Inequalities in Statistics and Probability
IMS Lecture Notes–Monograph Series Vol. 5 (1984), 141–149

STOCHASTIC ORDERING OF SPACINGS FROM DEPENDENT RANDOM VARIABLES

BY MOSHE SHAKED[1] and Y. L. TONG[2]
University of Arizona and University of Nebraska

Spacings (that is, the differences between successive order statistics) are useful in various applications in statistics. Many properties of the spacing are known when the spacings are constructed from a collection of independent identically distributed (i.i.d.) random variables. In this paper we study the spacings constructed from not necessarily i.i.d. random variables. We introduce models for which two sets of spacings, constructed from two sets of dependent random variables, can be stochastically ordered. Various examples will be given and applications for goodness-of-fit tests, tests for independence, density estimation and tests for outliers will be discussed.

1. Introduction. Let $\mathbf{X} = (X_1, \ldots, X_n)$ denote an n-dimensional random vector and let

$$X_{(1)} \leqslant \ldots \leqslant X_{(n)}$$

be the ordered components (order statistics) of $\mathbf{X}$. The nonnegative random variables

$$U_i = X_{(i+1)} - X_{(i)},\ i = 1, \ldots, n-1$$

are called the *spacings* and have various applications in statistics. For example, certain nonparametric test procedures depend on the maximum spacing or on linear combinations of spacings (see, e.g., Pyke (1965), Weiss (1965), Rao and Sethuraman (1970) and Kirmani and Alam (1974)); certain estimation and test procedures based on order statistics, such as those which depend on the range or midrange, involve spacings (David (1970), Ch. 6); and certain tests for slippage (Karlin and Truax (1960)) and outliers (Barnett and Lewis (1978), Ch. 3) also depend on spacings. For a comprehensive treatment of spacings see Pyke (1965, 1972).

In the literature the problem of spacings has been treated extensively under the assumption that $X_1, \ldots, X_n$ are independent and identically distributed (i.i.d.) random variables. In certain applications which involve a mixture of experiments, a (random) change of scale or a random shift in location may take place; then the random variables $X_1, \ldots, X_n$ are no longer independent. In this paper we study how the degree of dependence affects the distribution of the spacings. In the case when $X_1, \ldots, X_n$ are interchangeable, it follows from our main result that (in the model under consideration) the spacings vector $\mathbf{U} = (U_1, \ldots, U_{n-1})$ becomes stochastically smaller if $X_1, \ldots, X_n$ are more positively dependent (that is, when $X_1, \ldots, X_n$ have more tendency to "hang together").

After stating the model and proving the main result in Section 2, we apply the result to an additive, a multiplicative and a ratio model. In Section 3, after combining results given in Shaked and Tong (1985), we obtain a partial ordering property for the spacings which correspond to a number of important multivariate distributions, such as the multivariate normal, multivariate stable, multivariate beta and the Dirichlet distribution. For all these distributions the corresponding spacings vector $\mathbf{U}$ can be partially ordered through the degree of dependence of the components $X_1, \ldots, X_n$ of $\mathbf{X}$.

---

[1] Research partially supported by the NSF Grant MCS-8200098.
[2] Research partially supported by the NSF Grant MCS-8100775.
AMS 1980 subject classifications. Primary 60E15; secondary 62G30.
Key words and phrases. Stochastic ordering of spacings, dependent random variables.

In Section 4 we give applications and study the monotonicity properties of certain well-known procedures concerning goodness-of-fit tests, tests for independence, density estimation and slippage tests for outliers, which all depend on spacings.

2. The Model. Let $\mathbf{X} = (X_1, \ldots, X_n)$ and $\mathbf{Y} = (Y_1, \ldots, Y_n)$ denote two n-dimensional random vectors and let

$$X_{(1)} \leq \ldots \leq X_{(n)}, Y_{(1)} \leq \ldots \leq Y_{(n)}$$

be their ordered components. Define the $(n-1)$-dimensional spacings vectors by

$$\mathbf{U} = (U_1, \ldots, U_{n-1}) \text{ where } U_i = X_{(i+1)} - X_{(i)}, i = 1, \ldots, n-1.$$
$$\mathbf{V} = (V_1, \ldots, V_{n-1}) \text{ where } V_i = Y_{(i+1)} - Y_{(i)}, i = 1, \ldots, n-1.$$

The stochastic ordering of $\mathbf{U}$ and $\mathbf{V}$ will be developed under the following model:

Model A. There exist a random vector $\mathbf{Z} = (Z_1, \ldots, Z_n)$, random vectors (of any dimension) $\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3, \mathbf{W}_4$ and Borel-measurable functions ϕ, δ_1 and δ_2 such that

$$(X_1, \ldots, X_n) \stackrel{d}{=} (Z_1\phi(\mathbf{Z}, \mathbf{W}_1) + \delta_1(\mathbf{Z}, \mathbf{W}_3), \ldots, Z_n\phi(\mathbf{Z}, \mathbf{W}_1) + \delta_1(\mathbf{Z}, \mathbf{W}_3))$$
$$(Y_1, \ldots, Y_n) \stackrel{d}{=} (Z_1\phi(\mathbf{Z}, \mathbf{W}_2) + \delta_2(\mathbf{Z}, \mathbf{W}_4), \ldots, Z_n\phi(\mathbf{Z}, \mathbf{W}_2) + \delta_2(\mathbf{Z}, \mathbf{W}_4)).$$

Moreover, the following conditions are satisfied:

A1. $Z_1, \ldots, Z_n$ are i.i.d., $\mathbf{W}_i$ is independent of $\mathbf{Z}$, $i = 1,2$.

A2. $\phi(\mathbf{z},\mathbf{w})$, $\delta_1(\mathbf{z},\mathbf{w})$ and $\delta_2(\mathbf{z},\mathbf{w})$ are permutation symmetric functions of $\mathbf{z} = (z_1, \ldots, z_n)$ for every fixed $\mathbf{w}$, and $\phi > 0$ for all $(\mathbf{z},\mathbf{w})$ whenever $\mathbf{z}$ is in the support of $\mathbf{Z}$ and $\mathbf{w}$ is in the support of $\mathbf{W}_1$ or of $\mathbf{W}_2$.

A3. Either $\phi(\mathbf{z},\mathbf{w})$ is nondecreasing (componentwise) in $\mathbf{w}$ for every $\mathbf{z}$ and $\mathbf{W}_1 \stackrel{st}{\geq} \mathbf{W}_2$, or $\phi(\mathbf{z},\mathbf{w})$ is nonincreasing in $\mathbf{w}$ for every $\mathbf{z}$ and $\mathbf{W}_1 \stackrel{st}{\leq} \mathbf{W}_2$.

THEOREM 1. *Assume that* $\mathbf{X}$ *and* $\mathbf{Y}$ *have the representation of Model* A *and that* A1, A2 *and* A3 *are satisfied. Then, for all k and all constants* c_{ij}, $i = 1, \ldots, n$, $j = 1, \ldots, k$, *such that* $\sum_{i=1}^n c_{ij} = 0, j = 1, \ldots, k$,

$$E\psi(|\Sigma_{i=1}^n c_{i1}X_{(i)}|, \ldots, |\Sigma_{i=1}^n c_{ik}X_{(i)}|)$$
$$\geq E\psi(|\Sigma_{i=1}^n c_{i1}Y_{(i)}|, \ldots, |\Sigma_{i=1}^n c_{ik}Y_{(i)}|)$$

holds for all ψ *which are componentwise nondecreasing such that the expectations exist. Consequently,*

$$\mathbf{U} \stackrel{st}{\geq} \mathbf{V}. \tag{2.1}$$

Proof. Let $Z_{(1)} \leq \ldots \leq Z_{(n)}$ denote the order statistics of $\mathbf{Z} = (Z_1, \ldots, Z_n)$. Since ϕ, δ_1, and δ_2 are permutation symmetric in $z_1, \ldots, z_n$ for every fixed $\mathbf{w}$, we must have, a.s.,

$$\phi(\mathbf{Z},\mathbf{W}_j) = \phi(Z_{(1)}, \ldots, Z_{(n)}, \mathbf{W}_j)$$
$$\delta_j(\mathbf{Z},\mathbf{W}_{2+j}) = \delta_j(Z_{(1)}, \ldots, Z_{(n)}, \mathbf{W}_{2+j})$$

for $j = 1,2$. This implies that

$$Z_i\phi(\mathbf{Z},\mathbf{W}_j) + \delta_j(\mathbf{Z},\mathbf{W}_{j+2}) \leq Z_{i'}\phi(\mathbf{Z},\mathbf{W}_j) + \delta_j(\mathbf{Z},\mathbf{W}_{j+2})$$

holds if and only if $Z_i \leq Z_{i'}$. Consequently,

$$\begin{aligned}&(X_{(1)}, \ldots, X_{(n)})\\ &\stackrel{d}{=} (Z_{(1)}\phi(\mathbf{Z},\mathbf{W}_1) + \delta_1(\mathbf{Z},\mathbf{W}_3), \ldots, Z_{(n)}\phi(\mathbf{Z},\mathbf{W}_1) + \delta_1(\mathbf{Z},\mathbf{W}_3))\\ &= \phi(\mathbf{Z},\mathbf{W}_1)(Z_{(1)}, \ldots, Z_{(n)}) + (\delta_1(\mathbf{Z},\mathbf{W}_3), \ldots, \delta_1(\mathbf{Z},\mathbf{W}_3)).\end{aligned}$$

Hence, for all c_{ij}, $i = 1, \ldots, n, j = 1, \ldots, k$, such that $\Sigma_{i=1}^n c_{ij} = 0$,

$$(\Sigma_{i=1}^n c_{i1}X_{(i)} \ldots, \Sigma_{i=1}^n c_{ik}X_{(i)}) \stackrel{d}{=} \phi(\mathbf{Z},\mathbf{W}_1)(\Sigma_{i=1}^n c_{i1}Z_{(i)}, \ldots, \Sigma_{i=1}^n c_{ik}Z_{(i)}).$$

Without loss of generality assume that ϕ is nondecreasing in $\mathbf{w}$ and that $\mathbf{W}_1 \stackrel{st}{\geqslant} \mathbf{W}_2$. Then

$$\begin{aligned} E\psi(|\Sigma_{i=1}^n c_{i1}X_{(i)}|, \ldots, |\Sigma_{i=1}^n c_{ik}X_{(i)}|) &= E\psi(\phi(\mathbf{Z},\mathbf{W}_1)|\Sigma_{i=1}^n c_{i1}Z_{(i)}|, \ldots, \phi(\mathbf{Z},\mathbf{W}_1)|\Sigma_{i=1}^n c_{ik}Z_{(i)}|) \\ &= E\zeta_1(\mathbf{Z}) \end{aligned}$$

where $\zeta_1(\mathbf{z})$ is the conditional expectation of ψ, over the distribution of $\mathbf{W}_1$, given $\mathbf{Z} = \mathbf{z}$. Let $\zeta_2(\mathbf{z})$ denote the similar conditional expectation of ψ over the distribution of $\mathbf{W}_2$. Since ϕ is a nondecreasing function of $\mathbf{w}$ it follows that $\psi(\phi(\mathbf{z},\mathbf{w})|\Sigma_{i=1}^n c_{i1}z_{(i)}|, \ldots, \phi(\mathbf{z},\mathbf{w})|\Sigma_{i=1}^n c_{ik}z_{(i)}|)$ is also a nondecreasing function of $\mathbf{w}$ for every fixed $\mathbf{z}$. Thus, $\mathbf{W}_1 \stackrel{st}{\geqslant} \mathbf{W}_2$ implies that

$$\zeta_1(\mathbf{Z}) \geqslant \zeta_2(\mathbf{Z}) \text{ a.s.}$$

The proof is now completed by applying the equality

$$E\psi(|\Sigma_{i=1}^n c_{i1}Y_{(i)}|, \ldots, |\Sigma_{i=1}^n c_{ik}Y_{(i)}|) = E\zeta_2(\mathbf{Z}).$$

The last statement of the theorem follows by defining $k = n-1$, $c_{i+1,i} = 1$, $c_{i,i} = -1$, $c_{i,j} = 0, j \neq i, i+1$ where $i = 1, \ldots, n-1$.

In the following we consider special forms of Model A.

(a) (An additive model). Assume that there exist constants $a \geq 0$, b and d (b and d have the same sign) and independent random variables $Z_1, \ldots, Z_n$ and W such that

$$(2.2) \qquad \begin{aligned} (X_1, \ldots, X_n) &\stackrel{d}{=} (aZ_1 + bW, \ldots, aZ_n + bW) \\ (Y_1, \ldots, Y_n) &\stackrel{d}{=} (cZ_1 + dW, \ldots, cZ_n + dW). \end{aligned}$$

Note that without loss of generality we can assume that $b \geqslant 0$ and $d \geqslant 0$ because otherwise one can replace W by $-W$.

Letting W_1 and W_2 be degenerate at $a > 0$ and $c > 0$, respectively, letting $W_3 \stackrel{d}{=} W_4 \stackrel{d}{=} W$, setting $\phi(\mathbf{z},w) = w$, $\delta_1(\mathbf{z},w) = bw$ and $\delta_2(\mathbf{z},w) = dw$ and assuming $a \geq c$, it is easy to see that $\mathbf{X}$ and $\mathbf{Y}$ of (2.2) have the representation of Model A and satisfy A1, A2 and A3 provided $Z_1, \ldots, Z_n$ are i.i.d.

In some applications (see Shaked and Tong (1985)) $\mathbf{X}$ and $\mathbf{Y}$ have the same marginals, that is,

$$(2.3) \qquad X_i \stackrel{d}{=} Y_i, \; i = 1, \ldots, n.$$

In other applications the following condition, which is weaker than (2.3), holds:

$$(2.4) \qquad \Sigma_{i=1}^n EX_i = \Sigma_{i=1}^n EY_i$$

For example, (2.4) holds if $EZ_1 = \ldots = EZ_n = EW = 0$ or if $EZ_1 = \ldots = EZ_n = EW$ and $a + b = c + d$.

Let $U_i = X_{(i+1)} - X_{(i)}$ and $V_i = Y_{(i+1)} - Y_{(i)}$, $i = 1, \ldots, n-1$. From Theorem 1 it follows that if $\mathbf{X}$ and $\mathbf{Y}$ satisfy (2.2) with $a \geqslant c$ then

$$(2.5) \qquad (U_1, \ldots, U_{n-1}) \stackrel{st}{\geqslant} (V_1, \ldots, V_{n-1}).$$

Shaked and Tong (1985) considered the condition

$$(2.6) \qquad (EX_{(1)}, \ldots, EX_{(n)}) \succ (EY_{(1)}, \ldots, EY_{(n)}).$$

that is,

$$\Sigma_{i=1}^k EX_{(i)} \leqslant \Sigma_{i=1}^k EY_{(i)}, k = 1, \ldots, n-1$$

and

$$\Sigma_{i=1}^{n} EX_{(i)} = \Sigma_{i=1}^{n} EY_{(i)}.$$

They denoted the relation (2.6) by $\mathbf{X} \underset{D}{\succ} \mathbf{Y}$ and discussed some applications. Following their arguments it follows that if $\mathbf{X}$ and $\mathbf{Y}$ satisfy (2.5) and (2.4) then $\mathbf{X} \underset{D}{\succ} \mathbf{Y}$. Thus

PROPOSITION 1. *If* $\mathbf{X}$ *and* $\mathbf{Y}$ *satisfy* (2.2) *with* $a \geq c$ *and* (2.4), *then* $\mathbf{X} \underset{D}{\succ} \mathbf{Y}$.

The special case where $\mathbf{X}$ and $\mathbf{Y}$ have the representation

$$(2.7) \qquad \begin{aligned} (X_1, \dots, X_n) &\stackrel{d}{=} (1-\rho)^{1/\alpha}(Z_1, \dots, Z_n) + \rho^{1/\alpha}(W, \dots, W), \\ (Y_1, \dots, Y_n) &\stackrel{d}{=} (1-\eta)^{1/\alpha}(Z_1, \dots, Z_n) + \eta^{1/\alpha}(W, \dots, W), \end{aligned}$$

where $Z_1, \dots, Z_n$ and W are as in (2.2), $0 \leqslant \rho < \eta \leqslant 1$ and $\alpha > 0$, is Model 4.1 in Shaked and Tong (1985).

Note that in this special case we can actually have a stronger statement concerning the distribution of $\mathbf{U}$ and $\mathbf{V}$ corresponding to $\mathbf{X}$ and $\mathbf{Y}$. That is,

$$(2.8) \qquad \begin{aligned} (U_1, \dots, U_{n-1}) &\stackrel{d}{=} (1-\rho)^{1/\alpha}(Z_{(2)}-Z_{(1)}, \dots, Z_{(n)}-Z_{(n-1)}) \\ (V_1, \dots, V_{n-1}) &\stackrel{d}{=} (1-\eta)^{1/\alpha}(Z_{(2)}-Z_{(1)}, \dots, Z_{(n)}-Z_{(n-1)}). \end{aligned}$$

Thus we have

$$(2.9) \qquad (U_1, \dots, U_{n-1}) \stackrel{d}{=} \left(\frac{1-\rho}{1-\eta}\right)^{1/\alpha}(V_1, \dots, V_{n-1}).$$

Consequently

$$(2.10) \qquad \Sigma_{i=1}^{n-1}\lambda_i U_i \stackrel{d}{=} \left(\frac{1-\rho}{1-\eta}\right)^{1/\alpha}\Sigma_{i=1}^{n-1}\lambda_i V_i$$

for all $\lambda_1, \dots, \lambda_{n-1}$.

(b) (A multiplicative model). Assume $\mathbf{X}$ and $\mathbf{Y}$ have the representation

$$\begin{aligned} (X_1, \dots, X_n) &\stackrel{d}{=} W_1(Z_1, \dots, Z_n) + (W_3, \dots, W_3), \\ (Y_1, \dots, Y_n) &\stackrel{d}{=} W_2(Z_1, \dots, Z_n) + (W_4, \dots, W_4), \end{aligned}$$

where W_1 and $\mathbf{Z}$ are independent, W_2 and $\mathbf{Z}$ are independent and $Z_1, \dots, Z_n$ are i.i.d. If W_1 and W_2 are nonnegative a.s. and $W_1 \stackrel{st}{\geqslant} W_2$ then this is a special case of Model A and Theorem 1 applies.

(c) (A ratio model). In certain situations $X_1, \dots, X_n$ have the representation

$$(X_1, \dots, X_n) \stackrel{d}{=} (Z_1 / (\Sigma_{i=1}^{n} h(Z_i) + W_1), \dots, Z_n / (\Sigma_{i=1}^{n} h(Z_i) + W_1)),$$

where $Z_1, \dots, Z_n$ and W_1 are independent and $Z_1, \dots, Z_n$ are i.i.d. [When $Z_i \geqslant 0$, $W_1 > 0$ a.s. and h is the identity function, then

$$(X_1, \dots, X_n) \stackrel{d}{=} (Z_1/(\Sigma_{i=1}^{n} Z_i + W_1), \dots, Z_n/(\Sigma_{i=1}^{n} Z_i + W_1)).]$$

In this model if

$$(Y_1, \dots, Y_n) \stackrel{d}{=} (Z_1 / (\Sigma_{i=1}^{n} h(Z_i) + W_2), \dots, Z_n / (\Sigma_{i=1}^{n} h(Z_i) + W_2)),$$

$h \geqslant 0$ and $W_2 > 0$ a.s., then $W_2 \stackrel{st}{\geqslant} W_1$ implies $\mathbf{U} \stackrel{st}{\geqslant} \mathbf{V}$. This follows from Theorem 1 because in Model A one can take

$$\begin{aligned} \phi(\mathbf{z}, w) &= [\Sigma_{i=1}^{n} h(z_i) + w]^{-1} \\ \delta_1(\mathbf{z}, w) &= \delta_2(\mathbf{z}, w) = 0. \end{aligned}$$

We note in passing that it is easy to show that $\mathbf{U} \stackrel{st}{\geqslant} \mathbf{V}$ also for spacings vectors constructed from $\mathbf{X}$ and $\mathbf{Y}$ which satisfy the following model:

Model B. There exist a random vector $\mathbf{Z} = (Z_1, \ldots, Z_n)$, a random vector $\mathbf{W}$ and Borel-measurable functions ϕ_1, ϕ_2, δ_1 and δ_2 such that

$$(X_1, \ldots, X_n) \overset{d}{=} (Z_1\phi_1(\mathbf{Z},\mathbf{W}) + \delta_1(\mathbf{Z},\mathbf{W}), \ldots, Z_n\phi_1(\mathbf{Z},\mathbf{W}) + \delta_1(\mathbf{Z},\mathbf{W})),$$
$$(Y_1, \ldots, Y_n) \overset{d}{=} (Z_1\phi_2(\mathbf{Z},\mathbf{W}) + \delta_2(\mathbf{Z},\mathbf{W}), \ldots, Z_n\phi_2(\mathbf{Z},\mathbf{W}) + \delta_2(\mathbf{Z},\mathbf{W})).$$

Moreover, the following conditions are satisfied:

B1. $Z_1, \ldots, Z_n$ are i.i.d. and $\mathbf{W}$ and $\mathbf{Z}$ are independent.

B2. $\phi_1(\mathbf{z},\mathbf{w})$, $\phi_2(\mathbf{z},\mathbf{w})$, $\delta_1(\mathbf{z},\mathbf{w})$, and $\delta_2(\mathbf{z},\mathbf{w})$ are permutations symmetric functions of $z_1, \ldots, z_n$ for every fixed $\mathbf{w}$, and $\phi_1 \geqslant \phi_2 > 0$ over the support of $(\mathbf{Z},\mathbf{W})$.

The main difference between Models A and B is that in Model B we have two functions ϕ_1 and ϕ_2 compared to the single function ϕ of Model A. But, in Model B, ϕ_1 and ϕ_2 are not required to be monotone.

We end this section by showing that Model 4.2 (unlike Model 4.1) of Shaked and Tong (1985), which involves positive dependence by mixture, does not necessarily imply the basic relation (2.1).

A random vector $\mathbf{Y}$ is called positively dependent by mixture (PDM) if there exists a random vector $\mathbf{W}$ such that, given $\mathbf{W} = \mathbf{w}$, $Y_1, \ldots, Y_n$ are conditionally i.i.d. Shaked (1977) and Shaked and Tong (1985) showed that, in some respects, a PDM random vector $\mathbf{Y}$ is more positively dependent than a random vector $\mathbf{X}$ of i.i.d. random variables where $\mathbf{X}$ and $\mathbf{Y}$ have the same marginals. One can expect that the spacings vectors $\mathbf{U}$ and $\mathbf{V}$, corresponding to $\mathbf{X}$ and $\mathbf{Y}$ satisfy $\mathbf{U} \overset{st}{\geqslant} \mathbf{V}$. The following example shows that this is not necessarily the case.

Example. Let Y_1 and Y_2 have the joint probabilities

$y_1 \backslash y_2$	1	2	3
1	1/6	0	1/6
2	0	1/3	0
3	1/6	0	1/6

and let X_1 and X_2 be i.i.d. such that $X_i \overset{d}{=} Y_i$, $i = 1,2$, that is $P[X_1 = 1] = P[X_1 = 2] = P[X_1 = 3] = 1/3$. Then $P[V_1 \geqslant 2] = P[Y_{(2)} - Y_{(1)} = 2] = 1/3$ whereas $P[U_1 \geqslant 2] = P[X_{(2)} - X_{(1)} = 2] = 2/9$. Hence it is not true that $U_1 \overset{st}{\geqslant} V_1$ although (Y_1, Y_2) is PDM.

It follows that if $\mathbf{X}$ and $\mathbf{Y}$ satisfy Model 4.2 of Shaked and Tong (1985) then it is not necessarily true that the corresponding spacings satisfy $\mathbf{U} \overset{st}{\geqq} \mathbf{V}$.

3. Examples. In this section we describe some examples of well-known distributions for which the results of Section 2 apply.

(a) *Exchangeable normal variables*. Let $\mathbf{X}$ be a multivariate normal random vector with means μ_x, variances σ^2 and correlations ρ; let $\mathbf{Y}$ be another multivariate normal random vector with means μ_y, variances σ^2 and correlations η. If $0 \leqslant \rho < \eta \leqslant 1$ then $\mathbf{U} \overset{st}{\geqslant} \mathbf{V}$. This follows from (2.7) with $\alpha = 2$ where $Z_1, \ldots, Z_n$ and W are i.i.d. normal random variables with mean 0 and variance σ^2. Note that adding μ_x to all X_i's and μ_y to all Y_i's does not change the distributions of $\mathbf{U}$ and $\mathbf{V}$.

(b) ***Multivariate Cauchy and stable variables***. It is shown in Shaked and Tong (1985) that some exchangeable stable random vectors **X** and **Y** have the representation (2.7). Hence Theorem 1 applies.

(c) *Multivariate Dirichlet and beta variables*. Let **X** have the distribution defined by

$$(X_1, \ldots, X_n) = (Z_1 /(\Sigma_{i=1}^n Z_i + W), \ldots, Z_n /(\Sigma_{i=1}^n Z_i + W))$$

where $Z_1, \ldots, Z_n$ are i.i.d. gamma random variables (for Dirichlet) or i.i.d. chi-squared random variables (for multivariate beta), W is a gamma random variable or a chi-squared random variable (with the same shape parameter but possibly with different scale parameter) and **Z** and W are independent. In this case, as is shown in Section 2, a partial ordering of the spacings can be obtained via the value of the shape parameter of W.

Note that in this case $X_1, \ldots, X_n$ are not positively dependent. Actually they are negatively correlated. The result of Theorem 1 can be interpreted here by saying that the less negatively dependent are $X_1, \ldots, X_n$ the smaller stochastically are the corresponding spacings.

4. Applications

4.1. Goodness of fit tests. Let $Z_1, \ldots, Z_n$ be random variables, let $Z_{(1)}, \ldots, Z_{(n)}$ be the corresponding order statistics and let $U_i = Z_{(i+1)} - Z_{(i)}$, $i = 1, \ldots, n-1$, be the corresponding spacings. Statistics like the largest spacing, the kth smallest spacing, partial sums of ordered spacings, etc., have been used in statistical literature to construct tests of goodness of fit and related hypotheses (see, e.g., Rao and Sobel (1980) and references there). Pyke (1965) discussed statistics of the form $\Sigma_{i=1}^{n-1} g(U_i)$ where g is some monotone function. A general form for all the statistics mentioned above is $\Sigma_{i=1}^{n-1} g_i(U_i)$ where the g_i's are all monotone in the same direction (see Weiss (1957)).

In most applications the Z_i's are i.i.d. with a common distribution F, say, and one is interested to test $H_0 : F = F_0$ where F_0 is a given distribution (which may or may not depend on some unknown parameters). The hypothesis H_0 is then rejected if $\Sigma_{i=1}^{n-1} g_i(U_i)$ is large; tables of critical values have been prepared for various choices of the g_i's.

In some practical situations it may happen that the assumption of independence of the observations is not valid. For example, a random shift of all the observations combined with a change of scale may transform the Z_i's into dependent random variables with the same (or with different) marginals as the original Z_i's (we denote these dependent random variables then by Y_i's). For example, assume that the Z_i's are i.i.d. normal random variables with mean μ and variance σ^2. Define

$$Y_i = (1-p)^{1/2} Z_i + p^{1/2} Z, \; i = 1, \ldots, n.$$

where Z is a random shift, independent of $Z_1, \ldots, Z_n$ and having a normal distribution with mean λ and variance σ^2. Then each Y_i is a normal random variable with mean $(1-p)^{1/2} \mu + \rho^{1/2} \lambda$ and variance σ^2, but now the Y_i's are not independent. In most applications $\mu = \lambda = 0$ (the condition $\lambda = 0$ means that, on the average, the random shift is zero), so that the Y_i's have the same common marginal distribution as the original Z_i's, but they are not independent. It follows from Theorem 1, then, that if one uses the test statistic $\Sigma_{i=1}^{n-1} g_i(U_i)$, where the g_i's are nondecreasing, then one has a smaller probability of rejection of H_0 than intended. Actually, Theorem 1 shows that the more dependent the Y_i's are, the smaller is the probability of rejection of H_0. The opposite is true if the g_i's are nonincreasing.

Of course the above analysis is valid whenever the Y_i's are distributed according to any multivariate stable distribution (see Section 3).

Relation (2.10) is particularly useful in this setting. To see this let $\rho = 0$ in (2.7) and

$\eta \in (0,1)$. Then $(X_1, \ldots, X_n) = (Z_1, \ldots, Z_n)$, i.e. $X_1, \ldots, X_n$ are i.i.d. whereas $Y_1, \ldots, Y_n$ are not independent. The corresponding spacings $\mathbf{U}$ and $\mathbf{V}$, constructed from $\mathbf{X}$ and $\mathbf{Y}$, respectively, satisfy (2.8) and hence (2.9) and (2.10). If η is known and if the test statistic is $\Sigma_{i=1}^{n-1} \lambda_i U_i$ then the critical values for testing H_0 (mentioned above) can be obtained from existing tables by multiplying the tabulated values by $(1-\eta)^{1/\alpha}$ [recall that here we take $\rho = 0$].

We remark in passing that comments which are similar to the above apply to any statistic which is a monotone function of the spacings. In particular, they apply to the statistics discussed in del Pino (1979) which are monotone functions of the k-spacings: $Z_{(k+1)}-Z_{(1)}, \ldots, Z_{(n)}-Z_{(n-k)}$.

4.2. *Tests for independence*. The discussion in Section 4.1 shows that if observations are dependent in the sense Model A instead of being independent then the significance levels of many tests, which use these observations, are different than the desired ones. In particular, if the g_i's are nondecreasing then the probability of rejection decreases as the observations become more positively dependent.

One way to interpret this discussion is to observe that the statistics $\Sigma_{i=1}^{n-1} g_i(U_i)$ actually yield unbiased tests for the hypothesis which claims that the random variables are independent versus the alternative which claims, for example, that they are positive dependent in the sense (2.2) with $a \geq c$ and $d \geq b = 0$.

We note in passing that the resulting tests are not necessarily optimal in any sense. We do not try to derive here any optimality property for any test. We remark, however, that one advantage of the above tests is that existing tables of critical values of statistics of the form $\Sigma_{i=1}^{n-1} g_i(U_i)$ and existing results about their asymptotic distributions (see Pyke (1965) and references there) may be applied for testing the hypothesis of independence mentioned above.

4.3. *Empirical distributions and quantile function estimates*. Let $Z_1, \ldots, Z_n$ be identically distributed random variables with a common distribution F. Let $Z_{(1)} \leq \ldots \leq Z_{(n)}$ be the corresponding order statistics. The empirical distribution function, $\hat{F}$, is (denoting $Z_{(n+1)} = \infty$)

$$\begin{aligned} \hat{F}(z) &= 0 \quad \text{if } z < Z_{(1)} \\ &= i/n \quad \text{if } z \in [Z_{(i)}, Z_{(i+1)}),\ i = 1, \ldots, n, \end{aligned}$$

that is, $\hat{F}$ is constant on intervals whose length are the spacings associated with the Z_i's.

In most applications the Z_i's are independent and then $\hat{F}$ is an estimate of F whose properties are well known. However, in some applications the Z_i's are dependent in the sense of Model A [then we denote them by $Y_1, \ldots, Y_n$] although marginally $Y_i \stackrel{st}{=} Z_i$. In that case $\hat{F}$ is not necessarily an unbiased estimator of F. By Theorem 1, the more positively dependent the Y_i's are, the shorter (stochastically) the corresponding spacings are and thus, the shorter (stochastically) the range of the support of $\hat{F}$ is. Geometrically, if $\mathbf{X}$ and $\mathbf{Y}$ satisfy (2.1) then the graph of the $\hat{F}$ based on the Y_i's will be (stochastically) steeper than the graph of the $\hat{F}$ based on the X_i's.

Thus, various statistics which are functions of $\hat{F}$ can be compared stochastically. This is the case if these statistics are nondecreasing functions of the underlying spacings. For example, various measures of dispersion (such as the range, the interquartile range, the sample variance, etc.) computed from $\hat{F}$ based on the X_i's are stochastically larger than the same computed from $\hat{F}$ based on the Y_i's.

The inverse of F,

$$Q(u) = \inf\{x : F(x) \geq u\}$$

and its density q (if it exists), are called, respectively, the *quantile* function and the *quantile-density* function. Parzen (1979) discusses various estimators $\hat{Q}$ and $\hat{q}$ of Q and q. Graphically, one of the estimators, $\hat{Q}$, is obtained by "inverting" $\hat{F}$ (that is, flipping the graph of $\hat{F}$ around the main diagonal of the bivariate plane). Parzen (1979) also suggests various "sensible" estimates of q which are obtained by by differentiating "smooth" versions of $\hat{Q}$. For example, one estimator of q is given by

$$\hat{q}(u) = n(Z_{(i)}-Z_{(i-1)}) \text{ for } u \in (\frac{i-1}{n}, \frac{i}{n}),\ i = 2, \ldots, n.$$

The comments about the influence of positive dependence on $\hat{Q}$ and $\hat{q}$ are similar to the ones made above about $\hat{F}$. Various monotone functionals of $\hat{Q}$ are discussed in Parzen (1979). Thus, one can stochastically compare various statistics based on a $\hat{Q}$ which was constructed from X_i's to similar statistics based on a $\hat{Q}$ which was constructed from Y_i's where $\mathbf{X}$ and $\mathbf{Y}$ satisfy (2.1).

4.4. Tests for outliers. For $i = 1, \ldots, n$, let Z_i be a normal random variable with mean μ_i and variance σ^2. Consider the null hypothesis $H_0: \mu_1 = \ldots = \mu_n$ and the following possible alternatives which state that one or k of the Z_i's are outliers:

A : $\mu_1 = \ldots = \mu_{i-1} = \mu_{i+1} = \ldots = \mu_n < \mu_i$ for some $i \in 1, \ldots, n$ (one of the Z_i's is an outlier caused by a slippage to the right),

A$'$: $\mu_1 = \ldots = \mu_{i-1} = \mu_{i+1} = \ldots = \mu_n > \mu_i$ for some $i \in 1, \ldots, n$ (one of the Z_i's slipped to the left),

A$''$: $\mu_1 = \ldots = \mu_{i-1} = \mu_{i+1} = \ldots = \mu_n \neq \mu_i$ for some $i \in 1, \ldots, n$ (one of the Z_i's is an outlier),

B_k:$n-k$ μ_i's are equal to an unknown μ and the other μ_i's are larger than μ

(there are k outliers caused by slippages to the right). Similarly B'_k and B''_k can be defined.

Various tests have been proposed for testing these and similar alternatives when the Z_i's are assumed to be independent (Barnett and Lewis (1978, pp. 89–115)). For example, if σ^2 is known then one can test A [respectively, B_k] by rejecting H_0 if

$$T_A \equiv \sigma^{-1}\phi_1(\mathbf{Z}) \equiv \sigma^{-1}(Z_{(n)} - \bar{Z}) > c \text{ for some } c$$

[respectively,

$$T_{B_k} \equiv \sigma^{-1}\phi_k(\mathbf{Z}) \equiv \sigma^{-1}(Z_{(n)} + \ldots + Z_{(n-k+1)} - k\bar{Z}) > c \text{ for some } c];$$

here $\bar{Z} = n^{-1}(Z_1 + \ldots + Z_n)$. Similarly, A' [respectively B'_k] can be tested by rejecting H_0 when

$$T_{A'} \equiv \sigma^{-1}\phi'_1(\mathbf{Z}) \equiv \sigma^{-1}(\bar{Z} - Z_{(1)}) > c \text{ for some } c$$

[respectively,

$$T_{B'_k} \equiv \sigma^{-1}\phi'_k(\mathbf{Z}) \equiv \sigma^{-1}(k\bar{Z} - Z_{(1)} - \ldots - Z_{(k)}) > c \text{ for some } c].$$

The two-sided alternative A'' may be tested by rejecting H_0 when

$$T_{A''} \equiv \sigma^{-1}\psi(\mathbf{Z}) \equiv \sigma^{-1}\max(Z_{(n)} - \bar{Z}, \bar{Z} - Z_{(1)}) > c \text{ for some } c$$

and B''_k may be tested by rejecting H_0 when

$$T_{B''_k} \equiv \sigma^{-2}\tilde{\psi}(\mathbf{Z}) \equiv \sigma^{-2}\Sigma_{i=1}^{n}(Z_i - \bar{Z})^2 > c \text{ for some } c$$

(see Dixon (1950, p. 490)) or by rejecting H_0 when

$$\tilde{T}_{B''_k} \equiv \sigma^{-1}\phi(\mathbf{X}) \equiv \sigma^{-1}(Z_{(n)}-Z_{(1)}) > c \text{ for some } c.$$

If σ^2 is unknown, but an independent estimate S_υ^2 of σ^2 is available, then one can test the various alternatives by replacing σ by S_υ in the above statistics (see details in Barnett and Lewis (1978, pp. 89–115)).

Note that all the above test statistics are nondecreasing functions of the spacings (for example $Z_{(n)}-\bar{Z} = n^{-1}\sum_{i=1}^{n-1}(Z_{(n)}-Z_{(i)}) = n^{-1}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n-1}U_j$). It follows from Theorem 1 that if the observations are not independent but instead that a random shift and a rescaling in the sense Model A have been applied to the Z_i's [denote them then by Y_i's] leaving the marginals unchanged, then the significance level of each of the above tests may be smaller than the desired one.

Of course, the same analysis applies also to Z_i's which have distributions other than normal (see Section 3 for examples).

REFERENCES

BARNETT, V. and LEWIS, T. (1978). *Outliers in Statistical Data,* Wiley, New York.

DAVID, H. A. (1970). *Order Statistics*. Wiley, New York.

DIXON, W. J. (1950). Analysis of extreme values, *Ann. Math. Statist. 21* 488–506.

KARLIN, S. and TRUAX, D. R. (1960). Slippage problems. *Ann. Math. Statist. 31* 296–324.

KIRMANI, S. N. U. A. and ALAM, S. N. (1974). On goodness of fit tests based on spacings. *Sankhya Ser. A 36* 197–203.

PARZEN, E. (1979). Nonparametric statistical data modelling. *J. Amer. Statist. Assoc. 74* 105–121.

DEL PINO, G. E. (1979). On the symptotic distribution of k-spacings with applications to goodness-of-fit tests. *Ann. Statist. 7* 1058–1065.

PYKE, R. (1965). Spacings. *J. Roy. Statist. Soc. Ser. B 27* 395–449.

PYKE, R. (1972). Spacings revisited. *Sixth Berkeley Symp. on Math. Statist. and Prob., Vol. I* (eds. L. M. Le Cam, J. Neyman and E. L. Scott) University of California Press, Berkeley, 417–427.

RAO, J. S. and SETHURAMAN, J. (1970). Pitman efficiencies of tests based on spacings. *In Nonparametrics Techniques in Statistical Inference* (ed. M. L. Puri) Cambridge University Press, Cambridge, 267–273.

RAO, J. S. and SOBEL, M. (1980). Incomplete Dirichlet integrals with applications to ordered uniform spacings. *J. Multivariate Anal. 10* 603–610.

SHAKED, M. (1977). A concept of positive dependence for exchangeable random variables. *Ann. Statist. 5* 505–515.

SHAKED, M. and TONG, Y. L. (1985). Some partial orderings of exchangeable random variables by positive dependence. To appear in *J. Multivariate Anal.*

WEISS, L. (1957). Tests of fit in the presence of nuisance location and scale parameters. *Ann. Math. Statist. 28* 1016–1020.

WEISS, L. (1965). On asymptotic sampling theory for distributions approaching the uniform distribution. *Z. wahrscheinlichkeitstheorie und Gebiate 4* 217–221.