

DISTRIBUTIONS DETERMINED BY CUTTING A SIMPLEX WITH HYPERPLANES¹

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

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

for $0 \leq x \leq c_1$, where r is the largest positive integer such that $x \leq c_r$. 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, \dots, 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, \dots, Z_{n+1} are $n + 1$ independent random variables each with the standard exponential density function $\exp(-z)$ for $0 \leq z < \infty$. 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

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

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

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

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 $\text{cf}(\log X) \cdot \text{cf}(\log Y) = \text{cf}(\log Z)$, where cf denotes the characteristic function. Thus, the

Received 25 July 1967.

¹This 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 $z \geq 0$ while Y had the gamma density $y^n \exp(-y)/n!$ for $y \geq 0$, it would follow that X had the scaled beta density $nc^{-1}(1 - x/c)^{n-1}$ for $0 \leq x \leq c$. 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

$$(1.4) \quad \begin{aligned} f_j(x) &= nc_j^{-1}(1 - x/c_j)^{n-1} && \text{for } 0 \leq x \leq c_j \\ &= 0 && \text{otherwise.} \end{aligned}$$

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 $c_1 - X = c_1 X_{n+1} + (c_1 - c_n)X_n + \dots + (c_1 - c_2)X_2$ 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 $\sum_1^{n+1} d_i X_i$ is also given by (1.1), where X_1, X_2, \dots, X_{n+1} are uniformly distributed over the simplex $x_i \geq 0$ for $i = 1, 2, \dots, n + 1$ and $\sum_1^{n+1} x_i = 1$, and where constants d_i satisfy the condition

$$(1.5) \quad d_1 > d_2 > \dots > d_r \geq x > d_{r+1} > \dots > d_{n+1}.$$

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

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

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

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

$$(2.2) \quad a_1 > a_2 > \dots > a_r \geq 1 > a_{r+1} > \dots > a_n > 0.$$

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

$$(2.3) \quad \begin{aligned} H_j^+ : x_j &\geq 0, & H_j^- : x_j < 0, & \text{for } j = 1, 2, \dots, n, \\ A^+ : \sum_1^n a_j x_j &\leq 1, & A^- : \sum_1^n a_j x_j &> 1, \\ I^+ : \sum_1^n x_j &\leq 1, & \text{and } I^- : \sum_1^n x_j &> 1. \end{aligned}$$

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

$$(2.4) \quad S = I^+ \cap H_1^+ \cap H_2^+ \cap \dots \cap H_n^+,$$

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

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

it remains only to find

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

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

LEMMA 1. *The regions*

$$(2.7) \quad A^- \cap I^+ \cap_{i=1}^j H_i^- \cap_{i=j+1}^n H_i^+$$

for $j = 1, 2, \dots, r - 1$ and

$$(2.8) \quad A^+ \cap I^- \cap_{i=1}^{n-j} H_i^+ \cap_{i=n-j+1}^n H_i^-$$

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

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

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

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

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

Now let

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

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*

$$(2.12) \quad A^- \cap I^+ \cap_{i=1}^r H_i^- \cap_{i=r+1}^n H_i^+$$

and

$$(2.13) \quad A^+ \cap I^- \cap_{i=1}^r H_i^+ \cap_{i=r+1}^n H_i^-$$

are empty.

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

LEMMA 3.

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

PROOF. The region $I^+ \cap_{i=1}^n H_i^+$ is the union of the disjoint regions

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

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

Similarly,

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

and from (2.15) and (2.16)

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

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

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

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

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

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, \dots, X_n are distributed as in Section 1, and consider the problem of computing $P(X^{(1)} \leq x^{(1)}, X^{(2)} \leq x^{(2)}, \dots)$ where $X^{(k)} = \sum_1^n c_i^{(k)} X_i$ for $k = 1, 2, \dots$. 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 $P(X^{(1)} \leq x^{(1)})$. The additional condition $X^{(2)} \leq x^{(2)}$ 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 $P(X^{(1)} \leq x^{(1)}, X^{(2)} \leq x^{(2)}, \dots)$ 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, \dots, U_n are independent and identically distributed on the interval $(0, 1)$ and suppose that $U_{(1)} \leq$

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

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

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 $U_{(i)} = X_1 + X_2 + \dots + X_i$ for $i = 1, 2, \dots, n$

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

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 $i = 1, 2, \dots, n$, the cdf given by Olds can be written in the form

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

where $cr < x \leq c(r + 1)$. Thus, when $c_i = c(n - i + 1)$ for $i = 1, 2, \dots, n$, 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 $i = 1, 2, \dots, n$, 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, \dots, 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

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

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

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

where $n + 1 = (N - 1)/2$, ${}_L \lambda_1, {}_L \lambda_2, \dots, {}_L \lambda_{n+1}$ are the distinct latent roots of the matrix of the quadratic form in the numerator of (5.1), and Z_1, Z_2, \dots, Z_{n+1} are the independent, exponentially distributed random variables of Section 1. Thus it is clear that if ${}_L C_j$ denotes the j th largest latent root, the marginal distribution of ${}_L R_N$ is given by (1.1) with c_j replaced by ${}_L C_j - {}_L C_{n+1}$ and x by $x - {}_L C_{n+1}$. Furthermore, by applying the method discussed in Section 3, the joint distribution of ${}_1 R_N, {}_2 R_N, \dots, {}_L R_N, L < n + 1$, 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.