The Annals of Probability
1983, Vol. 11, No. 3, 823–826

A CHARACTERIZATION OF RECTANGULAR DISTRIBUTIONS

BY GEORGE R. TERRELL[1]

Rice University

It is well known that the smaller and the larger of a random sample of size two are positively correlated. The coefficient of correlation is at most one-half, and the upper bound is attained only for rectangular distributions.

1. Introduction. There has been considerable interest in recent years in alternate characterizations of distributions in terms of properties of independent samples from them. The most famous of these is the characterization of a normal distribution as one whose measures of location and scale are independently distributed. A number of these characterizations are phrased in terms of properties of order statistics, where the ith order statistic $X_{(i)}$ of a sample of n is that sample point with $i-1$ smaller sample points and $n-1$ larger sample points. For example, an exponential distribution is one for which $X_{(1)}$, $X_{(2)}-X_{(1)}, X_{(3)}-X_{(2)}, \cdots, X_{(n)}-X_{(n-1)}$ are mutually independent. The main result of this note may be interpreted as a simple condition for a continuous distribution with finite variance to be a rectangular distribution. For other characterizations of this kind, see e.g. Ferguson (1967) and Driscoll (1978).

Bartoszyński (1980) proposed that a result of this type might exist in connection with a problem in cell division. Since the two daughter cells cannot always be distinguished later, the times till their further division can only be recorded as the earlier event and the later event. The correlation between these ordered pairs thus may provide the only information on the independence of the two events.

2. A bound for correlations between ordered sample pairs.

THEOREM. *Let $X_{(1)} \le X_{(2)}$ be an independent sample of two from a continuous distribution F with finite variance. Then $X_{(1)}, X_{(2)}$ are positively correlated, and $\rho_{12} \le 1/2$ with equality if and only if F is a rectangular distribution.*

PROOF. Let $F^{-1}(Y)$ be the inverse of the cumulative distribution function. It is clear from a change of variables that X having a variance is equivalent to the condition:

$$\int_0^1 (F^{-1}(Y))^2\, dY$$

is finite. Thus F^{-1} is a member of the class of nondecreasing square-integrable functions on (0, 1). Now, functions in this class may be approximated arbitrarily well by polynomials, using the L^2-metric on (0, 1). Let $\{L_i(Y)\}$ $i = 0, 1 \cdots$ be the normalized Legendre polynomials on [0, 1] with the properties

$$\int_0^1 L_i(Y)L_j(Y)\, dY = \delta_{ij} \quad \text{and} \quad L_i(1) > 0 \text{ for all } i, j.$$

Then we can expand

$$F^{-1}(Y) = \sum_{i=0}^{\infty} a_i L_i(Y)$$

Received June 1982; revised July 1982.

[1] Work partially supported by Contract NAS9-15800 to the National Aeronautics and Space Administration.

AMS 1980 *subject classifications.* Primary 60E15; secondary 62E10.

Key words and phrases. Rectangular distribution, uniform distribution, order statistics, Legendre polynomials, coefficient of correlation.

where

$$\int_0^1 L_i(Y)F^{-1}(Y)\,dY = a_i,$$

and thus

$$\int_0^1 F^{-1}(Y)^2\,dY = \textstyle\sum_{i=0}^{\infty} a_i^2 < \infty$$

and our polynomial series converges in the L^2 metric. We will need some particular properties of the Legendre polynomials in the sequel:

LEMMA. (i) $$\int_{0\le X\le Y\le 1} 2L_i(X)L_i(Y)\,dX\,dY = 0, \quad i > 0$$

(ii) $$\int_{0\le X\le Y\le 1} 2L_{i-1}(Y)L_i(X)\,dY\,dX = \frac{-1}{\sqrt{(2i-1)(2i+1)}}, \quad i > 0$$

$$\int_{0\le X\le Y\le 1} 2L_{i-1}(X)L_i(Y)\,dX\,dY = \frac{1}{\sqrt{(2i-1)(2i+1)}}, \quad i > 0$$

(iii) $$\int_0^1 2XL_i^2(X)\,dX = 1, \quad i \ge 0$$

(iv) $$\int_0^1 2XL_i(X)L_{i-1}(X)\,dX = \frac{1}{\sqrt{(2i-1)(2i+1)}}, \; i > 0$$

PROOF. (i) and (iii) follow from the alternate oddness and eveness of L_i's with respect to ½. (ii) and (iv) follow from the observation that $\int P_i(X)L_i(X)$ (where P_i is a polynomial of degree i) depends only on the coefficient of X^i, since L_i is orthogonal to terms of lower degree. □

Let $X_{(1)} \le X_{(2)}$ be a sample of two from the distribution with quartile function F^{-1}. Then

$$\rho_{12} = \frac{\mathrm{Cov}(X_{(1)}, X_{(2)})}{\sqrt{\mathrm{Var}(X_{(1)})\mathrm{Var}(X_{(2)})}}$$

$$\begin{aligned}
\mathrm{Cov}(X_{(1)}, X_{(2)}) &= \int_{0\le X_1\le X_2\le 1} 2F^{-1}(X_{(1)})F^{-1}(X_{(2)})\,dX_{(1)}\,dX_{(2)} \\
&\quad - \int_0^1 2X_{(2)}F^{-1}(X_{(2)})\,dX_2 \int_0^1 2(1-X_{(1)})F^{-1}(X_{(1)})\,dX_1 \\
&= \textstyle\sum_{i=0}^{\infty} \displaystyle\int_{0\le X_{(1)}\le X_{(2)}\le 1} 2a_i^2 L_i(X_{(1)})L_i(X_{(2)})\,dX_{(1)}\,dX_{(2)} \\
&\quad + \textstyle\sum_{i=0}^{\infty} \displaystyle\int_{0\le X_{(1)}\le X_{(2)}\le 1} 2a_i a_{i+1} L_i(X_{(1)})L_{i+1}(X_{(2)})\,dX_{(1)}\,dX_{(2)} \\
&\quad + \textstyle\sum_{i=0}^{\infty} \displaystyle\int_{0\le X_{(1)}\le X_{(2)}\le 1} 2a_i a_{i+1} L_i(X_{(2)})L_{i+1}(X_{(1)})\,dX_{(1)}\,dX_{(2)} \\
&\quad - \left(a_0 + \frac{a_1}{\sqrt{3}}\right)\left(a_0 - \frac{a_1}{\sqrt{3}}\right)
\end{aligned}$$

The other cross product terms are zero by orthogonality of the L_i's. Now, applying the Lemma allows us to cancel almost all terms, leaving

$$\int_{0\le X_{(1)}\le X_{(2)}\le 1} 2a_1^2 L_0(X_{(1)})L_0(X_{(2)}) - a_0^2 + \frac{a_1^2}{3} = a_0^2 - a_0^2 + \frac{a_1^2}{3} = \frac{a_1^2}{3}.$$

Thus, the covariance of an ordered sample of two depends only on the second (or linear) Legendre coefficient.

Now

$$\mathrm{Var}(X_{(1)}) = \int_0^1 2(1 - X_{(1)})(F^{-1}(X_{(1)}))^2\, dX_{(1)} - \left[\int_0^1 2(1 - X_{(1)})F^{-1}(X_{(1)})\, dX_{(1)}\right]^2$$

$$= 2\textstyle\sum_{i=0}^{\infty} a_i^2 - \sum_{i=0}^{\infty} \displaystyle\int_0^1 2a_i^2 X_{(1)} L_i(X_{(1)})^2\, dX_{(1)}$$

$$- \textstyle\sum_{i=0}^{\infty} \displaystyle\int_0^1 4a_i a_{i+1} X_{(1)} L_i(X_{(1)}) L_{i+1}(X_{(1)})\, dX_{(1)} - \left(a_0 - \frac{a_1}{\sqrt{3}}\right)^2,$$

where once again a great many terms disappear because of the orthogonality of the L_i's to lower degree polynomials. And, once again applying the Lemma

$$= \textstyle\sum_{i=0}^{\infty} a_i^2 - \sum_{i=0}^{\infty} \displaystyle\frac{2(i+1)a_i a_{i+1}}{\sqrt{(2i+1)(2i+3)}} - \left(a_0 - \frac{a_1}{\sqrt{3}}\right)^2$$

These terms can be reorganized by completing squares:

$$\mathrm{Var}(X_{(1)}) = \textstyle\sum_{i=1}^{\infty} \displaystyle\left(\sqrt{\frac{i+1}{2i+1}}\, a_i - \sqrt{\frac{i+1}{2i+3}}\, a_{i+1}\right)^2$$

Similarly

$$\mathrm{Var}(X_{(2)}) = \textstyle\sum_{i=1}^{\infty} \displaystyle\left(\sqrt{\frac{i+1}{2i+1}}\, a_i + \sqrt{\frac{i+1}{2i+3}}\, a_{i+1}\right)^2$$

Let

$$c_i = \sqrt{\frac{i+1}{2i+1}}\, a_i - \sqrt{\frac{i+1}{2i+3}}\, a_{i+1}, \quad d_i = \sqrt{\frac{i+1}{2i+1}}\, a_i + \sqrt{\frac{i+1}{2i+3}}\, a_{i-1}$$

Then

$$\mathrm{Var}\, X_1 \mathrm{Var}\, X_2 = \textstyle\sum_{i=1}^{\infty} c_i^2 \sum_{i=1}^{\infty} d_i^2 \ge \left(\sum_{i=1}^{\infty} c_i\, d_i\right)^2$$

by Cauchy's inequality. But

$$c_i\, d_i = \frac{i+1}{2i+1}\, a_i^2 - \frac{i+1}{2i+3}\, a_{i+1}^2.$$

So

$$\mathrm{Var}\, X_{(1)} \mathrm{Var}\, X_{(2)} \ge \left(\frac{2}{3}\, a_1^2 + \textstyle\sum_{i=2}^{\infty} \displaystyle\frac{a_i^2}{2i+1}\right)^2 \ge \left(\frac{2}{3}\, a_1^2\right)^2$$

Thus

$$\rho_{12} = \frac{\mathrm{Cov}(X_{(1)}, X_{(2)})}{\sqrt{\mathrm{Var}\, X_{(1)} \mathrm{Var}\, X_{(2)}}} \le \frac{a_1^2/3}{2a_1^2/3} = \frac{1}{2}.$$

Further, equality is attained precisely when $a_2^2 = a_3^2 = \cdots = a_i^2 = \cdots = 0$. It is readily seen that this is the case only for a rectangular distribution. This completes the proof of the theorem.

3. Conclusions. Despite the computational nature of the preceeding proof, it is fairly easy to interpret the theorem. The covariance of a pair of sample values depends only on the second Legendre coefficients. The correlation is largest when the higher coefficients are zero; i.e., for a rectangular distribution. A simpler result of the same class states: The variance of a random variable is at least the square of the second Legendre coefficient, with equality only for rectangular distributions. Thus, the theorem is related to Bessel's inequality.

Acknowledgments. The author would like to thank Robert Bartoszyński and David Scott who proposed the problem of this paper and conjectured the theorem.

REFERENCES

BARTOSZYŃSKI, R. (1980). Personal communication.

DRISCOLL, M. F. (1978). On pairwise and mutual independence characterizations of rectangular distributions. *J. Amer. Statist. Assoc.* **73** 432–433.

FERGUSON, T. S. (1967). On characterizing distributions by properties of order statistics. *Sankhya, Series A* **29** 265–278.

MATHEMATICAL SCIENCES DEPARTMENT
RICE UNIVERSITY
BOX 1892
HOUSTON, TEXAS 77251