SOME CONCEPTS OF DEPENDENCE[1]

E. L. Lehmann

University of California, Berkeley

1. Summary and introduction. Problems involving dependent pairs of variables (X, Y) have been studied most intensively in the case of bivariate normal distributions and of 2×2 tables. This is due primarily to the importance of these cases but perhaps partly also to the fact that they exhibit only a particularly simple form of dependence. (See Examples 9(i) and 10 in Section 7.) Studies involving the general case center mainly around two problems: (i) tests of independence; (ii) definition and estimation of measures of association. In most treatments of these problems, there occurs implicitly a concept which is of importance also in other contexts (for example, the evaluation of the performance of certain multiple decision procedures), the concept of *positive* (or *negative*) dependence or association. Tests of independence, for example those based on rank correlation, Kendall's t-statistic, or normal scores, are usually not omnibus tests (for a discussion of such tests see [4], [15] and [17], but designed to detect rather specific types of alternatives, namely those for which large values of Y tend to be associated with large values of X and small values of Y with small values of X (positive dependence) or the opposite case of negative dependence in which large values of one variable tend to be associated with small values of the other. Similarly, measures of association are typically designed to measure the degree of this kind of association.

The purpose of the present paper is to give three successively stronger definitions of positive dependence, to investigate their consequences, explore the strength of each definition through a number of examples, and to give some statistical applications.

2. Quadrant dependence. For a first definition, we compare the probability of any quadrant $X \leqq x$, $Y \leqq y$ under the distribution F of (X, Y) with the corresponding probability in the case of independence. We say that the pair (X, Y) or its distribution F is *positively quadrant dependent* if

$$P(X \leqq x, Y \leqq y) \geqq P(X \leqq x)P(Y \leqq y) \qquad \text{for all } x, y. \tag{2.1}$$

The dependence is strict if inequality holds for at least some pair (x, y). The family of all distributions F satisfying (2.1) will be denoted by $\mathfrak{F}_1$. Similarly, (X, Y) or F is *negatively quadrant dependent* if (2.1) holds with the inequality sign reversed, and the totality of negatively quadrant dependent distributions

Received 3 September 1965; revised 2 February 1966.

[1] Prepared with the partial support of the Office of Naval Research, Contract Nonr-222-43. This paper in whole or in part may be reproduced for any purpose of the United States Government.

will be denoted by $\mathcal{G}_1$. To simplify the notation we shall write $(X, Y) \varepsilon \mathfrak{F}$ to mean that the distribution of (X, Y) belongs to $\mathfrak{F}$.

LEMMA 1.

(i) $(X, X) \varepsilon \mathfrak{F}_1$ *for all* X

(ii) $(X, Y) \varepsilon \mathfrak{F}_1 \Leftrightarrow (X, -Y) \varepsilon \mathcal{G}_1$

(iii) $(X, Y) \varepsilon \mathfrak{F}_1$ *implies* $(r(X), s(Y)) \varepsilon \mathfrak{F}_1$ *for all non-decreasing functions* r *and* s. *The concept of positive quadrant dependence is thus invariant under non-decreasing transformations (and similarly under nonincreasing transformations) of both variables.*

(iv) *The set of inequalities* (2.1) *is equivalent to that obtained by replacing one or both of the inequalities* $X \leqq x$ *or* $Y \leqq y$ *by the corresponding* $X < x$ *or* $Y < y$.

(v) *The set of inequalities* (2.1) *is equivalent to each of the following, where again the equality signs inside the probabilities are optional:*

$$P(X \leqq x, Y \geqq y) \leqq P(X \leqq x)\, P(Y \geqq y) \tag{2.1$'$}$$

$$P(X \geqq x, Y \leqq y) \leqq P(X \geqq x)\, P(Y \leqq y) \tag{2.1$''$}$$

$$P(X \geqq x, Y \geqq y) \geqq P(X \geqq x)\, P(Y \geqq y). \tag{2.1$'''$}$$

PROOF. (i), (ii) and (iii) are obvious.

(iv) To see that (2.1) implies the corresponding inequalities with one or both of the equal signs omitted, replace x and/or y by $x - 1/n$ and/or $y - 1/n$. To go in the other direction replace x and/or y by $x + 1/n$ and/or $y + 1/n$.

(v) The equivalence of (2.1) and (2.1$'$) follows from (iv) and the fact that $(X \leqq x, Y \geqq y) = (X \leqq x) - (X \leqq x, Y < y)$. The equivalence with (2.1$''$) follows analogously, and the equivalence with (2.1$'''$) from the fact that the sets involved in the left hand sides of the four inequalities (with some of the equality signs omitted) add up to the whole space.

An important class of examples of distributions with quadrant dependence is furnished by Theorem 1 below. Before stating this theorem, it is convenient to introduce the following definition. We shall say that two real-valued functions r and s of n arguments are *concordant* for the ith coordinate if, considered as functions of the ith coordinate (with all other coordinates held fixed), they are monotone in the same direction, i.e. either both non-decreasing or both non-increasing. Similarly r and s will be called *discordant* for the ith coordinate if they are monotone in opposite directions.

THEOREM 1. *Let* $(X_1, Y_1), \cdots, (X_n, Y_n)$ *be independent pairs of random variables with joint distributions* $F_1, \cdots, F_n$. *Let* r *and* s *be functions of* n *variables and let*

$$X = r(X_1, \cdots, X_n), \qquad Y = s(Y_1, \cdots, Y_n). \tag{2.2}$$

Then (i) $(X, Y) \varepsilon \mathfrak{F}_1$ *if for each* i *one (not necessarily the same for all* i*) of the following conditions hold:* (a) $F_i \varepsilon \mathfrak{F}_1$ *and* r, s *are concordant for the* ith *coordinate or* (b) $F_i \varepsilon \mathcal{G}_1$ *and* r, s *are discordant for the* ith *coordinate;*

(ii) *similarly* $(X, Y) \varepsilon \mathcal{G}_1$ *if for each* i *either* $F_i \varepsilon \mathfrak{F}_1$ *and* r, s *are discordant, or* $F_i \varepsilon \mathcal{G}_1$ *and* r, s *are concordant.*

(iii) *Let* U, V *be independent and independent of* $(X_1, Y_1), \cdots, (X_n, Y_n)$ *and let*

$$(2.3) \qquad X = r(U, X_1, \cdots, X_n), \qquad Y = s(V, Y_1, \cdots, Y_n).$$

Then the conclusions of (i) *and* (ii) *continue to hold under the assumptions made there without any assumptions concerning the behaviour of* r *and* s *as functions of* U *and* V *respectively.*

Proof. This will be proved as a consequence of Theorem 2 in the next section.

Example 1. The following are some pairs of random variables (X, Y) with positive quadrant dependence; the property in each case follows from Theorem 1 and the fact that $(X, X) \varepsilon \mathfrak{F}_1$ for all X.

(i) $X, Y = s(X)$ for any random variable X and any non-decreasing function s.

(ii) $X = U + aZ$, $Y = V + bZ$ for any independent random variables U, V, Z if a and b have the same sign.

(iii) $X, Y = X + V$ for any independent X, V

(iv) $X = r(U, Z)$, $Y = s(V, Z)$ where U, V, Z are independent and r and s are non-decreasing in Z but otherwise arbitrary.

Part (ii) of the example can be used to show that any bivariate normal distribution with positive correlation coefficient is in $\mathfrak{F}_1$. We shall later give different proofs of this result in Example 5 (Section 5) and Example 10 (i) (Section 8).

3. A property of quadrant dependent distributions. An important property of distributions in $\mathfrak{F}_1$ is given by Theorem 2 below, which is a generalization of an inequality of Chebyshev (cf. Hardy, Littlewood and Polya (1943), p. 43). I am grateful to Professors W. J. Hall and W. Hoeffding for providing me with a much simpler proof of the basic inequality (3.2) on which this theorem rests than my original proof. They point out that it is, in fact, an immediate consequence of the following Lemma 2, due to Hoeffding (1940), the proof of which was also communicated to me by Professor Hoeffding.

Lemma 2 (Hoeffding). *If* F *denotes the joint and* F_X *and* F_Y *the marginal distributions of* X *and* Y, *then*

$$(3.1) \quad E(XY) - E(X)E(Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [F(x, y) - F_X(x)F_Y(y)]\, dxdy$$

provided the expectations on the left hand side exist.

Proof. Let (X_1, Y_1), (X_2, Y_2) be independent, each distributed according to F. Then

$$2[E(X_1Y_1) - E(X_1)\, E(Y_1)] = E[(X_1 - X_2)(Y_1 - Y_2)]$$
$$= E \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [I(u, X_1) - I(u, X_2)][I(v, Y_1) - I(v, Y_2)]\, dudv$$

where $I(u, x) = 1$ if $u \leq x$ and $= 0$ otherwise. Since $E|XY|$, $E|X|$ and $E|Y|$ are assumed finite, we can take the expectation under the integral sign, and the resulting expression is seen to reduce to twice the right hand side of (3.1). This completes the proof.

A simple consequence is the following

LEMMA 3. *If* $(X, Y) \varepsilon \mathfrak{F}_1$ *and if the expectations in* (3.2) *exist, then*

$$E(XY) \geqq E(X)\,E(Y) \tag{3.2}$$

with equality holding if and only if X *and* Y *are independent.*

PROOF. That (3.2) holds is obvious from (3.1). Suppose now that (2.1) holds and that equality holds in (3.2). Then $F(x, y) = F_X(x)\,F_Y(y)$ except possibly on a set of Lebesgue measure zero. From the fact that cumulative distribution functions are continuous on the right, it is easily seen that if two distributions agree almost everywhere with respect to Lebesgue measure, they must agree everywhere. Thus X and Y are independent, and this completes the proof.

Lemma 3 is a special case of the following theorem.

THEOREM 2. *Throughout Theorem* 1, *the conclusion* $(X, Y) \varepsilon \mathfrak{F}_1$ *can be replaced by* (3.2), *and the conclusion* $(X, Y) \varepsilon \mathcal{G}_1$ *by* (3.2) *with the inequality sign reversed, provided the expectations in* (3.2) *exist.*

PROOF. (i)

1. We begin by proving the theorem for the case that X and Y are given by (2.2) and that $n = 1$. Suppose that r and s are both non-decreasing and that (X_1, Y_1) is in $\mathfrak{F}_1$. Then $(r(X_1), s(Y_1))$ is in $\mathfrak{F}_1$ by Lemma 1 (iii), and the result follows from Lemma 3.

If instead of non-decreasing, the functions r and s are non-increasing, the result follows if we replace them by $-r$ and $-s$ respectively. If one of them, say r, is increasing and the other decreasing and (X_1, Y_1) is in $\mathcal{G}_1$ we put $s'(y_1) = s(-y_1)$ and apply the result already proved to r and s'. This completes the proof of (i) for the case $n = 1$.

2. To prove (i) for general n we proceed by induction. From the result for $n = 1$ it follows that for all fixed $x_2, \cdots, x_n$ and $y_2, \cdots, y_n$,

$$E[r(X_1, x_2, \cdots, x_n)\,s(Y_1, y_2, \cdots, y_n)] \geqq E[r(X_1, x_2, \cdots, x_n)]\,E[s(Y_1, y_2, \cdots, y_n)].$$

Taking expectation of both sides we get

$$E[r(X_1, X_2, \cdots, X_n)\,s(Y_1, Y_2, \cdots, Y_n)) \geqq E[r^*(X_2, \cdots, X_n)\,s^*(Y_2, \cdots, Y_n)]$$

where $r^*(x_2, \cdots, x_n) = E\,r(X_1, x_2, \cdots, x_n)$ and s^* is defined correspondingly. Now r^* and s^* have the same monotonicity properties in $x_2, \cdots, x_n$ and $y_2, \cdots, y_n$ as do the functions r and s, and the result therefore follows by induction.

(ii) This is seen by putting $s'(y_1, \cdots, y_n) = s(-y_1, \cdots, -y_n)$ and applying (i) to r and s'.

(iii) This is proved by applying (i) or (ii) to $r(u, X_1, \cdots, X_n)$ and

$s(v, Y_1, \cdots, Y_n)$ for fixed u and v and then taking expectations on both sides, and using the independence of U and V.

We are now in a position to prove Theorem 1.

PROOF OF THEOREM 1. Apply Theorem 2 with (X, Y) replaced by the variables (X', Y'): $X' = I(U, X_1, \cdots, X_n)$, $Y' = J(V, Y_1, \cdots, Y_n)$, where I and J indicate the events

$$r(U, X_1, \cdots, X_n) \leqq x \quad \text{and} \quad s(V, Y_1, \cdots, Y_n) \leqq y$$

respectively. If r and s satisfy the assumptions of Theorem 1, so do I and J, and this completes the proof.

For $n = 1$, Part (i) of Theorem 2 shows that if $(X, Y) \varepsilon \mathfrak{F}_1$ and r and s are non-decreasing, then

$$E\, r(X)\, s(Y) \geqq E\, r(X)\, E\, s(Y). \tag{3.3}$$

Equality in (3.3) may hold even when X and Y are not independent, since it requires only the independence of $r(X)$ and $s(Y)$. As an example, let $X = I_1 T$, $Y = I_2 T$ where I_1, I_2, T are independent, I_1 and I_2 take on the values ± 1 with probability $\frac{1}{2}$ each, and T is any positive random variable. Then $(X, Y) \varepsilon \mathfrak{F}_1$ by Example 1 (iv), and X and Y are dependent. However, if $r(x) = s(x) = \operatorname{sgn}(x)$, it follows that $r(X) = I_1$ and $s(Y) = I_2$ are independent, so that equality holds in (3.3).

Lemma 3 states that if $(X, Y) \varepsilon \mathfrak{F}_1$ and the covariance of X and Y exists, then $\operatorname{Cov}(X, Y) \geqq 0$. Covariance is only one of a number of measures of association that have been proposed in the literature. (For a discussion of such measures cf. [14] and [16].) We shall now prove that several other such measures are non-negative for all distributions in $\mathfrak{F}_1$.

COROLLARY 1. *If F is positively quadrant dependent, then Kendall's τ, Spearman's ρ_s and the quadrant measure q discussed by Blomqvist* (1950) *are all non-negative.*

PROOF. (i) Since Kendall's τ is the covariance of $X = \operatorname{sgn}(X_2 - X_1)$ and $Y = \operatorname{sgn}(Y_2 - Y_1)$ it is by Lemma 2 enough to show that $(X, Y) \varepsilon \mathfrak{F}_1$. But this follows from Theorem 1 (i).

(ii) Since $\rho_s/3$ is the covariance of $X = \operatorname{sgn}(X_2 - X_1)$ and $Y = \operatorname{sgn}(Y_3 - Y_1)$ the result follows from Lemma 2 and Theorem 1 (i) by putting $n = 3$, $r(x_1, x_2, x_3) = \operatorname{sgn}(x_2 - x_1)$ and $s(y_1, y_2, y_3) = \operatorname{sgn}(y_3 - y_1)$.

(iii) Let μ and ν denote the medians of the marginal distributions of X and Y, and let $r(X)$ and $s(Y)$ indicate the events $X > \mu$ and $Y > \nu$ respectively. Then

$$\begin{aligned} q &= E[rs + (1 - r)(1 - s) - r(1 - s) - s(1 - r)] \\ &= E[1 + 4rs - 2r - 2s]. \end{aligned}$$

By Theorem 2(i), $E(rs) \geqq E(r)\, E(s)$ and hence

$$q \geqq [1 - 2E(r)]\,[1 - 2E(s)] \geqq 0.$$

4. Application to slippage problems. The basic idea of this section is due to Cochran (1941). It was further developed by Paulson (1952), Halperin, Greenhouse, Cornfield and Zalokar (1955), Hartley (1955), H. A. David (1956). Doornbos (1956) and by Doornbos and Prins (1956 and 1958).

Consider the testing of a hypothesis H against a number of different sets of alternatives K_i ($i = 1, \cdots, s$). Suppose H is rejected in favor of K_i if

$$T_i \geqq C_i \tag{4.1}$$

at level α_i and that the tests are similar so that

$$P_H(T_i \geqq C_i) = \alpha_i . \tag{4.2}$$

Typical examples are the so-called slippage problems in which it is assumed of s parameters $\theta_1, \cdots, \theta_s$ that they are either all equal (H) or that exactly one of them has "slipped", i.e. is different from the others, and where then K_i represents the possibility that θ_i has slipped. In such situations, rather than controlling the individual error probabilities (4.2), it is frequently of interest to control the experimentwise error rate, that is, the probability P of falsely rejecting H in favor of any of the alternatives K_i. Applying Bonferroni's inequalities to the events (4.1), we obtain for P the inequalities

$$\sum \alpha_i - \sum_{i<j} P(T_i \geqq C_i, T_j \geqq C_j) \leqq P \leqq \sum \alpha_i . \tag{4.3}$$

Suppose now that every pair (T_i, T_j) is negatively quadrant dependent, so that

$$P(T_i \geqq C_i, T_j \geqq C_j) \leqq \alpha_i \alpha_j .$$

Then it follows from (4.3) that

$$\sum \alpha_i - \sum_{i<j} \alpha_i \alpha_j \leqq P \leqq \sum \alpha_i . \tag{4.4}$$

To see how close these bounds are, note that if $\alpha = \sum \alpha_i$, then $\sum_{i<j} \alpha_i \alpha_j \leqq \frac{1}{2}(1 - s^{-1})\alpha^2$ so that

$$\alpha - \tfrac{1}{2}\alpha^2(1 - s^{-1}) \leqq P \leqq \alpha . \tag{4.5}$$

This shows α to be an excellent approximation for P whenever α is small. In the following examples, we shall consider some cases of negative quadrant dependence that have arisen in this context.

EXAMPLE 2. *The problem of m rankings.* A number of objects, say n, are ranked independently by m observers. The hypothesis H to be tested is that there is no preference among the objects, so that each observer ranks them at random. Under the alternative K_i, the ith object is typically preferred but there is no preference among the remaining objects.

Let R_{ik} be the rank assigned to the ith object by the kth observer, and let $R_i = \sum R_{ik}$ denote the rank sum for the ith object. Then H is rejected in favor of K_i if R_i is sufficiently large. In [8], it is shown by induction over m that R_i, R_j are negatively quadrant dependent, the result being obvious for $m = 1$ from the formula

$$(4.7) \qquad P(R_i \leqq x, R_j \leqq y) = [xy - \min(x, y)]/n(n-1).$$

Instead of proving the result by induction, we here only need note that $R_i = \sum R_{ik}$, $R_j = \sum R_{jk}$, the pairs (R_{ik}, R_{jk}) $k = 1, \cdots, n$ are independent, and each pair is in $\mathcal{G}_1$ by (4.7). The result now follows from Theorem 1.

EXAMPLE 3. In connection with the Poisson slippage problem, Doornbos and Prins (1958) proved that $(U_i, U_j) \varepsilon \mathcal{G}_1$ where U_i is the number of trials, in a sequence of n multinomial trials with s possible outcomes, resulting in outcome i. Their method is closely related to that of the next section. We note here that if $U_{ik} = 1$ when the kth trial results in outcome i and $U_{ik} = 0$ otherwise, a simple calculation shows that $(U_{ik}, U_{jk}) \varepsilon \mathcal{G}_1$ and since $U_i = \sum_{k=1}^n U_{ik}$, $U_j = \sum_{k=1}^n U_{jk}$, it follows from Theorem 1 that also $(U_i, U_j) \varepsilon \mathcal{G}_1$.

EXAMPLE 4. Consider testing the nonparametric hypothesis of the equality of s distributions $H: F_1 = \cdots = F_s$ against the alternatives K_i that F_i has slipped to the right, on the basis of samples X_{ik} $(k = 1, \cdots, n_i)$ from F_i. If the $N = \sum n_i$ observations are ranked, and R_{ik} denotes the rank of X_{ik} and $R_i = \sum R_{ik}$, a distribution free test rejects H in favor of K_i when R_i is sufficiently large. In this connection Doornbos and Prins [8] give a proof by Kesten of the fact that $(R_i, R_j) \varepsilon \mathcal{G}_1$. An examination of the proof shows that it carries over to the case in which an arbitrary set of numbers $v_1, \cdots, v_N$ (rather than the set of integers $1, \cdots, N$) is divided at random into s groups of $n_1, \cdots, n_s$ elements respectively, with R_i denoting the sum of the v-values in the ith group. (Kesten's proof is by induction over N. The population size is reduced by conditioning on which of the groups contains the integer N. If in the extension the largest v is not unique, the proof applies if one conditions on the position of any one of the largest v's.) Another interesting special case of the generalized result is that in which all v's are zero or one. This proves negative quadrant dependence for the components of a multiple hypergeometric distribution, which also follows from Example 10 (iii) of Section 8.

5. Regression dependence. Definition (2.1) can be rewritten as

$$(5.1) \qquad P(Y \leqq y \mid X \leqq x) \geqq P(Y \leqq y)$$

and in this form clearly expresses the fact that knowledge of X being small increases the probability of Y being small. It may be felt that the intuitive concept of positive dependence is better represented by the stronger condition.

$$(5.2) \qquad P(Y \leqq y \mid X \leqq x) \geqq P(Y \leqq y \mid X \leqq x') \qquad \text{for all} \quad x < x' \quad \text{and all} \quad y.$$

Rather than (5.2), we shall here consider the still stronger condition

$$(5.3) \qquad P(Y \leqq y \mid X = x) \quad \text{is non-increasing in } x,$$

which was discussed earlier by Tukey (1958) and Lehmann (1959). If (5.3) holds, we shall say that Y is *positively regression dependent* on X; the family of all distributions F of (X, Y) for which (5.3) holds will be denoted by $\mathfrak{F}_2$. Similarly, Y is *negatively regression dependent* on X if $P(Y \leqq y \mid X = x)$ is non-decreasing in x; the associated family of distributions F will be denoted by $\mathcal{G}_2$.

EXAMPLE 5. Let $Y = \alpha + \beta X + U$, where X and U are independent. Then Y is positively or negatively regression dependent on X as $\beta \geqq 0$ or $\leqq 0$. This is obvious since the conditional distribution of Y given $X = x$ is that of $\alpha + \beta x + U$ and hence is clearly stochastically increasing in x if $\beta > 0$. In particular, it follows from this example that the components of a bivariate normal distribution are positively or negatively regression dependent as $\rho \geqq 0$ or $\rho \leqq 0$.

Before discussing further examples, we state formally the relationship between the different definitions of positive dependence given up to this point.

LEMMA 4. *The definitions* (5.1) *to* (5.3) *are connected by the implications* (5.3) $\Rightarrow$ (5.2) $\Rightarrow$ (5.1) *and hence* $\mathfrak{F}_1 \subset \mathfrak{F}_2$.

PROOF. That (5.2) $\Rightarrow$ (5.1) is obvious. To show that (5.3) $\Rightarrow$ (5.2), let $h(u) = P(Y \leqq y \mid X = u)$, so that

$$P(Y \leqq y, X \leqq x) = \int_{-\infty}^{x} h(u)\, dF_X(u)$$

where F_X denotes the marginal distribution of X. Under the assumption that h is non-increasing, we must therefore show that

$$\int_{(-\infty, x]} h(u)\, dF_X(u)/P(X \leqq x) \geqq \int_{(-\infty, x']} h(u)\, dF_X(u)/P(X \leqq x') \qquad \text{for} \quad x < x',$$

which is obviously true.

Lemma 4 frequently provides a convenient method of proving membership in $\mathfrak{F}_1$. As a first example, let us consider once more the dependence of two components from a multinomial distribution.

EXAMPLE 3 (continued). If $(U_1, \cdots, U_s)$ have a multinomial distribution corresponding to n trials and success probabilities $(p_1, \cdots, p_s)$, the conditional distribution of U_j given $U_i = x$ is a binomial distribution with success probability $p_i/(1 - p_j)$ and corresponding to $n - x$ trials. Since $P(U_j \leqq y \mid U_i = x)$ is thus a decreasing function of x, it follows that (U_i, U_j) is in $\mathcal{G}_2$ and therefore in $\mathcal{G}_1$. (The proof of quadrant dependence for this case in [8] uses related ideas.) The case of a multiple hypergeometric distribution mentioned in Example 4 can be treated similarly.

EXAMPLE 6. An interesting example in which negative quadrant dependence was first stated by Cochran (1941) and proved by Doornbos (1956) permits a very simple treatment by the present method. The problem is that of testing H: $\sigma_1 = \cdots = \sigma_s$ in the case of s normal distributions $N(\xi_i, \sigma_i^2)$, on the basis of samples of equal size, say n. The hypothesis is rejected in favor of the alternative K_i that the ith variance has slipped to the right if

$$A_i = S_i^2 / \sum_{k=1}^{s} S_k^2 > C$$

where $S_k^2 = \sum (X_{kl} - X_{k\cdot})^2/(n - 1)$. Cochran showed that the joint density of A_i and A_j is proportional to

$$a_1^{f_1} a_2^{f_2}(1 - a_1 - a_2)^{f_3} \quad \text{for} \quad 0 \leqq a_1, a_2; \quad a_1 + a_2 \leqq 1 \tag{5.4}$$

with $f_1 = f_2 = \frac{1}{2}(n - 1) - 1$ and $f_3 = \frac{1}{2}(s - 2)(n - 1) - 1$. The conditional distribution of A_2 given $A_1 = a_1$ under (5.4) is that of $(1 - a_1)B$ where B has a beta distribution with density $C\, b^{f_2} (1 - b)^{f_3}$. Since $(1 - a_1)\, B$ is clearly stochastically decreasing as a_1 increases, this shows that $(A_1, A_2) \varepsilon\, \mathcal{G}_2$ for the general case (5.4).

While in many standard examples of positive dependence, the distributions satisfy not only (5.1) but also (5.3), the second of these conditions is much stronger than the first. The following are some examples, in which there is a strong intuitive feeling of positive or negative dependence, where there is quadrant dependence, but where the stronger condition of regression dependence is not satisfied.

Example 4 (continued). As was stated in Section 4, the variables (R_i, R_j) of Example 4 are in $\mathcal{G}_1$ for all values $v_1, \cdots, v_N$. While they are also in $\mathcal{G}_2$ for the special case that the v's are all zero or one, this is not true in general, as is seen by putting $v_1 = 3$, $v_2 = 3$, $v_3 = 5$, $v_4 = 15$; $n_1 = 2$, $n_2 = 1$; and $s = 2$. It is easily checked that in this case $P(R_2 \leqq 1 \mid R_1 \leqq x)$ is not increasing in x, and that therefore $X = R_1$, $Y = R_2$ do not even satisfy (5.2).

This example also demonstrates the asymmetry in x, y of Definitions (5.2) and (5.3). For while $P\{R_2 \leqq y \mid R_1 \leqq x\}$ and $P\{R_2 \leqq y \mid R_1 = x\}$ are not increasing functions of x for all y, $P\{R_1 \leqq x \mid R_2 \leqq y\}$ and $P\{R_1 \leqq x \mid R_2 = y\}$ are nondecreasing functions of y for all x. In fact, consider any $v_1, \cdots, v_N$ from which we draw first a sample of size n_1 with sample sum R_1 and then a sample of size $n_2 = 1$ with sample-value R_2. Then $P\{R_1 \leqq x \mid R_2 = y\}$ is an increasing function of y. This is easily seen by constructing from a random variable with the distribution of R_1 given y, a variable with the distribution of R_1 given $y' > y$, which is always at least as large as the first variable and with positive probability is actually larger.

As a definition of positive dependence one might wish to symmetrize (5.2) or (5.3) by adding to it the dual condition in which X and Y are interchanged. On the other hand, the asymmetric form is preferable as a tool for proving quadrant dependence since it is then enough to prove either (5.2) (or (5.3)) or the dual condition.

It may seem surprising that a pair of random variables, which intuitively appears to exhibit such strong negative dependence as that of the preceding example, does not belong to $\mathcal{G}_2$. However, the following example exhibits the same phenomenon.

Example 7. If U, V are independent, then $X = U + V$, $Y = U$ belong to $\mathfrak{F}_1$ by Example 1 (iii). That they do not necessarily satisfy (5.2) or (5.3) is shown by the case in which both variables take on the values 0, 2, 3 with probabilities p, q, r, $(p + q + r = 1)$. For it is then easily checked that

$$P(Y \leqq 2 \mid X \leqq 3) < P(Y \leqq 2 \mid X \leqq 4).$$

Again, the example also exhibits asymmetry. Although (X, Y) does not even satisfy (5.2), we see that (Y, X) satisfies (5.3) since $P(X \leqq x \mid Y = y) = P(V \leqq x - y)$ which decreases as y increases.

It was shown by Efron (1965) that $(U + V, U) \varepsilon \mathfrak{F}_2$ if U, V are independently distributed according to Pólya frequency functions of order 2. A simple proof of this result will be given in Example 12 (Section 8).

EXAMPLE 8 (Tukey). Let $X_1, \cdots, X_n$ be independently distributed with distribution F, let the ordered sample be denoted by $X^{(1)} < \cdots < X^{(n)}$, and suppose that F is subexponential to the right, i.e. that $[1 - F(x + a)]/[1 - F(x)]$ is an increasing function of x for each fixed $a > 0$. Then it was shown by Tukey (1958) that $X^{(s)} - X^{(r)}$ is negatively regression dependent on $X^{(r)}$ if $r < s$.

For the sake of completeness, we conclude this section with an example in which (X, Y) satisfies (5.2) but not (5.3).

EXAMPLE 9. Let the distribution of (X, Y) be given by the following 3×3 table

		y: 1	2	3
x	1	p	0	0
	2	0	q	0
	3	s	0	r

Then it is easily checked that (X, Y) satisfies (5.2) provided

$$qs \leqq pr; \tag{5.5}$$

on the other hand (X, Y) satisfies (5.3) only if $s = 0$.

6. Unbiasedness of tests of independence. Let $(X_1, Y_1), \cdots, (X_n, Y_n)$ be independently distributed with a common distribution F. Suppose that the marginal distributions of F are continuous (so that the X's and the Y's will be distinct with probability one) and consider the problem of testing the hypothesis H that X_i and Y_i are independent (without any further specification of F) against alternatives of positive dependence. Then any similar test (and hence any unbiased test) $\varphi(x_1, y_1; \cdots; x_n, y_n)$ is a permutation test. (See [18], Chap. 5). Suppose without loss of generality that in addition φ is invariant under permutations of the pairs $(x_1, y_1), \cdots, (x_n, y_n)$.

THEOREM 3. *If under the above assumptions and for any real-valued non-decreasing functions* $f_1, \cdots, f_n$ *satisfying*

$$f_1(v) \leqq f_2(v) \leqq \cdots \leqq f_n(v) \tag{6.1}$$

it is true that

$$x_1 < \cdots < x_n \quad \textit{and} \quad y_i' = f_i(y_i) \qquad \textit{for} \quad i = 1, \cdots, n \tag{6.2}$$

implies

$$\varphi(x_1, y_1; \cdots; x_n, y_n) \leqq \varphi(x_1, y_1'; \cdots; x_n, y_n') \tag{6.3}$$

then φ *is unbiased against all alternatives* F *in* $\mathfrak{F}_2$.

PROOF. If α is the size of the test, φ satisfies

$$\alpha = E_H(\varphi(X_1, Y_1; \cdots; X_n, Y_n) \mid x_1, \cdots, x_n). \tag{6.4}$$

This implies that for any fixed $x_1, \cdots, x_n$

$$\alpha = E\,\varphi(x_1, Y_1; \cdots; x_n, Y_n)$$

whenever $Y_1, \cdots, Y_n$ are identically, independently distributed. Consider now any alternative $F \varepsilon \mathfrak{F}_2$. Since then the conditional distributions of Y given x_i are stochastically increasing in i, there exist (see [18], p. 73) functions f_i satisfying (6.1) and independent, identically distributed random variables V_i such that $Y_i' = f_i(V_i)$ has the conditional distribution of Y given x_i. Since α is the conditional expectation of φ given $x_1, \cdots, x_n$ when $Y_i = V_i$ independent of X_i, it follows—if $\beta(x_1, \cdots, x_n)$ denotes the conditional expectation of φ under F—that

$$\beta(x_1, \cdots, x_n) = E\,\varphi(x_1, Y_1'; \cdots, x_n, Y_n' \mid x_1, \cdots, x_n)$$

and hence by (6.1) that

$$\alpha \leqq \beta(x_1, \cdots, x_n) \qquad \text{for all} \quad x_1, \cdots, x_n.$$

If β denotes the unconditional power of the test against φ, then

$$\beta = E_F\,\beta(X_1, \cdots X_n)$$

and hence $\alpha \leqq \beta$ as was to be proved.

COROLLARY 1. *The theorem remains valid if* φ *satisfies* (6.3) *for all pairs of points* $(x_1, y_1; \cdots; x_n, y_n)$ *and* $(x_1, y_1'; \cdots; x_n, y_n')$ *satisfying*

$$x_1 < \cdots < x_n, \qquad i < j, \qquad y_i < y_j \Rightarrow y_i' < y_j'. \tag{6.5}$$

PROOF. It is easily seen that (6.2) and (6.5) are equivalent.

COROLLARY 2. *The following one-sided tests of independence are unbiased against all alternatives belonging to* $\mathfrak{F}_2$. (*For a discussion of these tests see for example* [14].)

(i) *The difference sign covariance test based on*

$$t = \textstyle\sum_{i \neq j} \operatorname{sgn}(x_i - x_j) \operatorname{sgn}(y_i - y_j)/n(n-1)$$

or equivalently on the number of concordant pairs (x_i, y_i), (x_j, y_j) (i.e. *pairs satisfying the condition* $(x_i - x_j)(y_i - y_j) > 0$).

(ii) *The test based on the unbiased grade correlation*

$$R' = \sum \operatorname{sgn}(x_i - x_j) \operatorname{sgn}(y_i - y_k)/n(n-1)(n-2)$$

where the summation extends over all different subscripts i, j, k, *or equivalently on the number of triplets* (x_i, y_i), (x_j, y_j), (x_k, y_k) *with at least two concordant pairs.*

(iii) *The test based on the rank correlation coefficient*

$$R = [(n-2)R' + 3t]/(n+1).$$

(iv) *The quadrant test discussed by Blomqvist* (1950), *which is based on the number Q of points among* $(x_1, y_1), \cdots, (x_n, y_n)$ *lying in quadrants* I *or* III *when the origin is taken as* (m_X, m_Y) *where m_X is the median of the X's and m_Y the median of the Y's (so that Q is the number of pairs (x_i, y_i) concodrant with the pair* (m_X, m_Y)).

Proof. (i) to (iii). Since (6.5) states that

$$(6.6) \qquad (x_i, y_i), \ (x_j, y_j) \text{ concordant} \Rightarrow (x_i, y_i'), \ (x_j, y_j') \text{ concordant}$$

it is obvious that (6.5) implies $t(x, y) \leqq t(x, y')$ and $R'(x, y) \leqq R'(x, y')$ and hence also $R(x, y) \leqq R(x, y')$ where (x, y) stands for $(x_1, y_1), \cdots, (x_n, y_n)$. Thus in all three cases (6.5) implies (6.3) as was to be proved.

(iv) We shall consider only the case $n = 2m$. Let $(x_1, y_1), \cdots, (x_n, y_n)$ and $(x_1, y_1'), \cdots, (x_n, y_n')$ satisfy (6.5), let m_Y' denote the median of the y_i', and let k be the number of points among $(x_1, y_1), \cdots, (x_n, y_n)$ that are in quadrant I, so that there are also k points in III. Suppose that in moving from the unprimed to the primed y's, at least one point (x_i, y_i) moved from III to II'. Then $y_i' > m_Y'$ and hence the y' coordinates of all points originally in I are greater than m_Y' (since they are greater than y_i'). Thus all k points which originally were in I will be in I'. Suppose now that exactly $r \geqq 1$ points have risen from III to II'. By definition of m_Y' there must still be m points below m_Y'. Since none of the points from I have dropped to IV', at least r must have dropped from II to III'. Thus the number of points in III' is $\geqq$ the number of points in III, and hence $Q(x, y') \geqq Q(x, y)$.

7. Unbiasedness of tests of independence based on normal and other scores. In the present section we shall apply Theorem 3 to prove unbiasedness, against alternatives in $\mathfrak{F}_2$ with continuous marginals, of a class of rank tests for the hypothesis H of Section 6 proposed by Bhuchongkul (1964). These tests are based on statistics of the form $\sum A(r_i) B(s_i)$, where $(r_1, \cdots, r_n)$ denote the ranks of $(X_1, \cdots, X_n)$ and $(s_1, \cdots, s_n)$ the ranks of $(Y_1, \cdots, Y_n)$. Let $(t_1, \cdots, t_n)$ be the ranks of $(Y_{i_1}, \cdots, Y_{i_n})$ where $(i_1, \cdots i_n)$ is the permutation for which $X_{i_1} < \cdots < X_{i_n}$ so that the ranks of $(X_{i_1}, \cdots, X_{i_n})$ are $(1, \cdots, n)$. Then the test statistics can be written equivalently as

$$(7.1) \qquad T_n = \textstyle\sum_{i=1}^n A(i) B(t_i).$$

The purpose of the present section is to prove the following theorem.

Theorem 4. *The tests with rejection region $T_n \geqq C$, where T_n is given by* (7.1), *are unbiased against all alternatives belonging to $\mathfrak{F}_2$ provided the functions A and B are non-decreasing.*

Some important special cases are

(i) $A(i) = B(i) = i$, in which case the test reduces to the test based on the rank correlation coefficient considered in Corollary 2 (iii) of the preceding section.

(ii) $A(i) = B(i) = E(U^{(i)})$, where $U^{(1)} < \cdots < U^{(n)}$ denotes an ordered sample from a normal distribution. This leads to the normal scores test studied by Bhuchongkul.

(iii) More generally $A(i) = E[V^{(i)}]$, $B(i) = E[W^{(i)}]$, where $V^{(1)} < \cdots < V^{(n)}$ and $W^{(1)} < \cdots < W^{(n)}$ are ordered samples from any two continuous distributions J and K.

To prove Theorem 4, we shall define the following partial ordering of permutations and establish some properties of this ordering. Let

$$\{i\} = (i_1, \cdots, i_n) \quad \text{and} \quad \{j\} = (j_1, \cdots, j_n)$$

be two permutations of $(1, \cdots, n)$. We shall say that $\{i\}$ is better ordered than $\{j\}$ if for all a, b

$$a < b \quad \text{and} \quad j_a < j_b \Rightarrow i_a < i_b . \tag{7.2}$$

A more usual comparison defines $\{i\}$ to be better ordered than $\{j\}$ if it is possible to transform $\{j\}$ into $\{i\}$ by a number of steps, each of which consists in correcting an inversion. This definition neither implies nor is it implied by (7.2) as is seen by the following two examples:

(a) (132) is better ordered than (312) according to the second definition but not according to (7.2).

(b) (1324) is better ordered than (2413) according to (7.2) but not according to the second definition.

We shall here restrict use of the term "better ordered" to (7.2). It is clear that (7.2) does define a partial ordering among the $n!$ permutations of $(1, \cdots, n)$. The basic property required is contained in the following theorem.

THEOREM 5. *If* $(i_1, \cdots, i_n)$ *is better ordered than* $(j_1, \cdots, j_n)$ *and if for any* $m \leqq n$

$$i'_{1m} < \cdots < i'_{mm} \quad \text{and} \quad j'_{1m} < \cdots < j'_{mm} \tag{7.3}$$

denote the ordered m-tuples $(i_1, \cdots, i_m)$ *and* $(j_1, \cdots, j_m)$, *then*

$$i'_{km} \leqq j'_{km} \quad \text{for all} \quad 1 \leqq k \leqq m \quad \text{and all} \quad 1 \leqq m \leqq n. \tag{7.4}$$

PROOF. The proof will be by induction over m.

We note first the following. Let the rank of i_k among $(i_1, \cdots, i_m)$ be r and that of j_k among $(j_1, \cdots, j_m)$ be s. (For simplicity we suppress the subscripts km on r and s.) Then the number of i's to the right of i_m and less than i_k is $(i_k - 1) - (r - 1) = i_k - r$; the number of j's to the right of j_m and less than j_k is $j_k - s$. If $(i_1, \cdots, i_n)$ is better ordered than $(j_1, \cdots, j_n)$, it follows that

$$i_k - r \leqq j_k - s. \tag{7.5}$$

Applying (7.5) to $m = k = 1$ (so that also $r = s = 1$) it follows that $i_1 \leqq j_1$ and hence $i'_{11} \leqq j'_{11}$, so that the theorem is correct for $m = 1$.

Suppose now that the result holds for $m = 1, \cdots, a - 1$ and consider the situation for $m = a$. Let r and s denote the ranks of i_a and j_a among $(i_1, \cdots, i_a)$ and $(j_1, \cdots, j_a)$ respectively. Suppose first that $r \leqq s$. Then it is an easy consequence of the induction hypothesis that $i'_{ka} \leqq j'_{ka}$ for all $k = 1, \cdots, a$. Consider therefore the case $r > s$. Then it again follows easily from the induction hypothesis that $i'_{ka} \leqq j'_{ka}$ for $k < s$ and $k > r$. It only remains to consider the case $s \leqq k \leqq r$. Since $i_a = i'_{rm}$, it follows that i_a must exceed i'_{km} by at least $r - k$ so that $i_a \geqq i'_{km} + (r - k)$ and analogously $j_a \leqq j'_{km} + (k - s)$. Hence,

$$\begin{aligned} j'_{km} - i'_{km} &\geqq [j_a + (k - s)] - [i_a - (r - k)] \\ &= (j_a - s) - (i_a - r). \end{aligned}$$

The right hand side is $\geqq 0$ by (7.5), and this completes the proof.

COROLLARY 1. *If $\{i\}$ is better ordered than $\{j\}$ and if h is non-decreasing, then*

$$\textstyle\sum_{k=1}^{m} h(i_k) \leqq \sum_{k=1}^{m} h(j_k) \qquad \text{for all} \quad 1 \leqq m \leqq n. \tag{7.6}$$

PROOF. This is obvious since

$$\textstyle\sum_{k=1}^{m} h(i_k) = \sum_{k=1}^{m} h(i'_{km}) \leqq \sum_{k=1}^{m} h(j'_{km}) = \sum_{k=1}^{m} h(j_k).$$

(Although we do not need it here, it is in fact easily seen that (7.3) is equivalent to (7.6) holding for all non-decreasing functions h.)

COROLLARY 2. *If $\{i\}$ is better ordered than $\{j\}$, if $a_1 < \cdots < a_n$ and h is non-decreasing, then*

$$\textstyle\sum_{k=1}^{n} a_k h(i_k) \geqq \sum_{k=1}^{n} a_k h(j_k). \tag{7.7}$$

PROOF. Since the inequality is not affected by the addition of a constant to h, we may assume without loss of generality that h is positive, and then further that $\sum_{i=1}^{n} h(i) = 1$.

Let us now interpret $h(i_1), \cdots, h(i_n)$ and $h(j_1), \cdots, h(j_n)$ as probability distributions over the integers $1, \cdots, n$. Then a random variable with the first of these distributions is, by Corollary 1, stochastically larger than one with the second distribution, and the result follows from a well-known property of stochastically ordered distributions. (See, for example, [18], Chapter 3, Lemma 2 and Problem 11.)

PROOF OF THEOREM 4. Consider the application of Corollary 1 of Theorem 3 to a rank test. If $(t_1', \cdots, t_n')$ and $(t_1'', \cdots, t_n'')$ denote the ranks of $(y_1', \cdots, y_n')$ and $(y_1'', \cdots, y_n'')$, condition (6.5) states that $(t_1', \cdots, t_n')$ is better ordered than $(t_1'', \cdots, t_n'')$. Thus Theorem 4 now follows from Theorem 5 and Corollary 1 of Theorem 3.

8. Likelihood ratio dependence. Condition (5.3) which defined regression dependence requires the conditional variable Y given x to be stochastically increasing. An even stronger condition is obtained by requiring the conditional density of Y given x to have monotone likelihood ratio. (For a discussion of the relation between these two conditions, see [18], p. 74). We shall then say that (X, Y) or its distribution F shows *positive likelihood ratio dependence*. The family of all distributions F satisfying this condition will be denoted by $\mathfrak{F}_\ell$. If $f(x, y)$ is the joint density of X and Y with respect to some product measure, the condition may be written formally as

$$(8.1)\quad f(x, y')\,f(x', y) \leqq f(x, y)\,f(x', y') \qquad \text{for all} \quad x < x', \quad y < y'.$$

If the inequality is reversed, F is negatively likelihood ratio dependent and belongs to $\mathfrak{G}_3$. Like conditions (2.1) but unlike (5.2) and (5.3), condition (8.1) is symmetric in X and Y. Checking (8.1) is frequently the easiest way of proving positive dependence.

EXAMPLE 10. The following are some bivariate densities with monotone likelihood ratio dependence:

(i) a bivariate normal density is in $\mathfrak{F}_3$ or $\mathfrak{G}_3$ as $\rho \geqq 0$ or $\rho \leqq 0$;

(ii) any two components of a multinomial distribution (considered earlier in Example 3) are in $\mathfrak{G}_3$;

(iii) the Dirichlet distribution with density (5.4) is in $\mathfrak{G}_3$. In each of these cases it is easy to verify that the density satisfies the required inequality.

EXAMPLE 11. Consider the joint distribution F of any two dependent indicator variables I_0 and I_1 with, say, $p_{ij} = P(I_0 = i, I_1 = j)$, $i = 0, 1; j = 0, 1$. (Such a distribution defines the random structure of a two by two table if I_0 and I_1 indicate the occurrence of the two characteristics in question). Then it is seen that F is in $\mathfrak{F}_3$ (and hence in $\mathfrak{F}_1$ and $\mathfrak{F}_2$) if

$$(8.2)\qquad p_{00}p_{11} \geqq p_{01}\,p_{10}$$

and in $\mathfrak{G}_3$ if the inequality is reversed. As in the normal case, all distributions in this family therefore show likelihood ratio dependence.

EXAMPLE 12. Let U, V be independently distributed with densities g and h respectively, and let $X = U$, $Y = U + V$. Then we saw earlier in Example 7 that Y is always positively regression dependent on X but not necessarily X on Y. Now the joint density of X and Y is $g(x)\,h(y - x)$ and Condition (8.1) therefore reduces to

$$(8.3)\qquad h(y - x')/h(y - x) \leqq h(y' - x')/h(y' - x).$$

This condition is satisfied provided $-\log h$ is convex (see [18], p. 330). For such densities (X, Y) belongs to $\mathfrak{F}_3$ and hence not only (X, Y) but also (Y, X) belongs to $\mathfrak{F}_2$. This last result was proved earlier by Efron (1965).

EXAMPLE 13. Let $Z_1, \cdots, Z_n$ be independently distributed according to a univariate distribution G, and let X and Y be two order statistics of the Z's, say $X = Z^{(r)}$ and $Y = Z^{(s)}$ with $r < s$. Then the joint distribution F of X and Y has density

$$f(x, s) = [G(x)]^{r-1}\,[G(y) - G(x)]^{s-r-1}\,[1 - G(y)]^{n-s}, \qquad x < y$$

with respect to the product measure $G \times G$. This density satisfies (8.1) so that $F \,\varepsilon\, \mathfrak{F}_3$. That $F \,\varepsilon\, \mathfrak{F}_2$ and hence Cov $(Z^{(r)}, Z^{(s)}) \geqq 0$ was proved by Bickel (1965).

EXAMPLE 14. Let $f_\theta(x)$, $g_\theta(y)$ be two families of univariate densities, each with monotone (increasing) likelihood ratio. Then for any mixing distribution λ, the distribution with density.

$$h(x, y) = \int f_\theta(x)\, g_\theta(y)\, d\lambda(\theta)$$

belongs to $\mathfrak{F}_3$.

PROOF. The inequality (8.1) in the present case may be written as

$$\iint f_\theta(x)\, g_\theta(y)\, f_\eta(x')\, g_\eta(y')\, d\lambda(\theta)\, d\lambda(\eta)$$
$$\geqq \iint f_\theta(x)\, g_\theta(y')\, f_\eta(x')\, g_\eta(y)\, d\lambda(\theta)\, d\lambda(\eta).$$

Since the integrands on the two sides are equal for $\theta = \eta$, it is enough to integrate over the region $\theta \neq \eta$. Consider separately the contributions for $\theta < \eta$ and $\theta > \eta$. In the latter integrals interchange the variables of integration θ and η, so that all integrals extend over the region $\theta < \eta$. On combining all four integrals into a single integral on the left hand side, the integrand is seen to factor into

$$[f_\theta(x)f_\eta(x') - f_\eta(x)f_\theta(x')][g_\theta(y)g_\eta(y') - g_\eta(y)g_\theta(y')]$$

which is non-negative. This completes the proof.

REFERENCES

[1] BHUCHONGKUL, S. (1964). A class of nonparametric tests for independence in bivariate populations. *Ann. Math. Statist.* **35** 138–149.

[2] BICKEL, PETER J. (1965). Some contributions to the theory of order statistics. Submitted.

[3] BLOMQVIST, NILS (1950). On a measure of dependence between two random variables. *Ann. Math. Statist.* **21** 593–600.

[4] BLUM, J. R., KIEFER, J. AND ROSENBLATT, M. (1961). "Distribution free tests of independence based on the sample distribution," *Ann. Math. Statist.* **32** 485–498.

[5] COCHRAN, W. G. (1941). The distribution of the largest of a set of estimated variances as a fraction of their total. *Ann. Eugenics* **11** 47–52.

[6] DAVID, H. A. (1956). On the application to statistics of an elementary theorem in probability. *Biometrika* **43** 85–91.

[7] DOORNBOS, R. (1956). Significance of the smallest of a set of estimated normal variances. *Statistica Neerlandica*, **10** 17–26.

[8] DOORNBOS, R. AND PRINS, H. J. (1958). On slippage tests. *Indag. Math.* **20** 38–55 and 438–447.

[9] EFRON, BRADLEY (1965). Increasing properties of Pólya frequency functions. *Ann. Math. Statist.* **36** 272–279.

[10] HALPERIN, M., GREENHOUSE, S. W., CORNFIELD, J. AND ZALOKAR, JULIA (1955). Tables of percentage points for the studentifed maximum absolute deviate in normal samples. *J. Amer. Statist. Assoc.* **50** 185–195.

[11] HARDY, G. H., LITTLEWOOD, J. E. AND PÓLYA, G. (1934). *Inequalities*. Cambridge Univ. Press.

[12] HARTLEY, H. O. (1955). Some recent developments in analysis of variance. *Comm. Pure Appl. Math.* **8** 47–72.

[13] HÖFFDING, WASSILY (1940). Masstabinvariante Korrelations-theorie. *Schriften Math. Inst. Univ. Berlin* **5** 181–233.

[14] HÖEFFDING, WASSILY (1948). A class of statistics with asymptotically normal distributions. *Ann. Math. Statist.* **19** 293–375.

[15] HÖEFFDING, WASSILY (1948). A nonparametric test of independence. *Ann. Math. Statist.* **19** 546–557.

[16] KRUSKAL, WILLIAM H. (1958). Ordinal measures of association. *J. Amer. Statist. Assoc.* **53** 814–861.

[17] LEHMANN, E. L. (1951). Consistency and unbiasedness of certain nonparametric tests. *Ann. Math. Statist.* **22** 165–179.

[18] LEHMANN, E. L. (1959). *Testing statistical hypotheses*. Wiley, New York.

[19] PAULSON, EDWARD (1952). On the comparison of several experimental categories with a control. *Ann. Math. Statist.* **23** 239–246.

[20] PAULSON, EDWARD (1952). An optimum solution to the k-sample slippage problem for the normal distribution. *Ann. Math. Statist.* **23** 610–616.

[21] TUKEY, J. W. (1958). A problem of Berkson, and minimum variance orderly estimators. *Ann. Math. Statist.* **29** 588–592.