A note on the Fisher information matrix of the Birnbaum–Saunders distribution

Artur J. Lemonte

Departamento de Estatística, Universidade Federal de Pernambuco
Recife, Pernambuco, 50749-540, Brazil
arturlemonte@gmail.com

Received 8 December 2014
Accepted 5 May 2015

We show that the Fisher information matrix of the Birnbaum–Saunders distribution can present numerical problems for some values of the shape parameter. To overcome this problem, we provide a simple analytical approximation which works very well.

Keywords: Birnbaum–Saunders distribution; Fatigue life distribution; Fisher information matrix; Lifetime data; Fatigue life problems.

2000 Mathematics Subject Classification: 60E05

1. Introduction

Fatigue is a structural damage which occurs when a material is exposed to stress and tension fluctuations. When the effect of vibrations on material specimens and structures is studied, the first point to be considered is the mechanism that could cause fatigue of these materials. The fatigue process (fatigue life) begins with an imperceptible fissure, the initiation, growth, and propagation of which produces a dominant crack in the specimen due to cyclic patterns of stress, whose ultimate extension causes the rupture or failure of this specimen. The failure occurs when the total extension of the crack exceeds a critical threshold for the first time. The partial extension of a crack produced by fatigue in each cycle is modeled by a random variable which depends on the type of material, the magnitude of the stress, and the number of previous cycles, among other factors.

Motivated by problems of vibration in commercial aircraft that caused fatigue in the materials, [5, 6] proposed a family of two-parameter distributions to model failure time due to fatigue under cyclic loading and the assumption that failure follows from the development and growth of a dominant crack. This distribution is known as the two-parameter Birnbaum–Saunders (‘BS’ for short) distribution or as the fatigue life distribution. The BS distribution is an attractive alternative to the Weibull, gamma, and log-normal models, since its derivation considers the basic characteristics of the fatigue process. It was later derived by [11] using a biological model which followed from relaxing some of the assumptions originally made by [5]. The relationship between the BS distribution and the inverse Gaussian distribution was investigated by [12], who demonstrated that the BS
distribution is an equal-weight mixture of an inverse Gaussian distribution and its complementary reciprocal. For book treatments of inverse Gaussian and BS distributions and their relationships, see [25, Chapter 13] and especially [32, Chapter 10].

The BS distribution has received significant attention over the last few years by many researchers and there has been much theoretical developments with respect to this distribution. The developments have covered many aspects of the distribution. Improved frequentist inference as well as interval estimation are discussed in [22, 24] and [34], while Bayes estimation is presented in [2]. Goodness-of-fit tests are considered in [26]. The shape of the hazard function of the BS distribution is discussed in [19], who showed that the hazard function is not monotone and it is unimodal for all ranges of the parameter values. Some generalizations and extensions of the BS distribution are proposed in [7–9, 13, 15–18, 21, 29], among others. Other important works about the BS distribution are [3, 4, 23], among many others.

The BS distribution has also received wide ranging applications. Some recent applications include: modeling of hourly $SO_2$ concentrations at ten monitoring stations located in different zones in Santiago [20]; modeling of diameter at breast height distributions of near-natural complex structure silver fir-European beech forests [30]; modeling of hourly dissolved oxygen (DO) concentrations observed at four monitoring stations located at different areas of Santiago [33]; statistical analysis of redundant systems with one warm stand-by unit [27]. Because of the widespread study and applications of the BS distribution, it is very important to have a Fisher information matrix which works for all the values of the model parameters. In the next section we show that the Fisher information matrix of the BS distribution can present numerical problems for some values of the shape parameter. The aim of this note is to provide a Fisher information matrix which works for all the values of the model parameters of the BS distribution.

2. The BS distribution

The cumulative distribution function of a random variable $T$ with BS distribution, say $T \sim BS(\alpha, \beta)$, is $F(t) = \Phi(v)$, with $t > 0$, where $\Phi(\cdot)$ is the standard normal cumulative function, $v = \ln(t) = \ln(t/\beta)/\alpha$, $\rho(z) = z^{1/2} - z^{-1/2}$, and $\alpha > 0$ and $\beta > 0$ are the shape and scale parameters, respectively. The scale parameter is also the median of the distribution, since $F(\beta) = 1/2$. For any $k > 0$, it follows that $kT \sim BS(\alpha, k\beta)$. It is noteworthy that the reciprocal property holds for the BS distribution: $T^{-1} \sim BS(\alpha, \beta^{-1})$. The probability density function of $T \sim BS(\alpha, \beta)$ is $f(t) = \kappa(\alpha, \beta) t^{-3/2}(t + \beta) \exp[-\tau(t/\beta)/(2\alpha^2)]$, where $\kappa(\alpha, \beta) = \exp(\alpha^2)/(2\alpha \sqrt{2\pi\beta})$ and $\tau(z) = z + z^{-1}$. The BS density function is positively skewed and the asymmetry of the distribution decreases with $\alpha$. As $\alpha$ decreases, the distribution becomes more symmetric around $\beta$, the median. The expected value, variance, skewness and kurtosis are, respectively, $E(T) = \beta(1 + \alpha^2/2)$, $\var(T) = (\alpha\beta)^2(1 + 5\alpha^2/4)$, $\mu_3 = 16\alpha(11\alpha^2 + 6)/(5\alpha^2 + 4)^3$ and $\mu_4 = 3 + 6\alpha^2(93\alpha^2 + 41)/(5\alpha^2 + 4)^2$.

The log-likelihood function $\ell(\alpha, \beta)$ for a single observation $t$ of $T \sim BS(\alpha, \beta)$, except for a constant term, is given by $\ell(\alpha, \beta) = \log[\kappa(\alpha, \beta)] + \log(t + \beta) - \tau(t/\beta)/(2\alpha^2)$. By taking partial derivatives of $\ell(\alpha, \beta)$ with respect to $\alpha$ and $\beta$, we obtain the unit score functions: $\partial \ell(\alpha, \beta)/\partial \alpha = -(1 + 2\alpha^{-2})/\alpha + \tau(t/\beta)/\alpha^3$ and $\partial \ell(\alpha, \beta)/\partial \beta = -1/(2\beta) + (t + \beta)^{-1} + \rho(t^2/\beta^2)/(2\alpha^2\beta)$. For interval estimation and hypotheses test inference on the model parameters, we require the $2 \times 2$ per observation expected information matrix $K(\alpha, \beta)$, namely

$$K(\alpha, \beta) = \begin{bmatrix} K_{\alpha\alpha} & K_{\alpha\beta} \\ K_{\beta\alpha} & K_{\beta\beta} \end{bmatrix}.$$
whose elements are

\[ K_{\alpha\alpha} = \frac{2}{\alpha^2}, \quad K_{\alpha\beta} = K_{\beta\alpha} = 0, \quad K_{\beta\beta} = \frac{1 + \alpha (2\pi)^{-1/2} h(\alpha)}{(\alpha \beta)^2}, \]

where

\[ h(\alpha) = \alpha \left( \frac{\pi}{2} \right)^{1/2} - \pi e^{2/\alpha^2} \left[ 1 - \Phi \left( \frac{2}{\alpha} \right) \right]. \quad (2.1) \]

Note that \( K_{\alpha\beta} = K_{\beta\alpha} = 0 \) and hence \( \alpha \) and \( \beta \) are orthogonal [10]. The algebraic development to obtain \( K(\alpha, \beta) \) is provided in the Appendix.

In what follows, we show that \( K_{\beta\beta} \) (i.e. the function \( h(\alpha) \)) can present some numerical problems for small values of the shape parameter \( \alpha \). A simple R function [31] to compute \( h(\alpha) \) is given in the Appendix. Table 1 lists some values of \( h(\alpha) \) obtained by using the R function provided in the Appendix. From the figures in Table 1, it is evident that the function \( h(\alpha) \) presents numerical problems for small values of \( \alpha \), once it delivers NaN (Not a Number) for small values of the shape parameter \( \alpha \). It happens because the product

\[ e^{2/\alpha^2} \times \left[ 1 - \Phi \left( \frac{2}{\alpha} \right) \right] \]

delivers an undefined real number for small values of \( \alpha \); that is, for a small value of \( \alpha \), we have that \( e^{2/\alpha^2} = \infty \), whereas \( 1 - \Phi(2/\alpha) = 0 \). So, zero times infinity is an undefined real number. Figure 1 plots the function \( h(\alpha) \) for \( \alpha \in (0, 0.3) \). From this figure, notice that \( h(\alpha) \) behaves very unstable for \( \alpha \in (0, 0.3) \). In the next section, we shall derive a simple analytical approximation for \( h(\alpha) \) that works very well for small values of the shape parameter \( \alpha \).

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>( h(\alpha) )</th>
<th>( \alpha )</th>
<th>( h(\alpha) )</th>
<th>( \alpha )</th>
<th>( h(\alpha) )</th>
<th>( \alpha )</th>
<th>( h(\alpha) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.001</td>
<td>NaN</td>
<td>0.002</td>
<td>NaN</td>
<td>0.003</td>
<td>NaN</td>
<td>0.004</td>
<td>NaN</td>
</tr>
<tr>
<td>0.013</td>
<td>NaN</td>
<td>0.025</td>
<td>NaN</td>
<td>0.037</td>
<td>NaN</td>
<td>0.049</td>
<td>NaN</td>
</tr>
<tr>
<td>0.014</td>
<td>NaN</td>
<td>0.026</td>
<td>NaN</td>
<td>0.038</td>
<td>NaN</td>
<td>0.050</td>
<td>NaN</td>
</tr>
<tr>
<td>0.015</td>
<td>NaN</td>
<td>0.027</td>
<td>NaN</td>
<td>0.039</td>
<td>NaN</td>
<td>0.051</td>
<td>NaN</td>
</tr>
<tr>
<td>0.016</td>
<td>NaN</td>
<td>0.028</td>
<td>NaN</td>
<td>0.040</td>
<td>NaN</td>
<td>0.052</td>
<td>NaN</td>
</tr>
<tr>
<td>0.017</td>
<td>NaN</td>
<td>0.029</td>
<td>NaN</td>
<td>0.041</td>
<td>NaN</td>
<td>0.053</td>
<td>NaN</td>
</tr>
<tr>
<td>0.018</td>
<td>NaN</td>
<td>0.030</td>
<td>NaN</td>
<td>0.042</td>
<td>NaN</td>
<td>0.054</td>
<td>0.06768</td>
</tr>
<tr>
<td>0.019</td>
<td>NaN</td>
<td>0.031</td>
<td>NaN</td>
<td>0.043</td>
<td>NaN</td>
<td>0.055</td>
<td>0.06893</td>
</tr>
<tr>
<td>0.020</td>
<td>NaN</td>
<td>0.032</td>
<td>NaN</td>
<td>0.044</td>
<td>NaN</td>
<td>0.056</td>
<td>0.07019</td>
</tr>
<tr>
<td>0.021</td>
<td>NaN</td>
<td>0.033</td>
<td>NaN</td>
<td>0.045</td>
<td>NaN</td>
<td>0.057</td>
<td>0.07144</td>
</tr>
<tr>
<td>0.022</td>
<td>NaN</td>
<td>0.034</td>
<td>NaN</td>
<td>0.046</td>
<td>NaN</td>
<td>0.058</td>
<td>0.07269</td>
</tr>
<tr>
<td>0.023</td>
<td>NaN</td>
<td>0.035</td>
<td>NaN</td>
<td>0.047</td>
<td>NaN</td>
<td>0.059</td>
<td>0.07395</td>
</tr>
<tr>
<td>0.024</td>
<td>NaN</td>
<td>0.036</td>
<td>NaN</td>
<td>0.048</td>
<td>NaN</td>
<td>0.060</td>
<td>0.07520</td>
</tr>
</tbody>
</table>

Notes: NaN: Not a Number.

For a random sample \( t = (t_1, \ldots, t_n) \) of size \( n \) from the BS(\( \alpha, \beta \)) model, the total log-likelihood function takes the form \( \ell_n(\alpha, \beta) = \sum_{i=1}^{n} \ell^{(i)}(\alpha, \beta) \), where \( \ell^{(i)}(\alpha, \beta) \) is the log-likelihood function for the \( i \)th observation \( (i = 1, \ldots, n) \) as given before. The total score function is \( U_n(\alpha, \beta) = (\sum_{i=1}^{n} \partial \ell^{(i)}(\alpha, \beta) / \partial \alpha, \sum_{i=1}^{n} \partial \ell^{(i)}(\alpha, \beta) / \partial \beta)^\top \), where \( \partial \ell^{(i)}(\alpha, \beta) / \partial \alpha \) and \( \partial \ell^{(i)}(\alpha, \beta) / \partial \beta \), for \( i = 1, \ldots, n \), have the forms given earlier. The total Fisher information matrix is \( K_n(\alpha, \beta) = nK(\alpha, \beta) \).
The maximum likelihood estimates $\hat{\alpha}$ and $\hat{\beta}$ of $\alpha$ and $\beta$, respectively, are obtained by setting $U_n(\hat{\alpha}, \hat{\beta}) = 0$. There is no closed-form expression for the maximum likelihood estimator and its computation has to be performed numerically using a nonlinear optimization algorithm, such as Newton-Raphson or Fisher scoring algorithms. Under conditions that are fulfilled for parameters in the interior of the parameter space, we have that

$$
\sqrt{n} \left( \frac{\alpha}{\beta} \right) \overset{\text{d}}{\sim} N_2 \left( \begin{pmatrix} \alpha^2 \\ 2 \\ 0 \\ \frac{\beta}{\beta}^2 \\ 0 \\ \frac{1}{1 + \alpha (2\pi)^{-1/2} h(\alpha)} \end{pmatrix}, \begin{bmatrix} \frac{(\alpha \beta)^2}{(1 + \alpha (2\pi)^{-1/2} h(\alpha))^2} & \alpha \beta^2 \\ \alpha \beta^2 & \frac{1}{1 + \alpha (2\pi)^{-1/2} h(\alpha)} \end{bmatrix} \right),
$$

when $n$ is large, “$\overset{\text{d}}{\sim}$” denoting approximately distributed. Note that $\hat{\alpha}$ and $\hat{\beta}$ are asymptotically independent.

The above asymptotic bivariate normal distribution can be used to construct approximate confidence intervals and confidence regions for the parameters. The asymptotic $100(1-\eta)\%$ confidence intervals for $\alpha$ and $\beta$ are given, respectively, by

$$
\left[ \hat{\alpha} \left( 1 - \frac{z_{1-\eta/2}}{\sqrt{2n}} \right), \hat{\alpha} \left( 1 + \frac{z_{1-\eta/2}}{\sqrt{2n}} \right) \right],
$$

$$
\left[ \hat{\beta} \left( 1 - \frac{z_{1-\eta/2} \hat{\alpha}}{n[1 + \alpha (2\pi)^{-1/2} h(\hat{\alpha})]} \right)^{1/2}, \hat{\beta} \left( 1 + \frac{z_{1-\eta/2} \hat{\alpha}}{n[1 + \alpha (2\pi)^{-1/2} h(\hat{\alpha})]} \right)^{1/2} \right],
$$

Published by Atlantis Press
Copyright: the authors

199
where $z_{1-\eta/2}$ is the quantile $(1-\eta/2)$ of the standard normal distribution. Notice that $h(\alpha)$ needs to be computed correctly, since the asymptotic variance of $\hat{\beta}$ (which is used to define the confidence interval for $\beta$) depends on it. However, as showed before, this function can present numerical problems for small values of the shape parameter. Therefore, an approximation for $h(\alpha)$ which works for small values of $\alpha$ has to be derived. It will be done in the next section.

3. A simple approximation for $h(\alpha)$

First, we have that

$$
\Phi(x) = \frac{1}{2} + \frac{1}{2} \text{erf}\left(\frac{x}{\sqrt{2}}\right),
$$

where erf$(\cdot)$ is the error function and it is defined as

$$
\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_{0}^{x} e^{-z^2} dz.
$$

It follows that

$$
1 - \Phi\left(\frac{2}{\alpha}\right) = \frac{1}{2} - \frac{1}{2} \text{erf}\left(\frac{\sqrt{2}}{\alpha}\right),
$$

and hence $h(\alpha)$ can be reduced to

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

For small values of $\alpha$, it can be shown that [1, p. 298]

$$
e^{2/\alpha^2} \left[1 - \text{erf}\left(\frac{\sqrt{2}}{\alpha}\right)\right] = \frac{\alpha}{\sqrt{2\pi}} \left(1 - \frac{\alpha^2}{4} + \frac{3\alpha^4}{16} - \frac{5\alpha^6}{64}\right) + O(\alpha^9).
$$

Therefore, for small values of $\alpha$, the expression for $h(\alpha)$ in (2.1) can be approximated by

$$
h_1(\alpha) = \frac{\alpha}{2} \left(\frac{\pi}{2}\right)^{1/2} \left(1 + \frac{\alpha^2}{4} - \frac{3\alpha^4}{16} + \frac{5\alpha^6}{64}\right) + O(\alpha^9). \quad (3.1)
$$

Equation (3.1) has a very simple form and it does not depend on any complicated function. It is a simple polynomial in the shape parameter $\alpha$.

In Table 2, we present some values of $h_1(\alpha)$ given in equation (3.1) obtained by using the R function provided in the Appendix. We consider exactly the same values for the parameter $\alpha$ of Table 1. As expected, the figures in Table 2 reveal that no problem happens in computing $h_1(\alpha)$ presented in (3.1) for small values of the shape parameter $\alpha$. Figure 2 plots the functions $h(\alpha)$ and $h_1(\alpha)$ for $\alpha \in (0,0.3)$. Notice that $h_1(\alpha)$ is very stable for $\alpha \in (0,0.3)$, whereas $h(\alpha)$ is very unstable. Therefore, we strongly recommend to use the approximation (3.1) to make inference (confidence intervals and hypothesis testing) on the model parameters of the BS distribution when $\alpha < 0.3$.

Acknowledgments

We gratefully acknowledge grants from CNPq (Brazil) and FACEPE (Pernambuco, Brazil).
Table 2. Function $h_1(\alpha)$ for some values of $\alpha$. 

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$h_1(\alpha)$</th>
<th>$\alpha$</th>
<th>$h_1(\alpha)$</th>
<th>$\alpha$</th>
<th>$h_1(\alpha)$</th>
<th>$\alpha$</th>
<th>$h_1(\alpha)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.001</td>
<td>0.00063</td>
<td>0.013</td>
<td>0.00815</td>
<td>0.025</td>
<td>0.01567</td>
<td>0.037</td>
<td>0.02319</td>
</tr>
<tr>
<td>0.002</td>
<td>0.00125</td>
<td>0.014</td>
<td>0.00877</td>
<td>0.026</td>
<td>0.01630</td>
<td>0.038</td>
<td>0.02382</td>
</tr>
<tr>
<td>0.003</td>
<td>0.00188</td>
<td>0.015</td>
<td>0.00940</td>
<td>0.027</td>
<td>0.01692</td>
<td>0.039</td>
<td>0.02445</td>
</tr>
<tr>
<td>0.004</td>
<td>0.00251</td>
<td>0.016</td>
<td>0.01003</td>
<td>0.028</td>
<td>0.01755</td>
<td>0.040</td>
<td>0.02508</td>
</tr>
<tr>
<td>0.005</td>
<td>0.00313</td>
<td>0.017</td>
<td>0.01065</td>
<td>0.029</td>
<td>0.01818</td>
<td>0.041</td>
<td>0.02570</td>
</tr>
<tr>
<td>0.006</td>
<td>0.00376</td>
<td>0.018</td>
<td>0.01128</td>
<td>0.030</td>
<td>0.01880</td>
<td>0.042</td>
<td>0.02633</td>
</tr>
<tr>
<td>0.007</td>
<td>0.00439</td>
<td>0.019</td>
<td>0.01191</td>
<td>0.031</td>
<td>0.01943</td>
<td>0.043</td>
<td>0.02696</td>
</tr>
<tr>
<td>0.008</td>
<td>0.00501</td>
<td>0.020</td>
<td>0.01253</td>
<td>0.032</td>
<td>0.02006</td>
<td>0.044</td>
<td>0.02759</td>
</tr>
<tr>
<td>0.009</td>
<td>0.00564</td>
<td>0.021</td>
<td>0.01316</td>
<td>0.033</td>
<td>0.02069</td>
<td>0.045</td>
<td>0.02821</td>
</tr>
<tr>
<td>0.010</td>
<td>0.00627</td>
<td>0.022</td>
<td>0.01379</td>
<td>0.034</td>
<td>0.02131</td>
<td>0.046</td>
<td>0.02884</td>
</tr>
<tr>
<td>0.011</td>
<td>0.00689</td>
<td>0.023</td>
<td>0.01442</td>
<td>0.035</td>
<td>0.02194</td>
<td>0.047</td>
<td>0.02947</td>
</tr>
<tr>
<td>0.012</td>
<td>0.00752</td>
<td>0.024</td>
<td>0.01504</td>
<td>0.036</td>
<td>0.02257</td>
<td>0.048</td>
<td>0.03010</td>
</tr>
</tbody>
</table>

Fig. 2. Functions $h(\alpha)$ and $h_1(\alpha)$ for $\alpha \in (0, 0.3)$.

Appendix

It should be mentioned that the asymptotic joint distribution of $\hat{\alpha}$ and $\hat{\beta}$ was firstly derived in [14]; see also [28, p. 286]. However, the expression for the Fisher information matrix derived by [14] involves an integral that requires numerical solution, unlike the expression for the Fisher information matrix presented in Section 2, which only involves numerical integration through the evaluation of the standard normal distribution function $\Phi(\cdot)$ (or the error function $\text{erf}(\cdot)$). So, the main advantage of using the Fisher information matrix presented in Section 2 over the Fisher information matrix obtained in [14] is that any integral that requires numerical solution has to be computed.

Published by Atlantis Press
Copyright: the authors
201
In the following, we provide a detailed algebraic development to obtain $h(\alpha)$. The second derivatives of the log-likelihood function $\ell(\alpha, \beta) = \log|\chi(\alpha, \beta)| + \log(t + \beta) - \tau(t/\beta)/(2\alpha^2)$ are

$$
\frac{\partial^2 \ell(\alpha, \beta)}{\partial \alpha^2} = \frac{1}{\alpha^2} + \frac{6}{\alpha^4} - \frac{3}{\alpha^4} \left( \frac{t}{\beta} + \frac{\beta}{t} \right), \quad \frac{\partial^2 \ell(\alpha, \beta)}{\partial \alpha \partial \beta} = -\frac{1}{\alpha^2} \left( \frac{t}{\beta} - \frac{\beta}{t} \right),
$$

$$
\frac{\partial^2 \ell(\alpha, \beta)}{\partial \beta^2} = \frac{1}{2\beta^2} - \frac{1}{\alpha^2} \left( \frac{t}{(t+\beta)^2} - \frac{t}{\alpha^2} \right).
$$

Let $\theta = (\alpha, \beta)^T$ be the parameter vector. Under some regularity conditions, it follows that $K(\alpha, \beta) = -E(\partial^2 \ell(\alpha, \beta)/\partial \theta \partial \theta^T)$. Since $E(T) = \beta(1 + \alpha^2/2)$ and $E(T^{-1}) = \beta^{-1}(1 + \alpha^2/2)$, we have that

$$
K_{aa} = -E\left( \frac{\partial^2 \ell(\alpha, \beta)}{\partial \alpha^2} \right) = \frac{2}{\alpha^2} \quad \Rightarrow \quad K_{bb} = K_{a\beta} = -E\left( \frac{\partial^2 \ell(\alpha, \beta)}{\partial \alpha \partial \beta} \right) = 0,
$$

$$
K_{b\beta} = -E\left( \frac{\partial^2 \ell(\alpha, \beta)}{\partial \beta^2} \right) = \frac{1 + \alpha^2 \beta^2 / (\alpha \beta)^2}
$$

where

$$
\ell_{\beta\beta} = \frac{1}{2\alpha \beta \sqrt{2\pi}} \int_0^\infty \frac{1}{(t+\beta)^2} \left[ \left( \frac{\beta}{t} \right)^{1/2} + \left( \frac{t}{\beta} \right)^{3/2} \right] \exp \left\{ -\frac{1}{2\alpha^2} \left( \frac{t}{\beta} + \frac{\beta}{t} - 2 \right) \right\} dt.
$$

Making the change of variable $t = \beta u$ in the previous integral, we obtain

$$
\ell_{\beta\beta} = \frac{1}{2\alpha \beta \sqrt{2\pi}} \int_0^\infty \frac{1}{(u+1)^2} \exp \left\{ -\frac{1}{2\alpha^2} (u^{1/2} - u^{-1/2})^2 \right\} du.
$$

Now, making the change of variable $u = v^2$ in the above integral, we have that

$$
\ell_{\beta\beta} = \frac{h(\alpha)}{\alpha \beta \sqrt{2\pi}}
$$

where

$$
h(\alpha) = \int_0^\infty \frac{1}{v^2 (1 + v^2)} \exp \left\{ -\frac{1}{2\alpha^2} (v-v^{-1})^2 \right\} dv.
$$

Making the change of variable $u = v - v^{-1}$ it follows that

$$
h(\alpha) = \frac{1}{2} \int_{-\infty}^\infty \frac{u^2 + 2 - u \sqrt{u^2 + 4}}{u^2 + 4} \exp \left( -\frac{u^2}{2\alpha^2} \right) du
$$

$$
= \frac{1}{2} \int_{-\infty}^\infty \exp \left( \frac{u^2}{2\alpha^2} \right) du - \int_{-\infty}^\infty \frac{1}{u^2 + 4} \exp \left( -\frac{u^2}{2\alpha^2} \right) du
$$

$$
= \frac{1}{2} \int_{-\infty}^\infty \frac{u}{\sqrt{u^2 + 4}} \exp \left( -\frac{u^2}{2\alpha^2} \right) du
$$

$$
= \alpha \sqrt{\frac{\pi}{2}} - \frac{1}{\sqrt{\alpha}} \frac{1}{\sqrt{u^2 + 4}} \exp \left( -\frac{u^2}{2\alpha^2} \right) du.
$$

Let

$$
g(\xi) = \int_{-\infty}^\infty \frac{1}{u^2 + 4} \exp \left( -\frac{\xi^2 u^2}{2} \right) du,
$$

where $\xi = 1/\alpha$ for notational convenience. Then,

$$
g'(\xi) = -\xi \int_{-\infty}^\infty \frac{u^2}{u^2 + 4} \exp \left( -\frac{\xi^2 u^2}{2} \right) du
$$

$$
= -t \int_{-\infty}^\infty \exp \left( -\frac{\xi^2 u^2}{2} \right) du + 4t \int_{-\infty}^\infty \frac{1}{u^2 + 4} \exp \left( -\frac{\xi^2 u^2}{2} \right) du
$$

$$
= -\sqrt{2\pi} + 4t g(\xi)
$$

Published by Atlantis Press
Copyright: the authors
202
that is,
\[ g'(\xi) = 4t g(\xi) - \sqrt{2\pi}. \]  
(1)

The solution of the differential equation (1) is given by

\[ g(\xi) = e^{2\xi} [C - \pi \Phi(2\xi)]. \]

Notice that \( g(0) = \pi/2 \) and then \( C = \pi \). Thus

\[ \int_{-\infty}^{\infty} \frac{1}{u^2 + 4} \exp \left( -\frac{\xi^2 u^2}{2} \right) du = g(1/\alpha) = \pi e^{2/\alpha^2} [1 - \pi \Phi(2/\alpha)]. \]

Therefore, we have that

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

and hence

\[ K_{\beta\beta} = \frac{1 + \alpha (2\pi)^{-1/2} h(\alpha)}{(\alpha \beta)^2}. \]

By a similar algebraic development (but of independent interest), we can show that

\[ \int_{0}^{\infty} \frac{1}{v^2(1 + v^2)^2} \exp \left\{ -\frac{1}{2\alpha^2} (v - v^{-1})^2 \right\} dv = \alpha \left( \frac{\pi}{2} \right)^{1/2} - \frac{3\pi}{2} e^{2/\alpha^2} \left[ 1 - \Phi \left( \frac{2}{\alpha} \right) \right]. \]

Finally, the R functions used in this paper to compute \( h(\alpha) \) and \( h_1(\alpha) \) in equations (2.1) and (3.1), respectively, are given by

\begin{verbatim}
fh <- function(alpha) {
  h0 <- alpha*sqrt(pi/2)
  h1 <- pi*exp(2/(alpha^2))
  h2 <- 1 - pnorm(2/alpha)
  return(h0 - h1*h2)
}
fh1 <- function(alpha) {
  return((alpha/2)*sqrt(pi/2)*( 1 + (alpha^2)/4 - 3*(alpha^4)/16 + 5*(alpha^6)/64 ))
}
\end{verbatim}

References


Birnbaum–Saunders distribution


