The Annals of Mathematical Statistics
1968, Vol. 39, No. 5, 1473–1478

DISTRIBUTIONS DETERMINED BY CUTTING A SIMPLEX WITH HYPERPLANES[1]

By A. P. Dempster and Robert M. Kleyle

Harvard University and the University of Massachusetts

1. Introduction and summary. Suppose that $X_1, X_2, \ldots, X_n$ are random variables uniformly distributed over the simplex of points $x_1, x_2, \cdots, x_n$ such that $x_i \geqq 0$ for $i = 1, 2, \cdots, n$ and $\sum_1^n x_i \leqq 1$. The distribution of $X = \sum_1^n c_i X_i$ for constants c_i satisfying $c_1 > c_2 > \cdots > c_n > 0$ is easily seen to be given by

$$(1.1)\quad P(X \leqq x) = 1 - \sum_{j=1}^{r} (c_j - x)^n [c_j \prod_{i \neq j} (c_j - c_i)]^{-1}$$
$$= x^n [\prod_{i=1}^n c_i]^{-1} - \sum_{j=r+1}^n (x - c_j)^n [c_j \prod_{i \neq j} (c_i - c_j)]^{-1}$$

for $0 \leqq x \leqq c_1$, where r is the largest positive integer such that $x \leqq c_r$. A geometric derivation is given in Section 2 which relies on a principle of inclusion and exclusion and identifies the terms in (1.1) as volumes of various simplices. It is remarked in Section 3 that the argument of Section 2 extends in principle to give the joint distribution of a set of different linear combinations of $X_1, X_2, \cdots, X_n$. Finally, the relations to the theory of order statistics from a uniform distribution and to the theory of serial correlation are noted.

A simple derivation of (1.1) will now be sketched. Suppose that $Z_1, Z_2, \cdots, Z_{n+1}$ are $n + 1$ independent random variables each with the standard exponential density function $\exp(-z)$ for $0 \leqq z < \infty$. By a well-known argument involving the partial fraction expansion of the characteristic function (cf, Box (1954) Theorem 2.4), the density of $Z = \sum_1^n c_j Z_j$ is found to be $\sum_1^n (w_j/c_j) \exp(-z/c_j)$ where

$$(1.2)\qquad w_j = c_j^{n-1} [\prod_{i \neq j} (c_j - c_i)]^{-1}.$$

Now a set of random variables $X_1, X_2, \cdots X_n$ uniformly distributed over the simplex as described above may be represented as $X_j = Z_j/Y$ for $j = 1, 2, \cdots, n$ where $Y = \sum_1^{n+1} Z_j$, and the set $X_1, X_2, \cdots, X_n$ thus created is distributed independently of Y. Thus $X = \sum_1^n c_j X_j$ may be represented as

$$(1.3)\qquad X = Z/Y$$

where X and Y are independent. In a relation like (1.3) the marginal distributions of Z and Y together with the fact of independence of X and Y are sufficient to determine uniquely the distribution of X. To see this, note that cf $(\log X) \cdot$ cf $(\log Y)$ = cf $(\log Z)$, where cf denotes the characteristic function. Thus, the

Received 25 July 1967.
[1] This research was supported by the Office of Naval Research. Distribution of the document is unlimited.

characteristic function of $\log X$ is determined by those of $\log Y$ and $\log Z$. Now if Z were known to have the scaled exponential density $c^{-1} \exp(-z/c)$ for $z \geqq 0$ while Y had the gamma density $y^n \exp(-y)/n!$ for $y \geqq 0$, it would follow that X had the scaled beta density $nc^{-1}(1 - x/c)^{n-1}$ for $0 \leqq x \leqq c$. But the density of Z is, in fact, a mixture of scaled exponentials. Thus, the density of X is the same mixture of scaled betas and is given by $\sum_1^n w_j f_j(x)$, where the w_j's are defined by (1.2) and

$$f_j(x) = nc_j^{-1}(1 - x/c_j)^{n-1} \quad \text{for} \quad 0 \leqq x \leqq c_j \qquad (1.4)$$
$$= 0 \qquad \text{otherwise.}$$

The result just given yields immediately the first line of (1.1). To derive the second line of (1.1), set $X_{n+1} = 1 - \sum_1^n X_j$ and note that $c_1 - X = c_1X_{n+1} + (c_1 - c_n)X_n + \cdots + (c_1 - c_2)X_2$ has a distribution of the same type as X. Applying the first line of (1.1) to $c_1 - X$ yields the second line of (1.1).

The distribution of $\sum_1^{n+1} d_iX_i$ is also given by (1.1), where $X_1, X_2, \cdots, X_{n+1}$ are uniformly distributed over the simplex $x_i \geqq 0$ for $i = 1, 2, \cdots, n + 1$ and $\sum_1^{n+1} x_i = 1$, and where constants d_i satisfy the condition

$$d_1 > d_2 > \cdots > d_r \geqq x > d_{r+1} > \cdots > d_{n+1}. \qquad (1.5)$$

To see this, let $c_i = d_i - d_{n+1}$ and notice that $\sum_1^{n+1} d_iX_i = X + d_{n+1}$.

2. The geometric derivation. The event whose probability is expressed in (1.1) may be described as $\{\sum_1^n a_jX_j \leqq 1\}$ where

$$a_j = c_j/x \qquad (2.1)$$

for $j = 1, 2, \cdots, n$ and

$$a_1 > a_2 > \cdots > a_r \geqq 1 > a_{r+1} > \cdots > a_n > 0. \qquad (2.2)$$

Consider the following half-spaces of the ordinary Euclidean n-dimensional space with typical point $(x_1, x_2, \cdots x_n)$:

$$H_j^+ : x_j \geqq 0, \qquad H_j^- : x_j < 0, \quad \text{for} \quad j = 1, 2, \cdots, n,$$
$$A^+ : \textstyle\sum_1^n a_jx_j \leqq 1, \qquad A^- : \sum_1^n a_jx_j > 1, \qquad (2.3)$$
$$I^+ : \textstyle\sum_1^n x_j \leqq 1, \quad \text{and} \quad I^- : \sum_1^n x_j > 1.$$

Since $X_1, X_2, \cdots, X_n$ are uniformly distributed over the simplex

$$S = I^+ \cap H_1^+ \cap H_2^+ \cap \cdots \cap H_n^+, \qquad (2.4)$$

the desired probability (1.1) may be described as $\langle S \cap A^+\rangle/\langle S\rangle$ where $\langle\cdots\rangle$ denotes n-dimensional Euclidean volume. Since

$$\langle S\rangle = (n!)^{-1}, \qquad (2.5)$$

it remains only to find

$$\langle S \cap A^+\rangle = \langle A^+ \cap I^+ \cap H_1^+ \cap H_2^+ \cap \cdots \cap H_n^+\rangle. \qquad (2.6)$$

The following three lemmas provide the desired volume as a linear expression in terms of the volumes of simplices.

LEMMA 1. *The regions*

$$A^- \cap I^+ \bigcap_{i=1}^{j} H_i^- \bigcap_{i=j+1}^{n} H_i^+ \tag{2.7}$$

for $j = 1, 2, \cdots, r-1$ *and*

$$A^+ \cap I^- \bigcap_{i=1}^{n-j} H_i^+ \bigcap_{i=n-j+1}^{n} H_i^- \tag{2.8}$$

for $j = 1, 2, \cdots, n-r-1$ *are nonempty.*

PROOF. From (2.2) we may pick p such that

$$0 < p < (a_{j+1} - 1)(a_1 - a_{j+1})^{-1} \tag{2.9}$$

for given j selected from $1 \leqq j \leqq r-1$. Since (2.9) implies that $(1 + pa_1)/a_{j+1} < 1 + p$, we may further pick q such that

$$(1 + pa_1)/a_{j+1} < q < 1 + p. \tag{2.10}$$

Now let

$$\begin{aligned} x_i &= -p/j \quad \text{for} \quad i = 1, 2, \cdots, j, \\ x_{j+1} &= q, \qquad \text{and} \\ x_i &= 0 \qquad \text{for} \quad i = j+2, \cdots, n. \end{aligned} \tag{2.11}$$

It is easily checked that the point (2.11) belongs to the region (2.7), as required. The second part of the lemma may be proved similarly.

LEMMA 2. *The regions*

$$A^- \cap I^+ \bigcap_{i=1}^{r} H_i^- \bigcap_{i=r+1}^{n} H_i^+ \tag{2.12}$$

and

$$A^+ \cap I^- \bigcap_{i=1}^{r} H_i^+ \bigcap_{i=r+1}^{n} H_i^- \tag{2.13}$$

are empty.

PROOF. To satisfy (2.12), the point $(x_1, x_2, \cdots, x_n)$ must have $x_i < 0$ for $i = 1, 2, \cdots, r$ which together with (2.2) yields $\sum_1^r a_i x_i \leqq \sum_1^r x_i$. Similarly (2.12) yields $x_i \geqq 0$ for $i = r+1, \cdots, n$ which together with (2.2) yields $\sum_{r+1}^n a_i x_i < \sum_{r+1}^n x_i$. Combining we find $\sum_1^n a_i x_i \leqq \sum_1^n x_i$ which contradicts $A^- \cap I^+$. The result for (2.13) follows similarly.

LEMMA 3.

$$\begin{aligned} \langle A^+ \cap I^+ \textstyle\bigcap_{i=1}^{n} H_i^+ \rangle &= \langle I^+ \textstyle\bigcap_{i=1}^{n} H_i^+ \rangle \\ &\quad + \textstyle\sum_{j=1}^{r} (-1)^j \langle A^- \cap I^+ \bigcap_{i=1}^{j-1} H_i^- \bigcap_{i=j+1}^{n} H_i^+ \rangle \\ &= \langle A^+ \textstyle\bigcap_{i=1}^{n} H_i^+ \rangle \\ &\quad + \textstyle\sum_{j=1}^{n-r} (-1)^j \langle A^+ \cap I^- \bigcap_{i=1}^{n-j} H_i^+ \bigcap_{i=n-j+2}^{n} H_i^- \rangle. \end{aligned} \tag{2.14}$$

PROOF. The region $I^+ \bigcap_1^n H_i^+$ is the union of the disjoint regions

$A^+ \cap I^+ \bigcap_1^n H_i^+$ and $A^- \cap I^+ \bigcap_1^n H_i^+$, so that

$$(2.15)\quad \langle A^+ \cap I^+ \textstyle\bigcap_{i=1}^n H_i^+ \rangle = \langle I^+ \bigcap_{i=1}^n H_i^+ \rangle - \langle A^- \cap I^+ \bigcap_{i=1}^n H_i^+ \rangle.$$

Similarly,

$$(2.16)\quad \langle A^- \cap I^+ \textstyle\bigcap_{i=1}^n H_i^+ \rangle = \langle A^- \cap I^+ \bigcap_{i=2}^n H_i^+ \rangle - \langle A^- \cap I^+ \cap H_1^- \bigcap_{i=2}^n H_i^+ \rangle,$$

and from (2.15) and (2.16)

$$(2.17)\quad \begin{aligned} \langle A^+ \cap I^+ \textstyle\bigcap_{i=1}^n H_i^+ \rangle = \langle I^+ \bigcap_{i=1}^n H_i^+ \rangle &- \langle A^- \cap I^+ \textstyle\bigcap_{i=2}^n H_i^+ \rangle \\ &+ \langle A^- \cap I^+ \cap H_1^- \textstyle\bigcap_{i=2}^n H_i^+ \rangle. \end{aligned}$$

Continuing by induction, it can be shown that for any $k \leqq r$,

$$(2.18)\quad \begin{aligned} \langle A^+ \cap I^+ \textstyle\bigcap_{i=1}^n H_i^+ \rangle &= \langle I^+ \textstyle\bigcap_{i=1}^n H_i^+ \rangle \\ &\quad + \textstyle\sum_{j=1}^k (-1)^j \langle A^- \cap I^+ \bigcap_{i=1}^{j-1} H_i^- \bigcap_{i=j+1}^n H_i^+ \rangle \\ &\quad + (-1)^{k+1} \langle A^- \cap I^+ \textstyle\bigcap_{i=1}^k H_i^- \bigcap_{i=k+1}^n H_i^+ \rangle. \end{aligned}$$

Now from Lemmas 1 and 2, it is clear that the first time

$$A^- \cap I^+ \textstyle\bigcap_{i=1}^k H_i^- \bigcap_{i=k+1}^n H_i^+$$

is empty occurs when $k = r$. Therefore, the summation stops at r, and the first line of (2.14) holds.

The proof of the second line of (2.14) follows in a completely analogous way.

The two lines of (1.1) follow term by term from the two lines of (2.14) after dividing through by $n!$. It remains only to identify the vertices of the simplices whose volumes appear in (2.14) and then to check that the expressions in (1.1) are essentially the volumes of these simplices. This computation is left to the reader.

3. Multivariate extensions. Suppose that $X_1, X_2, \cdots, X_n$ are distributed as in Section 1, and consider the problem of computing $P(X^{(1)} \leqq x^{(1)}, X^{(2)} \leqq x^{(2)}, \cdots)$ where $X^{(k)} = \sum_1^n c_i^{(k)} X_i$ for $k = 1, 2, \cdots$. In principle, such probabilities follow easily from the geometric approach of Section 2, and numerically they may be easily found. Consider the case of $X^{(1)}$ and $X^{(2)}$. Formulas (1.1) or (2.14) may be used to find $P(X^{(1)} \leqq x^{(1)})$. The additional condition $X^{(2)} \leqq x^{(2)}$ partitions each simplex in (2.14) into 2 pieces (one of which may be empty). Finding the corresponding partitions of volume may be reduced in each case to an application of (2.14). Continuing in this way, it is clear that repeated applications of (2.14) yield bivariate, trivariate, etc., extensions of (1.1).

Explicit formulas for $P(X^{(1)} \leqq x^{(1)}, X^{(2)} \leqq x^{(2)}, \cdots)$ do not appear to be illuminating, due to the multitude of cases created by the relative orders of the $c_i^{(k)}$ and $x^{(k)}$. Some special cases are given by Kleyle (1967).

4. Relation to order statistics. Suppose that $U_1, U_2, \cdots, U_n$ are independent and identically distributed on the interval (0, 1) and suppose that $U_{(1)} \leqq$

$U_{(2)} \leqq \cdots \leqq U_{(n)}$ denote the ordered values of $U_1, U_2, \cdots, U_n$. Then $X_1 = U_{(1)}, X_2 = U_{(2)} - U_{(1)}, \cdots, X_n = U_{(n)} - U_{(n-1)}$ have the distribution treated in Section 1. Thus it is clear that

$$(4.1) \qquad X = (c_1 - c_2)U_{(1)} + (c_2 - c_3)U_{(2)} + \cdots + c_n U_{(n)},$$

and formula (1.1) gives directly the distribution of $\sum_1^n d_i U_{(i)}$ with each $d_i > 0$. Actually, by simple reordering and shifting one can handle the general case if $d_i \neq 0$. In principle, therefore, the joint distribution function of several linear combinations of order statistics may be directly written down.

Note also that since $U_{(i)} = X_1 + X_2 + \cdots + X_i$ for $i = 1, 2, \cdots, n$

$$(4.2) \qquad \sum_1^n U_i = \sum_1^n U_{(i)} = \sum_1^n (n - i + 1)X_i.$$

Olds (1952) has given the distribution of $\sum_1^n c_i U_i$, and in the special case where $c_i = c > 0$ for $i = 1, 2, \cdots n$, the cdf given by Olds can be written in the form

$$(4.3) \qquad F_n(x) = (c^n n!)^{-1} \sum_{j=0}^r (-1)^j \binom{n}{j} (x - cj)^n,$$

where $cr < x \leqq c(r + 1)$. Thus, when $c_i = c(n - i + 1)$ for $i = 1, 2, \cdots n$, the distribution of X is given by (4.3). This result can be easily checked by setting $c_i = c(n - i + 1)$ in the second line of (1.1).

When the c_i's are not all the same for $i = 1, 2, \cdots, n$, it is no longer clear that $\sum_1^n c_i U_i$ can be written as a linear combination of $X_1, X_2, \cdots, X_n$. Thus, in general, the cdf given by (1.1) and that given by Olds are not equivalent.

5. Relation to serial correlation. The circular serial correlation coefficient with lag L is given by

$$(5.1) \qquad {}_LR_N = \sum_1^N (X_j - \bar{X})(X_{j+L} - \bar{X}) / \sum_1^N (X_j - \bar{X})^2,$$

where $X_{N+j} = X_j$ and $L <$ sample size N. It has been shown that if N is odd, and if $X_1, X_2, \cdots, X_N$ are assumed to be independent normal variates with zero means and unit variances,

$$(5.2) \qquad {}_LR_N = \sum_1^{n+1} {}_L\lambda_j Z_j / \sum_1^{n+1} Z_j,$$

where $n + 1 = (N - 1)/2$, ${}_L\lambda_1, {}_L\lambda_2, \cdots, {}_L\lambda_{n+1}$ are the distinct latent roots of the matrix of the quadratic form in the numerator of (5.1), and $Z_1, Z_2, \cdots, Z_{n+1}$ are the independent, exponentially distributed random variables of Section 1. Thus it is clear that if ${}_LC_j$ denotes the jth largest latent root, the marginal distribution of ${}_LR_N$ is given by (1.1) with c_j replaced by ${}_LC_j - {}_LC_{n+1}$ and x by $x - {}_LC_{n+1}$. Furthermore, by applying the method discussed in Section 3, the joint distribution of ${}_1R_N, {}_2R_N, \cdots, {}_LR_N, L < n + 1$, can be found.

Anderson (1942) has given the marginal distributions of the circular serial correlation coefficient, while the joint distribution has been derived by Watson (1956). The method used by Watson is a generalization of the derivation given in Section 1. Although Watson does point out some of the geometrical implications of his proof, neither he nor Anderson attempt a direct geometrical argument.

The authors of this paper are indebted to J. Durbin for pointing out the relationship between their results and the distributions of the serial correlation coefficients.

REFERENCES

[1] ANDERSON, R. L. (1942). Distribution of the serial correlation coefficient. *Ann. Math. Statist.* **13** 1–13.

[2] BOX, G. E. P. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effect of inequality of variance in the one-way classification. *Ann. Math. Statist.* **25** 290–302.

[3] KLEYLE, ROBERT. (1967). An application of the direct probability system of inference. Ph.D. thesis, Department of Statistics, Harvard Univ.

[4] OLDS, EDWIN G. (1952). A note on the convolution of uniform distributions. *Ann. Math. Statist.* **23** 282–285.

[5] WATSON, G. S. (1956). On the joint distribution of the circular serial correlation coefficients. *Biometrika* **43** 161–168.