table 7. Note that the corrected estimates for γ is greater than the MLE, while the corrected estimates for δ are smaller than the MLE. An additional Monte Carlo study taking $n = 20$, $\gamma = -1.4908$ and $\delta = 1.4424$ as the true values for γ and δ suggests that the maximum likelihood estimates is underestimating γ and overestimating δ. It is also interesting to note that the bias-corrected MLEs provide lower standard errors, which means more accurate estimates than the MLE. Standard errors were obtained by parametric Bootstrap as suggest by Efron and Tibshirani (1986).

To make a comparison, we also considered the Beta and Kumaraswamy distributions, since they are widely used to modeling data on the unit interval. Their corresponding probability density functions are, respectively, written as

$$f(x|\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\ \Gamma(\beta)}\ x^{\alpha-1}\ (1-x)^{\beta-1} \quad \text{and} \quad f(x|\alpha,\beta) = \alpha\ \beta\ x^{\alpha-1}\ (1-x^{\alpha})^{\beta-1}$$

where $\alpha>0$ and $\beta>0$ are shape parameters.

Table 8. MLEs and bias-corrected MLEs (Bootstrap standard-error) for Beta and Kumaraswamy.

	Beta		Kumaraswamy	
Estimators	A	β	α	β
MLE	7.3157 (3.1147)	2.9099 (1.1781)	6.3476 (1.5868)	4.4894 (3.5662)
BCE	6.2191 (2.7972)	2.5032 (1.0716)	5.8336 (1.5296)	3.4762 (2.4866)
PBE	6.0055 (2.6130)	2.4329 (0.9956)	5.7712 (1.5053)	3.0718 (1.9123)

Table 9. Likelihood-based statistics and goodness-of-fit measures evaluated at the analytical bias-corrected MLEs.

	Distribution		
Statistics	Johnson S_B	Beta	Kumaraswamy
AIC	−23.4840	−20.8640	−22.5436
AIC$_c$	−22.7781	−20.1581	−21.8378
CAIC	−20.4925	−17.8725	−19.5522
HQIC	−23.0952	−20.4752	−22.1549
KS	0.2144 (0.2751)	0.2365 (0.1814)	0.1901 (0.4132)
CvM	0.2324 (0.2131)	0.2843 (0.1496)	0.1949 (0.2788)
AD	1.3220 (0.2251)	1.5434 (0.1667)	1.3092 (0.2292)

The maximum likelihood estimates were obtained by optim function R Core Team (2017). Moreover, we also computed the bias estimates along with the Bootstrap standard errors using the coxsnell.bc function available in the mle.tools package (Mazucheli 2017) of the @@@R environment (see the @@@R code in appendix). The results are given in Table 8.

Table 9 shows the values for likelihood-based statistics (Akaike's Information Criterion (AIC), corrected Akaike's Information Criterion (AIC_c), consistent Akaike's Information Criterion ($CAIC$) and Hannan–Quinn Information Criterion ($HQIC$)) and for goodness-of-fit measures (Kolmogorov–Smirnov statistic (KS), Anderson–Darling statistic (AD) and Cramér–von Mises statistic (CvM)) evaluated at the analytical bias-corrected MLEs. The best model is the one which provides the minimum values of those criteria.

Based on the results from Table 9 we can conclude that the data may have been modeled by the three distributions, since we cannot reject the null hypothesis of the goodness-of-fit tests. Nevertheless, it should be point that the Johnson S_B distribution fits the current data better than the others distributions, since it has the lowest values of AIC, AIC_c, $CAIC$ and $HQIC$ as well as the values of the goodness-of-fit measures.

6. Conclusion

In this article, we derived closed-form expressions for the second-order biases of the MLEs of the parameter which indexes the Johnson S_B distribution with is supported on the unit interval. The biases of the proposed estimators are of order $\mathcal{O}(n^{-2})$, while for the MLEs they are $\mathcal{O}(n^{-1})$ (Cordeiro and Cribari-Neto 2014). Thus, the biases of the newly proposed estimators converge to zero faster than those of the MLEs. In addition, we also considered an alternative bias-correction using the parametric Bootstrap. The numerical results showed that the bias-correcting schemes are generally effective, even when the sample size is small. Therefore, we strongly recommended that the corrected estimators proposed in this article should be used instead of the MLEs, since they are quite efficient in bringing the corrected estimates closer to their true values. Finally, it should be pointed out that the analytical bias estimators have a great advantage on the Bootstrap estimators, once they do not require data resampling, being available in closed form.

Acknowledgments

The authors thank the Editor, the Associate Editor, and the referee for careful reading and comments which greatly improved the article.

ORCID

Josmar Mazucheli http://orcid.org/0000-0001-6740-0445

References

Bacon-Shone, J. 1985. Fitting five parameter Johnson S_B curves by moments. *Journal of the Royal Statistical Society. Series C* 34 (1):95–100.
Barndorff-Nielsen, O., and B. Jørgensen. 1991. Some parametric models on the simplex. *Journal of Multivariate Analysis* 39 (1):106–16.
Bukac, J. 1972. Fitting S_B curves using symmetrical percentile points. *Biometrika* 59 (3):688–90.
Bury, K. 1999. *Statistical distributions in engineering*. Cambridge: Cambridge University Press.
Castellares, F., and A. J. Lemonte. 2016. On the Marshall–Olkin extended distributions. *Communications in Statistics—Theory and Methods* 45 (15):4537–55.
Condino, F., and F. Domma. 2016. A new distribution function with bounded support: the reflected generalized Topp–Leone Power series distribution. *METRON* 75:1–18.
Cook, D. O., R. Kieschnick, and B. McCullough. 2008. Regression analysis of proportions in finance with self selection. *Journal of Empirical Finance* 15 (5):860–7.
Cordeiro, G. M., and F. Cribari-Neto. 2014. *An introduction to Bartlett correction and bias reduction. Briefs in statistics*. New York: Springer.
Cordeiro, G. M., and A. J. Lemonte. 2012. The McDonald arcsine distribution: A new model to proportional data. *Statistics* 48 (1):182–99.
Cordeiro, G. M., A. J. Lemonte, and A. K. Campelo. 2016. Extended arcsine distribution to proportional data: Properties and applications. *Studia Scientiarum Mathematicarum Hungarica* 53 (4):440–66.
Cordeiro, G. M., E. C. D. Rocha, J. G. C. D. Rocha, and F. Cribari-Neto. 1997. Bias-corrected maximum likelihood estimation for the beta distribution. *Journal of Statistical Computation and Simulation* 58 (1):21–35.
Cox, D. R., and E. J. Snell. 1968. A general definition of residuals. *Journal of the Royal Statistical Society, Series B* 30 (2):248–75.
Cribari-Neto, F., and K. L. P. Vasconcellos. 2002. Nearly unbiased maximum likelihood estimation for the beta distribution. *Journal of Statistical Computation and Simulation* 72 (2):107–18.
Davison, A. C., and D. V. Hinkley. 1997. *Bootstrap methods and their applications*. Cambridge: Cambridge University Press, New York.
Doornik, J. A. 2007. *Object-oriented matrix programming using ox*, 3rd ed. London: Timberlake Consultants Press and Oxford.
Edwards, A. W. F. 1992. *Likelihood (expanded edition)*. Baltimore: Johns Hopkins University Press.
Efron, B., 1982. The Jackknife, the bootstrap and other resampling plans. Vol. 38. SIAM.

Efron, B., and R. Tibshirani. 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science* 1 (1):54–75.

Efron, B., and R. J. Tibshirani. 1993. *An introduction to the bootstrap*. Vol. 57 of monographs on statistics and applied probability. New York: Chapman and Hall.

Ferrari, S. L., and F. Cribari-Neto. 1998. On bootstrap and analytical bias corrections. *Economics Letters* 58 (1):7–15.

George, F., 2007. Johnson's system of distributions and microarray data analysis. Ph.D. thesis, University of South Florida.

Giles, D. E. 2012a. Bias reduction for the maximum likelihood estimators of the parameters in the half-Logistic Distribution. *Communication in Statistics—Theory and Methods* 41 (2):212–22.

Giles, D. E., 2012b. A note on improved estimation for the Topp–Leone distribution. Tech. rep., Department of Economics, University of Victoria.

Giles, D. E., and Feng, H., 2009. Bias of the maximum likelihood estimators of the two-parameter Gamma distribution revisited. Tech. Rep., Department of Economics, University of Victoria.

Giles, D. E., H. Feng, and R. T. Godwin. 2013. On the bias of the maximum likelihood estimator for the two-parameter Lomax distribution. *Communications in Statistics—Theory and Methods* 42 (11):1934–50.

Gómez-Déniz, E., M. A. Sordo, and E. Calderín-Ojeda. 2013. The log-Lindley distribution as an alternative to the beta regression model with applications in insurance. *Insurance: Mathematics and Economics* 54:49–57.

Grassia, A. 1977. Family of distributions with argument between 0 and 1 obtained by transformation of the gamma distribution and derived compound distributions. *Australian Journal of Statistics* 19 (2):108–14.

Gupta, A. K., and S. Nadarajah. 2004. *Handbook of beta distribution and its applications*. Boca Raton: CRC Press.

Hill, I. D., R. Hill, and R. L. Holder. 1976. Fitting Johnson curves by moments. *Journal of the royal statistical society. Series C (Applied Statistics)* 25 (2):180–9.

Jiang, R. 2013. A new bathtub curve model with a finite support. *Reliability Engineering & System Safety* 119:44–51.

Johnson, N. L. 1949. Systems of frequency curves generated by methods of translation. *Biometrika* 36 (Pt. 1–2):149–76.

Johnson, N. L. 1955. Systems of frequency curves derived from the first law of Laplace. *Trabajos De Estadistica* 5 (3):283–91.

Johnson, N. L., S. Kotz, and N. Balakrishnan. 1995. *Continuous univariate distributions*. vol. 2. 3rd ed. New York: John Wiley & Sons Inc.

Kottegoda, N. T. 1987. Fitting Johnson S_B curve by the method of maximum likelihood to annual maximum daily rainfalls. *Water Resources Research* 23 (4):728–32.

Kotz, S., and J. R. van Dorp. 2004. *Beyond beta: Other continuous families of distributions with bounded support and applications*. Singapore: World Scientific.

Kumaraswamy, P. 1980. A generalized probability density function for double-bounded random processes. *Journal of Hydrology* 46 (1–2):79–88.

Lagos-Àlvarez, B., M. D. Jiménez-Gamero, and V. Alba-Fernández. 2011. Bias correction in the type I generalized logistic distribution. *Communications in Statistics—Simulation and Computation* 40 (4):511–31.

Lehmann, E. 1999. *Elements of large-sample theory*. New York: Springer-Verlag.

Lemonte, A. J. 2011. Improved point estimation for the Kumaraswamy distribution. *Journal of Statistical Computation and Simulation* 81 (12):1971–82.

Lemonte, A. J., W. Barreto-Souza, and G. M. Cordeiro. 2013. The exponentiated Kumaraswamy distribution and its log-transform. *Brazilian Journal of Probability and Statistics* 27 (1):31–53.

Lemonte, A. J., F. Cribari-Neto, and K. L. Vasconcellos. 2007. Improved statistical inference for the two-parameter Birnbaum–Saunders distribution. *Computational Statistics & Data Analysis* 51 (9):4656–81.

Ling, X., and D. E. Giles. 2014. Bias reduction for the maximum likelihood estimator of the parameters of the generalized Rayleigh family of distributions. *Communications in Statistics—Theory and Methods* 43 (8):1778–92.
Mage, D. T. 1980. An explicit solution for S_B parameters using four percentile points. *Technometrics* 22 (2):247–51.
Mazucheli, J., 2017. mle.tools: expected/observed fisher information and bias-corrected maximum likelihood estimate(s). R package version 1.0.0.
Mazucheli, J., and S. Dey. 2018. Bias-corrected maximum likelihood estimation of the parameters of the generalized half-normal distribution. *Journal of Statistical Computation and Simulation* 88 (6):1027–38.
Mazucheli, J., A. F. B. Menezes, and S. Dey. 2018a. Bias-corrected maximum likelihood estimators of the parameters of the inverse Weibull distribution. *Communications in Statistics—Simulation and Computation*.
Mazucheli, J., A. F. B. Menezes, and S. Dey. 2018b. Improved maximum likelihood estimators for the parameters of the unit-Gamma distribution. *Communications in Statistics – Theory and Methods* 47 (15):3767–78.
Mazucheli, J., A. F. B. Menezes, and S. Nadarajah. 2017. mle.tools: An R package for maximum likelihood bias correction. *The R Journal* 9 (2):268–90. https://journal.r-project.org/archive/2017/RJ-2017-055/index.html.
McDonald, J. B. 1984. Some generalized functions for the size distribution of income. *Econometrica* 52 (3):647–63.
Millar, R. B. 2011. *Maximum likelihood estimation and inference*. West Sussex, UK: John Wiley & Sons Ltd.
Nadar, M., A. Papadopoulos, and Kızılaslan K. 2013. Statistical analysis for Kumaraswamy's distribution based on record data. *Statistical Papers* 54 (2):355–69.
Papke, L. E., and J. M. Wooldridge. 1996. Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics* 11 (6):619–32.
Pawitan, Y. 2001. *In all likelihood: Statistical modelling and inference using likelihood*. Oxford: Oxford University Press.
Pourdarvish, A., S. M. T. K. Mirmostafaee, and K. Naderi. 2015. The exponentiated Topp–Leone distribution: Properties and application. *Journal of Applied Environmental and Biological Sciences* 5 (7):251–6.
R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.
Reath, J., J. Dong, and M. Wang. 2018. Improved parameter estimation of the log-Logistic distribution with applications. *Computational Statistics* 33 (1):339–56.
Saha, K., and S. Paul. 2005. Bias-corrected maximum likelihood estimator of the negative binomial dispersion parameter. *Biometrics* 61 (1):179–85.
Schwartz, J., and D. E. Giles. 2016. Bias-reduced maximum likelihood estimation of the zero-inflated Poisson distribution. *Communications in Statistics—Theory and Methods* 45 (2):465–78.
Schwartz, J., R. T. Godwin, and D. E. Giles. 2013. Improved maximum-likelihood estimation of the shape parameter in the Nakagami distribution. *Journal of Statistical Computation and Simulation* 83 (3):434–45.
Sharma, D., and T. K. Chakrabarty. 2016. On size biased Kumaraswamy distribution. *Statistics, Optimization & Information Computing* 4 (3):252–64.
Shuaib, M. K., K. Robert, and H. I. Lena. 2016. Transmuted Kumaraswamy distribution. *Statistics in Transition. New Series* 17 (2):183–210.
Siekierski, K. 1992. Comparison and evaluation of three methods of estimation of the Johnson S_B distribution. *Biometrical Journal* 34 (7):879–95.
Singh, A. K., A. Singh, and D. J. Murphy. 2015. On bias corrected estimators of the two parameter Gamma distribution. Paper presented at the 12th International Conference on Information Technology—New Generations (ITNG), pp. 127–32.

Slifker, J. F., and S. S. Shapiro. 1980. The Johnson system: Selection and parameter estimation. *Technometrics* 22 (2):239–46.

Tadikamalla, P. R., and N. L. Johnson. 1982. Systems of frequency curves generated by transformations of logistic variables. *Biometrika* 69 (2):461–5.

Teimouri, M., and S. Nadarajah. 2013. Bias corrected MLEs for the Weibull distribution based on records. *Statistical Methodology* 13:12–24.

Teimouri, M., and S. Nadarajah. 2016. Bias corrected MLEs under progressive type-II censoring scheme. *Journal of Statistical Computation and Simulation* 86 (14):2714–26.

Topp, C. W., and F. C. Leone. 1955. A family of J-shaped frequency functions. *Journal of the American Statistical Association* 50 (269):209–19.

van Drop, J., and S. Kotz. 2006. Modeling income distributions using elevated distributions on a bounded domain. In *Distribution models theory*, 1–25. London: World Scientific.

Wang, B. X., X. K. Wang, and K. Yu. 2017. Inference on the Kumaraswamy distribution. *Communications in Statistics—Theory and Methods* 46 (5):2079–90.

Wang, M., and W. Wang. 2017. Bias-corrected maximum likelihood estimation of the parameters of the weighted Lindley distribution. *Communications in Statistics—Theory and Methods* 46 (1):530–45.

Wheeler, R. E. 1980. Quantile estimators of Johnson curve parameters. *Biometrika* 67 (3):725–8.

Zhang, G., and R. Liu. 2015. Bias-corrected estimators of scalar skew normal. *Communications in Statistics—Simulation and Computation* 46 (2):831–9.

Zhou, B., and J. P. McTague. 1996. Comparison and evaluation of five methods of estimation of the Johnson system parameters. *Canadian Journal of Forest Research* 26 (6):928–35.

Appendix

In this appendix we present the R code (R Core Team 2017) in order to check the validity of the analytic expressions using the coxsnell.bc function, available in mle.tools package (Mazucheli 2017). For the coxsnell.bc function we had to provide an @@@R expression with the probability density function and its logarithm, the sample size, the parameter names, the maximum likelihood estimates and the support of the distribution. In line 30, below, are the evaluated bias corrected estimates presented in Table 7.

```
1  library(mle.tools)
2
3  loglike.jsb < function(mle)
4  {
5    n < length(x)
6    gamma < mle[1]; delta < mle[2]
7    n * log(delta) − 0.5 * sum((gamma + delta * log(x /
         (1 − x)))∘2) − sum(log((sqrt(2 * pi) * x * (1 − x))))
8  }
9
10 x < c(0.3389, 0.7680, 0.4319, 0.8435, 0.7599, 0.7874, 0.7246,
       0.8499, 0.7576, 0.696, 0.8116, 0.8423, 0.7853, 0.8287,
       0.7837, 0.5802, 0.8156, 0.4307, 0.8474, 0.7426)
11
12 mle < optim(par = c(1,1), fn = loglike.jsb, method = "BFGS",
   hessian = T, control = list(fnscale = −1))$par
13
14 pdf < quote (delta / (sqrt (2 * pi ) * x * (1 - x ) ) * exp (-0.5 * (gam-
       ma + delta * log ( x / (1- x)))∘2)
15
```

```
lpdf < quote (log (delta) * 0.5 * (gamma+delta * log ( x / (1 -
        x )))∘2)

coxsnell . bc ( density=pdf , logdensity=lpdf, n=20, parms=c
       ("gamma" , " delta"), mle=c (-1.4908 , 1.4424), lower =0 ,
       upper =1)

$mle
  gamma    delta
-1.4910   1.4420

$varcov
         gamma      delta
gamma   0.10556  -0.05376
 delta -0.05376   0.0501

$mle.bc
 gamma    delta
 -1.3976   1.3522

$varcov.bc
         gamma       delta
gamma   0.09883   -0.04725
 delta  -0.04725   0.04571

$bias
  gamma       delta
 -0.09317     0.09015
```