mle.tools: An R Package for Maximum Likelihood Bias Correction

by Josmar Mazucheli, André Felipe B. Menezes and Saralees Nadarajah

Abstract Recently, Mazucheli (2017) uploaded the package mle.tools to CRAN. It can be used for bias corrections of maximum likelihood estimates through the methodology proposed by Cox and Snell (1968). The main function of the package, coxsnell.bc(), computes the bias corrected maximum likelihood estimates. Although in general, the bias corrected estimators may be expected to have better sampling properties than the uncorrected estimators, analytical expressions from the formula proposed by Cox and Snell (1968) are either tedious or impossible to obtain. The purpose of this paper is twofolded: to introduce the mle.tools package, especially the coxsnell.bc() function; secondly, to compare, for thirty one continuous distributions, the bias estimates from the coxsnell.bc() function and the bias estimates from analytical expressions available in the literature. We also compare, for five distributions, the observed and expected Fisher information. Our numerical experiments show that the functions are efficient to estimate the biases by the Cox-Snell formula and for calculating the observed and expected Fisher information.

Introduction

Since it was proposed by Fisher in a series of papers from 1912 to 1934, the maximum likelihood method for parameter estimation has been employed to several issues in statistical inference, because of its many appealing properties. For instance, the maximum likelihood estimators, hereafter referred to as MLEs, are asymptotically unbiased, efficient, consistent, invariant under parameter transformation and asymptotically normally distributed (Edwards, 1992; Lehmann, 1999). Most properties that make the MLEs attractive depend on the sample size, hence such properties as unbiasedness, may not be valid for small samples or even moderate samples (Kay, 1995). Indeed, the maximum likelihood method produces biased estimators, i.e., expected values of MLEs differ from the true parameter values providing systematic errors. In particular, these estimators typically have biases of order $O(n^{-1})$, thus these errors reduce as sample size increases (Cordeiro and Cribari-Neto, 2014).


In general, the Cox-Snell methodology is efficient for bias corrections. However, obtaining analytical expressions for some probability distributions, mainly for those indexed by more than two parameters, can be notoriously cumbersome or impossible. Stošič and Cordeiro (2009) presented Maple and Mathematica scripts that may be used to calculate closed form analytic expressions for bias corrections using the Cox-Snell formula. They tested the scripts for 20 two-parameter continuous probability distributions, and the results were compared with those published in earlier works. In the same direction, researchers from the University of Illinois, at Urbana-Champaign, have developed a Mathematica program, entitled “CSCK MLE Bias Calculation” (Johnson et al., 2012b) that enables the user to calculate the analytic Cox-Snell MLE bias vectors for various probability distributions with up to four unknown parameters. It is important to mention that both, Maple (Maple, 2017) and Mathematica (Wolfram Research, Inc., 2010), are commercial softwares.

In this paper, our objective is to introduce a new contributed R (R Core Team, 2016) package, namely mle.tools that computes the expected/observed Fisher information and the bias corrected estimates by the methodology proposed by Cox and Snell (1968). The theoretical background of the methodology is presented in Section Overview of the Cox-Snell methodology. Details about the mle.tools package are described in Section The mle.tools package details. Closed form solutions of bias corrections are collected from the literature for a large number of distributions and compared to the output from the coxsnell.bc() function, see Section Comparative study. In Section Additional Applications, we compare various estimates of Fisher’s information, considering a real application from the literature. Finally, Section Concluding Remarks contains some concluding remarks and directions for future research.
Overview of the Cox-Snell methodology

Let $X_1, \ldots, X_n$ be $n$ independent random variables with probability density function $f(x_i \mid \theta)$ depending on a $p$-dimensional parameter vector $\theta = (\theta_1, \ldots, \theta_p)$. Without loss of generality, let $l = l(\theta \mid x)$ be the log-likelihood function for the unknown $p$-dimensional parameter vector $\theta$ given a sample of $n$ observations. We shall assume some regularity conditions on the behavior of $l(\theta \mid x)$ (Cox and Hinkley, 1979).

The joint cumulants of the derivatives of $l$ are given by:

$$
\kappa_{ij} = \mathbb{E} \left[ \frac{\partial^2 l}{\partial \theta_i \partial \theta_j} \right],
$$

(1)

$$
\kappa_{ijl} = \mathbb{E} \left[ \frac{\partial^3 l}{\partial \theta_i \partial \theta_j \partial \theta_l} \right],
$$

(2)

$$
\kappa_{ij,l} = \mathbb{E} \left[ \left( \frac{\partial^2 l}{\partial \theta_i \partial \theta_j} \right) \left( \frac{\partial l}{\partial \theta_l} \right) \right],
$$

(3)

$$
\kappa_{ij}^{(l)} = \frac{\partial \kappa_{ij}}{\partial \theta_l}
$$

(4)

for $i, j, l = 1, \ldots, p$.

The bias expression of the $s$th element of $\hat{\theta}$, the MLEs of $\theta$, when the sample data are independent, but not necessarily identically distributed, was proposed by Cox and Snell (1968):

$$
B \left( \hat{\theta}_s \right) = \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{l=1}^{p} x_i^j x_l^{ij} \left[ 0.5 \kappa_{ijl} + \kappa_{ij,l} \right] + \mathcal{O} \left( n^{-2} \right),
$$

(5)

where $s = 1, \ldots, p$ and $x_{ijl}$ is the $(i, j)$th element of the inverse of the negative of the expected Fisher information.

Thereafter, Cordeiro and Klein (1994) noticed that equation (5) holds even if the data are non-independent, and it can be re-expressed as:

$$
B \left( \hat{\theta}_s \right) = \sum_{i=1}^{p} x_i^s \sum_{j=1}^{p} \sum_{l=1}^{p} \left[ \kappa_{ij}^{(l)} - 0.5 \kappa_{ijl} \right] x_l^{ij} + \mathcal{O} \left( n^{-2} \right).
$$

(6)

Defining $a_{ij}^{(l)} = \kappa_{ij}^{(l)} - 0.5 \kappa_{ijl}$, $A^{(l)} = \{ a_{ij}^{(l)} \}$ and $K = [ -x_{ij} ]$, the expected Fisher information matrix for $i, j, l = 1, \ldots, n$, the bias expression for $\hat{\theta}$ in matrix notation is:

$$
B \left( \hat{\theta} \right) = K^{-1} \text{vec} \left( K^{-1} \right) + \mathcal{O} \left( n^{-2} \right),
$$

(7)

where vec$(K^{-1})$ is the vector obtained by stacking the columns of $K^{-1}$ and $A = \{ A^1 | \cdots | A^p \}$.

Finally, the bias corrected MLE for $\hat{\theta}_s$ can be obtained as:

$$
\hat{\theta}_s = \hat{\theta}_s - \tilde{B} \left( \hat{\theta}_s \right).
$$

(8)

Alternatively, using matrix notation the bias corrected MLEs can be expressed as Cordeiro and Klein (1994):

$$
\hat{\theta} = \hat{\theta} - \tilde{K}^{-1} \text{vec} \left( \tilde{K}^{-1} \right),
$$

(9)

where $\tilde{K} = K |_{\theta=\hat{\theta}}$ and $\tilde{A} = A |_{\theta=\hat{\theta}}$.

The mle.tools package details

The current version of the mle.tools package, uploaded to CRAN in February, 2017, has implemented three functions — observed.varcov(), expected.varcov() and coxsnell1.bc() — which are of great interest in data analysis based on MLEs. These functions calculate, respectively, the observed Fisher information, the expected Fisher information and the bias corrected MLEs using the bias formula in (5). The above mentioned functions can be applied to any probability density function whose terms
are available in the derivatives table of the \( \mathcal{D}(\) function (see “deriv.c” source code for further details). Integrals, when required, are computed numerically via the \( \text{integrate()} \) function. Below are some mathematical details of how the returned values from the three functions are calculated.

Let \( X_1, \ldots, X_n \) be independent and identical random variables with probability density function \( f(x_i \mid \theta) \) depending on a \( p \)-dimensional parameter vector \( \theta = (\theta_1, \ldots, \theta_p) \). The \((j,k)\)th element of the observed, \( H_{jk} \), Fisher information are calculated, respectively, as

\[
H_{jk} = \left. -\sum_{i=1}^{n} \frac{\partial^2}{\partial \theta_j \partial \theta_k} \log f(x_i \mid \theta) \right|_{\theta = \hat{\theta}}
\]

and

\[
I_{jk} = -n \times E \left( \frac{\partial^2}{\partial \theta_j \partial \theta_k} \log f(x \mid \theta) \right) = -n \times \int_{X} \frac{\partial^2}{\partial \theta_j \partial \theta_k} \log f(x \mid \theta) \times f(x \mid \theta) \, dx \bigg|_{\theta = \hat{\theta}},
\]

where \( j, k = 1, \ldots, p, \hat{\theta} \) is the MLE of \( \theta \) and \( X \) denotes the support of the random variable \( X \).

The observed \( \text{varcov()} \) function is as follows:

function (logdensity, X, parms, mle)

where \( \text{logdensity} \) is an R expression of the log of the probability density function, \( X \) is a numeric vector containing the observations, \( \text{parms} \) is a character vector of the parameter name(s) specified in the logdensity expression and \( \text{mle} \) is a numeric vector of the parameter estimate(s). This function returns a list with two components (i) \( \text{mle} \): the inputed MLEs and (ii) \( \text{varcov} \): the observed variance-covariance evaluated at the inputed MLE argument. The elements of the Hessian matrix are calculated analytically.

The functions expected \( \text{varcov()} \) and \( \text{coxsnell.bc()} \) have the same arguments and are as follows:

function (density, logdensity, n, parms, mle, lower = "-Inf", upper = "Inf", ...)

where \( \text{density} \) and \( \text{logdensity} \) are R expressions of the probability density function and its logarithm, respectively, \( n \) is a numeric scalar of the sample size, \( \text{parms} \) is a character vector of the parameter name(s) specified in the density and log-density expressions, \( \text{mle} \) is a numeric vector of the parameter estimates, \( \text{lower} \) is the lower integration limit (\(-\text{Inf} \) is the default), \( \text{upper} \) is the upper integration limit (\(\text{Inf} \) is the default) and \( \ldots \) are additional arguments passed to the \( \text{integrate()} \) function. The \( \text{expected.varcov()} \) function returns a list with two components:

\$\text{mle} \) the inputed MLEs and
\$\text{varcov} \) the expected covariance evaluated at the inputed MLEs.

The \( \text{coxsnell.bc()} \) function returns a list with five components:

\$\text{mle} \) the inputed MLEs,
\$\text{varcov} \) the expected variance-covariance evaluated at the inputed MLEs,
\$\text{mle.bc} \) the bias corrected MLEs,
\$\text{varcov.bc} \) the expected variance-covariance evaluated at the bias corrected MLEs
\$\text{bias} \) the bias estimate(s).

Furthermore, the bias corrected MLE of \( \theta_s, s = 1, \ldots, p \) denoted by \( \hat{\theta}_s \) is calculated as \( \hat{\theta}_s = \hat{\theta} - \mathcal{B} \left( \hat{\theta} \right) \), where \( \hat{\theta} \) is the MLE of \( \theta_s \) and

\[
\mathcal{B} \left( \hat{\theta} \right) = \sum_{j=1}^{p} \sum_{k=1}^{p} \chi^{jk} \left[ 0.5 \kappa_{jk} + \kappa_{jk, l} \right] \bigg|_{\theta = \hat{\theta}},
\]

where \( \chi^{jk} \) is the \((j,k)\)th element of the inverse of the negative of the expected Fisher information,

\[
\kappa_{jk, l} = n \int_{X} \frac{\partial^3}{\partial \theta_j \partial \theta_k \partial \theta_l} \log f(x \mid \theta) f(x \mid \theta) \, dx \bigg|_{\theta = \hat{\theta}},
\]

\[
\kappa_{jk} = n \int_{X} \frac{\partial^2}{\partial \theta_j \partial \theta_k} \log f(x \mid \theta) \frac{\partial}{\partial \theta_l} \log f(x \mid \theta) f(x \mid \theta) \, dx \bigg|_{\theta = \hat{\theta}},
\]
and $X'$ denotes the support of the random variable $X$.

It is important to emphasize that first, second and third-order partial log-density derivatives are analytically calculated via the $D()$ function, while integrals are computed numerically, using the $\text{integrate()}$ function. Furthermore, if numerical integration fails and/or the expected/observed information is singular, an error message is returned.

**Comparative study**

In order to evaluate the robustness of the $\text{coxsnell.bc()}$ function, we compare, through real applications, the estimated biases obtained from the package and from the analytical expressions for a total of thirty one continuous probability distributions. The analytical expressions for each distribution, named as $\text{distname.bc()}$, can be found in the supplementary file “analyticalBC.R”. For example, the entry $\text{lindley.bc(n, mle)}$ evaluates the bias estimates locally at $n$ and $mle$ values.

In the sequel, the probability density function, the analytical Cox-Snell expressions and the bias estimates are provided for: Lindley, inverse Lindley, inverse Exponential, Shanker, inverse Shanker, Topp-Leone, Lévy, Rayleigh, inverse Rayleigh, Half-Logistic, Half-Cauchy, Half-Normal, Normal, inverse Gaussian, Log-Normal, Log-Logistic, Gamma, inverse Gamma, Lomax, weighted Lindley, generalized Rayleigh, Weibull, inverse Weibull, generalized Half-Normal, inverse generalized Half-Normal, Marshall-Olkin extended Exponential, Beta, Kumaraswamy, inverse Beta, Birnbaum-Saunders and generalized Pareto distributions.

It is noteworthy that analytical bias corrected expressions are not reported in the literature for the Lindley, Shanker, inverse Shanker, Lévy, inverse Rayleigh, half-Cauchy, inverse Weibull, inverse generalized half-normal and Marshall-Olkin extended exponential distributions.

According to all the results presented below, we observe concordance between the bias estimates given by the $\text{coxsnell.bc()}$ function and the analytical expression(s) for 28 out the 31 distributions. The distributions which did not agree with the $\text{coxsnell.bc()}$ function were the beta, Kumaraswamy and inverse beta distributions. Perhaps there are typos either in our typing or in the analytical expressions reported by Cordeiro et al. (1997), Lemonte (2011) and Stošič and Cordeiro (2009). Having this view, we recalculated the analytical expressions for the biases. For the beta and inverse beta distributions, our recalculated analytical expressions agree with the results returned by the $\text{coxsnell.bc()}$ function, so there are actually typos in the expression of Cordeiro et al. (1997) and Stošič and Cordeiro (2009). For the Kumaraswamy, we could not evaluate the analytical expression given by the author but we compare the results from $\text{coxsnell.bc()}$ function with a numerical evaluation in Maple (Maple, 2017) and the results are exactly equals.

1. One-parameter Lindley distribution with scale parameter $\theta$

$$f(x \mid \theta) = \frac{\theta^2}{1 + \theta} (1 + x) \exp(-\theta x), \quad x > 0.$$  

- Bias expression (not previously reported in the literature):

$$B(\hat{\theta}) = \frac{(\theta^3 + 6 \theta^2 + 6 \theta + 2) (\theta + 1)}{n \left(\theta^2 + 4 \theta + 2\right)^2}.$$  

Using the data set from Ghitany et al. (2008) we have $n = 100, \hat{\theta} = 0.1866$ and $\hat{\theta} = 0.0133$. Evaluating the analytical expression (10) and the $\text{coxsnell.bc()}$ function, we have, respectively,

```
1. Inverse Lindley distribution with scale parameter $\theta$

$$f(x \mid \theta) = \frac{\theta^2}{1 + \theta} \left(\frac{1 + x}{x^3}\right) \exp\left(-\frac{\theta}{x}\right), \quad x > 0.$$
```
• Bias expression (Wang, 2015):

$$B(\hat{\theta}) = \frac{(\theta + 1)\theta (\theta^3 + 6\theta^2 + 6\theta + 2)}{n(\theta^2 + 4\theta + 2)^2}.$$  \hspace{1cm} (11)

Using the data set from Sharma et al. (2015) we have \( n = 58, \hat{\theta} = 60.0016 \) and \( \hat{\text{se}}(\hat{\theta}) = 7.7535 \).

Evaluating the analytical expression (11) and the coxsnell.bc() function, we have, respectively,

```r
invlindley.bc(n = 58, mle = 60.0016)
## theta
## 1.017
def < quote(theta^2 / (theta + 1) * ((1 + x) / x^3) * exp(-theta / x))
lpdf < quote(2* log(theta) - log(1 + theta) - theta / x)
coxsnell.bc(density = pdf, logdensity = lpdf, n = 58,
parms = c("theta"), mle = 60.0016, lower = 0)$bias
## theta
## 1.017
```

3. Inverse exponential distribution with rate parameter \( \theta \)

$$f(x \mid \theta) = \frac{\theta}{x^2} \exp\left(\frac{-\theta}{x}\right), \quad x > 0.$$  

• Bias expression (Johnson et al., 2012b):

$$B(\hat{\theta}) = \frac{\theta}{n}. \hspace{1cm} (12)$$

Using the data set from Lawless (2011), we have \( n = 30, \hat{\theta} = 11.1786 \) and \( \hat{\text{se}}(\hat{\theta}) = 2.0409 \).

Evaluating the analytical expression (12) and the coxsnell.bc() function, we have, respectively,

```r
invel.expbc(n = 30, mle = 11.1786)
## theta
## 0.3726
def < quote(theta / x^2 * exp(- theta / x))
lpdf < quote(log(theta) - theta / x)
coxsnell.bc(density = pdf, logdensity = lpdf, n = 30,
parms = c("theta"), mle = 11.1786, lower = 0)$bias
## theta
## 0.3726
```

4. Shanker distribution with scale parameter \( \theta \)

$$f(x \mid \theta) = \frac{\theta^2}{\theta^2 + 1} (\theta + x) \exp(-\theta x), \quad x > 0.$$  

• For bias expression (not previously reported in the literature, see the “analyticalBC.R” file.

Using the data set from Shanker (2015), we have \( n = 31, \hat{\theta} = 0.0647 \) and \( \hat{\text{se}}(\hat{\theta}) = 0.0082 \).

Evaluating the analytical expression and the coxsnell.bc() function, we have, respectively,

```r
shanker.bc(n = 31, mle = 0.0647)
## theta
## 0.001035
def < quote(theta^2 / (theta^2 + 1) * (theta + x) * exp(-theta * x))
lpdf < quote(2* log(theta) - log(theta^2 + 1) + log(theta + x) -
theta * x)
coxsnell.bc(density = pdf, logdensity = lpdf, n = 31,
parms = c("theta"), mle = 0.0647, lower = 0)$bias
## theta
## 0.001035
```

5. Inverse Shanker distribution with scale parameter \( \theta \)

$$f(x \mid \theta) = \frac{\theta^2}{1+\theta^2} \left(\frac{1+\theta x}{x^3}\right) \exp\left(-\frac{\theta}{x}\right), \quad x > 0.$$  

The R Journal Vol. 9/2, December 2017

ISSN 2073-4859
• Bias expression (not previously reported in the literature):

\[
B (\hat{\theta}) = \frac{\theta^3 + 2 \theta}{n (\theta^2 + 1)},
\]

(13)

Using the data set from Sharma et al. (2015), we have \( n = 58, \hat{\theta} = 59.1412 \) and \( \hat{se}(\hat{\theta}) = 7.7612 \).

Evaluating the analytical expression (13) and the coxsnell.bc() function, we have, respectively,

```r
invshanker.bc(n = 58, mle = 59.1412)
## theta
## 1.02
```

```r
pdf <- quote(theta^2 / (theta^2 + 1) * (theta * x + 1) / x^3 * exp(-theta / x))
lpdf <- quote(log(theta) - 2 * log(x) - theta / x)
coxsnell.bc(density = pdf, logdensity = lpdf, n = 58, 
params = c("theta"), mle = 59.1412, lower = 0)$bias
## theta
## 1.02
```

6. Topp-Leone distribution with shape parameter \( \nu \)

\[
f(x \mid \nu) = 2 \nu (1 - x)^{\nu - 1} (2 - x)^{\nu - 1}, \quad 0 < x < 1.
\]

• Bias expression (Giles, 2012a):

\[
B (\hat{\nu}) = \frac{\nu}{n}.
\]

(14)

Using the data set from Cordeiro and dos Santos Brito (2012), we have \( n = 107, \hat{\nu} = 2.0802 \) and \( \hat{se}(\hat{\nu}) = 0.2011 \).

Evaluating the analytical expression (14) and the coxsnell.bc() function, we have, respectively,

```r
toppleone.bc(n = 107, mle = 2.0802)
## nu
## 0.01944
```

```r
pdf <- quote(2 * nu * x^(nu - 1) * (1 - x) * (2 - x)^(nu - 1))
lpdf <- quote(log(nu) + nu * log(x) + log(1 - x) + (nu - 1) * log(2 - x))
coxsnell.bc(density = pdf, logdensity = lpdf, n = 107, 
params = c("nu"), mle = 2.0802, lower = 0, upper = 1)$bias
## nu
## 0.01944
```

7. One-parameter Lévy distribution with scale parameter \( \sigma \)

\[
f(x \mid \sigma) = \sqrt{\frac{\sigma}{2 \pi}} x^{-\frac{3}{2}} \exp \left( -\frac{\sigma^2}{2x} \right), \quad x > 0.
\]

• Bias expression (not previously reported in the literature):

\[
B (\hat{\sigma}) = \frac{2 \sigma^2}{n}.
\]

(15)

Using the data set from Achcar et al. (2013), we have \( n = 361, \hat{\sigma} = 4.4460 \) and \( \hat{se}(\hat{\sigma}) = 0.3309 \).

Evaluating the analytical expression (15) and the coxsnell.bc() function, we have, respectively,
8. Rayleigh distribution with scale parameter $\sigma$

$$f(x | \sigma) = \frac{x}{\sigma^2} \exp \left( -\frac{x^2}{2\sigma^2} \right), \quad x > 0.$$

- Bias expression (Xiao and Giles, 2014):

$$B(\hat{\sigma}) = -\frac{\sigma}{8n}. \tag{16}$$

Using the data set from Bader and Priest (1982), we have $n = 69$, $\hat{\sigma} = 1.2523$ and $\hat{\sigma} \pm (\hat{\sigma}) = 0.0754$. Evaluating the analytical expression (16) and the coxsnell.bc() function, we have, respectively,

```r
rayleigh.bc(n = 69, mle = 1.2522)
## sigma
## -0.002268
```

```r
density <- quote(x / sigma^2 * exp(- 0.5 * (x / sigma)^2))
lpdf <- quote(- 2 * log(sigma) - 0.5 * x^2 / sigma^2)
coxsnell.bc(density = pdf, logdensity = lpdf, n = 69, 
            parms = c("sigma"), mle = 1.2522, lower = 0)$bias
## sigma
## -0.002268
```

9. Inverse Rayleigh distribution with scale parameter $\sigma$

$$f(x | \sigma) = \frac{2\sigma^2}{x^3} \exp \left( -\frac{\sigma}{x^2} \right), \quad x > 0.$$

- Bias expression (not previously reported in the literature):

$$B(\hat{\sigma}) = \frac{3\sigma}{8n}. \tag{17}$$

Using the data set from Bader and Priest (1982), we have $n = 63$, $\hat{\sigma} = 2.8876$ and $\hat{\sigma} \pm (\hat{\sigma}) = 0.1819$. Evaluating the analytical expression (17) and the coxsnell.bc() function, we have, respectively,

```r
invrayleigh.bc(n = 63, mle = 2.8876)
## sigma
## 0.01719
```

```r
density <- quote(2 * sigma^2 / x^3 * exp(-sigma^2 / x^2))
lpdf <- quote(2 * log(sigma) - sigma^2 / x^2)
coxsnell.bc(density = pdf, logdensity = lpdf, n = 63, 
            parms = c("sigma"), mle = 2.8876, lower = 0)$bias
## sigma
## 0.01719
```

10. Half-logistic distribution with scale parameter $\sigma$

$$f(x | \sigma) = \frac{2 \exp \left( -\frac{x}{\sigma} \right)}{\sigma \left[ 1 + \exp \left( -\frac{x}{\sigma} \right) \right]^2}, \quad x > 0.$$

- Bias expressions (Giles, 2012b):

$$B(\hat{\sigma}) = -\frac{0.05256766607 \sigma}{n}. \tag{18}$$

Using the data set from Bhaumik et al. (2009), we have $n = 34$, $\hat{\sigma} = 1.3926$ and $\hat{\sigma} \pm (\hat{\sigma}) = 0.2056$. Evaluating the analytical expression (17) and the coxsnell.bc() function, we have, respectively,

```r
halflogistic.bc(n = 34, mle = 1.3925)
## sigma
## -0.002153
```

```r
density <- quote((2/sigma) * exp(-x / sigma) / (1 + exp(-x / sigma))^2))
lpdf <- quote(-log(sigma) - x / sigma - 2 * log(1 + exp(-x / sigma)))
coxsnell.bc(density = pdf, logdensity = lpdf, n = 34, 
            parms = c("sigma"), mle = 1.3925, lower = 0)$bias
## sigma
## -0.002153
```
11. Half-Cauchy distribution with scale parameter $\sigma$

$$f(x \mid \sigma) = \frac{2}{\pi} \frac{\sigma}{\sigma^2 + x^2}, \quad x > 0.$$  

- Bias expression (not previously reported in the literature):

$$B(\hat{\sigma}) = -\frac{\sigma}{n}. \quad (19)$$

Using the data set from Alzaatreh et al. (2016), we have $n = 64$, $\hat{\sigma} = 28.3345$ and $\hat{se}(\hat{\sigma}) = 4.4978$.

```r
halfcauchy.bc(n = 64, mle = 28.3345)
## sigma
## 0.4427

pdf <- quote(2 / pi * sigma / (x^2 + sigma^2))
lpdf <- quote(log(sigma) - log(x^2 + sigma^2))
coxsnell.bc(density = pdf, logdensity = lpdf, n = 64, 
            parms = c("sigma"), mle = 28.3345, lower = 0)$bias
## sigma
## 0.4456
```

12. Half-normal distribution with scale parameter $\sigma$

$$f(x \mid \sigma) = \sqrt{\frac{2}{\pi}} \frac{1}{\sigma} \exp \left( -\frac{x^2}{2\sigma^2} \right), \quad x > 0.$$  

- Bias expressions (Xiao and Giles, 2014):

$$B(\hat{\sigma}) = -\frac{\sigma}{4n}. \quad (20)$$

Using the data set from Raqab et al. (2008), we have $n = 69$, $\hat{\sigma} = 1.5323$ and $\hat{se}(\hat{\sigma}) = 0.1304$.

```r
halfnormal.bc(n = 69, mle = 1.5323)
## sigma
## -0.005552

pdf <- quote(sqrt(2) / (sqrt(pi) * sigma) * exp(-x^2 / (2 * sigma^2)))
lpdf <- quote(-log(sigma) - x^2 / sigma^2 / 2)
coxsnell.bc(density = pdf, logdensity = lpdf, n = 69, 
            parms = c("sigma"), mle = 1.5323, lower = 0)$bias
## sigma
## -0.005552
```

13. Normal distribution with mean $\mu$ and standard deviation $\sigma$

$$f(x \mid \mu, \sigma) = \frac{1}{\sqrt{2\pi} \sigma} \exp \left[ -\frac{(x-\mu)^2}{2\sigma^2} \right], \quad x \in (-\infty, \infty).$$  

- Bias expressions (Stošić and Cordeiro, 2009):

$$B(\hat{\mu}) = 0 \quad \text{and} \quad B(\hat{\sigma}) = -\frac{3\sigma}{4n}. \quad (21)$$

Using the data set from Kundu (2005), we have $n = 23$, $\hat{\mu} = 4.1506$, $\hat{\sigma} = 0.5215$, $\hat{se}(\hat{\mu}) = 0.1087$ and $\hat{se}(\hat{\sigma}) = 0.0769$.

```r
normal.bc(n = 23, mle = c(4.1506, 0.5215))
## mu sigma
## 0.00000 -0.01701

pdf <- quote(1 / (sqrt(2 * pi) * sigma) * 
             exp(-0.5 / sigma^2 * (x - mu)^2))
lpdf <- quote(-log(sigma) - 0.5 / sigma^2 * (x - mu)^2)
coxsnell.bc(density = pdf, logdensity = lpdf, n = 23, 
            parms = c("mu", "sigma"), mle = c(4.1506, 0.5215))$bias
## mu sigma
## -4.071e-13 -1.701e-02
```
14. Inverse Gaussian distribution with mean $\mu$ and shape $\lambda$

\[
 f(x \mid \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left[-\frac{\lambda(x - \mu)^2}{2x\mu^2}\right], \quad x > 0.
\]

- Bias expressions (Stošić and Cordeiro, 2009):

\[
 B(\hat{\mu}) = 0 \quad \text{and} \quad B\left(\hat{\lambda}\right) = \frac{3\lambda}{n}. \tag{22}
\]

Using the data set from Chhikara and Folks (1977), we have $n = 46$, $\hat{\mu} = 3.6067$, $\hat{\lambda} = 1.6584$, $\hat{\sigma}(\hat{\mu}) = 0.7843$ and $\hat{\sigma}\left(\hat{\lambda}\right) = 0.3458$. Evaluating the analytical expressions (22) and the coxsnell.bc() function, we have, respectively,

```r
invgaussian.bc(n = 46, mle = c(3.6065, 1.6589))
## mu lambda
## 0.0000 0.1082
```

```r
density <- quote(sqrt(lambda / (2 * pi * x^3)) * exp(-lambda * (x - mu)^2 / (2 * mu^2 * x)))
lpdf <- quote(0.5 * log(lambda) - lambda * (x - mu)^2 / (2 * mu^2 * x))
coxsnell.bc(density = density, logdensity = lpdf, n = 46, parms = c("mu", "lambda"), mle = c(3.6065, 1.6589), lower = 0)$bias
## mu lambda
## 3.483e-07 1.082e-01
```

15. Log-normal distribution with location $\mu$ and scale $\sigma$

\[
 f(x \mid \mu, \sigma) = \frac{1}{\sqrt{2\pi} \sigma x} \exp\left[-\frac{(\log x - \mu)^2}{2\sigma^2}\right], \quad x > 0.
\]

- Bias expressions (Stošić and Cordeiro, 2009):

\[
 B(\hat{\mu}) = 0 \quad \text{and} \quad B\left(\hat{\sigma}\right) = -\frac{3\sigma^4}{4n}. \tag{23}
\]

Using the data set from Kumagai et al. (1989), we have $n = 30$, $\hat{\mu} = 2.164$, $\hat{\sigma} = 1.1765$, $\hat{\sigma}(\hat{\mu}) = 0.2148$ and $\hat{\sigma}(\hat{\sigma}) = 0.1519$. Evaluating the analytical expressions (23) and the coxsnell.bc() function, we have, respectively,

```r
lognormal.bc(n = 30, mle = c(2.1643, 1.1765))
## mu sigma
## 0.00000 -0.02941
```

```r
density <- quote(1 / (sqrt(2 * pi) * x * sigma) * exp(-0.5 * (log(x) - mu)^2 / sigma^2))
lpdf <- quote(-log(sigma) - 0.5 * (log(x) - mu)^2 / sigma^2)
coxsnell.bc(density = density, logdensity = lpdf, n = 30, parms = c("mu", "sigma"), mle = c(2.1643, 1.1765), lower = 0)$bias
## mu sigma
## -5.952e-09 -2.941e-02
```

16. Log-logistic distribution with shape $\beta$ and scale $\alpha$

\[
 f(x \mid \alpha, \beta) = \frac{(\beta / \alpha) (x / \alpha)^{\beta-1}}{[1 + (x / \alpha)^{\beta}]^2}, \quad x > 0.
\]

- For bias expressions, see Reath (2016).

From Reath (2016) we have $n = 19$, $\hat{\alpha} = 6.2542$, $\hat{\beta} = 1.1732$, $\hat{\sigma}(\hat{\alpha}) = 2.1352$, $\hat{\sigma}(\hat{\beta}) = 0.0239$, $\hat{B}(\hat{\alpha}) = 0.3585$ and $\hat{B}(\hat{\beta}) = 0.0789$. Evaluating the coxsnell.bc() function, we have:

```r
pdf <- quote((beta / alpha) * (x / alpha)^{beta-1} / [1 + (x / alpha)^beta]^2)
lpdf <- quote(log(beta) - log(alpha) + (beta - 1) * log(x / alpha) -...}
```
2 * log(1 + (x / alpha)*beta))
coxsnell.bc(density = pdf, logdensity = lpdf, n = 19,
  parms = c("alpha", "beta"), mle = c(6.2537, 1.1734),
  lower = 0)$bias
## alpha beta
## 0.35854 0.07883

17. Gamma distribution with shape \( \alpha \) and rate \( \lambda \)

\[
f(x \mid \alpha, \lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\lambda x), \quad x > 0.
\]

- Bias expressions (Giles and Feng, 2009):
  \[
  B(\hat{\alpha}) = \frac{\alpha \left[ \Psi'(\alpha) - a \Psi''(\alpha) \right] - 2}{2n[\alpha\Psi'(\alpha) - 1]^2}
  \tag{24}
  \]
  and
  \[
  B(\hat{\lambda}) = \frac{\lambda \left[ 2\alpha \left( \Psi'(\alpha) \right)^2 - 3\Psi'(\alpha) - a \Psi''(\alpha) \right]}{2n[\alpha\Psi'(\alpha) - 1]^2}.
  \tag{25}
  \]

Using the data set from Delignette-Muller et al. (2008), we have \( n = 254, \hat{\alpha} = 4.0083, \hat{\lambda} = 0.0544, \hat{se}(\hat{\alpha}) = 0.3413 \) and \( \hat{se}(\hat{\lambda}) = 0.0049. \) Evaluating the analytical expressions (24), (25) and the coxsnell.bc() function, we have, respectively,

\[
\text{gamma.bc}(n = 254, mle = c(4.0082, 0.0544))
\]
## alpha lambda
## 0.0448278 0.0006618

\[
\text{pdf} \leftarrow \text{quote}(\text{lambda}^\text{alpha} / \text{gamma}(\text{alpha}) * x^{(\text{alpha} - 1)} * \exp(-\text{lambda} * x))
\]
\[
\text{lpdf} \leftarrow \text{quote}(\text{alpha} * \log(\text{lambda}) - \text{lgamma}(\text{alpha}) + \text{alpha} * \log(x) - \text{lambda} / x)
\]
\[
\text{coxsnell.bc}(\text{density} = \text{pdf}, \text{logdensity} = \text{lpdf}, \text{n} = 254,
  \text{parms} = \text{c("alpha", "lambda"), mle = c(4.0082, 0.0544),
  \text{lower} = 0})$bias
\]
## alpha lambda
## 0.0448278 0.0006618

18. Inverse gamma distribution with shape \( \alpha \) and scale \( \beta \)

\[
f(x \mid \alpha, \beta) = \frac{1}{\Gamma(\alpha) \beta^\alpha} x^{\alpha-1} \exp\left(\frac{-x}{\beta}\right), \quad x > 0.
\]

- Bias expressions (Stošić and Cordeiro, 2009):
  \[
  B(\tilde{\alpha}) = \frac{-0.5 \alpha^2 \Psi''(\alpha) + 0.5 \Psi'(\alpha) \alpha - 1}{n \alpha \left( \Psi'(\alpha) - 1 \right)^2}
  \tag{26}
  \]
  and
  \[
  B(\tilde{\beta}) = \frac{\beta \left( -0.5 \Psi''(\alpha) - 1.5 \Psi'(\alpha) + (\Psi'(\alpha))^2 \alpha \right)}{n \left( \Psi'(\alpha) \alpha - 1 \right)^2}.
  \tag{27}
  \]

Using the data set from Kumagai and Matsunaga (1995), we have \( n = 31, \tilde{\alpha} = 1.0479, \tilde{\beta} = 5.491, \hat{se}(\tilde{\alpha}) = 0.2353 \) and \( \hat{se}(\tilde{\beta}) = 1.5648. \) Evaluating the analytical expressions (26), (27) and the coxsnell.bc() function, we have, respectively,

\[
\text{invgamma.bc}(n = 31, mle = c(5.4901, 1.0479))
\]
## beta alpha
## 0.60849 0.08388

\[
\text{pdf} \leftarrow \text{quote}(\text{beta}^\text{alpha} / \text{gamma}(\text{alpha}) * x^{-(\text{alpha} - 1)} * \exp(-\text{beta} / x))
\]
\[
\text{lpdf} \leftarrow \text{quote}(\text{alpha} * \log(\text{beta}) - \text{lgamma}(\text{alpha}) - \text{alpha} * \log(x) - \text{beta} / x)
\]
19. Lomax distribution with shape \( \alpha \) and scale \( \beta \)

\[
f(x \mid \alpha, \beta) = \alpha \beta (1 + \beta x)^{-(\alpha + 1)}, \quad x > 0.
\]

- **Bias expressions** (Giles et al., 2013):

\[
B(\hat{\alpha}) = \frac{2\alpha (\alpha + 1) (\alpha^2 + \alpha - 2)}{(\alpha + 3) n} \tag{28}
\]

and

\[
B(\hat{\beta}) = -\frac{2\beta (\alpha + 1.6485) (\alpha + 0.3934) (\alpha - 1.5419)}{n \alpha (\alpha + 3)}. \tag{29}
\]

Using the data set from Tahir et al. (2016), we have \( n = 179, \hat{\alpha} = 4.9103, \hat{\beta} = 0.0028, \hat{\sigma}(\hat{\alpha}) = 0.6208 \) and \( \hat{\sigma}(\hat{\beta}) = 3.4803 \times 10^{-4} \). Evaluating the analytical expressions (28), (29) and the `coxsnell.bc()` function, we have, respectively,

```r
coxsnell.bc(density = pdf, logdensity = lpdf, n = 179, 
parms = c("alpha", "beta"), mle = c(4.9103, 0.0028), 
lower = 0)$bias
```

\[
\begin{array}{cc}
\text{alpha} & \text{beta} \\
1.281e+00 & -9.439e-05 \\
\end{array}
\]

20. Weighted Lindley distribution with shape \( \alpha \) and scale \( \theta \)

\[
f(x \mid \alpha, \theta) = \frac{\theta^{\alpha + 1}}{(\theta + \alpha) \Gamma(\alpha)} x^{\alpha - 1} (1 + x) \exp(-\theta x), \quad x > 0.
\]

- **For bias expressions**, see (Wang and Wang, 2017):

Using the data set from Ghitany et al. (2013), we have \( n = 69, \hat{\alpha} = 22.8889, \hat{\theta} = 9.6246, \hat{\sigma}(\hat{\alpha}) = 3.9507 \) and \( \hat{\sigma}(\hat{\theta}) = 1.6295 \). Evaluating the analytical expressions and the `coxsnell.bc()` function, we have, respectively,

```r
wlindley.bc(n = 69, mle = c(22.8889, 9.6246))
```

\[
\begin{array}{cc}
\text{alpha} & \text{theta} \\
1.0070 & 0.4167 \\
\end{array}
\]

21. Generalized Rayleigh with shape \( \alpha \) and scale \( \theta \)

\[
f(x \mid \beta, \mu) = \frac{2 \theta^{\alpha + 1}}{\Gamma(\alpha + 1)} x^{2\alpha + 1} \exp \left(-\theta x^2\right), \quad x > 0.
\]

- **For bias expressions**, see (Xiao and Giles, 2014):
Using the data set from Gomes et al. (2014), we have $n = 384$, $\hat{\theta} = 0.5195$, $\hat{\alpha} = 0.0104$, $\hat{\text{se}}(\hat{\theta}) = 0.2184$ and $\hat{\text{se}}(\hat{\alpha}) = 0.0014$. Evaluating the analytical expressions and the coxsnell.bc() function, we have, respectively,

\[
\text{generalizedrayleigh.bc}(n = 384, \text{mle} = c(0.5195, 0.0104))
\]
\[
\begin{array}{c|c}
\text{alpha} & \text{theta} \\
\hline
1.035e-02 & 8.865e-05 \\
\end{array}
\]
\begin{verbatim}
pdf <- quote(2 * theta^(alpha + 1) / gamma(alpha + 1) * x^(2 * alpha + 1) * exp(-theta * x^2 ))
lpdf <- quote((alpha + 1) * log(theta) - lgamma(alpha + 1) + 2 * alpha * log(x) - theta * x^2)
coxsnell.bc(density = pdf, logdensity = lpdf, n = 384, 
parms = c("alpha", "theta"), mle = c(0.5195, 0.0104), 
lower = 0)$bias
\end{verbatim}
\[
\begin{array}{c|c}
\text{alpha} & \text{theta} \\
\hline
1.035e-02 & 8.865e-05 \\
\end{array}
\]

22. Weibull distribution with shape $\beta$ and scale $\mu$

\[
f(x \mid \beta, \mu) = \frac{\beta}{\mu^\beta} x^{\beta - 1} \exp\left(-\frac{x}{\mu}\right)^\beta, \quad x > 0.
\]

• Bias expressions (the expressions below differs from Stošćić and Cordeiro (2009)):

\[
B(\hat{\beta}) = \mu \frac{(0.5543324495 - 0.3698145397 \beta)}{n \beta^2} \tag{30}
\]

and

\[
B(\hat{\mu}) = \frac{1.379530692 \beta}{n} \tag{31}
\]

From Datta and Datta (2013), we have $n = 50$, $\hat{\mu} = 2.5752$, $\hat{\beta} = 38.0866$, $\hat{\text{se}}(\hat{\mu}) = 0.2299$ and $\hat{\text{se}}(\hat{\beta}) = 2.2299$. Evaluating the analytical expression (30), (31) and the coxsnell.bc() function, we have, respectively,

\[
\text{weibull.bc}(n = 50, \text{mle} = c(38.0866, 2.5751))
\]
\[
\begin{array}{c|c}
\text{mu} & \text{beta} \\
\hline
-0.04572 & 0.07105 \\
\end{array}
\]
\begin{verbatim}
pdf <- quote(beta / mu^beta * x^(beta - 1) * exp(-(x / mu)^beta))
lpdf <- quote(log(beta) - beta * log(mu) + beta * log(x) - (x / mu)^beta)
coxsnell.bc(density = pdf, logdensity = lpdf, n = 50, 
parms = c("mu", "beta"), mle = c(38.0866, 2.5751), 
lower = 0)$bias
\end{verbatim}
\[
\begin{array}{c|c}
\text{mu} & \text{beta} \\
\hline
-0.04572 & 0.07105 \\
\end{array}
\]

23. Inverse Weibull distribution with shape $\beta$ and scale $\mu$

\[
f(x \mid \beta, \alpha) = \beta \mu^\beta x^{-(\beta + 1)} \exp\left[-\left(\frac{\mu}{x}\right)^\beta\right], \quad x > 0.
\]

• Bias expressions (not previously reported in the literature):

\[
B(\hat{\beta}) = \frac{1.379530690 \beta}{n} \tag{32}
\]

and

\[
B(\hat{\mu}) = \mu \frac{(0.3698145391 \beta + 0.5543324494)}{n \beta^2} \tag{33}
\]

Using the data set from Nichols and Padgett (2006), we have $n = 100$, $\hat{\beta} = 1.769$, $\hat{\mu} = 1.8917$, $\hat{\text{se}}(\hat{\beta}) = 0.1119$ and $\hat{\text{se}}(\hat{\mu}) = 0.1138$. Evaluating the analytical expressions (32), (33) and the coxsnell.bc() function, we have, respectively,
inverseweibull.bc(n = 100, mle = c(1.7690, 1.8916))
## beta mu
## 0.024404 0.007305

pdf <- quote(beta * mu^beta * x^(-beta - 1) * exp(-(mu / x)^beta))

lpdf <- quote(log(beta) + beta * log(mu) - beta * log(x) - (mu / x)^beta)

coxsnell.bc(density = pdf, logdensity = lpdf, n = 100,
parms = c("beta", "mu"), mle = c(1.7690, 1.8916),
lower = 0)$bias
## beta mu
## 0.024404 0.007305

24. Generalized half-normal distribution with shape $\alpha$ and scale $\theta$

\[ f(x \mid \alpha, \theta) = \sqrt{\frac{2}{\pi \alpha^\theta}} x^{\alpha - 1} \exp\left[ -\frac{1}{2} \left( \frac{x^\alpha}{\theta} \right)^2 \right], \quad x > 0. \]

- Bias expressions (Mazucheli and Dey, 2017):
\[ B(\hat{\alpha}) = 1.48379456 \frac{\alpha}{n} \]

and
\[ B(\hat{\theta}) = (0.2953497661 - 0.3665611957 \alpha) \frac{\theta}{n \alpha^2}. \]

Using the data set from Nadarajah (2008a), we have $n = 119$, $\hat{\alpha} = 3.8096$, $\hat{\theta} = 4.9053$, $\hat{\sigma}(\hat{\alpha}) = 0.2758$ and $\hat{\sigma}(\hat{\theta}) = 0.0913$. Evaluating the analytical expressions (34), (35) and the `coxsnell.bc()` function, we have, respectively,
genhalfnormal.bc(n = 119, mle = c(3.8095, 4.9053))
## alpha theta
## 0.047500 -0.003127

pdf <- quote(sqrt(2 / pi) * alpha / theta^alpha * x^(alpha - 1)* exp(- 0.5 * (x / theta)^(2 * alpha) ))

lpdf <- quote(log(alpha) - alpha * log(theta) + alpha * log(x) - 0.5 * (x / theta)^(2 * alpha))

coxsnell.bc(density = pdf, logdensity = lpdf, n = 119,
parms = c("alpha", "theta"), mle = c(3.8095, 4.9053),
lower = 0)$bias
## alpha theta
## 0.047500 -0.003127

25. Inverse generalized half-normal distribution with shape $\alpha$ and scale $\theta$

\[ f(x \mid \alpha, \theta) = \sqrt{\frac{2}{\pi \alpha^\theta}} \left( \frac{1}{\theta x} \right)^\alpha \exp\left[ -\frac{1}{2} \left( \frac{1}{\theta x} \right)^{2 \alpha} \right], \quad x > 0. \]

- For bias expressions (not previously reported in the literature, see the “analyticalBC.R” file.

Using the data set from Nadarajah et al. (2011), we have $n = 20$, $\hat{\alpha} = 3.0869$, $\hat{\theta} = 0.6731$, $\hat{\sigma}(\hat{\alpha}) = 0.5534$ and $\hat{\sigma}(\hat{\theta}) = 0.0379$. Evaluating the analytical expressions and the `coxsnell.bc()` function, we have, respectively,
invgenhalfnormal.bc(n = 20, mle = c(3.0869, 0.6731))
## alpha theta
## 0.229016 -0.002953

pdf <- quote(sqrt(2) * pi^(-0.5) * alpha * x^(-alpha - 1) * exp(-0.5 * x^(-2 * alpha) * (1 / theta)^(2 * alpha) * theta^(-alpha))

lpdf <- quote(log(alpha) - alpha * log(theta) - 0.5e0 / (x^alpha)^2* theta^(-2 * alpha) - alpha * log(theta))

coxsnell.bc(density = pdf, logdensity = lpdf, n = 20,
parms = c("alpha", "theta"), mle = c(3.0869, 0.6731),
lower = 0)$bias
26. Marshall-Olkin extended exponential distribution with shape $\alpha$ and rate $\lambda$

$$f(x | \alpha, \lambda) = \frac{\lambda \alpha \exp(-\lambda x)}{[1 - (1 - \alpha) \exp(-\lambda x)]^2}, \quad x > 0.$$ 

- For bias expressions (not previously reported in the literature, see the “analyticalBC.R” file. Using the data set from Linhart and Zucchini (1986), we have $n = 20$, $\hat{\alpha} = 0.2782$, $\hat{\lambda} = 0.0078$, $\hat{\text{se}}(\hat{\alpha}) = 0.2321$ and $\hat{\text{se}}(\hat{\lambda}) = 0.0049$. Evaluating the analytical expressions and the coxsnell.bc() function, we have, respectively,

```r
moeexp.bc(n = 20, mle = c(0.2781, 0.0078))
## alpha lambda
## 0.2109 0.0037
```

27. Beta distribution with shapes $\alpha$ and $\beta$

$$f(x | \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}, \quad 0 < x < 1.$$ 

- For bias expressions, see (Cordeiro et al., 1997).

Using the data set from Javanshiri et al. (2015), we have $n = 48$, $\hat{\alpha} = 5.941$, $\hat{\beta} = 21.2024$, $\hat{\text{se}}(\hat{\alpha}) = 1.1812$ and $\hat{\text{se}}(\hat{\beta}) = 4.3462$. Evaluating the analytical expressions in Cordeiro et al. (1997), our analytical expressions and the coxsnell.bc() function, we have, respectively,

```r
beta.gauss.bc(n = 48, mle = c(5.941, 21.2024))
## alpha beta
## -4.784 -4.125
```

28. Kumaraswamy distribution with shapes $\alpha$ and $\beta$

$$f(x | \alpha, \beta) = \alpha \beta x^{\alpha-1} (1-x^{\beta})^{\beta-1}, \quad 0 < x < 1.$$ 

- For bias expressions, see (Lemonte, 2011).

Using the data set from Wang et al. (2017), we have $n = 20$, $\hat{\alpha} = 6.3478$, $\hat{\beta} = 4.4898$, $\hat{\text{se}}(\hat{\alpha}) = 1.5576$ and $\hat{\text{se}}(\hat{\beta}) = 2.0414$. Evaluating the analytical expressions and the coxsnell.bc() function, we have, respectively,
28. Inverse beta distribution with shapes $\alpha$ and $\beta$

\[
 f(x \mid \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha - 1} (1 + x)^{-(\alpha + \beta)}, \quad x > 0.
\]

- For bias expressions, see (Stošić and Cordeiro, 2009).

Using the data set from Nadarajah (2008b), we have $n = 116$, $\hat{\alpha} = 28.5719$, $\hat{\beta} = 1.3782$, $\hat{se}(\hat{\alpha}) = 4.0367$ and $\hat{se}(\hat{\beta}) = 0.1637$. Evaluating the analytical expressions and the `coxsnell.bc()` function, we have, respectively,

\[
 \text{invbeta.bc}(n = 116, \text{mle} = c(28.5719, 1.3782))
\]

<table>
<thead>
<tr>
<th>alpha</th>
<th>beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>534.26</td>
<td>17.73</td>
</tr>
</tbody>
</table>

```r
pdf <- quote(gamma(alpha + beta) * x^(alpha - 1) * (1 + x)^(- alpha - beta) / gamma(alpha)/gamma(beta))
lpdf <- quote(lgamma(alpha + beta) + alpha * log(x) - (alpha + beta) * log(1 + x) - lgamma(alpha) - lgamma(beta))
coxsnell.bc(density = pdf, logdensity = lpdf, n = 116, parms = c("alpha", "beta"), mle = c(28.5719, 1.3782), lower = 0)\$bias
```

<table>
<thead>
<tr>
<th>alpha</th>
<th>beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8025</td>
<td>0.0306</td>
</tr>
</tbody>
</table>

30. Birnbaum-Saunders distribution with shape $\alpha$ and scale $\beta$

\[
 f(x \mid \alpha, \beta) = \frac{1}{2 \alpha \beta \sqrt{2 \pi}} \left[ \left( \frac{\beta}{x} \right)^{1/2} + \left( \frac{\beta}{x} \right)^{3/2} \right] \exp \left[ - \frac{1}{2 \alpha^2} \left( \frac{x}{\beta} + \frac{\beta}{x} - 2 \right) \right], \quad x > 0.
\]

- Bias expressions (Lemonte et al., 2007):

\[
 B(\hat{\alpha}) = -\frac{\alpha}{4n} \left(1 + \frac{2 + \alpha^2}{\alpha} h(\alpha) + 1\right) \tag{36}
\]

and

\[
 B(\hat{\beta}) = \frac{\beta^2 a^2}{2\pi \left[ \alpha (2\pi)\right]^{1/2} h(\alpha) + 1}, \tag{37}
\]

where

\[
 h(\alpha) = \alpha \sqrt{\frac{\pi}{2}} - \pi e^{2/a^2} \left[ 1 - \Phi \left( \frac{2}{\alpha} \right) \right].
\]

Using the data set from Gross and Clark (1976), we have $n = 20$, $\hat{\alpha} = 0.3149$, $\hat{\beta} = 1.8105$, $\hat{se}(\hat{\alpha}) = 0.0498$ and $\hat{se}(\hat{\beta}) = 0.1259$. Evaluating the analytical expressions (36), (37) and the `coxsnell.bc()` function, we have, respectively,

\[
 \text{birnbaumsaunders.bc}(n = 20, \text{mle} = c(0.3148, 1.8104))
\]

<table>
<thead>
<tr>
<th>alpha</th>
<th>beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.011991</td>
<td>0.004374</td>
</tr>
</tbody>
</table>

```r
pdf <- quote(1 / (2 * alpha * beta * sqrt(2 * pi)) * ((beta / x)^0.5 + (beta / x)^1.5) * exp((-1/(2 * alpha^2) * (x / beta + beta/ x - 2))))
lpdf <- quote(-log(alpha) - log(beta) - 1 / (2 * alpha^2) * (x / beta + beta/ x - 2) + log((beta / x)^0.5 +
```
\[(\beta / x)^{1.5})\]

\[\text{coxsnell.bc(density = pdf, logdensity = lpdf, n = 20,}
\]  
\[\text{parms = c("alpha", "beta"), mle = c(0.3148, 1.8104),}
\]  
\[\text{lower = 0)$bias}
\]  
\[
\begin{array}{cc}
\text{alpha} & \text{beta} \\
-0.011991 & 0.004374
\end{array}
\]

31. Generalized Pareto distribution with shape \(\xi\) and scale \(\sigma\)

\[f(x \mid \xi, \sigma) = \frac{1}{\sigma} \left(1 + \frac{\xi x}{\sigma}\right)^{-\left(1/\xi + 1\right)}, \quad x > 0, \xi \neq 0.
\]

- Bias expressions (Giles et al., 2016):

\[B(\hat{\xi}) = -\sigma \left(3 + 5\xi + 4\xi^2\right) \frac{n}{n + 3\xi}.
\]

Using the data set from Ross and Lott (2003), we have \(n = 58, \hat{\xi} = 0.736, \hat{\sigma} = 1.709, \hat{se}(\hat{\xi}) = 0.223\) and \(\hat{se}(\hat{\sigma}) = 0.41\). Evaluating the analytical expressions (38), (39) and the \text{coxsnell.bc()} function, we have, respectively,

\[\text{genpareto.bc(n = 58, mle = c(0.736, 1.709))}
\]  
\[
\begin{array}{cc}
\text{xi} & \text{sigma} \\
-0.03486 & 0.08126
\end{array}
\]

\[\text{pdf <- quote(1 / sigma * (1 + xi * x / sigma)^(-1/4))}
\]  
\[\text{Rlpdf <- quote(-log(sigma) - (1 + 1 / xi) * log(1 + xi * x / sigma))}
\]

\[\text{coxsnell.bc(density = pdf, logdensity = lpdf, n = 58,}
\]  
\[\text{parms = c("xi", "sigma"), mle = c(0.736, 1.709),}
\]  
\[\text{lower = 0)$bias}
\]  
\[
\begin{array}{cc}
\text{xi} & \text{sigma} \\
-0.03486 & 0.08126
\end{array}
\]

### Additional Applications

In this section, we present additional numerical results returned by \text{coxsnell.bc()}, \text{observed.varc()} and \text{expected.varcov()}. For the data describing the times between successive electric pulses on the surface of isolated muscle fiber (Cox and Lewis, 1966; Jørgensen, 1982), we fitted the exponentiated Weibull, Marshall-Olkin extended Weibull, Weibull, Marshall-Olkin extended exponential and exponential distributions. These distributions were also fitted by Cordeiro and Lemonte (2013). There are 799 observations and for each distribution we report the MLEs, the bias corrected MLEs, the observed variance-covariance obtained from the numerical Hessian \(H_1^{-1}(\hat{\theta})\), the observed variance-covariance obtained from the analytical Hessian \(H_2^{-1}(\hat{\theta})\), the expected variance-covariance \(I^{-1}(\hat{\theta})\) and the expected variance-covariance evaluated at the bias corrected MLEs \(I^{-1}(\hat{\theta})\). The MLEs and the \(H_1^{-1}(\hat{\theta})\) matrix were obtained by the \text{fitdistrplus} package (Delignette-Muller et al., 2017). The R codes used to obtain the numerical results are available in the supplementary material.

It is important to emphasize that for the Marshall-Olkin extended Weibull and exponentiated Weibull distributions, it is not possible to obtain analytical expressions for bias corrections. The exponentiated-Weibull family was proposed by Mudholkar and Srivastava (1993). Its probability density function is:

\[f(x \mid \lambda, \beta, \alpha) = \alpha \beta \lambda x^{\beta - 1} e^{-\lambda x^{\beta}} \left(1 - e^{-\lambda x^{\beta}}\right)^{\alpha - 1},\]

where \(\lambda > 0\) is the scale parameter and \(\beta > 0\) and \(\alpha > 0\) are the shape parameters. The Marshall-Olkin extended Weibull distribution was introduced by Marshall and Olkin (1997). Its probability density
function is:
\[ f(x \mid \lambda, \beta, \alpha) = \frac{\alpha \beta \lambda x^{\beta - 1} e^{-\lambda x^\beta}}{(1 - \pi e^{-\lambda x^\beta})^{2}} \]
where \( \lambda > 0 \) is the scale parameter, \( \beta > 0 \) is the shape parameter, \( \alpha > 0 \) is an additional shape parameter and \( \pi = 1 - \alpha \).

The fitted parameter estimates and their bias corrected estimates are shown in Table 1. We see that the bias corrected MLEs for \( \alpha \) and \( \lambda \) of the MOE-Weibull and exp-Weibull distributions are quite different from the original MLEs.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>( \hat{\alpha} )</th>
<th>( \hat{\beta} )</th>
<th>( \hat{\lambda} )</th>
<th>( \hat{\lambda} )</th>
<th>( \hat{\beta} )</th>
<th>( \hat{\lambda} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOE-Weibull</td>
<td>0.3460</td>
<td>1.3247</td>
<td>0.0203</td>
<td>0.3283</td>
<td>1.3240</td>
<td>0.0188</td>
</tr>
<tr>
<td>exp-Weibull</td>
<td>1.9396</td>
<td>0.7677</td>
<td>0.2527</td>
<td>1.8973</td>
<td>0.7625</td>
<td>0.2461</td>
</tr>
<tr>
<td>Weibull</td>
<td>–</td>
<td>1.0829</td>
<td>0.0723</td>
<td>–</td>
<td>1.0811</td>
<td>0.0723</td>
</tr>
<tr>
<td>MOE-exponential</td>
<td>1.1966</td>
<td>–</td>
<td>0.0998</td>
<td>1.1820</td>
<td>–</td>
<td>0.0994</td>
</tr>
<tr>
<td>exponential</td>
<td>–</td>
<td>–</td>
<td>0.0913</td>
<td>–</td>
<td>–</td>
<td>0.0912</td>
</tr>
</tbody>
</table>

It is important to assess the accuracy of MLEs. The two common ways for this are through the inverse observed Fisher information and the inverse expected Fisher information matrices. The results below show large differences between the observed \( H^{-1} \) and expected \( I^{-1} \) information matrices. As demonstrated by Cao (2013), the \( I^{-1} \) outperforms the \( H^{-1} \) under a mean squared error criterion, hence with \texttt{mle.tools} the researchers may choose one of them and not use the easier. Furthermore, in general, we observe that the bias corrected MLEs decrease the variance of estimates.

- Exponentiated Weibull distribution:

  \( H^{-1}_1(\hat{\theta}) = \begin{bmatrix} 0.00004 & -0.00036 & 0.00052 \\ -0.00036 & 0.00361 & -0.00430 \\ 0.00052 & -0.00430 & 0.00748 \end{bmatrix}, \quad H^{-1}_2(\hat{\theta}) = \begin{bmatrix} 0.00005 & -0.00047 & 0.00068 \\ -0.00047 & 0.00468 & -0.00582 \\ 0.00068 & -0.00582 & 0.00967 \end{bmatrix} \),

  \( I^{-1}_1(\hat{\theta}) = \begin{bmatrix} 0.00006 & -0.00056 & 0.00082 \\ -0.00056 & 0.00542 & -0.00699 \\ 0.00082 & -0.00699 & 0.01146 \end{bmatrix}, \quad I^{-1}_2(\hat{\theta}) = \begin{bmatrix} 0.00005 & -0.00051 & 0.00072 \\ -0.00051 & 0.00526 & -0.00651 \\ 0.00072 & -0.00651 & 0.01030 \end{bmatrix} \).

- Marshall-Okin extended Weibull distribution:

  \( H^{-1}_1(\hat{\theta}) = \begin{bmatrix} 0.00004 & -0.00018 & 0.00018 \\ -0.00018 & 0.00086 & 0.00018 \end{bmatrix}, \quad H^{-1}_2(\hat{\theta}) = \begin{bmatrix} 0.00004 & -0.00018 & 0.00087 \\ -0.00018 & 0.00087 & 0.00004 \end{bmatrix} \),

  \( I^{-1}_1(\hat{\theta}) = \begin{bmatrix} 0.00004 & -0.00018 & 0.00089 \\ -0.00018 & 0.00089 & 0.00004 \end{bmatrix}, \quad I^{-1}_2(\hat{\theta}) = \begin{bmatrix} 0.00004 & -0.00018 & 0.00089 \\ -0.00018 & 0.00089 & 0.00004 \end{bmatrix} \).

- Weibull distribution:

  \( H^{-1}_1(\hat{\theta}) = \begin{bmatrix} 0.00004 & 0.00081 & 0.02022 \end{bmatrix}, \quad H^{-1}_2(\hat{\theta}) = \begin{bmatrix} 0.00004 & 0.00081 \\ 0.00081 & 0.02023 \end{bmatrix} \),

  \( I^{-1}_1(\hat{\theta}) = \begin{bmatrix} 0.00004 & 0.00083 \\ 0.00083 & 0.02094 \end{bmatrix}, \quad I^{-1}_2(\hat{\theta}) = \begin{bmatrix} 0.00004 & 0.00082 \\ 0.00082 & 0.02047 \end{bmatrix} \).
Concluding Remarks

As pointed out by several works in the literature, the Cox-Snell methodology, in general, is efficient for reducing the bias of the MLEs. However, the analytical expressions are either notoriously cumbersome or even impossible to deduce. To the best of our knowledge, there are only two alternatives to obtain the analytical expressions automatically, those presented in Stoˇ csi´ c and Cordeiro (2009) and Johnson et al. (2012a). They use the commercial softwares Maple (Maple, 2017) and Mathematica (Wolfram Research, Inc., 2010).

In order to calculate the bias corrected estimates in a simple way, Mazucheli (2017) developed an R (R Core Team, 2016) package, uploaded to CRAN on 2 February, 2017. Its main function, coxsnell.bc(), evaluates the bias corrected estimates. The usefulness of this function has been tested for thirty one continuous probability distributions. Bias expressions, for most of them, are available in the literature.

It is well known that the Fisher information can be computed using the first or second order derivatives of the log-likelihood function. In our implementation, the functions expected.varcov() and coxsnell.bc() are using the second order derivatives, analytically returned by the D() function.

In a future work, we intend to check if there is any gain in calculating the Fisher information from the first order derivatives of the log-hazard rate function or from the first order derivatives of the log-reversed-hazard rate function. Efron and Johnstone (1990) showed that the Fisher information can be computed using the hazard rate function. Gupta et al. (2004) computed the Fisher information from the first order derivatives of the log-reversed-hazard rate function. In general, expressions of the first order derivatives of the log-hazard rate function (log-reversed-hazard rate function) are simpler than second order derivatives of the log-likelihood function. In this sense, the integrate() function can work better. It is important to point out that the hazard rate function and the reversed hazard rate function are given, respectively, by

\[
    h(x | \theta) = -\frac{d}{dx} \log [S(x | \theta)],
\]

and

\[
    h(x | \theta) = \frac{d}{dx} \log [F(x | \theta)],
\]

where \(S(x | \theta)\) and \(F(x | \theta)\) are, respectively, the survival function and the cumulative distribution function.

In the next version of mle.tools, we will include, using analytical first and second-order partial derivatives, the following:

- the MLEs of \(g(\theta)\) and \(\text{Var}[g(\theta)]\),
- the negative log likelihood value \(-2 \log(L)\),
- the Akaike’s information criterion \(-2 \log(L) + 2p\),
- the corrected Akaike’s information criterion \(-2 \log(L) + \frac{2np}{n-p-1}\),
- the Schwarz’s Bayesian information criterion \(-2 \log(L) + p \log(n)\),
- the Hannan-Quinn information criterion \(-2 \log(L) + 2 \log \log(n)p\),

where \(L\) is the value of the likelihood function evaluated at the MLEs, \(n\) is the number of observations, and \(p\) is the number of estimated parameters.

Also, the next version of the package will incorporate analytical expressions for the distributions studied in Section 2.4 implemented in the supplementary file “analyticalBC.R”.

Bibliography


P. H. Johnson, Y. Qi, and Y. C. Chueh. CSCK MLE bias calculation. Working papers, University of Illinois, Urbana-Champaign, Champaign, IL, 2012b. [p268, 272]


J. Mazucheli and S. Dey. Bias-corrected maximum likelihood estimation of the parameters of the
generalized half-normal distribution. submitted to Journal of Statistical Computation and Simulation,
2017. [p268, 280]

G. S. Mudholkar and D. K. Srivastava. Exponentiated Weibull family for analyzing bathtub failure-rate

S. Nadarajah. The model for fracture toughness. Journal of Mechanical Science and Technology, 22:
1255–1258, 2008a. [p280]

S. Nadarajah. A truncated inverted beta distribution with application to air pollution data. Stochastic

331–359, 2011. [p280]

M. D. Nichols and W. Padgett. A bootstrap control chart for Weibull percentiles. Quality and reliability
engineering international, 22(2):141–151, 2006. [p279]

R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical
Computing, Vienna, Austria, 2016. ISBN 3-900051-07-0. [p268, 285]

M. Z. Raqab, M. T. Madi, and D. Kundu. Estimation of P(Y < X) for the three-parameter generalized
[p275]

Michigan Technological University, 2016. [p268, 276]

Report 01, National Climatic Data Center, 2003. [p283]

K. Saha and S. Paul. Bias-corrected maximum likelihood estimator of the negative binomial dispersion

J. Schwartz and D. E. Giles. Bias-reduced maximum likelihood estimation of the zero-inflated Poisson

J. Schwartz, R. T. Godwin, and D. E. Giles. Improved maximum-likelihood estimation of the shape
434–445, 2013. [p268]


reliability model with application to head and neck cancer data. Journal of Industrial and Production

B. D. Stošić and G. M. Cordeiro. Using Maple and Mathematica to derive bias corrections for two
271, 275, 276, 277, 279, 282, 285]

M. Tahir, M. A. Hussain, G. M. Cordeiro, G. Hamedani, M. Mansoor, and M. Zubair. The Gumbel-
61–79, 2016. [p278]

M. Teimouri and S. Nadarajah. Bias corrected MLEs for the Weibull distribution based on records.
Statistical Methodology, 13:12–24, 2013. [p268]

M. Teimouri and S. Nadarajah. Bias corrected MLEs under progressive type-II censoring scheme.

B. X. Wang, X. K. Wang, and K. Yu. Inference on the Kumaraswamy distribution. Communications in

[p268, 278]


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

André Felipe Berdosco Menezes  
Department of Statistics  
Universidade Estadual de Maringá  
Maringá, Brazil  
andrefelipemaringa@gmail.com

Saralees Nadarajah  
School of Mathematics  
University of Manchester  
Manchester M13 9PL, United Kingdom  
mbbssn2@manchester.ac.uk