A MODIFIED BESSEL DISTRIBUTION OF THE SECOND KIND

Saralees Nadarajah

1. INTRODUCTION

Univariate Bessel function distributions are rapidly becoming distributions of first choice whenever "something" with heavier than Gaussian tails is observed in the data. They have been used to model: signal output processed by a radar receiver under various sets of conditions (McNolty, 1967) and vibrational amplitude on the surfaces of ultrasonic transducers (Kielczynski and Pajewski, 1993; Kielczynski *et al.*, 1993). They have also been used for modeling problems in applied physics (Salingaros, 1991).

The two standard Bessel function distributions are the Bessel function distribution of the first kind and the Bessel function distribution of the second kind. The latter has the probability density function (pdf) given by

$$f(x) \;=\; \frac{|x|^m}{\sqrt{\pi}\, 2^m\, b^{m+1}\Gamma(m+1/2)} K_m\left(\left|\frac{x}{b}\right|\right) \tag{1}$$

for $-\infty < x < \infty$, $b > 0$ and $m > 1$, where

$$K_m(x) = \frac{\sqrt{\pi}\, x^m}{2^m\, \Gamma(m+1/2)} \int_1^\infty (t^2 - 1)^{m-1/2} \exp(-xt) dt$$

is the modified Bessel function of the second kind. Many generalizations of (1) have been proposed in the literature. See Harris and Soms (1974), Thabane and Kibria (1999), Anh et al. (2005), Gupta and Nadarajah (2006) and Srivastava and Nadarajah (2006). However, most of these generalizations are motivated by mathematical arguments, and not by statistical or physical needs. For example, the pdf of the distribution proposed in Srivastava and Nadarajah (2006) takes the form $C|x|^\alpha K_m(c|x|)K_n(d|x|)$ which has attractive mathematical properties.

The aim of this note is to introduce a modification of (1) that has real statistical motivation and to study its mathematical properties. Particularly, the new distribution is motivated by a Bayesian inference of an inverse Gaussian sample. An inverse Gaussian distribution is given by the pdf

$$f(x) = \frac{\sqrt{\lambda}}{\sqrt{2\pi x^{3/2}}} \exp\left(-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right) \tag{2}$$

for $x > 0$, $\lambda > 0$ and $\mu > 0$, where λ is a reciprocal measure of dispersion and μ is a measure of location. Sometimes, it is more convenient to rewrite (2) as

$$f(x) = \frac{\sqrt{\lambda}\exp(\phi)}{\sqrt{2\pi x^{3/2}}} \exp\left(-\frac{\phi^2 x}{2\lambda} - \frac{\lambda}{2x}\right) \tag{3}$$

where $\phi = \lambda/\mu$. The parameter ϕ determines the shape of the distribution and the pdf is highly skewed for moderate values of ϕ. As ϕ increases the inverse Gaussian tends towards the normal law.

The inverse Gaussian distribution is one of the most applied in the sciences. Areas of application include: actuarial science, analysis of reciprocals, demography, histomorphometry, electric networks, hydrology, life tests, management science, meteorology, mental health, physiology, remote sensing, traffic noise intensity, market research, regression, slug lengths in pipelines, ecology, entomology, small area estimation, CUSUM, and plutonium estimation. For details on the theory and applications of the distribution, we refer the reader to Chhikara and Folks (1989) and Seshadri (1998).

Suppose now we have a random sample $\mathbf{x} = (x_1,...,x_m)$ from (3) with parameters (λ, ϕ). We wish to make inferences about the value of ϕ, e.g. $H_0 : \phi = c$ versus $H_1 : \phi \neq c$. The joint pdf of $\mathbf{x}$ is

$$f(\mathbf{x}\,|\,\lambda,\phi) \propto (2\pi)^{-m/2} \exp(m\phi)\lambda^{m/2} \prod_{i=1}^{m} x_i^{-3/2} \exp\left\{-\frac{\phi^2}{2\lambda}\sum_{i=1}^{m} x_i - \frac{\lambda}{2}\sum_{i=1}^{m}\frac{1}{x_i}\right\} \tag{4}$$

Consider the prior pdf

$$\pi(\lambda,\phi) \propto \lambda^{-1} \exp(-m\phi - n\phi^2)\,, \tag{5}$$

a diffuse prior in terms of λ and a normal in terms of ϕ. We obtain the joint posterior pdf as

$$\pi(\lambda,\phi\,|\,\mathbf{x}) \propto (2\pi)^{-m/2}\lambda^{m/2-1} \prod_{i=1}^{m} x_i^{-3/2} \exp\left\{-n\phi^2 - \frac{\phi^2}{2\lambda}\sum_{i=1}^{m} x_i - \frac{\lambda}{2}\sum_{i=1}^{m}\frac{1}{x_i}\right\}.$$

Integrating out λ using equation (2.3.16.1) in Prudnikov *et al.* (1986, volume 1), we obtain the marginal posterior pdf of ϕ as

$$\pi(\phi\,|\,\mathbf{x}) \propto \phi^{m/2} \exp(-n\phi^2) K_{m/2}\left(\phi\sum_{i=1}^{m} x_i \sum_{j=1}^{m}\frac{1}{x_j}\right). \tag{6}$$

Thus, given the joint density (4) and the prior (5), the marginal posterior density of ϕ takes the form of

$$f(\phi) = C \, |\phi|^m \, \exp(-p\phi^2) K_m\left(\frac{|\phi|}{b}\right) \tag{7}$$

for $-\infty < \phi < \infty$, $b > 0$, $p > 0$ and $m > 1$, where C denotes the normalizing constant. Application of equation (2.16.8.5) in Prudnikov *et al.* (1986, volume 2) shows that one can determine C as

$$\frac{1}{C} = \frac{\Gamma(m+1)}{b^m \, 2^{m+2} \, p^{m+1}} \exp\left(\frac{1}{4pb^2}\Gamma\right)\left(-m, \frac{1}{4pb^2}\right)$$

where $\Gamma(\cdot,\cdot)$ denotes the complementary incomplete gamma function defined by

$$\Gamma(a,x) = \int\limits_{x}^{\infty} t^{a-1} \exp(-t)dt$$

Hence, by studying the mathematical properties of (7) one can make Bayesian inferences about ϕ.

The rest of this note is organized as follows: various expressions for particular forms of (7) and its moments are derived in Sections 2 and 3, respectively; estimation procedures are considered in Section 4; and, an application is discussed in Section 5.

2. PARTICULAR CASES

When m takes half-integer values one can reduce (7) to elementary forms. Using the results in Appendix A, several particular forms of (7) can be obtained for half-integer values of m. For example, if $m = 3/2$ then (7) reduces to

$$f(x) = Cb^{3/2}\sqrt{\frac{\pi}{2}}\exp(-px^2)\exp\left(-\frac{x}{b}\right)\left(\frac{x}{b}+1\right)$$

with C given by

$$\frac{1}{C} = \frac{\sqrt{\pi}}{2\sqrt{2b}\,p^{3/2}}\left\{2b\sqrt{p} - \sqrt{x}(1-2b^2p)\exp\left(\frac{1}{4b^2p}\right)\text{erfc}\left(\frac{1}{2b\sqrt{p}}\right)\right\},$$

where $\text{erfc}(\cdot)$ denotes the complementary error function defined by

$$\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \exp(-t^2)dt \, .$$

If $m = 5/2$ then (7) reduces to

$$f(x) = Cb^{5/2} \sqrt{\frac{\pi}{2}} \exp(-px^2) \exp\left(-\frac{x}{b}\right)\left(\frac{x^2}{b^2} + \frac{3x}{b} + 3\right)$$

with C given by

$$\frac{1}{C} = \frac{\sqrt{\pi}}{4\sqrt{2}b^{3/2}p^3}\left\{2bp(6b^2p - 1) + \sqrt{p\pi}(1 - 4pb^2 + 12p^2b^4)\exp\left(\frac{1}{4b^2p}\right)\mathrm{erfc}\left(\frac{1}{2b\sqrt{p}}\right)\right\}.$$
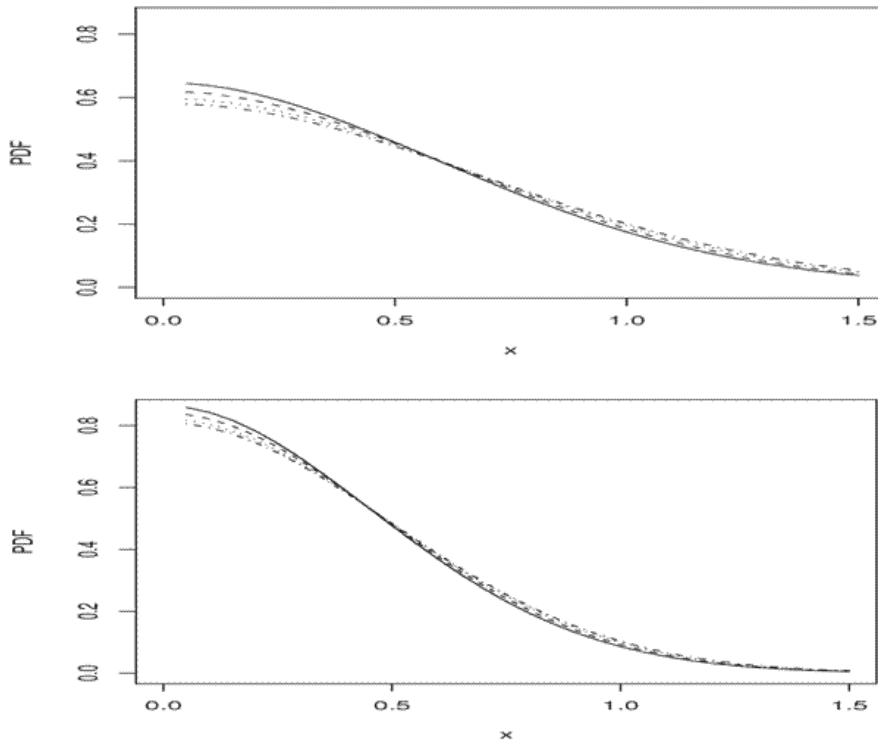

Figure 1 – Plots of the pdf (7) for $b = 1$, $p = 1$ and $m = 1.5$, 2, 3, 5 (top); and, $b = 1$, $p = 2$ and $m = 1.5$, 2, 3, 5 (bottom). The four curves in each plot are: the solid curve ($m = 1.5$), the curve of lines ($m = 2$), the curve of dots ($m = 3$), and the curve of lines and dots ($m = 5$).

Figure 1 illustrates possible shapes of the pdf (7) for selected values of m and p. The four curves in each plot correspond to selected values of m. The effect of the parameters is evident.

3. MOMENTS

Suppose X is a random variable with pdf (7). If n is odd then it is obvious that $E(X^n) = 0$. If n is even then one can write

$$E(X^n) = 2C \int_0^\infty x^{m+n} \exp(-px^2) K_m\left(\frac{x}{b}\right) dx .$$

By application of equation (2.16.8.4) in Prudnikov *et al.* (1986, volume 2), the above can be calculated as

$$E(X^n) = \frac{Cb}{p^{(m+n)/2}} \Gamma\left(m + \frac{n+1}{2}\right) \Gamma\left(\frac{n+1}{2}\right) \exp\left(\frac{1}{8b^2 p}\right) W_{-(m+n)/2, m/2}\left(\frac{1}{4b^2 p}\right) \quad (8)$$

where W denotes the Whittaker function defined by

$$W_{\lambda, \nu^{(x)}} = \frac{x^{\mu+1/2} \exp(-x/2)}{\Gamma(\mu - \lambda + 1/2)} \int_0^\infty \exp(-xt) t^{\mu-\lambda-1/2} (1+t)^{\mu+\lambda-1/2} dt .$$

Using special properties of the Whittaker function, one can obtain the simpler form

$$E(X^n) = \frac{C\Gamma((n+1)/2)\Gamma(m+(n+1)/2)}{2^m p^{m+(n+1)/2} b^m (2m+n-1)(1-m)_{m+(n-1)/2}}$$

$$\times \left(\frac{d}{dz}\right)^{m+(n-1)/2} \{z^{(n-1)/2} \exp(z) \Gamma(m, z)\}\Big|_{z=1/(4b^2 p)} , \quad (9)$$

for half integer values of m. The expressions in (8) and (9) can be used to obtain Bayes estimators of the parameters of (6).

4. MAXIMUM LIKELIHOOD ESTIMATION

The highest posterior density set corresponding to (6) cannot be obtained in closed form. However, an approximation method is possible via a normal approximation of the posterior density. The normal approximation requires finding the maximum likelihood estimates of the parameters (b, p, m) in (7).

Suppose $X_1, X_2, ..., X_n$ is a random sample from (7). Then the log likelihood function for (b, p, m) is

$$\log L(b, p, m) = n \log C + m \sum_{i=1}^n \log |X_i| - p \sum_{i=1}^n X_i^2 + \sum_{i=1}^n \log K_m\left(\left|\frac{X_i}{b}\right|\right).$$

The derivatives with respect to the three parameters are:

$$\frac{\partial \log L}{\partial b} = \frac{n \partial C}{C \partial b} + \sum_{i=1}^{n} \frac{|X_i|}{b^2} \left\{ \frac{K_{m+1}(|X_i|^{/b})}{K_m(|X_i|^{/b})} - \frac{mb}{|X_i|} \right\}, \tag{10}$$

$$\frac{\partial \log L}{\partial p} = \frac{n \partial C}{C \partial p} - \sum_{i=1}^{n} X_i^2 \tag{11}$$

and

$$\frac{\partial \log L}{\partial m} = \frac{n \partial C}{C \partial m} + \sum_{i=1}^{n} \log |X_i| + \sum_{i=1}^{n} \frac{\partial K_m(|X_i|^{/b})^{/\partial m}}{K_m(|X_i|^{/b})}. \tag{12}$$

The partial derivatives in (10)-(12) can be computed by using the facts that

$$\frac{\partial \Gamma(a,z)}{\partial z} = -z^{a-1} \exp(-z),$$

$$\frac{\partial \Gamma(a,z)}{\partial a} = \Gamma^2(a) z^a {}_2F_2(a,a;a+1,a+1;-z) - \gamma(a,z) \log z + \Gamma(a)\psi(a),$$

and

$$\frac{\partial K_\nu(z)}{\partial \nu} = \frac{\pi \csc(\pi \nu)}{2} [-2\cos(\pi \nu) K_\nu(z) - \log\left(\frac{z}{2}\right)\{I_{-\nu}(z) + I_\nu(z)\}$$

$$+ \sum_{K=0}^{\infty} \left\{ \frac{\psi(k-\nu+1)}{k!\Gamma(k-\nu+1)} \left(\frac{z}{2}\right)^{2k-\nu} + \frac{\psi(k+\nu+1)}{k!\Gamma(k+\nu+1)} \left(\frac{z}{2}\right)^{2k+\nu} \right\}],$$

where $\psi(z) = \partial \log \Gamma(z) / \partial z$ is the digamma function,

$$\gamma(a,x) = \int_0^x t^{a-1} \exp(-t) dt$$

is the incomplete gamma function,

$$I_m(x) = \frac{x^m}{\sqrt{\pi} 2^m \Gamma(m+1/2)} \int_{-1}^{1} (1-t^2)^{m-1/2} \exp(\pm xt) dt$$

is the modified Bessel function of the first kind, and

$$2^F 2^{(a,b;c,d;x)} = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k (d)_k}$$

is the hypergeometric function, where $(e)_k = e(e + 1) \cdots (e + k - 1)$ denotes the ascending factorial. The maximum likelihood estimators of (b,p,m) are the simultaneous solutions of the equations $\partial \log L / \partial b = 0, \partial \log L / \partial p = 0$ and $\partial \log L / \partial m = 0$.

5. APPLICATION

We use the data given by Section 5.7 of Chhikara and Folks (1989), reproduced in Table 1 below.

TABLE 1

Failure data of 23 ball bearings

17.88	28.92	33.00	41.52	42.12	45.60	48.48
51.84	51.96	54.12	55.56	67.80	68.64	68.64
68.88	84.12	93.12	98.64	105.12	105.84	127.92
128.04	173.40					

The data consist of the number of million revolutions before failure for each of 23 ball bearings used in a life test. The background details about the data can be found in Lieblein and Zelen (1956). Chhikara and Folks (1989) showed by means of the Kolmogorov-Smirnov test that the inverse Gaussian distribution in (3) provides a good fit to the data in Table 1. If we assume that the parameters (λ, ϕ) have the prior given by (5) with $n = 1/2$, corresponding to unit standard deviation, then the posterior for ϕ, the shape parameter, is

$$\pi(\phi \mid \mathbf{x}) \propto \phi^{23/2} \exp\left(-\frac{\phi^2}{2}\right) K_{23/2}\left(\phi \sum_{i=1}^{23} x_i \sum_{j=1}^{23} \frac{1}{x_j}\right).$$

Using the results in Section 3, we obtain the Bayes estimate of ϕ as

$$\frac{\int_0^{\infty} \phi^{25/2} \exp\left(-\frac{\phi^2}{2}\right) K_{23/2}\left(\phi \sum_{i=1}^{23} x_i \sum_{j=1}^{23} \frac{1}{x_j}\right) d\phi}{\int_0^{\infty} \phi^{23/2} \exp\left(-\frac{\phi^2}{2}\right) K_{23/2}\left(\phi \sum_{i=1}^{23} x_i \sum_{j=1}^{23} \frac{1}{x_j}\right) d\phi} = 0.005788836.$$

This estimate suggests that the data are highly skewed. The earliest and the most popular model used to describe failure data is the exponential distribution. So, it is of interest to see whether the exponential model provides an adequate fit

for the data. Since the exponential and inverse Gaussian distributions are not nested, we test the adequacy by comparing the skewness values. It is known that the skewness for the two distributions are 2 and $3/\sqrt{\phi}$, respectively. So, testing for adequacy of the exponential model amounts to testing $H_0 : \phi = 2.25$ versus $H_1 : \phi \neq 2.25$. Using the likelihood ratio, we obtain:

$$-2\{\log L(2.25 \,|\, \mathbf{x}) - \log L(0.006 \,|\, \mathbf{x}\} = 1505.880\,.$$

Since $1505.880 > 3.841 = \chi^2_{1,0.95}$ (we have assumed χ^2_1 distribution to be consistent with the classical likelihood inference although it may not hold exactly in the case of Bayesian inference), there is overwhelming evidence that the exponential model cannot provide an adequate fit for the data. This is consistent with findings in Lieblein and Zelen (1956).

School of Mathematics   SARALEES NADARAJAH
University of Manchester, UK

APPENDIX A: PARTICULAR FORMS OF $K_\nu(x)$

Note that

$$K_{5/2}(x) = \sqrt{\frac{\pi}{2}} \frac{\exp(-x)(x^2 + 3x + 3)}{x^{5/2}},$$

$$K_{3/2}(x) = \sqrt{\frac{\pi}{2}} \frac{\exp(-x)(x + 1)}{x^{3/2}},$$

$$K_{7/2}(x) = \sqrt{\frac{\pi}{2}} \frac{\exp(-x)(x^3 + 6x^2 + 15x + 15)}{x^{7/2}} \text{ and}$$

$$K_{9/2}(x) = \frac{\pi}{\sqrt{2}} \frac{\exp(-x)(x^4 + 10x^3 + 45x^2 + 105x + 105)}{x^{9/2}}.$$

More generally, if $\nu - 1/2 \geq 1$ is an integer then

$$K_\nu(x) = \sqrt{\pi} \exp(-x)\sqrt{2x} \sum_{j=0}^{[|\nu|-1/2]} \frac{(j + |\nu| - 1/2)!(2x)^{-j}}{j!(|\nu| - j - 1/2)!}.$$

ACKNOWLEDGMENTS

REFERENCES

V. V. ANH, R. MCVINISH, C. PESEE, (2005), *Estimation and simulation of the Riesz-Bessel distribution*, "Communications in Statistics-Theory and Methods", 34, pp. 1881-1897.

R. CHHIKARA, L. FOLKS, (1989), *The Inverse Gaussian distribution: theory, methodology, and applications*, Marcel Dekker, New York.

A.K. GUPTA, S. NADARAJAH, (2006), *Beta Bessel distributions*, "International Journal of Mathematics and Mathematical Sciences", Article ID 16156.

B. HARRIS, A.P. SOMS, (1974), *Properties of the generalized incomplete modified Bessel distributions with applications to reliability theory*, "Journal of the American Statistical Association", 69, pp. 259-263.

P. KIELCZYNSKI, W. PAJEWSKI, (1993), *Acoustic field of gaussian and Bessel transducers*, "Journal of the Acoustical Society of America", 94, pp. 1719-1721.

P. KIELCZYNSKI, W. PAJEWSKI, M. SZALEWSKI, (1993), *Ultrasonic transducers with Bessel–function distribution of vibrational amplitude on their surfaces*, "IEEE Transactions on Robotics and Automation", 9, pp. 732-739.

J. LIEBLEIN, M. ZELEN, (1956), *Statistical investigation of the fatigue life of deep-groove ball bearings*, "Journal Research of the National Bureau of Standards", 57, pp. 273-316.

F. MCNOLTY, (1967), *Applications of Bessel function distributions*, "Sankhyā", B, 29, pp. 235-248.

A.P. PRUDNIKOV, Y.A. BRYCHKOV, O.I. MARICHEV, (1986), *Integrals and series*, volumes 1, 2 and 3, Gordon and Breach Science Publishers, Amsterdam.

N.A. SALINGAROS, (1991), *Optimal current distribution for energy-storage in superconducting magnets*, "Journal of Applied Physics", 69, pp. 531-533.

V. SESHADRI, (1986), *The Inverse Gaussian distribution: statistical theory and applications*, Springer Verlag, New York.

H.M. SRIVASTAVA, S. NADARAJAH, (2006), *Some families of Bessel distributions and their applications*, "Integral Transforms and Special Functions", 17, pp. 65-73.

L. THABANE, B.M.G. KIBRIA, (1999), *On the generalized modified power Bessel distribution.* "Journal of Statistical Research", 33, pp. 9-22.

SUMMARY

A modified Bessel distribution of the second kind

Motivated by a Bayesian inference problem, a modification of the Bessel function distribution is introduced. Various particular cases, expressions for its moments and estimation procedures are derived. An application is illustrated to failure data.