Ann. Inst. Statist. Math.
Vol. 52, No. 3, 438–447 (2000)

ON THE BESSEL DISTRIBUTION AND RELATED PROBLEMS

LIN YUAN AND JOHN D. KALBFLEISCH

Department of Statistics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

(Received October 13, 1997; revised February 9, 1999)

Abstract. This article investigates basic properties of the Bessel distribution, a power series distribution which has not been fully explored before. Links with some well-known distributions such as the von Mises-Fisher distribution are described. A simulation scheme is also proposed to generate random samples from the Bessel distribution. This scheme is useful in Bayesian inferences and Monte Carlo computation.

Key words and phrases: Bessel distribution, randomized gamma distribution, von Mises-Fisher distribution, Bessel quotient, simulation, path integral.

1. Introduction

The interest on probability distributions involving Bessel functions can be traced back to the early work of McKay (1932) and Laha (1954) in which two classes of continuous distributions, called Bessel function distributions of type I and type II, are studied. Don McLeish (1982) further investigates the application of the Bessel function distributions of type II and some closely related results to create a robust alternative to the normal distribution. Devroye (1986) provides more discussion on Bessel function distributions.

Unlike the previous work, the Bessel distribution we study here is a power series distribution generated by the first type of (modified) Bessel function. A random variable Y, taking values from non-negative integers, is said to be a Bessel random variable with parameters $\nu > -1$ and $a > 0$ if

$$\text{(1.1)} \qquad \Pr(Y = n) = \frac{1}{I_\nu(a) n! \Gamma(n + \nu + 1)} \left(\frac{a}{2}\right)^{2n+\nu}, \qquad n = 0, 1, \dots,$$

where $I_\nu(x)$ denotes the first type of (modified) Bessel function given by

$$I_\nu(x) = \left(\frac{x}{2}\right)^\nu \sum_{n=0}^{\infty} \frac{1}{n! \Gamma(n + \nu + 1)} \left(\frac{x}{2}\right)^{2n}, \qquad x > 0, \quad \nu > -1.$$

It should be noted that this distribution (1.1) also arises in Pitman and Yor (1982); see Section 3 for discussion. For simplicity we use the notation $Bes(\nu, a)$ for the Bessel distribution with parameters ν and a.

The Bessel distribution (1.1) can be thought of as an inverse probability as is illustrated by the following example. Assigning a gamma prior to the mean of a Poisson distribution is a standard procedure in Bayesian statistics; in this example, however, we put a Poisson prior on a gamma distribution. Suppose we study the number of customers visiting a laundromat with reference to the power consumption. The observable total power consumption Y in a period T breaks into two parts: the customer consumption Y_1 and an independent base amount Y_2. We assume the power consumption of each customer is independently an exponential random variable with scale a, and the distribution of Y_2 is $G(\nu+1,a)$, the gamma distribution with shape parameter $\nu+1$ and scale a where $\nu > -1$. If the number of customers is r, the distribution of Y given r is $G(r+\nu+1,a)$. Suppose that customer arrivals are described by a Poisson process with rate λ. The prior distribution for r is then Poisson with mean λT. Given an observation $Y=y$, it follows that the posterior distribution of r is $Bes(\nu, 2\sqrt{a\lambda Ty})$.

The Bessel distribution is related to many distributions with Bessel functions involved in the density, the von Mises-Fisher distribution and the squared Bessel bridge, for example. These distributions are mixtures involving the Bessel distribution.

This article is organized as follows. Section 2 gives some basic results about the Bessel distribution. Section 3 introduces the gamma distribution randomized by a Bessel distribution, and its link to an appealing multivariate gamma distribution. Section 4 presents a simple scheme for simulating the Bessel distribution, and Section 5 discusses some applications.

2. Some properties of Bessel distributions

This section examines the Bessel distribution and describes links to distributions that are fimiliar to most readers.

(i) *The Bessel distribution as a conditional Poisson distribution.* Let $P(\lambda)$ denote the Poisson distribution with mean λ. When ν is an integer, it is easily seen that the Bessel distribution $Bes(\nu,a)$ is the conditional distribution of Y given $X-Y=\nu$, where $X \sim P(\lambda_1)$ and $Y \sim P(\lambda_2)$ are independent and $\lambda_1\lambda_2 = a^2/4$.

For the general case $\nu \geq 0$, X is generated from a randomized Poisson distribution. More specifically, $X \sim P(\lambda_1 - \eta)$ with η distributed as a $G(\nu-[\nu],1)$ but variate right truncated at λ_1 (to include the integer case we adopt the convention that $G(0,1)$ denotes the probability distribution concentrated on zero). Now, the density of η is proportional to $\eta^{\nu-[\nu]-1}e^{-\eta}I(0<\eta<\lambda_1)$, so that,

$$\begin{aligned}\Pr(X=k) &\propto \int \Pr(X=k\mid\eta)\eta^{\nu-[\nu]-1}e^{-\eta}I(0<\eta<\lambda_1)d\eta \\ &= \frac{e^{-\lambda_1}}{\Gamma(k+1)}\int_0^{\lambda_1}\eta^{\nu-[\nu]-1}(\lambda_1-\eta)^k d\eta \\ &\propto \frac{\lambda_1^{k+\nu-[\nu]}}{\Gamma(k+1+\nu-[\nu])}.\end{aligned}$$

Now, it can be seen that for $\nu \geq 0$, the conditional distribution of Y given $X-Y=[\nu]$ is $Bes(\nu,a)$.

(ii) *The Bessel distribution as a sum of Bernoulli variables.* It is well-known that the Bessel function satisfies the recurrence equation

$$I_\nu(x) = I_{\nu+2}(x) + \frac{2(\nu+1)}{x}I_{\nu+1}(x),$$

which implies a kind of relation between $Bes(\nu,a)$, $Bes(\nu+1,a)$ and $Bes(\nu+2,a)$. In fact, it is immediately seen that, the Bessel distribution $Bes(\nu,a)$ is a mixture of $Bes(\nu+1,a)$ and a right-shifted $Bes(\nu+2,a)$ produced by moving the mass at each integer k to $k+1$. The weights for this mixture are $2(\nu+1)R_\nu(a)/a$ and $R_\nu(a)R_{\nu+1}(a)$ respectively, where $R_\nu(a) = I_{\nu+1}(a)/I_\nu(a)$ is called the Bessel quotient. In the language of sampling, a random variable $Y \sim Bes(\nu,a)$ can be generated by first generating a Bernoulli random variable η with $\Pr(\eta = 1) = R_\nu(a)R_{\nu+1}(a)$ followed by $X \sim Bes(\nu + \eta + 1, a)$, and then $Y = X + \eta$.

From this property, we see that a Bessel random variable can be expressed as a sum of dependent Bernoulli variables: First, a random variable $Y \sim Bes(\nu, a)$ can be written as $Y = \eta_1 + X_1$ with η_1 a Bernoulli variable with parameter $R_\nu(a)R_{\nu+1}(a)$ and $X_1 \sim Bes(\nu + \eta_1 + 1, a)$. Then, X_1 can be written as $X_1 = \eta_2 + X_2$ with η_2 a Bernoulli variable with parameter $R_{\nu+\eta_1+1}(a)R_{\nu+\eta_1+2}(a)$ and $X_2 \sim Bes(\nu + \eta_1 + \eta_2 + 2, a)$. Since $Bes(\nu + k, a)$ can be treated as a point mass on zero for k large enough, we can express Y as an infinite sum of Bernoulli variables $\sum_{i=1}^\infty \eta_i$.

(iii) *Relationship to spherical distributions.* The von Mises distribution is an analogue of the normal distribution in circular statistics. Its density function is

$$\phi(\theta) = \frac{1}{2\pi I_0(\kappa)} \exp(\kappa \cos\theta), \qquad -\pi \le \theta < \pi, \quad \kappa > 0,$$

where κ is the concentration parameter. Detailed study and interesting applications can be found in Mardia (1972). The Bessel distribution offers a simple description for the von Mises distribution. If θ is a von Mises variable with concentration parameter κ, then the distribution of $\cos^2\theta$ is a randomized beta distribution $Beta(\xi + 1/2, 1/2)$ where $\xi \sim Bes(0, \kappa)$. However, given $\cos^2\theta = y$ there are still four possible values of θ in $[-\pi, \pi)$. Further analysis of the conditional distribution of θ given $\cos^2\theta = y$ we have the following relation. Suppose $Y \sim Beta(\xi + 1/2, 1/2)$ where $\xi \sim Bes(0, \kappa)$, and given Y, b_1 and b_2 are conditionally independent Bernoulli variables with parameters $1/2$ and $1/(1 + e^{-2\kappa\sqrt{Y}})$ respectively. Then,

$$\theta = (2b_1 - 1) \arccos[(2b_2 - 1)\sqrt{Y}]$$

is a von Mises variable with concentration parameter κ.

As a generalization the von Mises-Fisher distribution $M_r(\mu, \kappa)$ is a distribution on a r-dimensional hypersphere with density

$$\frac{\kappa^{(1/2)r-1}}{(2\pi)^{(1/2)r}} I_{(1/2)r-1}(\kappa) \exp(\kappa\langle\mu, y\rangle)$$

where $\langle\mu, x\rangle = \sum_{i=1}^r \mu_i y_i$ is the inner product and $|\mu| = 1$ and $\kappa > 0$ are parameters.

To describe the joint distribution of $(Y_1, \ldots, Y_r) \sim M_r(\mu, \kappa)$ we introduce a randomized multinomial distribution. Suppose that $(N_1, \ldots, N_r) \sim multinomial(N, p_1, \ldots, p_r)$ with $N \sim Bes(\nu, a)$. The joint distribution of $(N_1, \ldots, N_r)$ is then

$$\Pr(N_1 = n_1, \ldots, N_r = n_r) \propto \frac{1}{n_1! \cdots n_r! \Gamma(n_1 + \cdots + n_r + \nu + 1)} \left(\frac{a\sqrt{p_1}}{2}\right)^{2n_1} \cdots \left(\frac{a\sqrt{p_r}}{2}\right)^{2n_r}.$$

This distribution can be used to describe the von Mises-Fisher distribution $M_r(\mu, \kappa)$ on a r-dimensional hypersphere.

The joint density of $(Y_1^2, \ldots, Y_r^2)$ is proportional to

$$I\left(\sum_{i=1}^r y_i = 1\right) \prod_{i=1}^r \frac{\cosh \kappa\mu_i\sqrt{y_i}}{\sqrt{y_i}} = I\left(\sum_{i=1}^r y_i = 1\right) \sum_{n_1=0}^\infty \cdots \sum_{n_r=0}^\infty \prod_{i=1}^r \kappa^{2n_i} \mu_i^{2n_i} \frac{y_i^{n_i-(1/2)}}{(2n_i)!}.$$

Note that

$$\prod_{i=1}^{r} \kappa^{2n_i} \mu_i^{2n_i} \frac{y_i^{n_i-(1/2)}}{(2n_i)!} = \frac{\pi^{r/2}}{\Gamma\left(n_1 + \cdots + n_r + \frac{r}{2}\right)} \left[\prod_{i=1}^{r} \frac{1}{n_i!} \left(\frac{\kappa\mu_i}{2}\right)^{2n_i}\right]$$
$$\times \frac{\Gamma\left(n_1 + \cdots + n_r + \frac{r}{2}\right)}{\Gamma\left(n_1 + \frac{1}{2}\right) \cdots \Gamma\left(n_r + \frac{1}{2}\right)} y_1^{n_1-(1/2)} \cdots y_r^{n_r-(1/2)},$$

which is a mixture of Dirichlet distribution $D(n_1 + \frac{1}{2}, \ldots, n_r + \frac{1}{2})$ with $(n_1, \ldots, n_r)$ from a randomized multinomial distribution $multinomial(N, \mu_1^2, \ldots, \mu_r^2)$ where $N \sim Bes(\frac{r}{2} - 1, \kappa)$.

(iv) *Moments and mode.* The moments of the Bessel distribution can be expressed in terms of the Bessel quotient. For instance, if $Y \sim Bes(\nu, a)$, then

$$EY = \frac{1}{2} a R_\nu(a) \quad \text{and} \quad EY^2 = \frac{1}{4} a^2 R_\nu(a) R_{\nu+1}(a) + \frac{1}{2} a R_\nu(a). \tag{2.1}$$

The factorial moments,

$$EY(Y-1)\cdots(Y-k+1) = \left(\frac{a}{2}\right)^k R_\nu(a) \cdots R_{\nu+k-1}(a), \qquad k = 1, 2, \ldots,$$

are easily obtained and from these we can calculate the moment of any order.

The Bessel distribution has a unique mode, or two modes at consecutive integers. For convenience we make the convention that the mode of a Bessel distribution always refers to the larger one if there are two modes and, it then follows that the mode of $Bes(\nu, a)$ is the integer part of $m(\nu, a) = (\sqrt{a^2 + \nu^2} - \nu)/2$. This is useful in simulating the Bessel distributions.

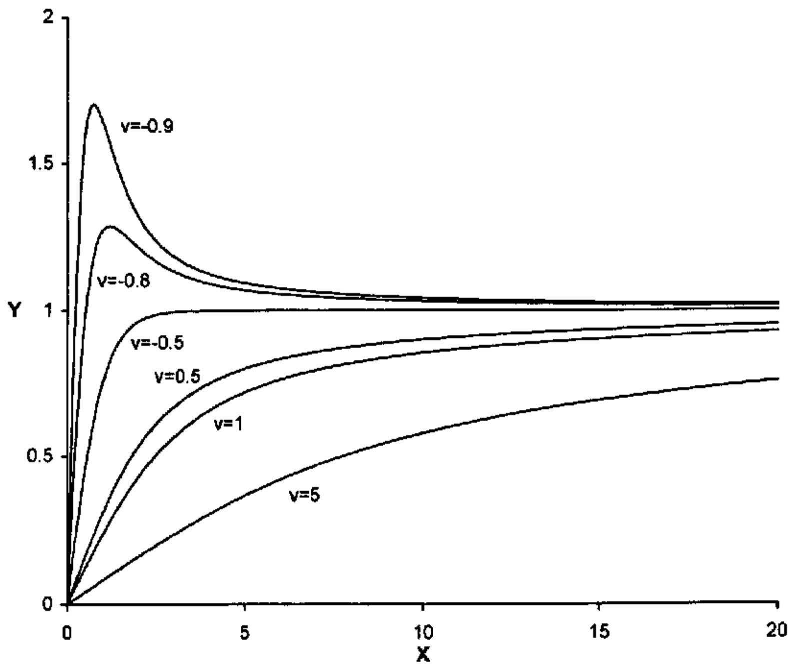

Fig. 1. The Bessel quotient for different ν values.

(v) *The Bessel quotient.* It should be noted that the evaluation of the Bessel functions is avoided due to the inefficiency of the existing numerical methods. The Bessel quotient is much more stable than the Bessel function and can be evaluated using a continued fraction (Amos (1974)).

As shown in Fig. 1, all curves of functions $R_\nu(x)$, $\nu > -1$ start from $(0,0)$ and share the same asymptote $y = 1$. The monotonicity is classified into two categories according to whether ν is larger than $-\frac{1}{2}$ or not. For $\nu > -\frac{1}{2}$, the function $R_\nu(x)$ is increasing over the whole interval $(0,\infty)$; while for $-1 < \nu < -\frac{1}{2}$, the function $R_\nu(x)$ is increasing first to reach a maximum and then decreasing. We also see that $R_\nu(x) \leq R_\mu(x)$ if $\nu > \mu$ and the inequality holds strictly in $(0,\infty)$. Further discussion of the Bessel quotient is given in the Appendix.

3. Multivariate and randomized gamma distributions

Several multivariate extensions of the gamma distribution are available in literatures. Prominent among these are the two basic approaches introduced in Johnson and Kotz (1970). The first approach is to create some kind of linear combinations of independent gamma variables. More complete references of this type of work can be found, for example, in Mathai and Moschopoulos (1991). The second approach works with the multivariate integral transforms such as Fourier or Laplace transforms; see for example, Richards (1986) and Pinky (1993).

We begin with a simple construction of a bivariate gamma distribution, which is a special case of the second approach as suggested by Johnson and Kotz (1970). Consider the joint Laplace transform of independent random variables $Y_1 \sim G(\alpha, \lambda_1)$ and $Y_2 \sim G(\alpha, \lambda_2)$,

$$Ee^{-t_1Y_1-t_2Y_2} = \left(1+\frac{t_1}{\lambda_1}\right)^{-\alpha}\left(1+\frac{t_2}{\lambda_2}\right)^{-\alpha} = [\det(I_2 + AT)]^{-\alpha},$$

where I_2 is a 2×2 identity matrix, A and T are 2×2 diagonal matrices with entries $1/\lambda_1$, $1/\lambda_2$ and t_1, t_2 respectively. One way to construct a bivariate gamma distribution is to alter the matrix A to a symmetric positive definite matrix. Therefore, a more general form of A can be obtained if we replace the two zeroes in A by $\sqrt{\rho/\lambda_1\lambda_2}$ with $0 \leq \rho < 1$. The corresponding Laplace transform

$$\left[1+\frac{t_1}{\lambda_1}+\frac{t_2}{\lambda_2}+\frac{(1-\rho)t_1t_2}{\lambda_1\lambda_2}\right]^{-\alpha} \tag{3.1}$$

possesses a closed-form inverse. Lengthy but straightforward calculation shows that the density function corresponding to the Laplace transform (3.1) is proportional to

$$(y_1y_2)^{(\alpha-1)/2}\exp\left(-\frac{\lambda_1y_1+\lambda_2y_2}{1-\rho}\right)I_{\alpha-1}\left(\sqrt{\frac{4\rho\lambda_1\lambda_2y_1y_2}{1-\rho}}\right), \quad y_1, y_2 > 0. \tag{3.2}$$

The density (3.2) retains the marginal distributions of Y_1 and Y_2, and the main point of interest is the dependence between Y_1 and Y_2. Specifically, we are interested in the conditional distribution of Y_1 given Y_2, or vice versa. It is seen that both conditional distributions are the Bessel function distributions of the first type, or randomized gamma distributions as named by Feller (1966).

Suppose that $Y \mid \eta \sim G(\alpha+\eta, \lambda)$ with α, $\lambda > 0$ and η a Poisson variable. Then the marginal distribution of Y is called a randomized gamma distribution of the first type. One can easily verify that the conditional distribution of Y_1 given $Y_2 = y_2$ is a randomized gamma distribution $G(\alpha+\eta, \lambda_1/(1-\rho))$ with $\eta \sim P(\rho\lambda_2 y_2)$. The constant ρ called coefficient of dependence measures the degree of association between Y_1 and Y_2, and a larger value of ρ indicates a stronger dependence.

Can we construct a multivariate gamma distribution in this way? The problem is that a closed form density may not always be available. We restrict our attention to a special case, namely, the multivariate gamma with Markov dependence where a closed form density does exist.

A random vector $(Y_1, \ldots, Y_n)$ is said to be Markov if the conditional distribution of Y_i given $Y_1, \ldots, Y_{i-1}$ is the same as that of Y_i given Y_{i-1}. Let ρ_i denote the dependence coefficient between Y_{i-1} and Y_i. If we wish that each marginal distribution of Y_i is $G(\alpha, \lambda_i)$, then under the assumption of Markov dependence, the joint density of $(Y_1, \ldots, Y_n)$ is proportional to

$$(3.3) \qquad (y_1 y_n)^{(\alpha-1)/2} \exp\left[-\sum_{i=1}^{n} \frac{(1-\rho_i\rho_{i+1})\lambda_i y_i}{(1-\rho_i)(1-\rho_{i+1})}\right] \prod_{i=1}^{n-1} I_{\alpha-1}\left(\sqrt{\frac{4\rho_{i+1}\lambda_i\lambda_{i+1}y_i y_{i+1}}{1-\rho_{i+1}}}\right), \quad y_i > 0,$$

where $\rho_1 = \rho_{n+1} = 0$.

The randomized gamma distribution of the second type is the mixture distribution $G(\alpha+\eta_1+2\eta_2, \lambda)$, α, $\lambda > 0$, where η_1 and η_2 are independent with Poisson and Bessel distributions respectively. This can be viewed as a generalization of the first type taking η_2, in a limiting case, as zero. For any positive numbers a, b, λ and α, the randomized gamma $G(\alpha+\eta_1+2\eta_2, \lambda)$ with $\eta_1 \sim P((a+b)/(4\lambda))$ and $\eta_2 \sim Bes(\alpha-1, \sqrt{ab}/(2\lambda))$ independent has a density function proportional to $e^{-\lambda y} I_{\alpha-1}(\sqrt{ay}) I_{\alpha-1}(\sqrt{by})$, $y > 0$.

A randomized gamma distribution of the second type arises from (3.3) when we consider the conditional distribution of Y_i given $Y_{i-1} = y_{i-1}$ and $Y_{i+1} = y_{i+1}$. The conditional density is proportional to

$$\exp\left[-\frac{(1-\rho_i\rho_{i+1})\lambda_i y_i}{(1-\rho_i)(1-\rho_{i+1})}\right] I_{\alpha-1}\left(\sqrt{\frac{4\rho_i\lambda_{i-1}\lambda_i y_{i-1} y_i}{1-\rho_i}}\right) I_{\alpha-1}\left(\sqrt{\frac{4\rho_{i+1}\lambda_i\lambda_{i+1}y_i y_{i+1}}{1-\rho_{i+1}}}\right), \quad y_i > 0,$$

exactly of the form given above.

More generally, we may consider the squared Bessel process and the squared Bessel bridge. For any $d > 0$, the d-dimensional squared Bessel process $\xi(t)$, $t \geq 0$ is a time homogeneous Markov process with transition density

$$(3.4) \qquad q(t, x, y) = \frac{1}{2t}\left(\frac{y}{x}\right)^{\nu/2} \exp\left(-\frac{x+y}{2t}\right) I_\nu\left(\frac{\sqrt{xy}}{t}\right), \qquad t > 0, \quad x, y \geq 0,$$

where $\nu + 1 = d/2$. We see from (3.4) that the conditional distribution of $\xi(t)$ given $\xi(0) = x$ is a randomized gamma distribution of the first type $G(\nu+\eta+1, 1/2t)$ with $\eta \sim P(x/2t)$.

When $\xi(0) = 0$ there is a correspondence between a multivariate gamma given by (3.3) and a re-scaled finite dimensional distribution of a Bessel process. Specifically, if we sample $\xi(t)$ at times $t_i = 1/(\rho_1 \cdots \rho_i)$, $i = 1, \ldots, n$, then the joint distribution of

$\xi(t_1)/(t_1\lambda_1),\ldots,\xi(t_n)/(t_n\lambda_n)$ is exactly the same as (3.3) provided $d = 2\alpha$ and $\lambda_i > 0$, $0 < \rho_i < 1$.

A standard squared Bessel bridge $\xi_{x_0,x_1}(t)$ is a stochastic process on $[0,1]$ generated by $\xi(t)$ with $\xi(0)$ and $\xi(1)$ tied at x_0 and x_1 respectively. The distribution of $\xi_{x_0,x_1}(t)$ is studied in detail by Pitman and Yor (1982). Its transition distribution is a randomized gamma distribution of the second type. Let $0 \le s < t \le 1$. If we know that $\xi_{x_0,x_1}(s) = x$, then, $Y = \xi_{x_0,x_1}(t)$ can be obtained by generating independent random variables

$$\eta_1 \sim P\left(\frac{1}{2(1-s)}\left[\frac{(1-t)}{(t-s)}x + \frac{(t-s)}{(1-t)}x_1\right]\right) \quad \text{and} \quad \eta_2 \sim Bes\left(\nu, \frac{\sqrt{xx_1}}{1-s}\right)$$

and then

$$Y \sim G\left(\nu + \eta_1 + 2\eta_2 + 1, \frac{1-s}{2(t-s)(1-t)}\right).$$

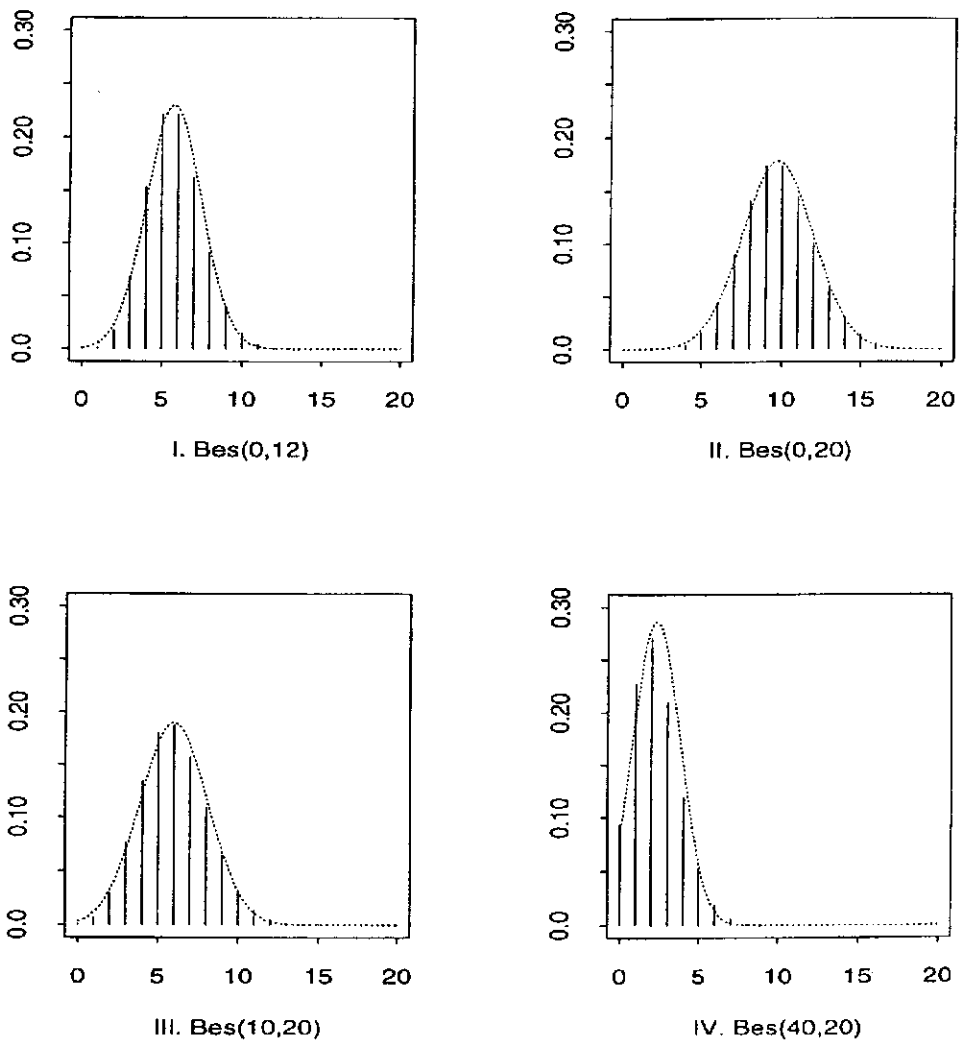

Fig. 2. The Bessel distributions and their normal approximations.

4. Simulations

Simulating a Bessel distribution is generally difficult. However, intensive numerical experiments show that a normal approximation is applicable to Bessel distributions with mode $m(\nu, a) \geq 6$. This is illustrated by Fig. 2 (I)–(IV).

We thus propose the following method for simulating a Bessel distribution. When $m(\nu, a) \geq 6$, a sample $Y \sim Bes(\nu, a)$ is drawn by generating $U \sim U(0,1)$ followed by

$$X = \mu + \sigma\Phi^{-1}[U + (1-U)\Phi(-\mu/\sigma)]$$

where the mean μ and variance σ^2 are calculated from (2.1) and then Y is set as the closest integer to X.

When $m(\nu, a) < 6$ the Bessel distribution has a short right tail. Therefore, a table sampling is appropriate for this case, which is easily implemented.

5. Discussion

The Bessel distribution links the spherical distributions such as the von Mises-Fisher distribution with well-known distributions on the real line. Therefore, simulating the von Mises-Fisher distribution is easily accomplished once we have an efficient way of simulating the Bessel distribution. Of course, there might be a better scheme for simulating the Bessel distribution and we leave this for future study.

A frequency approach to estimation in the von Mises-Fisher is made difficult by the complicated sampling distributions of some key statistics, whereas Bayesian inference (Mardia (1976)) can be simply carried out especially when the concentration parameter is known. In fact, a von Mises-Fisher prior for μ leads to the posterior of μ is again a von Mises-Fisher distribution. Although some numerical features of the posterior can be computed (Mardia (1976)), probability intervals are possible only when we can sample from the posterior distribution.

This study also facilitates simulation of the squared Bessel bridge and Monte Carlo computation for path integrals of the squared Bessel process. For example, the path integral

$$E_{x,y}\exp\left[-\int_0^t f(\xi(s))d\mu(s)\right] = E\left(\exp\left[-\int_0^t f(\xi(s))d\mu(s)\right] \,\middle|\, \xi(0)=x, \xi(t)=y\right)$$

where μ is a finite measure, has been discussed by Pitman and Yor (1982). For some special choices of f and μ the path integral can be expressed in a closed-from by solving a boundary problem in differential equations. More generally, however, this path integral has to be evaluated by numerical methods. For instance, we can simulate many sample paths of a squared Bessel bridge with $\xi(0) = x$ and $\xi(t) = y$, then the path integral can be approximated by an average.

Acknowledgements

The authors wish to thank the referees for helpful comments that improved the presentation.

Appendix

A recurrence formula for Bessel quotients arises immediately from the recurrence relation of Bessel functions:

$$R_\nu(x) = \frac{1}{\frac{2(\nu+1)}{x} + R_{\nu+1}(x)}. \tag{A.1}$$

Further, the following relation of Bessel functions

$$\frac{I'_\nu(x)}{I_\nu(x)} = \frac{\nu}{x} + R_\nu(x), \qquad \nu > -1$$

leads to a differential equation

$$\begin{aligned} y' &= 1 - y^2 - \frac{2\nu+1}{x} y, \\ y(0) &= 0 \end{aligned} \tag{A.2}$$

which has the unique solution $R_\nu(x)$. These equations are generally important in studying the Bessel quotient. On the other hand, some special Bessel quotients do have closed form expressions. For example,

$$I_{1/2}(x) = \sqrt{\frac{2}{\pi x}} \sinh x, \qquad \text{and} \qquad I_{3/2}(x) = \sqrt{\frac{2}{\pi x^3}} (x \cosh x - \sinh x).$$

It follows that $R_{1/2}(x) = \coth x - 1/x$ and, by (A.1), $R_{-1/2}(x) = \tanh x$.

The asymptotic expansion (Spain and Smith (1970)) of the Bessel function implies that $R_\nu(x) \to 1$ as $x \to \infty$, and thus all curves of functions $R_\nu(x)$, $\nu > -1$ start from $(0,0)$ and share the same asymptote $y = 1$. Further, we have

$$\lim_{x\to\infty} x[1 - R_\nu(x)] = 2\nu + 1 \tag{A.3}$$

which means the curve of $R_\nu(x)$ approaches the asymptote from above when $-1 < \nu < -\frac{1}{2}$ and from below when $\nu > -\frac{1}{2}$.

The monotonicity is also classified into two categories according to whether ν is larger than $-\frac{1}{2}$ or not. For $\nu > -\frac{1}{2}$, the function $R_\nu(x)$ is increasing over the whole interval $(0, \infty)$; while for $-1 < \nu < -\frac{1}{2}$, the function $R_\nu(x)$ is increasing first to reach a maximum and then decreasing. To verify this we differentiate (A.2) to obtain

$$y'' = \frac{2\nu+1}{x^2} y - \left(2y + \frac{2\nu+1}{x}\right) y'.$$

The sign of y'' at stationary points, where y' is zero, is the same as that of $2\nu+1$. Thus, $R_\nu(x)$ when $\nu > -\frac{1}{2}$ can have local minima only whereas if $\nu < -\frac{1}{2}$, $R_\nu(x)$ can have local maxima only.

From the equation (A.2), y' cannot change sign when $\nu > -\frac{1}{2}$. Otherwise, there must be a stationary point x_0 which is, from the discussion above, a local minimum at which y' changes from negative to positive. Since $y'(0) = 1/(2\nu+2) > 0$, it follows that there must be a local maximum between zero and x_0 which is impossible.

When $-1 < \nu < -\frac{1}{2}$, (A.3) indicates that y must be larger than one when x is large enough and $y(\infty) = 1$. Hence, there must be a point at which y' is negative. Since $y'(0) > 0$, there must be a stationary point x_0 which is a local maximum. Furthermore, $y(x_0) > 1$ is guaranteed by (A.2). We can show that, y' changes sign only once, otherewise we will have two local maxima between which there must be a local minimum, which is impossible. Hence, $y(x_0)$ is also a global maximum.

We now give some bounds for the Bessel quotient. In fact, the variance of the Bessel distribution must be non-negative which, from (2.1), implies

$$R_\nu(x) \le \frac{x}{\sqrt{x^2 + \nu^2} + \nu}. \tag{A.4}$$

On the other hand, the differential equation (A.2) suggests that, $R_\nu(x) \le R_\mu(x)$ if $\nu > \mu$ for these two functions have the same initial value but the former has a smaller derivative. Hence, $R_{\nu+1}(x) \le R_\nu(x)$ and this combined with (A.1) and (A.4) leads to

$$\frac{x}{\sqrt{x^2 + (\nu+1)^2} + (\nu+1)} \le R_\nu(x) \le \frac{x}{\sqrt{x^2 + \nu^2} + \nu}, \qquad \nu > -1. \tag{A.5}$$

For $\nu > -\frac{1}{2}$ a slightly sharper upper bound can be derived from the fact that $R'_\nu(x) \ge 0$. This kind of bound is also found by Amos (1974) for the case $\nu \ge 0$.

Numerical evaluation for the Bessel quotient was studied in detail by Amos (1974) using continued fraction. Actually, by repeating the recurrence relation, $R_\nu(x)$ can be written as a continued fraction

$$R_\nu(x) = \frac{1}{2(\nu+1)/x+}\ \frac{1}{2(\nu+2)/x+}\ \frac{1}{2(\nu+3)/x + \cdots}.$$

The upper bound in (A.5) implies that, for fixed x, $R_{\nu+k}(x) \to 0$ as $k \to \infty$ and it seems that $R_\nu(x)$ can be computed by iteration.

References

Amos, D. E. (1974). Computation of modified Bessel functions and their ratios, *Math. Comp.*, **28**, 239–251.

Devroye, L. (1986). *Non-uniform Random Variate Generation*, Springer, New York.

Feller, W. (1966). *An Introduction to Probability Theory and Its Applications*, Vol. II, Wiley, New York.

Johnson, N. and Kotz, S. (1970). *Distributions in Statistics*, Vol. 4, Wiley, New York.

Laha, R. G. (1954). On some properties of the Bessel function distributions, *Bull. Calcutta Math. Soc.*, **46**, 59–72.

Mardia, K. V. (1972). *Statistics of Directional Data*, Academic Press, London.

Mardia, K. V. (1976). Bayesian inference for the von Mises-Fisher distribution, *Biometrika*, **63**, 203–206.

Mathai, A. M. and Moschopoulos, P. G. (1991). On a multivariate gamma, *J. Multivariate Anal.*, **39**, 135–153.

McKay, A. T. (1932). A Bessel function distribution, *Biometrika*, **24**, 39–44.

McLeish, D. L. (1982). A robust alternative to the normal distribution, *Canad. J. Statist.*, **10**, 89–102.

Pinky, M. (1993). Fourier inversion of multidimensional characteristic functions, *J. Theoret. Probab.*, **6**, 187–193.

Pitman, J. and Yor, M. (1982). A decomposition of Bessel bridges, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, **59**, 425–457.

Richards, D. (1986). Positive definite symmetric functions on finite dimensional space, *J. Multivariate Anal.*, **19**, 280–285.

Spain, B. and Smith, M. G. (1970). *Functions of Mathematical Physics*, Van Nostrand Reinhold Company, London.