The Annals of Statistics
1977, Vol. 5, No. 2, 274–277

AN ORDERING THEOREM FOR CONDITIONALLY INDEPENDENT AND IDENTICALLY DISTRIBUTED RANDOM VARIABLES

BY Y. L. TONG

University of Nebraska

Let **a** and **b** be r-dimensional real vectors. It is shown that if **a** majorizes **b**, then $E(\prod_{j=1}^r X_j^{a_j}) \geqq E(\prod_{j=1}^r X_j^{b_j})$ holds for nonnegative random variables $X_1, \cdots, X_r$ whose joint pdf is permutation symmetric. If in addition the components of **a**, **b** are nonnegative integers, then for every Borel-measurable set A,

$$\prod_{j=1}^r P[\bigcap_{i=1}^{a_j} \{Z_i \in A\}] \geqq \prod_{j=1}^r P[\bigcap_{i=1}^{b_j} \{Z_i \in A\}]$$

holds for conditionally i.i.d. random variables Z_i. Applications are considered.

1. A moment inequality. For fixed positive integer $r \geqq 2$ and a real number k let

$$\mathbf{a} = (a_1, \cdots, a_r)', \qquad \mathbf{b} = (b_1, \cdots, b_r)' \tag{1.1}$$

be two real vectors such that $\sum_{j=1}^r a_j = \sum_{j=1}^r b_j = k$. It is understood that "$\mathbf{a} \succ \mathbf{b}$" means "**a** majorizes **b**" or equivalently, "**b** is majorized by **a**" (for definition, see [2], page 45). Let $X_1, \cdots, X_r$ have a joint pdf $f = f(x_1, \cdots, x_r)$, and let us define $\alpha(\mathbf{a}) = E(\prod_{j=1}^r X_j^{a_j})$. All moments under consideration are assumed to be finite.

THEOREM 1. *Assume that* $P[\bigcap_{j=1}^r \{X_j \geqq 0\}] = 1$ *and f is permutation symmetric. If* $\mathbf{a} \succ \mathbf{b}$, *then* $\alpha(\mathbf{a}) \geqq \alpha(\mathbf{b})$.

PROOF. By Muirhead's theorem (see Corollary 2 of [6]), for every ω in the sample space the inequality

$$\sum_L (\prod_{j=1}^r X_{l_j}^{a_j}(\omega)) \geqq \sum_L (\prod_{j=1}^r X_{l_j}^{b_j}(\omega)) \tag{1.2}$$

holds, where the summations are taken over all $L = (l_1, \cdots, l_r)'$, the permutations of $\{1, \cdots, r\}$. The conclusion follows by taking expections on both sides of (1.2) and by the symmetric property of f.

REMARKS. (1) It should be noted that in the above theorem a_j and b_j need not be nonnegative. (2) The conclusion of the theorem can be written as

$$\prod_{j=1}^r \mu_{a_j} \geqq \prod_{j=1}^r \mu_{b_j} \tag{1.3}$$

in the special case where μ_m is the mth moment of a nonnegative random variable X ($\mu_0 = 1$). Particular cases of this inequality include the well-known inequalities $\mu_k \geqq \mu_1^k$, $\mu_k \geqq \mu_{k-m}\mu_m$ for $m \in (0, k)$. It is also immediate that $\mu_{k-m}\mu_m$ is nonincreasing in $m \in [0, k/2]$.

Received October 1974; revised April 1976.

AMS 1970 *subject classifications.* Primary 26A86; Secondary 62H99.

Key words and phrases. Conditionally i.i.d. random variables, majorization, moment inequalities, probability inequalities for multivariate distributions, multiple decision problems.

2. An ordering theorem. For fixed k, let $Z_1, \cdots, Z_k$ be q-dimensional random vectors ($q \geqq 1$) with a joint distribution of the form

$$G = G(z_1, \cdots, z_k) = \int (\prod_{i=1}^{k} F(z_i))\, d\lambda(F)\,, \tag{2.1}$$

where F is a cdf and λ is a probability measure. (G is a mixture of distributions of i.i.d. random variables.) In case $q = 1$, Dykstra, Hewett and Thompson [1] call $Z_1, \cdots, Z_k$ "conditionally i.i.d." and Shaked [8] calls $Z_1, \cdots, Z_k$ "positively dependent by mixture." Note that if an infinite sequence of random variables $Z_1, Z_2, \cdots$ are exchangeable (according to the definition given in [4], page 364), then any finite subset $\{Z_{l_1}, \cdots, Z_{l_k}\}$ of $Z_1, Z_2, \cdots$ are conditionally i.i.d.; this follows because "the concept of exchangeability is equivalent to that of conditional independence with common cdf" ([4], page 365).

Now for given $\mathbf{a}$ and a Borel-measurable set A in R^q consider

$$\beta_1(\mathbf{a}) = \prod_{j=1}^{r} \int [(\int_A dF)^{a_j}]\, d\lambda(F)\,. \tag{2.2}$$

In the special case when the a_j's are nonnegative integers, the r.h.s. of (2.2) reduces to $\prod_{j=1}^{r} P[\bigcap_{i=1}^{a_j} \{Z_i \in A\}] = \beta_2(\mathbf{a})$ (say).

THEOREM 2. *If* $\mathbf{a} \succ \mathbf{b}$, *then* $\beta_1(\mathbf{a}) \geqq \beta_1(\mathbf{b})$ *holds for every Borel-measurable set* A. *In particular, if* $\mathbf{a} \succ \mathbf{b}$ *and* $Z_1, \cdots, Z_k$ *are conditionally* i.i.d., *then* $\beta_2(\mathbf{a}) \geqq \beta_2(\mathbf{b})$ *holds for every Borel-measurable set* A.

PROOF. The proof follows immediately from (1.3) with

$$\mu_m = \int [(\int_A dF)^m]\, d\lambda(F)\,.$$

REMARK. Special cases of Theorem 2 include the known inequalities $P[\bigcap_{i=1}^{k} \{Z_i \in A\}] \geqq (P[Z_1 \in A])^k$, which was given in [1]; and

$$P[\bigcap_{i=1}^{k} \{Z_i \in A\}] \geqq P[\bigcap_{i=1}^{k-m} \{Z_i \in A\}] \cdot P[\bigcap_{i=1}^{m} \{Z_i \in A\}] \quad \text{for } m < k\,, \tag{2.3}$$

which follows from an inequality of Kimball [3]. Also, it follows that the r.h.s. of (2.3) is nonincreasing in m for $m \leqq k/2$.

We now consider a special form of Theorem 2 which we find useful in applications. Let $X_1, \cdots, X_k$ be i.i.d. s_1-dimensional random variables, $Y_1, \cdots, Y_r$ be i.i.d. s_2-dimensional random variables and the X_i's and the Y_j's are independent. Let $g\colon R^{(s_1+s_2)} \to R^q$ be a Borel-measurable real-valued function. For fixed $\mathbf{a}$ with integer components, let $h_0 = 0$ and

$$h_j = \sum_{j'=1}^{j} a_{j'} \qquad \text{for } j = 1, \cdots, r\,.$$

We define, for $i = 1, \cdots, k$,

$$V_i = g(X_i, Y_j) \qquad \text{for } h_{j-1} + 1 \leqq i \leqq h_j\,. \tag{2.4}$$

Here the random variables $V_1, \cdots, V_k$ are obtained in such a way that the first a_1 of them depend on different X_i's and on the same Y_1, the next a_2 of them depend on different X_i's and on the same Y_2, etc. For a given Borel-measurable set A in R^q let us denote $\gamma(\mathbf{a}) = P[\bigcap_{i=1}^{k} \{V_i \in A\}] = \prod_{j=1}^{r} P[\bigcap_{i=h_{j-1}+1}^{h_j} \{V_i \in A\}]$.

If follows from Theorem 2 that if $\mathbf{a} \succ \mathbf{b}$, then $\gamma(\mathbf{a}) \geqq \gamma(\mathbf{b})$. This result covers a recent theorem of Šidák [9] as a special case.

3. Applications. In this section we give several examples of the applications of the results obtained in Section 2.

(a) *An application to multiple decision problems.* In a multiple comparison problem one wishes to compare k experimental populations with the same control simultaneously. In a ranking and selection problem one wishes to select the population associated with the largest parameter. It can be shown that in such problems the probability of correct decision under the slippage configuration can be expressed in the form given in (2.1). (Perhaps the most familiar form is $\int_{-\infty}^{\infty} \Phi^{k-1}(cz + d)\, d\Phi(z)$, where $c > 0$, $d > 0$ and Φ is the standard normal cdf.) Hence results obtained in Section 2 applies to such problems.

(b) *Applications to probability inequalities of multivariate distributions.* For $s_1 = s_2 = q$ let the g function in (2.4) be $g(x, y) = x + y$, and let $\mathbf{V} = (V_1, \cdots, V_k)'$. If $\mathbf{a} \succ \mathbf{b}$, then for every $c < d$ the inequality

$$P_{\mathbf{a}}[\bigcap_{i=1}^{k} \{c < V_i < d\}] \geqq P_{\mathbf{b}}[\bigcap_{i=1}^{k} \{c < V_i < d\}] \tag{3.1}$$

holds. This particular g function has applications to several special multivariate distributions ((i)—(iii)):

(i) *Multivariate normal.* If the X_i's are i.i.d. normal variables and the Y_j's are i.i.d. normal variables, then $\mathbf{V}$ has a multivariate normal distribution and (3.1) holds. This implies the following result: let $\mathbf{V} = (V_1, \cdots, V_k)'$ and $\mathbf{W} = (W_1, \cdots, W_k)'$ be two multivariate normal variables with a common mean vector and covariance matrices Σ_V and Σ_W, respectively, where $\Sigma_V = (R_{jj'})$, $\Sigma_W = (S_{jj'})$, $R_{jj'} = S_{jj'} = \mathbf{0}$ for $j \neq j'$, R_{jj} is $(a_j \times a_j)$, S_{jj} is $(b_j \times b_j)$, and the elements of R_{jj} and S_{jj} are 1 on the diagonal and ρ ($\rho \geqq 0$) otherwise. If $\mathbf{a} \succ \mathbf{b}$, then for every $c < d$ the inequality

$$P[\bigcap_{i=1}^{k} \{c < V_i < d\}] \geqq P[\bigcap_{i=1}^{k} \{c < W_i < d\}] \tag{3.2}$$

holds.

(ii) *Multivariate Poisson.* If the X_i's and the Y_j's are independent Poisson variables with parameters λ_x and λ_y, respectively, then $\mathbf{V}$ has a multivariate Poisson distribution and (3.1) holds.

(iii) *Multivariate Gamma.* If the X_i's and the Y_j's are independent Gamma variables with a common scale parameter, then $\mathbf{V}$ has a multivariate Gamma distribution and (3.1) holds.

(iv) *Multivariate t.* Let $\mathbf{V}$ have a multivariate normal distribution with mean vector $\mathbf{0}$, variances one and correlation matrix Σ; let S be independent of $\mathbf{V}$ and νS^2 have a χ^2-distribution with df ν. Then $\mathbf{t} = (t_1, \cdots, t_k)'$ has a multivariate t distribution with parameters ν and Σ, where $t_i = V_i/S$ $(i = 1, \cdots, k)$. Consider two multivariate t variables $\mathbf{t}$ and $\mathbf{u}$ with parameters ν, Σ_Z and ν, Σ_W, respectively. If Σ_V and Σ_W have the structures described in (i) and if $\mathbf{a} \succ \mathbf{b}$, then (3.2) holds with V_i and W_i replaced by t_i and u_i, respectively.

(v) *Multivariate exponential.* Let the X_i's and Y_j's be independent negative exponential variables with parameters λ_x and λ_y, respectively. If the g function is $g(x, y) = \min(x, y)$, then $\mathbf{V}$ has a multivariate exponential distribution and (3.1) holds.

4. Acknowledgments. I wish to thank Professor A. W. Marshall and a referee for their helpful comments and suggestions.

REFERENCES

[1] DYKSTRA, R. L., HEWETT, J. E. and THOMPSON, W. A., JR. (1973). Events which are almost independent. *Ann. Statist.* **1** 674–681.

[2] HARDY, G. H., LITTLEWOOD, J. E. and PÓLYA, G. (1959). *Inequalities* (2nd ed.). Cambridge Univ. Press.

[3] KIMBALL, A. W. (1951). On dependent tests of significance in the analysis of variance. *Ann. Math. Statist.* **22** 600–602.

[4] LOÈVE, M. (1963). *Probability Theory* (3rd ed.). Van Nostrand, Princeton.

[5] MARSHALL, A. W. and OLKIN, I. (1974). Majorization in multivariate distributions. *Ann. Statist.* **2** 1189–1200.

[6] MARSHALL, A. W. and PROSCHAN, F. (1965). An inequality for convex functions involving majorization. *J. Math. Anal. Appl.* **12** 87–90.

[7] PROSCHAN, F. and SETHURAMAN, J. (1977). Schur functions in statistics. I. The preservation theorem. *Ann. Statist.* **5** 256–262.

[8] SHAKED, M. (1975). A concept of positive dependence for exchangeable random variables. Technical report, Dept. of Statistics, Univ. of Rochester.

[9] ŠIDÁK, Z. (1973). A chain of inequalities for some types of multivariate distributions, with nine special cases. *Apl. Mat.* **18** 110–118.

[10] TONG, Y. L. (1970). Some probability inequalities of multivariate normal and multivariate *t*. *J. Amer. Statist. Assoc.* **65** 1243–1247.

DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF NEBRASKA
LINCOLN, NEBRASKA 68588