The Annals of Statistics
1989, Vol. 17, No. 1, 429–435

INEQUALITIES FOR A CLASS OF POSITIVELY DEPENDENT RANDOM VARIABLES WITH A COMMON MARGINAL[1]

By Y. L. Tong

Georgia Institute of Technology

This paper concerns a partial ordering of positive dependence of a class of random variables which have a common marginal distribution and are not necessarily exchangeable. The main theorem is obtained by applying a moment inequality via majorization. Inequalities for exchangeable random variables, for random variables whose marginal densities possess the semigroup property and for the multivariate normal distribution are then obtained as special cases.

1. Introduction and motivation. The general study of inequalities for positively dependent random variables has played a central role in the development of inequalities in statistics and probability and has yielded numerous useful results [see, e.g., Tong (1980), Chapters 2 and 5]. When the random variables have a common marginal distribution, an important case of interest has concerned exchangeable random variables. In particular, Rinott and Pollak (1980) proved an inequality for bivariate exchangeable variables and recently Shaked and Tong (1985) studied inequalities via the partial orderings of the "strength" of the positive dependence of exchangeable random variables. To illustrate how such partial orderings yield inequalities, simply consider exchangeable normal variables $Y_1, \dots, Y_n$ $(Z_1, \dots, Z_n)$ with means μ, variances σ^2 and correlations ρ_2 (ρ_1). If $\rho_2 > \rho_1 \geq 0$, then the Y_i's tend to hang together more. Consequently, one has $E\prod_{i=1}^{n}\phi(Y_i) \geq E\prod_{i=1}^{n}\phi(Z_i)$ for all Borel-measurable functions $\phi: \mathscr{R} \to [0, \infty)$ such that the expectations exist [Shaked and Tong (1985)]. For $n = 2$, one has $\operatorname{corr}(\phi(Y_1), \phi(Y_2)) \geq \operatorname{corr}(\phi(Z_1), \phi(Z_2))$ for all ϕ as previously shown by Rinott and Pollak (1980).

In this paper we consider the case when the random variables are not necessarily exchangeable but have a common marginal distribution. Our main result, given in Theorem 2.1, depends on a representation for such a class of random variables and an application of a majorization inequality. [For a complete treatment of theory of majorization see Marshall and Olkin (1979)]. The main theorem is then applied to yield moment and probability inequalities for exchangeable random variables and for random variables whose marginal densities possess the semigroup property and to the multivariate normal distribution as special cases.

2. The main results. For fixed $n \geq 2$ let $\mathbf{Y} = (Y_1, \dots, Y_n)$ and $\mathbf{Z} = (Z_1, \dots, Z_n)$ denote two random vectors. Extending the definition in Rinott and

Received February 1987; revised April 1988.
[1]Supported in part by NSF Grant DMS-85-02346.
AMS 1980 *subject classifications*. Primary 60E15; secondary 62H05.
Key words and phrases. Moment inequalities, probability inequalities, positive dependence, exchangeable random variables, mixture of distributions, de Finetti's theorem.

Pollak (1980) we have

DEFINITION 2.1. $\mathbf{Y}$ is said to be more positively dependent than $\mathbf{Z}$ (more precisely, the components of $\mathbf{Y}$ are said to be more positively dependent than the components of $\mathbf{Z}$), in symbols $\mathbf{Y} \geq_{pd} \mathbf{Z}$, if $Y_1, \ldots, Y_n, Z_1, \ldots, Z_n$ have a common marginal distribution and if

$$E \prod_{i=1}^{n} \phi(Y_i) \geq E \prod_{i=1}^{n} \phi(Z_i) \tag{2.1}$$

holds for all Borel-measurable functions $\phi: \mathscr{R} \to \mathscr{R}$ such that the expectations exist.

In certain applications ϕ may be assumed to be nonnegative. Thus Definition 2.1 can be modified as:

DEFINITION 2.2. $\mathbf{Y}$ is said to be more positively dependent than $\mathbf{Z}$ through nonnegative transformations ($\mathbf{Y} \geq_{pd+} \mathbf{Z}$) if $Y_1, \ldots, Y_n, Z_1, \ldots, Z_n$ have a common marginal distribution and if (2.1) holds for all Borel-measurable nonnegative functions ϕ such that the expectations exist.

Note that if $\mathbf{Y} \geq_{pd+} \mathbf{Z}$ then, by letting ϕ be the indicator function of a subset, the inequality

$$P[Y_1 \in B, \ldots, Y_n \in B] \geq P[Z_1 \in B, \ldots, Z_n \in B] \tag{2.2}$$

holds for all Borel-measurable subsets of the real line. Thus probability inequalities immediately follow once after the partial ordering of positive dependence is established.

To obtain sufficient conditions for such a partial ordering we consider a sequence of i.i.d. random variables $\{U_i\}_{i=1}^{n}$, another independent sequence of i.i.d. random variables $\{V_i\}_{i=1}^{n}$ and an independent random variable W as "building blocks." Then for a given Borel-measurable function $g: \mathscr{R}^3 \to \mathscr{R}$ and a fixed n-dimensional vector of nonnegative integers

$$\begin{gathered} \mathbf{k} = (k_1, \ldots, k_r, 0, \ldots, 0), \\ 1 \leq r \leq n,\ k_j \geq 1 \text{ for } j \leq r \quad \text{and} \quad \sum_{j=1}^{r} k_j = n, \end{gathered} \tag{2.3}$$

we define an n-dimensional random vector $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ given by

$$\begin{gathered} \xi_1 = g(U_1, V_1, W), \ldots, \xi_{k_1} = g(U_{k_1}, V_1, W), \\ \xi_{k_1+1} = g(U_{k_1+1}, V_2, W), \ldots, \xi_{k_1+k_2} = g(U_{k_1+k_2}, V_2, W), \ldots, \\ \xi_{k_1+\cdots+k_{r-1}+1} = g(U_{k_1+\cdots+k_{r-1}+1}, V_r, W), \ldots, \xi_n = g(U_n, V_r, W). \end{gathered} \tag{2.4}$$

That is, each of the ξ_i's depends on the common variable W and on a different variable U_i. Furthermore, the first k_1 of them depend on the common variable

V_1, the next k_2 of them depend on the common variable V_2 and so on. The vector $(\xi_1, \dots, \xi_n)$ will be denoted by $\boldsymbol{\xi}(\mathbf{k})$.

It is obvious that $\xi_1, \dots, \xi_n$ have a common marginal distribution. Furthermore, the vector $\mathbf{k}$ plays an important role in the positive dependence of $\boldsymbol{\xi}(\mathbf{k})$. To see two extreme cases, if one chooses (i) W to be a singular random variable and (ii) $\mathbf{k} = (1, 1, \dots, 1)$, then clearly $\xi_1, \dots, \xi_n$ are i.i.d. random variables; on the other hand if one chooses (i) $P[U_i = u] = 1$, $i = 1, \dots, n$, and (ii) $\mathbf{k} = (n, 0, \dots, 0)$, then $P[\xi_1 = \cdots = \xi_n] = 1$ so $\operatorname{corr}(\xi_i, \xi_{i'}) = 1$ for all $i \neq i'$. Thus, for given random variables $\{U_i\}$, $\{V_i\}$ and W, the strength of the positive dependence of the components of $\boldsymbol{\xi}(\mathbf{k})$ can be determined via the diversity of the components of $\mathbf{k}$. In the following we state such a result via the notion of majorization.

THEOREM 2.1. *For fixed $n \geq 2$ assume that* (i) $\{U_i\}_1^n$, $\{V_i\}_1^n$ *and W are stochastically independent, $U_1, \dots, U_n$ are i.i.d. and $V_1, \dots, V_n$ are i.i.d.,* (ii) *$g: \mathscr{R}^3 \to \mathscr{R}$ is any Borel-measurable function and* (iii) *$\mathbf{k}$ and $\mathbf{k}'$ are two real vectors of the form given in* (2.3). *Let $\boldsymbol{\xi}(\mathbf{k})$ and $\boldsymbol{\xi}(\mathbf{k}')$ be the random vectors as defined in* (2.4). *If $\mathbf{k} \succ \mathbf{k}'$ (that is, if $\mathbf{k}$ majorizes $\mathbf{k}'$), then $\boldsymbol{\xi}(\mathbf{k}) \geq_{pd+} \boldsymbol{\xi}(\mathbf{k}')$.*

PROOF. For notational convenience let $k_i > 0$ ($k_i' > 0$) for $i \leq r$ (for $i \leq r'$) for some r (r') and equal to 0 otherwise. Then for every given $\phi \geq 0$ such that the expectations exist one can write

$$\begin{aligned} E\prod_{i=1}^{n}\phi(\xi_i) &= E\prod_{j=1}^{r}E\left[E\left\{\prod_{i=1}^{k_j}\phi\big(g(U_i, V_j, W)\big)\big|(V_j, W)\right\}\Big|W\right] \\ &= E\prod_{j=1}^{r}E\left[\psi^{k_j}(V_j, W)|W\right], \end{aligned} \tag{2.5}$$

where

$$\psi(v_j, w) = E\left\{\phi\big(g(U_1, V_j, W)\big)\big|(V_j, W)\right\} \tag{2.6}$$

denotes the conditional expectation. Now for every given $W = w$ the random variables $\psi(V_1, w), \dots, \psi(V_n, w)$ are i.i.d. and are greater than or equal to 0 a.s. By defining μ_{k_j} to be the k_jth moment of $\psi(V_j, w)$ and applying the inequality in (1.3) of Tong (1977), it follows that $\prod_{j=1}^{r}E\psi^{k_j}(V_j, w) \geq \prod_{j=1}^{r'}E\psi^{k_j'}(V_j, w)$ holds true for every fixed w. Thus for $(\xi_1', \dots, \xi_n') = \boldsymbol{\xi}(\mathbf{k}')$ we have

$$E\prod_{i=1}^{n}\phi(\xi_i) = E\prod_{j=1}^{r}E\left[\psi^{k_j}(V_j, W)|W\right] \geq E\prod_{j=1}^{r'}E\left[\psi^{k_j'}(V_j, W)|W\right] = E\prod_{i=1}^{n}\phi(\xi_i').$$

□

In the following corollary we show that if the elements in $\mathbf{k}$ and $\mathbf{k}'$ are even integers (including 0), then the condition that $\phi \geq 0$ can be dropped.

COROLLARY 2.1. *Let $\{U_i\}_1^n$, $\{V_i\}_1^n$, W and g satisfy the conditions stated in Theorem* 2.1. *Let $\mathbf{k}, \mathbf{k}'$ be two real vectors such that their components are nonnegative even integers. If $\mathbf{k} \succ \mathbf{k}'$, then $\boldsymbol{\xi}(\mathbf{k}) \geq_{pd} \boldsymbol{\xi}(\mathbf{k}')$.*

PROOF. The proof follows as Theorem 2.1 since $\frac{1}{2}\mathbf{k} \succ \frac{1}{2}\mathbf{k}'$. □

In certain applications the special case $\mathbf{k}' = (1, \dots, 1)$ is of great interest. In the following corollary we show that if the vector $\mathbf{k}$ contains only even integers, then again the condition that $\phi \geq 0$ can be removed.

COROLLARY 2.2. *Let $\{U_i\}_1^n$, $\{V_i\}_1^n$, W and g satisfy the conditions stated in Theorem* 2.1 *and let $\mathbf{k}, \mathbf{k}'$ be two n-dimensional real vectors such that* $\mathbf{k}' = (1, \dots, 1)$. *If the components of $\mathbf{k}$ are nonnegative even integers such that $\sum_{i=1}^n k_i = n$, then $\xi(\mathbf{k}) \geq_{pd} \xi(\mathbf{k}')$.*

PROOF. For every fixed $W = w$ the function ψ defined in (2.6) satisfies [again by (1.3) of Tong (1977)]

$$\prod_{j=1}^{r} E\psi^{k_j}(V_j, w) \geq \left[E\psi^2(V_1, w)\right]^{n/2} \geq \left[E\psi(V_1, w)\right]^n.$$

The proof then follows by unconditioning. □

As a special consequence, we observe that:

COROLLARY 2.3. *Let $\{U_i\}_1^n$, $\{V_i\}_1^n$, W and g satisfy the conditions stated in Theorem* 2.1 *and let $\xi(\mathbf{k})$ be the random vector defined in* (2.4). *If* $\mathbf{k} = (n, 0, \dots, 0)$, $\mathbf{k}' = (1, \dots, 1)$, *and if n is a positive even integer, then $\xi(\mathbf{k}) \geq_{pd} \xi(\mathbf{k}')$.*

Note that if n is not an even integer, then the statement in Corollary 2.3 no longer holds true. A counterexample is easy to construct and is omitted.

In certain applications to be discussed in Section 3 we restrict our attention to a family of random variables such that $\xi(\mathbf{k})$ is obtained by choosing $\mathbf{k} = (s, 1, \dots, 1, 0, \dots, 0)$ in (2.4). For notational convenience we shall denote the random vector $\xi(\mathbf{k})$ (with such a $\mathbf{k}$ vector) by $\xi(s)$. Corollary 2.4 shows how the positive dependence of the components of $\xi(s)$ depends on s. Its proof follows immediately from Theorem 2.1 and is omitted.

COROLLARY 2.4. *Let $\{U_i\}_1^n$, $\{V_i\}_1^n$, W and g satisfy the conditions stated in Theorem* 2.1. *For given $s \geq 1$ let $\xi(s) = (\xi_1, \dots, \xi_n)$ be the random vector obtained according to* (2.4) *by choosing*

$$(2.7) \quad k_1 = s, \qquad k_2 = \cdots = k_{n-s+1} = 1, \qquad k_{n-s+2} = \cdots = k_n = 0.$$

Then (a) $\xi(s+1) \geq_{pd+} \xi(s)$ *holds for all n and all $s < n$ and* (b) $\xi(s+2) \geq_{pd} \xi(s)$ *holds for all nonnegative even integers $s \leq n - 2$ and positive even integers n.*

3. Applications to special families of random variables and distributions. In this section we apply the main results in section 2 to yield inequalities via partial ordering of positive dependence for several families of random variables and distributions. The applications will be given for illustrative purposes only and obviously are not exhaustive.

3.1. *Exchangeable random variables.* An infinite sequence of random variables $\{X_i\}_{i=1}^{\infty}$ is said to be exchangeable if $(X_{\pi_1}, \ldots, X_{\pi_n})$ and $(X_1, \ldots, X_n)$ are identically distributed for every subset $(\pi_1, \ldots, \pi_n)$ of $\{1, 2, \ldots\}$ and every finite n [Loève (1963), page 364]. Random variables $X_1, \ldots, X_n$ are said to be exchangeable if it is a finite subset of an infinite sequence of exchangeable random variables. It is well known that (by de Finetti's theorem) $X_1, \ldots, X_n$ are exchangeable if and only if they are positively dependent by mixture [as defined in Shaked (1977)]. Note that exchangeability is stronger than permutation symmetry. For more details see the discussion in Tong [(1980), pages 96–97].

Now for a finite n let us consider the random variables defined by, for $i = 1, \ldots, n$,

$$\xi_i = g(U_i, V_1, W), \qquad \xi_i' = g(U_i, V_i, W). \tag{3.1}$$

Then $\xi_1, \ldots, \xi_n$ are exchangeable and $\xi_1', \ldots, \xi_n'$ are exchangeable. But for $\mathbf{k} = (n, 0, \ldots, 0)$ and $\mathbf{k}' = (1, \ldots, 1)$ one has $\xi(\mathbf{k}) =_d (\xi_1, \ldots, \xi_n)$ and $\xi(\mathbf{k}') =_d (\xi_1', \ldots, \xi_n')$. Thus a partial ordering of positive dependence can be obtained by applying Theorem 2.1 and the related results given in Section 2.

When applying this result to the exchangeable normal, t, chi-square, gamma, F and exponential variables, many useful inequalities follow as special cases. The multivariate normal variables will be treated separately in this section. The exchangeable exponential variables can be obtained by taking $g(u, v, w) = \min(u, v, w)$ as considered previously by Marshall and Olkin (1967) and have an important application in reliability theory.

3.2. *Distributions with the semigroup property.* Let $\{f_\theta(x): \theta \in \Omega\}$ denote a family of density functions and assume that Ω is an interval of real numbers or an interval of integers. It is said to possess the semigroup property [see, e.g., Proschan and Sethuraman (1977)] if $\theta', \theta'' \in \Omega$ implies $\theta' + \theta'' \in \Omega$ and $f_{\theta'}(x) * f_{\theta''}(x) = f_{\theta'+\theta''}(x)$, where "$*$" denotes convolution.

APPLICATION 3.1. Let $X_{\theta,1}, \ldots, X_{\theta,n}$ denote i.i.d. random variables with density $f_\theta(x)$ and for fixed θ_0 and $\theta_0 - \theta \in \Omega$ let $X_{\theta_0-\theta}$ denote another independent random variable with density $f_{\theta_0-\theta}(x)$. Then define an n-dimensional random vector $\mathbf{X}(\theta) = (X_1, \ldots, X_n)$ such that $X_i = X_{\theta,i} + X_{\theta_0-\theta}$ for $i = 1, \ldots, n$. If $\{f_\theta(x): \theta \in \Omega\}$ possesses the semigroup property and if $\theta_1, \theta_2 \in \Omega$, $\theta_1 \neq \theta_2$ implies $|\theta_1 - \theta_2| \in \Omega$, then (a) $E_\theta \prod_{i=1}^n \phi(X_i)$ is a nonincreasing function of θ for $\theta < \theta_0$ for all Borel-measurable functions $\phi \geq 0$ (provided that the expectations exist); (b) $E_\theta \prod_{i=1}^n \phi(X_i)$ is a nonincreasing function of θ for $\theta < \theta_0$ for all positive even integers n and all Borel-measurable functions ϕ and (c) $P_\theta[X_1 \in B, \ldots, X_n \in B]$ is a nonincreasing function of θ for $\theta < \theta_0$ for all Borel-measurable subsets $B \subset \mathscr{R}$.

PROOF. For every fixed $\theta_1, \theta_2 \in \Omega$ such that $\theta_1 < \theta_2 < \theta_0$ define $U_i = X_{\theta_1, i}$, $V_i = X_{\theta_2 - \theta_1, i}$ $(i = 1, \dots, n)$ and $W = X_{\theta_0} - X_{\theta_2}$. The proof follows by applying Theorem 2.1. □

Note that Application 3.1 applies to the binomial, gamma and Poisson distributions, and Poisson processes and several other distributions.

3.3. *The multivariate normal distribution.* In this section we show how the positive dependence of a multivariate normal variable with a common marginal distribution can be partially ordered via their correlation coefficients.

APPLICATION 3.2. Let $0 \le \rho_1 < \rho_2 \le 1$ be arbitrary but fixed. Let $\mathbf{k}$ and $\mathbf{k}'$ be two vectors of nonnegative integers as given in (2.3) and define a correlation matrix $\mathbf{R} = \mathbf{R}(\mathbf{k}) = (\rho_{ij})$ to be such that (for $i \ne j$)

$$\rho_{ij} = \begin{cases} \rho_2, & \text{if } 1 \le i, j \le k_1,\ k_1 + 1 \le i, j \le k_1 + k_2, \dots, \sum_{m=1}^{r-1} k_m < i, j \le n, \\ \rho_1, & \text{otherwise.} \end{cases}$$

(That is, the random variables $X_1, \dots, X_n$ are partitioned into r groups with groups sizes $k_1, \dots, k_r$, respectively; the correlations of the variables within the same group are ρ_2 and the correlations between groups are ρ_1.) Let $\mathbf{X}(\mathbf{k}) \sim \mathscr{N}_n(\boldsymbol{\mu}, \sigma^2 \mathbf{R}(\mathbf{k}))$, where $\boldsymbol{\mu} = (\mu, \dots, \mu)$, and let $\mathbf{R}(\mathbf{k}')$ and $\mathbf{X}(\mathbf{k}')$ be defined similarly. (a) If $\mathbf{k} \succ \mathbf{k}'$, then $\mathbf{X}(\mathbf{k}) \ge_{pd+} \mathbf{X}(\mathbf{k}')$ and (b) if $\mathbf{k} \succ \mathbf{k}'$ and if the components of $\mathbf{k}, \mathbf{k}'$ are even integers, then $\mathbf{X}(\mathbf{k}) \ge_{pd} \mathbf{X}(\mathbf{k}')$.

PROOF. Immediate by choosing

$$g(u, v, w) = \mu + \sigma\left(\sqrt{1 - \rho_2}\, u + \sqrt{\rho_2 - \rho_1}\, v + \sqrt{\rho_1}\, w\right) \tag{3.2}$$

in Theorem 2.1 and Corollary 2.1. □

EXAMPLE 3.1. Let $\mathbf{X} = (X_1, X_2, X_3, X_4)$ have a multivariate normal distribution with equal means, equal variances and a correlation matrix $\mathbf{R}(\mathbf{k})$. Let $\mathbf{k}_4 = (4, 0, 0, 0)$, $\mathbf{k}_3 = (3, 1, 0, 0)$, $\mathbf{k}_2 = (2, 2, 0, 0)$, $\mathbf{k}_1 = (1, 1, 1, 1)$. Then $\mathbf{X}(\mathbf{k}_{i+1}) \ge_{pd+} \mathbf{X}(\mathbf{k}_i)$ for $i = 1, 2, 3$. Note that here all correlations in $\mathbf{R}(\mathbf{k}_4)$ are ρ_2 and all correlations in $\mathbf{R}(\mathbf{k}_1)$ are ρ_1.

As a special case of Application 3.2, we observe:

APPLICATION 3.3. Let $\mathbf{R}(s)$ denote the correlation matrix such that (for $i \ne j$)

$$\rho_{ij} = \begin{cases} \rho_2, & \text{for } 1 \le i, j \le s, \\ \rho_1, & \text{otherwise.} \end{cases}$$

If $0 \le \rho_1 < \rho_2 \le 1$, then $E_{\mathbf{R}(s)}\prod_{i=1}^n \phi(X_i)$ is a nondecreasing function of s for $\phi \ge 0$ (and for any ϕ and any even n and for $s = 0, 2, \dots, n$).

PROOF. Immediate from $(s+1, 1, \dots, 1, 0, \dots, 0) \succ (s, 1, 1, \dots, 1, 0, \dots, 0)$. □

Finally by taking $s = 0$ and $s = n$ in Application 3.3 we observe the following fact as a special consequence for exchangeable normal variables.

APPLICATION 3.4. If $(X_1, \dots, X_n)$ has a multivariate normal distribution with means μ, variances σ^2 and correlation coefficients $\rho \ge 0$, then (a) $E_\rho \prod_{i=1}^n \phi(X_i)$ is a nondecreasing function of ρ for all Borel-measurable functions $\phi \ge 0$ (and for all ϕ when n is even) such that the expectations exist and (b) the probability $P_\rho[X_1 \in B, \dots, X_n \in B]$ is a nondecreasing function of ρ for all Borel-measurable subsets $B \subset \mathscr{R}$, and it is strictly increasing in ρ unless $P[X_1 \in B]$ is 0 or 1.

This result was obtained previously by Rinott and Pollak (1980) for $n = 2$ and by Shaked and Tong (1985) for general n.

REFERENCES

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

MARSHALL, A. W. and OLKIN, I. (1967). A multivariate exponential distribution. *J. Amer. Statist. Assoc.* **62** 30–44.

MARSHALL, A. W. and OLKIN, I. (1979). *Inequalities: Theory of Majorization and Its Applications.* Academic, New York.

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

RINOTT, Y. and POLLAK, M. (1980). A stochastic ordering induced by a concept of positive dependence and monotonicity of asymptotic test sizes. *Ann. Statist.* **8** 190–198.

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. *J. Multivariate Anal.* **17** 333–349.

TONG, Y. L. (1977). An ordering theorem for conditionally independent and identically distributed random variables. *Ann. Statist.* **5** 274–277.

TONG, Y. L. (1980). *Probability Inequalities in Multivariate Distributions.* Academic, New York.

SCHOOL OF MATHEMATICS
GEORGIA INSTITUTE OF TECHNOLOGY
ATLANTA, GEORGIA 30332