Some Statistical Properties of a Family of Continuous Univariate Distributions

PETER McCULLAGH*

A new two-parameter family of continuous univariate distributions on the interval $(-1, 1)$ is introduced, and some properties are given. It is shown that parameters θ and ν are globally orthogonal in the Fisher information sense and that, to some extent, they have the properties of a location parameter and a precision parameter, respectively. A pivotal statistic is constructed whose distribution is independent of the location parameter. Finally, the connection with ultraspherical functions and Brownian motion is described.

KEY WORDS: Beta density; Brownian motion; Cumulants; Hypergeometric density; Orthogonal parameters; Pivotal statistic; Ultraspherical polynomial.

1. INTRODUCTION

I discuss several properties of the continuous univariate density

$$f_X(x; \theta, \nu) = \frac{(1 - x^2)^{\nu - 1/2}}{(1 - 2\theta x + \theta^2)^\nu B(\nu + \frac{1}{2}, \frac{1}{2})}, \quad -1 < x < 1, \quad (1)$$

and the related density

$$f_{X'}(x'; \theta, \nu) = \frac{(1 - x'^2)^{\nu - 1/2}(1 - \theta^2)}{(1 - 2\theta x' + \theta^2)^{\nu+1} B(\nu + \frac{1}{2}, \frac{1}{2})}, \quad -1 < x < 1. \quad (2)$$

The random variables are related via the transformation

$$X' - \theta = \frac{(X - \theta)(\theta^2 - 1)}{1 - 2\theta X + \theta^2}, \quad (3)$$

which for each $-1 < \theta < 1$ maps the interval $(-1, 1)$ onto itself. The same transformation applied to X' gives $X - \theta$. Both families are defined for $\nu > -\frac{1}{2}$ and $-1 < \theta < 1$; (1) is also defined for $\theta = \pm 1$.

Since the pair of families is connected via transformation (3), most of the discussion focuses on (1) to avoid duplication of calculations. For brevity, write $X \sim H(\theta, \nu)$ and $X' \sim H'(\theta, \nu)$ to distinguish the two families.

It is perhaps not immediately apparent that (1) defines a probability density for all parameter values in the indicated range. Nevertheless, a proof follows easily from the following properties of hypergeometric functions:

$$\begin{aligned} F(a, \tfrac{1}{2} + a; 1 + 2a; z) &= \frac{\Gamma(2a + 1)}{\Gamma^2(a + \frac{1}{2})} \int_0^1 \frac{t^{a - 1/2}(1 - t)^{a - 1/2}}{(1 - tz)^a}\, dt, \quad a > -\tfrac{1}{2} \\ &= 2^{2a}[1 + (1 - z)^{1/2}]^{-2a}. \end{aligned} \quad (4)$$

See Abramowitz and Stegun (1970, eqs. 15.1.13 and 15.3.1). The integral of $f_X(x; \theta, \nu)$ reduces by a change of variables to

$$\int_{-1}^1 f_X(x; \theta, \nu)\, dx = \frac{2^{2\nu}}{(1 + \theta)^{2\nu} B(\nu + \frac{1}{2}, \frac{1}{2})} \int_0^1 \frac{t^{\nu - 1/2}(1 - t)^{\nu - 1/2}}{(1 - 4\theta t/(1 + \theta)^2)^\nu}\, dt.$$

Application of (4) with $z = 4\theta/(1 + \theta)^2$ $(a = \nu)$ gives $\int_{-1}^1 f_X(x; \theta, \nu)\, dx = 1$ for all parameter values in the range indicated.

Note that if $z = 4\theta/(1 + \theta)^2$ it is necessary to choose

$$\begin{aligned} (1 - z)^{1/2} &= (1 - \theta)/(1 + \theta) \quad \text{for } -1 \le \theta \le 1 \\ &= (\theta - 1)/(1 + \theta) \quad \text{for } |\theta| > 1. \end{aligned}$$

Thus although the function $f_X(x; \theta, \nu)$ is well defined and has a finite integral for all θ, the total integral is unity only for $|\theta| \le 1$. For θ outside this range, the total integral of (1) is $|\theta|^{-2\nu}$. Therefore, it is possible to extend the domain of the parameter space by modifying the definition for $|\theta| > 1$ as follows:

$$f_X(x; \theta, \nu) = \frac{(1 - x^2)^{\nu - 1/2}\, |\theta|^{2\nu}}{(1 - 2\theta x + \theta^2)^\nu B(\nu + \frac{1}{2}, \frac{1}{2})}, \quad -1 \le x \le 1.$$

Density (2) can be extended in a similar way. In the discussion that follows, however, it is assumed that $|\theta| \le 1$. (For a physical interpretation of this discontinuity at $|\theta| = 1$ see Sec. 10.)

An important qualitative difference between families (1) and (2) and the beta family is that as θ varies in either (1) or (2), the order of contact at the terminals remains fixed. By contrast, the mean of the beta family can be changed only by adjusting the order of contact at the terminals. For example, there is only one member of the beta family that is finite and nonzero at both terminals, namely the uniform distribution. In (1) and (2), however, the density is finite and nonzero at ± 1 for all $|\theta| < 1$, provided $\nu = \frac{1}{2}$.

* Peter McCullagh is Professor, Department of Statistics, University of Chicago, Chicago, IL 60637. This work was supported in part by National Science Foundation (NSF) Grant DMS 88-01853. The manuscript was prepared using computer facilities supported in part by NSF Grants DMS 86-01732 and DMS 87-03942 to the University of Chicago, Department of Statistics. The author thanks John Kolassa for preparing the plots in Figure 1, and Colin Mallows and Norman Johnson for suggestions regarding integration of Density (1).

Journal of the American Statistical Association
March 1989, Vol. 84, No. 405, Theory and Methods

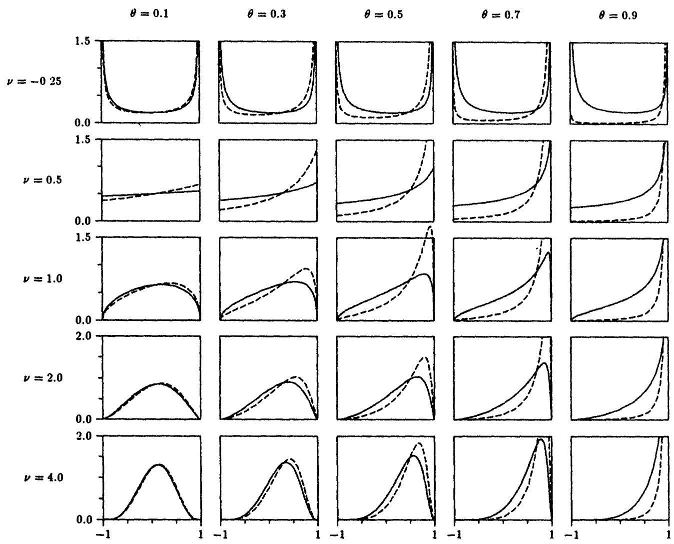

Figure 1. Densities (1) (——) and (2) (---) Plotted for Various Values of the Parameters.

Families (1) and (2) are plotted in Figure 1 for various values of ν and θ. It can be seen graphically, and is easily proved analytically, that both densities are asymptotically normal as $\nu \to \infty$ for each fixed θ in the interval $-1 < \theta < 1$. Evidently, the rate of convergence is slow for $|\theta|$ near 1 and fast for θ near 0.

2. SPECIAL CASES

The density function of $Y = \frac{1}{2}(X + 1)$, which is concentrated on the interval (0, 1), is given by

$$f_Y(y; \theta, \nu) = \frac{y^{\nu-1/2}(1 - y)^{\nu-1/2}2^{2\nu}}{\{(1 + \theta)^2 - 4\theta y\}^\nu B(\nu + \frac{1}{2}, \frac{1}{2})}, \quad 0 < y < 1. \tag{5}$$

For $\theta = -1$, 0, and 1, Y has the beta density with parameters $(\frac{1}{2}, \nu + \frac{1}{2})$, $(\nu + \frac{1}{2}, \nu + \frac{1}{2})$, and $(\nu + \frac{1}{2}, \frac{1}{2})$, respectively. For other values of θ the distribution is not a member of the beta family (unless $\nu = 0$). If $\theta = 0$, the distribution is symmetric about $x = 0$ or $y = \frac{1}{2}$: If in addition $\nu = \frac{1}{2}$, the distribution is uniform. If X has the distribution (1), then $-X$ has a distribution in the same family with parameters $-\theta$ and ν. Similarly, if Y has the distribution (5), then $1 - Y$ has the same distribution with θ replaced by $-\theta$. To some extent, therefore, θ behaves like a location parameter.

A peculiar aspect of the parameterization in (5) is that for $\nu = 0$ (5) is equal to the beta $(\frac{1}{2}, \frac{1}{2})$ distribution and does not depend on θ. This point thus constitutes a singularity of the likelihood. There is a different kind of singularity at $\theta = \pm 1$. See Section 7 for further details.

3. CUMULANTS

The moments of X and X' may be expressed in terms of the hypergeometric function with argument $4\theta/(1 + \theta)^2$. Such expressions are not particularly helpful for computational purposes, but the lower-order cumulants can be simplified drastically:

$$E(X) = \frac{\nu\theta}{\nu + 1},$$

$$\operatorname{var}(X) = \left[1 - \frac{\nu(\nu - 1)}{(\nu + 1)(\nu + 2)}\theta^2\right] \Big/ \{2(\nu + 1)\},$$

$$\kappa_3(X) = \frac{\nu}{2(\nu + 1)^2(\nu + 2)}\left[\frac{(\nu - 1)(3\nu - 1)\theta^3}{(\nu + 1)(\nu + 3)} - 3\theta\right],$$

and

$$\kappa_4(X) = \frac{-3}{4(\nu + 1)^2(\nu + 2)} \times \left[1 - \frac{4\nu(3\nu - 1)\theta^2}{(\nu + 1)(\nu + 3)} + \frac{\nu(\nu - 1)(11\nu^3 + 16\nu^2 - 17\nu + 2)\theta^4}{(\nu + 1)^2(\nu + 2)(\nu + 3)(\nu + 4)}\right].$$

The corresponding expressions for the cumulants of X' are simpler:

$$E(X') = \theta,$$

$$\operatorname{var}(X') = (1 - \theta^2)/\{2(\nu + 1)\},$$

$$\kappa_3(X') = \frac{-3\theta(1 - \theta^2)}{2\nu(\nu + 1)(\nu + 2)},$$

and

$$\kappa_4(X') = \frac{3(1 - \theta^2)(11\nu\theta^2 + 13\theta^2 - \nu - 3)}{4(\nu + 1)^2(\nu + 2)(\nu + 3)}.$$

These formulas have been derived algebraically and checked numerically. Higher-order cumulants can also be obtained using Gauss's recurrence formulas for contiguous hypergeometric functions (Abramowitz and Stegun 1970, eqs. 15.2.10–15.2.27). Such computations are unusually tedious, however. The method described in Section 8, using orthogonal polynomials, is simpler.

From the aforementioned formulas it appears that $\kappa_r(X)$ and $\kappa_r(X')$ are both $O(\nu^{-r+1})$. If true, this conjecture supports the assertion that ν is a precision parameter: In effect, the cumulants of X behave asymptotically like an average of ν independent random variables.

Note that the mean of X is an increasing function of θ only for $\nu > 0$ and is decreasing in θ for $-\frac{1}{2} < \nu < 0$. Similarly, for $\nu > 1$ the variance has a maximum at $\theta = 0$. For $\nu = 1$ the variance of X is equal to $\frac{1}{4}$ for all θ. For ν between 0 and 1, the variance is a minimum at $\theta = 0$; this pattern is reversed for $\nu < 0$.

4. PIVOTAL STATISTIC FOR θ

Define the random variable

$$T(\theta) = \frac{1 - X^2}{1 - 2\theta X + \theta^2} = \frac{1 - X'^2}{1 - 2\theta X' + \theta^2}. \tag{6}$$

Since $-1 \le X \le 1$ for all θ, it follows that $0 \le T(\theta) \le 1$

for all values of θ. Evidently, by direct integration using (4) we have

$$E\{T(\theta)^r\} = \frac{B(\nu + r + \frac{1}{2}, \frac{1}{2})}{B(\nu + \frac{1}{2}, \frac{1}{2})}$$

independently of θ. These are the moments of a beta random variable with parameters $(\nu + \frac{1}{2}, \frac{1}{2})$. Since the beta distribution is determined by its moments, it follows that $T(\theta)$ is a *pivotal statistic* having the $B(\nu + \frac{1}{2}, \frac{1}{2})$ distribution for all θ.

If ν is known, exact confidence intervals for θ based on X may be obtained directly from the pivotal statistic in the usual way.

For future reference, we note that the mean of log T is

$$\begin{aligned} E \log T(\theta) &= \frac{\partial}{\partial \nu} \log B(\nu + \tfrac{1}{2}, \tfrac{1}{2}) \\ &= \psi(\nu + \tfrac{1}{2}) - \psi(\nu + 1) \\ &\simeq \frac{1}{2\nu} + \frac{1}{8\nu^2} + O(\nu^{-3}). \end{aligned}$$

The latter approximation holds for large ν.

The pivotal statistic can be used as an intermediate step in the computer generation of pseudorandom variables having density (1) or (2). Given a supply of pseudorandom variables t_j having the $B(\nu + \frac{1}{2}, \frac{1}{2})$ distribution and a supply of independent uniform variables, we may generate a sequence $X_j \sim H(\theta, \nu)$ as follows. First, solve Equation (6) to obtain the two values of x for the given value of $t = t_j$ with $0 < t < 1$. These roots are $x_1 = \theta t + [(1 - t)(1 - \theta^2 t)]^{1/2}$ and $x_2 = \theta t - [(1 - t)(1 - \theta^2 t)]^{1/2}$. To obtain a random variable having Density (1), let $X_j = x_1$ with probability

$$\pi_1 = \left\{1 + \left|\frac{T_1'}{T_2'}\right| \frac{f(x_2)}{f(x_1)}\right\}^{-1},$$

where T_1' and T_2' are the derivatives of T with respect to x at the two roots. A similar scheme may be derived for the density (2) by modifying the selection probability. For details see Michael, Schucany, and Haas (1976).

5. LIKELIHOOD FUNCTION BASED ON (1)

Suppose that there is a single observation x from Density (1). The log-likelihood for (θ, ν) can then be written in the form

$$l(\theta, \nu) = \nu \log T(\theta) - \log B(\nu + \tfrac{1}{2}, \tfrac{1}{2}) - \tfrac{1}{2} \log(1 - x^2). \tag{7}$$

The derivative with respect to θ is

$$\frac{\partial l}{\partial \theta} = \nu \frac{T'(\theta)}{T(\theta)} = 2\nu \frac{x - \theta}{1 - 2\theta x + \theta^2}.$$

It then follows that

$$E\left(\frac{X - \theta}{1 - 2\theta X + \theta^2}\right) = 0 \tag{8}$$

for all θ, and that $\hat{\theta} = x$ is the maximum likelihood estimate of θ based on the single observation x. From a single observation it is not possible to estimate both θ and ν; this is consistent with the claim that θ is a location parameter and ν is a precision parameter. If θ is known, the log-likelihood (7) has the exponential family form in which ν is the canonical parameter, log $T(\theta)$ the canonical statistic, and log $B(\nu + \frac{1}{2}, \frac{1}{2})$ the cumulant function. Note that $T(\theta)$ is not a 1–1 function of x: Evidently, from (6) there are two values of x corresponding to each value of $T(\theta)$. When θ is known, $T(\theta)$ is a minimal sufficient statistic for ν and the maximum likelihood estimate of ν is obtained by equating the observed value of the canonical statistic to its expectation, giving $\log T(\theta) = \psi(\hat{\nu} + \frac{1}{2}) - \psi(\hat{\nu} + 1)$. An approximate solution for small values of $|\log T(\theta)|$ is $\hat{\nu}^{-1} \simeq -2 \log T(\theta)$, which is obtained using the approximation of Section 3.

For a simple random sample of observations, the log-likelihood is a sum of independent contributions, each of the form (7). The derivative with respect to θ is then

$$\frac{\partial l}{\partial \theta} = 2\nu \sum \frac{x_i - \theta}{1 - 2\theta x_i + \theta^2}.$$

Thus the maximum likelihood estimate of θ satisfies

$$\sum \frac{x_i - \hat{\theta}}{1 - 2\hat{\theta} x_i + \hat{\theta}^2} = 0, \tag{9}$$

whether or not ν is known. There may be multiple solutions to (9) for θ in $[-1, 1]$. Note that (9) can be written in the form

$$\sum \frac{x_i - \hat{\theta}}{1 - x_i^2 + (x_i - \hat{\theta})^2} = 0,$$

which is formally similar to the Cauchy estimating equation

$$\sum \frac{x_i - \hat{\theta}}{\tau_i^2 + (x_i - \hat{\theta})^2} = 0$$

in which X_i has the Cauchy distribution centered at θ with known scale parameter τ_i. An identical estimating equation occurs in the normal-theory problem of estimating a common mean from n samples of size 2 in which the n variances are unknown (e.g., see Cox and Reid 1987, sec. 4.2.1).

The maximum likelihood estimate of θ obtained from (9) is the same whether ν is known or unknown. By contrast, the maximum likelihood estimate of ν depends on θ. For a sample of iid observations, $\hat{\nu}$ satisfies

$$\frac{1}{n} \sum \log T(\hat{\theta}) = \psi(\hat{\nu} + \tfrac{1}{2}) - \psi(\hat{\nu} + 1),$$

where $\hat{\theta}$ is obtained from (9). The approximation $\hat{\nu}^{-1} \simeq -2 \sum \log T(\hat{\theta})/n$ may be helpful if $-2 \sum \log T(\theta)/n$ is small.

6. LIKELIHOOD RATIO STATISTIC

For a single observation X with known precision index ν, $-2\nu \log T(\theta_0)$ is the likelihood ratio statistic for testing the hypothesis that $\theta = \theta_0$. For large values of ν, it follows from the usual asymptotic theory that $-2\nu \log T(\theta_0) \sim \chi_1^2$ approximately under H_0. This claim can be verified directly from the observation in Section 3 that $T(\theta_0)$ has the beta distribution with parameters $(\nu + \frac{1}{2}, \frac{1}{2})$.

In the case of a simple random sample, consider the null hypothesis H_0: $\theta = \theta_0$ with ν given. The likelihood ratio statistic is then

$$-2\nu\{\sum \log T(\theta_0) - \sum \log T(\hat{\theta})\},$$

which is approximately distributed as χ_1^2 for large samples. The quantity

$$-2 \sum \log T(\hat{\theta}) = -2 \sum \log \left(\frac{1 - X_i^2}{1 - 2\hat{\theta}X_i + \hat{\theta}^2}\right)$$

is the deviance statistic in the sense of McCullagh and Nelder (1983).

7. FISHER INFORMATION

For a single observation from family (1), the second derivatives of the log-likelihood are

$$\frac{\partial^2 l}{\partial \theta^2} = 2\nu \frac{(x - \theta)^2 - (1 - x^2)}{\{(x - \theta)^2 + (1 - x^2)\}^2} = 2\nu \frac{1 - 2T}{D}, \tag{10}$$

where $D = 1 - 2\theta x + \theta^2$ is the denominator in the expression for T,

$$\frac{\partial^2 l}{\partial \nu \partial \theta} = 2 \frac{x - \theta}{1 - 2\theta x + \theta^2},$$

and $\partial^2 l/\partial \nu^2 = \psi'(\nu + 1) - \psi'(\nu + \frac{1}{2})$. It follows from (8) that $E(\partial^2 l/\partial \nu \partial \theta) = 0$, so the parameters are globally orthogonal. This property does not apply to the family $H'(\theta, \nu)$. In addition, $\partial l/\partial \theta$ and $\partial l/\partial \nu$ are uncorrelated, so

$$\operatorname{cov}\left(\log T, \frac{\partial l}{\partial \theta}\right) = \operatorname{cov}\left(\log \frac{1 - X^2}{1 - 2\theta X + \theta^2}, \frac{X - \theta}{1 - 2\theta X + \theta^2}\right) = 0.$$

Furthermore, $(\partial l/\partial \theta)^2 = 4\nu^2(1 - T)/D$. From (10) then, we have

$$4\nu^2 E\left(\frac{1 - T}{D}\right) = -2\nu E\left(\frac{1 - 2T}{D}\right).$$

This gives $2(\nu + 1)E(T/D) = (2\nu + 1)E(1/D)$, and since $E(T) = (2\nu + 1)/\{2(\nu + 1)\}$, it follows that T and $1/D$ are uncorrelated. Finally, from Abramowitz and Stegun (1970, eq. 15.1.13) it follows that $E(1/D) = 1/(1 - \theta^2)$, independently of ν. Thus the Fisher information for θ based on a single observation X is $i_{\theta\theta} = E(\partial l/\partial \theta)^2 = [2\nu^2/(\nu + 1)][1/(1 - \theta^2)]$.

Note that the Fisher information for θ is 0 if $\nu = 0$: this result is consistent with the fact that θ is indeterminate if $\nu = 0$. Moreover, if $\nu \neq 0$ the Fisher information tends to infinity as $\theta \to \pm 1$.

Similar calculations for family (2) give

$$i_{\theta\theta} = \frac{2}{(1 - \theta^2)^2}\left\{\nu + 1 - \frac{\nu(\nu - 1)\theta^2}{\nu + 2}\right\},$$

$$i_{\theta\nu} = 2\theta/\{(\nu + 1)(1 - \theta^2)\},$$

and

$$i_{\nu\nu} = \psi'(\nu + \tfrac{1}{2}) - \psi'(\nu + 1) \simeq 1/\{2(\nu + \tfrac{1}{4})^2\}.$$

8. COMPUTATION OF CUMULANTS

The coefficient $C_r^{(\nu)}(x)$ of θ^r in the Taylor expansion of $M(\theta) = (1 - 2\theta x + \theta^2)^{-\nu}$ is called the ultraspherical or Gegenbauer polynomial of degree r. These polynomials are orthogonal over $[-1, 1]$ with respect to the weight function $(1 - x^2)^{\nu - 1/2}$ (Appell and Kampé de Fériet 1926). It then follows that if $X \sim H(\theta, \nu)$, the mean of $C_r^{(\nu)}(X)$ is $E\{C_r^{(\nu)}(X); \theta\} = k_{\nu r}\theta^r$, where $k_{\nu r} = E\{(C_r^{(\nu)}(X))^2; 0\}$ is a constant independent of θ. It is evident that the rth moment of X must be a polynomial in θ of degree r, and likewise for the cumulants. In this respect, the cumulants of $H(\theta, \nu)$ behave like the cumulants of the binomial distribution.

Similar calculations apply to Density (2) because $(1 - \theta^2)(1 - 2\theta x + \theta^2)^{-\nu-1}$ has a Taylor expansion in which the coefficient of θ^r is a multiple of $C_r^{(\nu)}(x)$. Again, the rth cumulant is a polynomial in θ of degree r.

9. HYPERGEOMETRIC DENSITY FUNCTIONS

The hypergeometric function $F(a, b; c; z)$ may be defined for $-1 < \theta < 1$ via the integral

$$\int_0^1 \frac{x^{\alpha-1}(1 - x)^{\beta-1}\, dx}{\{(1 + \theta)^2 - 4\theta x\}^\gamma} = \frac{B(\alpha, \beta)}{(1 + \theta)^{2\gamma}} F\left(\gamma, \alpha; \alpha + \beta; \frac{4\theta}{(1 + \theta)^2}\right).$$

For brevity, the integral is denoted by $k(\alpha, \beta, \gamma; \theta)$. It is defined for $\alpha > 0$, $\beta > 0$, $-1 < \theta < 1$. The integrand is nonnegative for $0 \le x \le 1$. Hence

$$\frac{x^{\alpha-1}(1 - x)^{\beta-1}}{\{(1 + \theta)^2 - 4\theta x\}^\gamma k(\alpha, \beta, \gamma; \theta)}$$

defines a probability distribution over $(0, 1)$.

Apart from the special cases $\gamma = 0$ and $\theta \to \pm 1$, which correspond to the beta family, it is difficult to make much progress analytically with this density. The distribution (1) is obtained by taking $\alpha = \beta = \nu + \frac{1}{2}$ and $\gamma = \nu$. In that case (but not otherwise) the normalizing constant k does not depend on the location parameter θ.

10. CONNECTION WITH BROWNIAN MOTION

The following discussion describes how families (1) and (2) arise as exit distributions for Brownian motion in p space, privided that $p = 2\nu + 2$ is a positive integer not less than 2.

Suppose that $\mathbf{Z}(t)$ is the position of a particle at time t

undergoing Brownian motion in p dimensional space, starting from the origin at time $t = 0$. Let $X' = (X'_1, \ldots, X'_p)$ be the point at which the particle first hits the unit sphere $|Z| = 1$. Evidently, X' is uniformly distributed over the sphere and X'_1 has the symmetric beta distribution on $(-1, 1)$ with index $\nu = p/2 - 1$, that is, $X'_1 \sim H'(0, \nu)$.

Suppose, instead of starting at the origin, that the particle starts at the point $\theta = (\theta, 0, \ldots, 0)$ on the x_1 axis. If $-1 < \theta < 1$ the particle will eventually hit the unit sphere at a point X'. The distribution of X' with respect to Lebesgue measure on the unit sphere is given by

$$g'(X'; \theta, p) = \frac{1 - |\theta|^2}{A_p|X' - \theta|^p} = \frac{1 - \theta^2}{A_p(1 - 2\theta x'_1 + \theta^2)^{p/2}}$$

(Durrett 1984, sec. 1.10), where A_p is the surface area of the sphere in R^p. Consequently, the marginal distribution of X'_1 is given by (2). Hence the $H'(\theta, \nu)$ family is a natural noncentral version of the symmetric beta family.

The reflected exit point X is obtained by extending a chord from X' through the starting point θ to intersect the sphere at $X = (X_1, \ldots, X_p)$. It is a straightforward exercise to show that

$$X - \theta = -\frac{1 - |\theta|^2}{|X' - \theta|^2}(X' - \theta),$$

so X has the distribution

$$g(X; \theta, p) = \frac{1}{A_p|X' - \theta|^{p-2}} = \frac{1}{A_p(1 - 2\theta x'_1 + \theta^2)^{p/2-1}},$$

with respect to Lebesgue measure on the unit sphere. Note that if $p = 2$, X is uniformly distributed whatever the starting point θ. It then follows that X_1 has the distribution (1), which could also be described as a noncentral version of the symmetric beta family.

If $|\theta| > 1$, the probability of ever hitting the sphere is $|\theta|^{-2\nu}$. The extended definition given in Section 1 represents the conditional density of X_1, given that the particle eventually hits the sphere.

[*Received February 1988. Revised September 1988.*]

REFERENCES

Abramowitz, M., and Stegun, I. A. (1970), *Handbook of Mathematical Functions*, New York: Dover Press.

Appell, P., and Kampé de Fériet, J. (1926), *Fonctions Hypergéometriques et Hypersphériques, Polynomes d'Hérmite*, Paris: Gautier-Villars.

Cox, D. R., and Reid, N. (1987), "Orthogonal Parameters and Conditional Inference" (with discussion), *Journal of the Royal Statistical Society*, Ser. B, 49, 1–39.

Durrett, R. (1984), *Brownian Motion and Martingales in Analysis*, Belmont, CA: Wadsworth.

McCullagh, P., and Nelder, J. A. (1983), *Generalized Linear Models*, London: Chapman & Hall.

Michael, J. R., Schucany, W. R., and Haas, R. W. (1976), "Generating Random Variables Using Transformations With Multiple Roots," *The American Statistician*, 40, 88–91.