The Annals of Statistics
1980, Vol. 8, No. 1, 183–189

A NEW CHARACTERIZATION OF THE DIRICHLET DISTRIBUTION THROUGH NEUTRALITY

BY IAN R. JAMES AND JAMES E. MOSIMANN

C.S.I.R.O. South Melbourne, Victoria, Australia and National Institutes of Health, Bethesda, MD

A new characterization of the Dirichlet distribution using neutrality is given. This settles conjectures of Doksum and Mosimann. The characterization is contrasted with a parallel result for the lognormal distribution. Possible applications to random probabilities and prior distributions in survival analysis are noted.

1. Characterization of the Dirichlet distribution. Let $\mathbf{X} = (X_1, \cdots, X_n)$ be a random vector with $n \geqslant 2$ nonnegative coordinates satisfying $\Sigma_{i=1}^n X_i \leqslant 1$. Let $S_k = \Sigma_{i=1}^k X_i$, $k = 1, \cdots, n$. We shall suppose that none of the X_i nor $1 - S_n$ is degenerate at zero.

DEFINITION 1. $(X_1, \cdots, X_k)$, $k \leqslant n - 1$ is neutral in $\mathbf{X}$ if there exist nonnegative random variables $V_1, \cdots, V_n$ with $(V_1, \cdots, V_k)$ independent of $(V_{k+1}, \cdots, V_n)$ such that $\mathbf{X}$ and

$$(V_1, \cdots, V_k, V_{k+1}(1 - T_K), \cdots, V_n(1 - T_K))$$

have the same distribution (where $T_k = \Sigma_{i=1}^k V_i$).

Note that the neutrality of $(X_1, \cdots, X_k)$ implies that S_k and T_k have the same distribution. This neutrality states, in effect, that the vectors $(X_1, \cdots, X_k)$ and

$$(X_{k+1}/(1 - S_k), \cdots, X_n/(1 - S_k))$$

are independent, but is defined as above to avoid the possibility of division by zero (c.f. Doksum, 1974, page 186). The term neutrality was introduced by Connor and Mosimann (1969).

Neutrality of any vector of coordinates of $\mathbf{X}$, for example of any X_i, or $(X_j$: $j \neq i)$, is defined by obvious modification of Definition 1, and Fabius (1973) proved:

THEOREM 1. *The following assertions are equivalent*:

(i) X_i *is neutral in* $\mathbf{X}$ *for all* $i = 1, \cdots, n$;

(ii) $(X_j; j \neq i)$ *is neutral in* $\mathbf{X}$ *for all* $i = 1, \cdots, n$;

(iii) *The distribution of* $\mathbf{X}$ *is Dirichlet or a limit of Dirichlet distributions.*

Recall that $\mathbf{X}$ is Dirichlet with parameters $(\alpha_1, \cdots, \alpha_{n+1})$ if it has a density

---

Received May 1978.

AMS 1970 *subject classifications.* Primary 62E10; secondary 62H05, 62C10.

Key words and phrases. Neutral processes, size variables, neutrality, Dirichlet distribution, lognormal distribution, Bayesian survival analysis.

function of the form

$$f(\mathbf{x}) = K(\Pi_{i=1}^n x_i^{\alpha_i - 1})(1 - \Sigma_{i=1}^n x_i)^{\alpha_{n+1}-1}$$

where $x_i > 0$, $i = 1, \cdots, n$; $\Sigma_1^n x_i < 1$ and $K = \Gamma(\Sigma\alpha_i)/\Pi\Gamma(\alpha_i)$; $\alpha_i > 0$, $i = 1, \cdots, n+1$. In the limiting cases in (iii) $\mathbf{X}$ is either degenerate at a point, or its distribution is concentrated at the vertices of the simplex (Fabius, 1973).

The purpose of this note is to provide an alternative characterization of the Dirichlet distribution and its limits, thereby confirming the conjectures of Doksum (1974, page 193) and Mosimann (1975b, page 233).

DEFINITION 2. $\mathbf{X}$ is *completely neutral* if there exist mutually independent, nonnegative variables $W_1, \cdots, W_n$ such that $\mathbf{X}$ and

$$(W_1, W_2(1 - W_1), \cdots, W_n\Pi_{i=1}^{n-1}(1 - W_i))$$

have the same distribution. (Doksum, 1974, page 186).

If division by zero is avoided, complete neutrality states that the variables

$$X_1, X_2/(1 - S_1), \cdots, X_n/(1 - S_{n-1})$$

are mutually independent, and can be shown to be equivalent to the property "$(X_1, \cdots, X_i)$ is neutral in $\mathbf{X}$ for all $i = 1, \cdots, n-1$" (Connor and Mosimann, 1969). Note that the order of the coordinates of $\mathbf{X}$ is important in the definition of complete neutrality. We then have

THEOREM 2. *The following assertions are equivalent*:

(i) $\mathbf{X}$ *is completely neutral, and* X_n *is neutral in* $\mathbf{X}$;

(ii) *The distribution of* $\mathbf{X}$ *is Dirichlet or a limit of Dirichlet distributions.*

PROOF OF THEOREM 2. We need only prove that (i) implies (ii) since (ii) implies (i) is well known.

Consider the vectors

$$\mathbf{X}^j = (X_n, S_j), \qquad j = 1, \cdots, n-1.$$

Note that X_n, S_j are nonnegative and $X_n + S_j \leq 1$. From the complete neutrality of $\mathbf{X}$ there exist mutually independent nonnegative variables $W_1, \cdots, W_n$ such that S_j and

$$1 - \Pi_{i=1}^j(1 - W_i) = U_j, \qquad \text{say},$$

have the same distribution. Also X_n and the variable

$$W_n\Pi_{i=1}^{n-1}(1 - W_i) = W_n\left[\Pi_{i=j+1}^{n-1}(1 - W_i)\right](1 - U_j)$$
$$= R_j(1 - U_j), \qquad \text{say}, \qquad (\text{with } R_{n-1} = W_n)$$

have the same distribution. But for a given j, U_j and R_j are independent and therefore S_j is neutral in $\mathbf{X}^j$, $j = 1, \cdots, n-1$. We now use the condition that X_n is neutral in $\mathbf{X}$, so there exist nonnegative variables $V_1, \cdots, V_n$ such that $(V_1, \cdots, V_{n-1})$ is independent of V_n, where V_n and X_n have the same distribution.

Further X_i and $(1 - V_n)V_i$ have the same distribution, $i = 1, \cdots, n-1$, and therefore S_j and $(1 - V_n)\Sigma_{i=1}^{j} V_i = (1 - V_n)T_j$ have the same distribution, $j = 1, \cdots, n-1$. Since V_n and T_j are independent, then X_n is neutral in $\mathbf{X}^j$, $j = 1, \cdots, n-1$.

By Theorem 1 then for each $j = 1, \cdots, n-1$, $(X_n, S_j) = \mathbf{X}^j$ is Dirichlet or has a distribution which is a limit of Dirichlet distributions. If $\mathbf{X}^j$ is degenerate or discrete it is straightforward to show inductively that $\mathbf{X}$ is a limit of Dirichlet distributions. Thus we need only consider the case where $\mathbf{X}^j$ is Dirichlet $(\alpha_j, \beta_j, \gamma_j)$, say, for all $j = 1, \cdots, n-1$. It then follows that X_n has a beta distribution with parameters $(\alpha_j, \beta_j + \gamma_j), j = 1, \cdots, n-1$ so that

$$\alpha_j = \alpha \quad \text{(constant)}$$

$$\beta_j + \gamma_j = c \quad \text{(constant)}, \qquad j = 1, \cdots, n-1.$$

Further $1 - S_j$ is beta $(\alpha + \gamma_j, \beta_j)$ for all $j = 1, \cdots, n-1$, so that for any $r > 0$

$$E\left[(1 - S_j)^r\right] = \frac{\Gamma(\alpha + \gamma_j + r)\Gamma(\alpha + c)}{\Gamma(\alpha + \gamma_j)\Gamma(\alpha + c + r)}.$$

But we can write

$$1 - S_j = \left(1 - \frac{X_j}{1 - S_{j-1}}\right)(1 - S_{j-1}), \qquad j = 2, \cdots, n-1$$

where the independence of the two right-hand terms follows from the complete neutrality of $\mathbf{X}$. It then follows that

$$E\left[\left(1 - \frac{X_j}{1 - S_{j-1}}\right)^r\right] = \frac{\Gamma(\alpha + \gamma_j + r)}{\Gamma(\alpha + \gamma_j)} \frac{\Gamma(\alpha + \gamma_{j-1})}{\Gamma(\alpha + \gamma_{j-1} + r)}$$

for all $r > 0$ and $j = 2, \cdots, n-1$. Since $E(1 - S_{j-1}) > E(1 - S_j)$ then $\gamma_{j-1} > \gamma_j$, and so we let $\varepsilon_j = \gamma_{j-1} - \gamma_j > 0, j = 2, \cdots, n-1$. Consequently we can write

$$E\left[\left(1 - \frac{X_j}{1 - S_{j-1}}\right)^r\right] = \frac{\Gamma(\alpha + \gamma_j + r)\Gamma(\alpha + \gamma_j + \varepsilon_j)}{\Gamma(\alpha + \gamma_j)\Gamma(\alpha + \gamma_j + \varepsilon_j + r)}$$

for all $r > 0, j = 2, \cdots, n-1$, and therefore $X_j/(1 - S_{j-1})$ is beta $(\varepsilon_j, \alpha + \gamma_j)$, for each $j = 2, \cdots, n-1$. Further, $X_1 = S_1$ is beta $(\beta_1, \alpha + \gamma_1)$ and $X_n/(1 - S_{n-1})$ is beta (α, γ_{n-1}). Since

$$\alpha + \gamma_j = \alpha + \varepsilon_{j+1} + \gamma_{j+1}, \qquad j = 1, \cdots, n-2,$$

and since the mutual independence of the n variables,

$$X_1, X_j/(1 - S_{j-1}), \qquad j = 2, \cdots, n,$$

follows from the complete neutrality, application of a result of Connor and Mosimann (1969, page 200) shows that $\mathbf{X}$ is Dirichlet $(\beta_1, \varepsilon_2, \cdots, \varepsilon_{n-1}, \alpha, \gamma_{n-1})$. This completes the proof.

2. A parallel characterization of the lognormal distribution. The concept of neutrality for positive random vectors was extended to neutrality with respect to a regular sequence of size variables by Mosimann (1975a). Mosimann (1975b) also conjectured the truth of Theorem 2 of the present paper, and proved Theorem 3 below which is a parallel characterization of the lognormal distribution. In order to introduce general concepts of neutrality for positive random vectors and to compare our Theorem 2 with Theorem 3, we briefly introduce the notion of regular sequences of size variables.

Let P^1 be the set of positive real numbers and P^n the vectors of n positive, real, coordinates. A function G_n: $P^n \to P^1$ is a "size variable" if it has the homogeneity property $G_n(a\mathbf{x}) = aG_n(\mathbf{x})$, for all $\mathbf{x} \in P^n$, $a \in P^1$. Now consider $\mathbf{x}^{n+1} \in P^{n+1}$, $n > 0$, and, henceforth, let $\mathbf{x}_i$ denote its first i coordinates. Thus, for example, $\mathbf{x}_{n+1} = (\mathbf{x}_n; x_{n+1})$. Define R and S, both $P^{n+1} \to P^1$, by

$$R(\mathbf{x}_{n+1}) = G_{n+1}(\mathbf{x}_{n+1})/G_n(\mathbf{x}_n),$$

$$S(\mathbf{x}_{n+1}) = x_{n+1}/G_n(\mathbf{x}_n),$$

for all $\mathbf{x}_{n+1} \in P^{n+1}$. It can be shown that S is onto P^1, but that R is not generally onto P^1; $N =$ Image (R) need not equal P^1. The size variables G_n, G_{n+1} are "related" if there exists F: $P^1 \to N$, with inverse F^{-1}: $N \to P^1$, such that $FS = R$ (Mosimann, 1975a, page 203). If we now define size variables G_i: $P^i \to P^1$, $i = 1, \cdots, n+1$, then $G_1, \cdots, G_{n+1}$ is a "regular" sequence of size variables if G_i, G_{i+1} are related $i = 1, \cdots, n$.

Now call $\mathbf{X}_{n+1}$ a positive random vector if each of its coordinates is a positive (scalar) random variable. Assume $G_1, \cdots, G_{n+1}$ regular, and each G_i measurable so that $G_i(\mathbf{X}_i)$ is a random variable, $i = 1, \cdots, n+1$. We have

DEFINITION 3. $\mathbf{X}_{n+1}$ is completely neutral (to the left) with respect to $G_1, \cdots, G_{n+1}$ if the random ratios

$$R_{i+1} = \frac{G_{i+1}(\mathbf{X}_{i+1})}{G_i(\mathbf{X}_i)}, \qquad i = 1, \cdots, n$$

are mutually independent. (Mosimann, 1975a, page 211).

We say that $\mathbf{X}_{n+1}$ is completely neutral to the right if the permuted vector $(X_{n+1}, \cdots, X_1)$ is completely neutral to the left. (Definition 2 of the previous section gave neutrality to the right.) We shall also say that the last coordinate, X_{n+1} is neutral with respect to $G_1, \cdots, G_{n+1}$ in $\mathbf{X}_{n+1}$ if R_{n+1} is independent of $(R_2, \cdots, R_n)$.

Unlike Definition 2, these definitions apply to arbitrary positive $\mathbf{X}_{n+1}$, constrained or not. If we consider $\mathbf{X}_{n+1}$ constrained so that $G_{n+1} = 1$, then letting $G_j = \Sigma_{i=1}^j X_i = S_j$, $j = 1, \cdots, n+1$ we obtain the neutrality of this paper. A characterization of the lognormal distribution is obtained by considering the regular sequence $M_j = (\Pi_{i=1}^j X_i)^{1/j}$, $j = 1, \cdots, n+1$ and using neutrality with respect to $M_1, \cdots, M_{n+1}$:

THEOREM 3. (Mosimann 1975b, Theorem 4). *If* $\mathbf{X}_{n+1}$ *is a positive nondegenerate random vector constrained by* $M_{n+1} = 1$, *then the following two assertions are equivalent.*

(i) $\mathbf{X}_{n+1}$ *is completely neutral (to the right) and* X_{n+1} *is neutral in* $\mathbf{X}_{n+1}$.

(ii) $\mathbf{X}_{n+1}$ *is lognormal* $(\boldsymbol{\mu}, \mathbf{H})$.

(Here $\mathbf{H}$, $n + 1$ by $n + 1$ is defined by $h_{ij} = -v$, $i \neq j$; $h_{ii} = nv$, $i = 1, \cdots, n + 1$; $v > 0$, and $\mathbf{X}$ is lognormal $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ if $\mathbf{Y}$, with coordinates $Y_i = \log X_i$, is normal $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. This theorem is a direct parallel of our characterization of the Dirichlet distribution in Theorem 2. In fact, for nondegenerate positive $\mathbf{X}_{n+1}$ constrained so that $S_{n+1} = 1$, Theorem 3 holds exactly for neutrality with respect to $S_1, \cdots, S_{n+1}$ and substitution of "Dirichlet" for "lognormal $(\boldsymbol{\mu}, \mathbf{H})$".

To show that this is implied by Theorem 2 where X_n, not X_{n+1}, is neutral in $\mathbf{X}_{n+1}$, we note that $(X_1, \cdots, X_n, X_{n+1})$ is completely neutral with respect to $S_1, \cdots, S_{n+1}$ if and only if $(X_1, \cdots, X_{n+1}, X_n)$ is. Also $1 - R_{n+1}^{-1} = X_{n+1}$, and R_{n+1} independent of $R_2, \cdots, R_n$ is equivalent to the neutrality of X_{n+1} in the n-coordinate vector $(X_1, \cdots, X_{n-1}, X_{n+1})$ as defined in Section 1.

The lognormal $(\boldsymbol{\mu}, \mathbf{H})$ distribution can be generated by considering $\mathbf{X}_{n+1}$ lognormal $(\cdot, \sigma^2\mathbf{I})$. The distribution of $\mathbf{X}_{n+1}/M_{n+1}$ is then lognormal $(\boldsymbol{\mu}, \mathbf{H})$ where $v = \sigma^2/(n + 1)$. This reinforces the parallel with the Dirichlet distribution which is the distribution of $\mathbf{X}_{n+1}/S_{n+1}$ when $X_1, \cdots, X_{n+1}$ are independent gamma variables each with the same scale parameter.

Other points of interest are that no nondegenerate member of the lognormal family can exhibit neutrality with respect to $S_1, \cdots, S_{n+1}$ (Mosimann, 1975b, Theorem 1) and that for arbitrary positive $\mathbf{X}_{n+1}$ neutrality can occur with respect to *at most* one regular sequence of size variables (1975a, Theorem 3). Given this latter result, the characterization of distributions through neutrality is hardly surprising.

3. Discussion. Concepts of neutrality with respect to $S_1, \cdots, S_{n+1}$ have played an important role in recent years in the area of random probabilities. Thus if $F(t)$ is a random distribution function, it is said to be *neutral to the right* if the random proportions $(F(t_1), F(t_2) - F(t_1), \cdots, F(t_{n-1}) - F(t_{n-2}), 1 - F(t_n))$ are completely neutral for all $t_1 < t_2 < \cdots < t_n$. (Doksum 1972, 1974). This notion can be related to that of *tailfree* distributions (Freedman 1963), and may be defined in terms of independent increments processes (Doksum 1974, Theorem 3.1). Doksum shows that the posterior distribution of a random probability neutral to the right is also neutral to the right, and conjectures (page 193) that apart from the exceptional cases where the random probability measure P is

(i) degenerate at a given probability P_0,

(ii) concentrated on a random point,

or

(iii) concentrated on two nonrandom points, the only process which is both neutral to the right and neutral to the left is the Dirichlet process of Ferguson (1973). Neutrality to the left is defined analogously to neutrality to the right, except that the partitions are taken in the reverse order $(t_n, \infty), (t_{n-1}, t_n], \cdots, (t_1, t_2], (-\infty, t_1]$. It is straightforward to show that this implies (and is considerably stronger than) the property "$1-F(t_n)$ is neutral in

$$(F(t_1), F(t_2) - F(t_1), \cdots, F(t_{n-1}) - F(t_{n-2}), 1 - F(t_n))\text{''}.$$

Thus for each partition, Theorem 2 states that the proportions are Dirichlet or limits of Dirichlets, and Doksum's conjecture is proved as in Doksum (1972). Other characterizations of the Dirichlet process from neutrality properties in Theorem 1 are given by Doksum (1974).

The concept of complete neutrality has also been utilized recently in the assignment of prior distributions for the cell probabilities in Bayesian life-table analyses with grouped data (Lochner and Basu 1972, Lochner 1975). Thus suppose individuals are put on test at time 0 and followed for a maximum time t, during which they either fail, are lost to follow-up or are known to survive. Let p_i denote the probability of failure in $(t_{i-1}, t_i]$, $i = 1, \cdots, n$ where $t_0 = 0 < t_1 < \cdots < t_n = t$, $\Sigma_{i=1}^n p_i < 1$. From the likelihood function a seemingly natural assumption for a prior distribution for $\mathbf{p} = (p_1, \cdots, p_n)$ is that the conditional failure probabilities $p_1, p_i/(1 - \Sigma_{j=1}^{i-1} p_j)$, $i = 2, \cdots, n$ be independent; i.e. that $\mathbf{p}$ be completely neutral (Lochner and Basu 1972). According to Theorem 2, if we assume in addition that $\Sigma_1^n p_j$ is independent of $\{p_i/\Sigma_1^n p_j;\ i = 1, \cdots, n-1\}$, i.e. that the prior conditional cell probabilities given failure before t are distributed independently of the probability of such failure, then the prior for $\mathbf{p}$ must be Dirichlet (or degenerate).

General concepts of neutrality are related to concepts of isometry and relative growth in biology by Mosimann (1975a).

REFERENCES

Connor, R. J. and Mosimann, J. E. (1969). Concepts of independence for proportions with a generalization of the Dirichlet distribution. *J. Amer. Statist. Assoc.* **64** 194–206.

Doksum, K. A. (1971). Tailfree and neutral processes and their posterior distributions. ORC Report 71-72, Univ. California, Berkeley.

Doksum, K. A. (1974). Tailfree and neutral random probabilities and their posterior distributions. *Ann. Probability* **2** 183–201.

Fabius, J. (1973). Two characterizations of the Dirichlet distribution. *Ann. Statist.* **1** 583–587.

Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209–230.

Freedman, D. A. (1963). On the asymptotic behavior of Bayes estimates in the discrete case. *Ann. Math. Statist.* **34** 1386–1403.

Lochner, R. H. (1975). A generalized Dirichlet distribution in Bayesian life testing. *J. Roy. Statist. Soc. Ser. B* **37** 103–113.

Lochner, R. H. and Basu, A. P. (1972). Bayesian analysis of the two-sample problem with incomplete data. *J. Amer. Statist. Assoc.* **67** 432–438.

MOSIMANN, J. E. (1975A). Statistical problems of size and shape. I. Biological applications and basic theorems. In *Statistical Distributions in Scientific Work, Vol.* 2 (G. P. Patil, S. Kotz, J. K. Ord, eds.), 187–217. D. Reidel, Dordrecht.

MOSIMANN, J. E. (1975b). Statistical problems of size and shape. II. Characterizations of the lognormal, gamma and Dirichlet distributions. in *Statistical Distributions in Scientific Work*, Vol. 2 (G. P. Patil, S. Kotz, J. K. Ord, eds.), 219–239. D. Reidel, Dordrecht.

C.S.I.R.O.
SOUTH MELBOURNE
VICTORIA, AUSTRALIA

LAB OF STATISTICAL AND
MATHEMATICAL METHODOLOGY
DIVISION OF COMPUTER RESEARCH
AND TECHNOLOGY
N.I.H., BUILDING 12A, ROOM 3047
BETHESDA, MD 20014