Sankhyā : The Indian Journal of Statistics
1995, Volume 56, Series A, Pt. 2, pp. 365–368

A NOTE ON ORDER STATISTICS FOR NONDENTICALLY DISTRIBUTED VARIABLES

By SAYAJI HANDE
Indian Statistical Institute

SUMMARY. In this note we give simpler proofs and extensions of some results presented by Bapat and Beg (1989) on order statistics for nonidentically distributed variables using permanents.

Let $X_1, \ldots, X_n$ be independent random variables and let $Y_1 \leqslant \ldots \leqslant Y_n$ denote the associated ordered values. For the sake of completeness we state and prove the following well known result.

Theorem 1. *If X_i's are symmetric about zero then $-Y_r$ and Y_{n-r+1} are identically distributed.*

Proof. Note that $(X_1, \ldots, X_n)$ and $(-X_1, \ldots, -X_n)$ have the same distribution. Hence the r-th ordered value of X_i for $i = 1, \ldots, n$, has the same distribution as the r-th ordered value of $-X_i$ for $i = 1, \ldots, n$. This completes the proof. □

Remark 1. Note that if $X_1, \ldots, X_n$ are arbitrary random variables (not necessarily independent) such that $(X_1, \ldots, X_n)$ and $(-X_{i_1}, \ldots, -X_{i_n})$ have the same distribution for some permutation $\{i_1, i_2, \ldots, i_n\}$ then $-Y_r$ and Y_{n-r+1} have the same distribution for every $r = 1, 2, \ldots, n$.

Bapat and Beg (1989) [BB] proved a partial converse of the above theorem (Theorem 3.1 in their paper), where they assume absolute continuity of the distribution functions. In the following we give a simpler proof for Theorem 3.1 of BB, without assuming continuity of the distribution functions. In Theorem 3 we present a generalized version of Theorem 2, and in Theorem 4 we give a partial converse of the statement in Remark 1. Henceforth, we assume that $P(X_i > t) \in (0, 1)$ for all t and for all $i = 1, \ldots, n$.

Theorem 2. *Let $X_1, \ldots, X_n$ be independent random variables. Suppose X_i for $i = 2, \ldots, n$, are symmetric about zero. If $-Y_i$ and Y_{n-i+1} have the same distribution then X_1 is symmetric about zero.*

Proof. Note that

$$\begin{aligned}P(-Y_r \leqslant t) = P(Y_r \geqslant -t) &= P(X_i \geqslant -t \text{ for at least } n-r+1 \text{ indices } i \in \{1, \ldots, n\})\\ &= P(X_1 \geqslant -t)P(X_i \geqslant -t \text{ for at least } n-r \text{ indices } i \in \{2, \ldots, n\})\\ &\quad + P(X_1 < -t)P(X_i \geqslant -t \text{ for at least } n-r+1 \text{ indices } i \in \{2, \ldots, n\})\end{aligned}$$

and

$$\begin{aligned}P(Y_{n-r+1} \leqslant t) &= P(X_i > t \text{ for all most } r-1 \text{ indices } i \in \{1, \ldots, n\})\\ &= P(X_1 \leqslant t)P(X_i > t \text{ for at most } r-1 \text{ indices } i \in \{2, \ldots, n\})\\ &\quad + P(X_1 > t)P(X_i > t \text{ for at most } r-2 \text{ indices } i \in \{2, \ldots, n\}).\end{aligned}$$

From Theorem 1, we have

$$\begin{aligned}&P(X_i \geqslant -t \text{ for at least } n-r \text{ indices } i \in \{2, \ldots, n\})\\ &= P(X_i > t \text{ for at most } r-1 \text{ indices } i \in \{2, \ldots, n\}) \text{ and}\\ &P(X_i \geqslant -t \text{ for at least } n-r+1 \text{ indices } i \in \{2, \ldots, n\})\\ &= P(X_i > t \text{ for at most } r-2 \text{ indices } i \in \{2, \ldots, n\}).\end{aligned}$$

This shows that for all $1 \leqslant r \leqslant n$,

$$\begin{aligned}&[P(X_1 \geqslant -t) - P(X_1 \leqslant t)]P(X_i > -t \text{ for at least } n-r \text{ indices } i \in \{2, \ldots, n\})\\ &= [P(X_1 > t) - P(X_1 < -t)]P(X_i > -t \text{ for at least } n-r+1 \text{ indices for } i \in \{2, \ldots, n\}).\end{aligned}$$

Note that $P(X_1 \geqslant -t) - P(X_1 \leqslant t) = P(X_1 > t) - P(X_1 < -t)$ and

$$\begin{aligned}&P(X_i > -t \text{ for at least } n-r \text{ indices } i \in \{2, \ldots, n\})\\ &\neq P(X_i > -t \text{ for at least } n-r+1 \text{ indices for } i \in \{2, \ldots, n\}).\end{aligned}$$

Hence $P(X_1 \geqslant -t) - P(X_1 \leqslant t) = 0$. This completes the proof. □

In the light of Remark 1 it is clear that the strict converse of Theorem 1 is not true in general. Consider the following example. Let X_1 be a $N(3, 1)$, X_2 be a $N(-3, 1)$, X_3 be a χ_7^2 and X_4 be a $-\chi_7^2$. Let $Y_1 \leqslant Y_2 \leqslant Y_3 \leqslant Y_4$ be the ordered X_i's. From Remark 1 it follows that $-Y_r$ and Y_{n-r+1} possess the same distribution. Note that none of the X_i's are symmetric about zero.

Theorem 3. *Let $X_1, \ldots, X_n$ be independent random variables. Suppose X_i for $i = k+1, \ldots, n$ are symmetric about zero, $X_1, \ldots, X_k$ are i.i.d. If $-Y_r$ and Y_{n-r+1} have the same distribution then $X_1, \ldots, X_k$ are symmetric about zero.*

Proof. Let $F(.)$ be the c.d.f. of X_1 and let $C(F) = \{x : F(x) \text{ is continuous}\}$. For $t \in C(F)$, let

$$p(l) = P(X_i > -t \text{ for at least } l \text{ indices } i \in \{k+1, \ldots, n\}).$$

Note that

$$P(-Y_r \leqslant t)$$
$$= \sum_{j=0}^{k} P(X_i > -t \text{ for } j \text{ indices } i \in \{1, \ldots, k\})$$
$$\times P(X_i > -t \text{ for at least } n-r+1-j \text{ indices } i \in \{k+1, \ldots, n\})$$
$$= \sum_{j=0}^{k} P(X_i > -t \text{ for } j \text{ indices } i \in \{1, \ldots, k\})\, p(n-r+1-j).$$

Since $X_{k+1}, \ldots, X_n$ are symmetric about zero, by similar argument as in the proof of Theorem 2, it is easy to see that

$$P(Y_{n-r+1} \leqslant t)$$
$$= \sum_{j=0}^{k} P(X_i < t \text{ for } j \text{ indices } i \in \{1, \ldots, k\})\, p(n-r+1-j).$$

If $-Y_r$ and Y_{n-r+1} have the same distribution then as in the proof of Theorem 2, $P(-Y_r \leqslant t) - P(Y_{n-r+1} \leqslant t) = 0$ can be written as

$$\sum_{j=0}^{k} [P(Z_1 = j) - P(Z_2 = j)]\, p(n-r+1-j) = 0,$$

where Z_1 is a binomial random variable with k trials and probability $p_1 = 1 - P(X_1 < -t)$ and Z_2 is a binomial random variable with k trials and probability $p_2 = 1 - P(X_1 < t)$.

Note that the sequence $p(n-r+1-j)$ is increasing in j. Let $a_0 = p(n-r+1)$ and for $j = 1, \ldots, k$, define $a_j = p(n-r+1-j) - p(n-r+1-j+1)$. Hence $a_j > 0 \;\forall\, j$, and we have

$$\sum_{j=0}^{k} [P(Z_1 \geqslant j) - P(Z_2 \geqslant j)]\, a = 0.$$

If $p_1 \neq p_2$, then all the terms in the summation have the same sign, which is not possible. This completes the proof. □

It is well known that if $-Y_r$ and Y_{n-r+1} have the same distribution for some r then X_i's are symmetric about zero, provided X_i's are i.i.d (see for example David (1981)). Theorem 4 is in that spirit. Before we proceed, we need a definition.

Definition 1. The random variables X and Y are said to be stochastically ordered if

$$P(X > t) \geqslant P(Y > t) \;\forall t,$$

or

$$P(Y > t) \geqslant P(Y > t) \;\forall t.$$

If strict inequality holds above for all t then they are said to be strictly stochastically ordered.

Theorem 4. *Let X_i's be strictly stochastically ordered independent random variables. The following two statements are equivalent.*

[i] $(X_1, X_2, \ldots, X_n)$ *and* $-(X_{i_1}, X_{i_2}, \ldots, X_{i_n})$ *have the same distribution for some permutation* $i_1, i_2, \ldots, i_n$.

[ii] $-Y_r$ *and* Y_{n-r+1} *have the same distribution for every* $r = 1, 2, \ldots, n$. we need the following lemma, proof of which is obvious.

Lemma 1. *Let Z be a sum of n independent Bernoulli random variables with associated probabilities p_i, $i = 1, \ldots, n$ and Z' be a sum of n independent Bernoulli random variables with associated probabilities p'_i, $i = 1, \ldots, n$. ...If Z and Z' have the same distribution then*

$$(p_1, p_2, \ldots, p_n) = (p'_{i_1}, p'_{i_2}, \ldots, p'_{i_n})$$

for some permutation $i_1, i_2, \ldots, i_n$.

Proof of Theorem 4. Let $Z = \sum_{i=1}^{n} I(X_i \leqslant t)$ and $Z' = \sum_{i=1}^{n} I(-X_i \leqslant t)$, where $I(.)$ denotes the indicator function. From the assumption that $-Y_r$ and Y_{n-r+1} have the same distribution for every r, it follows that Z and Z' have the same distribution. Now the result follows from the lemma and from the fact that X_i's are strictly stochastically ordered. □

In the following result we prove log-concavity of the sequence $P(Y_r > t)$ for $r = 1, \ldots, n$ (Theorem 4.5 in BB). Log-concavity of the sequence $P(Y_r < t)$ for $r = 1, \ldots, n$ can be proved similarly.

Theorem 5. *The sequence $P(Y_r > t)$ for $r = 1, \ldots, n$, is log-concave.*

Proof. Let $Z_i = I(X_i \leqslant t)$ for $i = 1, \ldots, n$, and $Z = \sum_{i=1}^{n} Z_i$. Note that Z is sum of n independent Bernoulli variables hence it is strongly unimodal (see for example Joag-Dev and Dharmadhikari 1988, pp 109). Hence the sequences $P(Z = r)$ and $P(Z \leqslant r)$ are log-concave.

Now the proof follows from the fact that

$$P(Y_r > t) = P(Z \leqslant n - r + 1). \quad \square$$

Acknowledgements. The author is thankful to the Editor for the comments which have improved the presentation of this paper.

References

Bapat, R. B. and Beg. M. I. (1989). Order statistics for nonidentically distributed variables and permanents. *Sankhyā A*, **51**, 79-93.

David, H. A. (1981). *Order Statistics*, John Wiley and sons, New York.

Dharmadhikari, S. and Joag-Dev, K. (1988). *Unimodality, Convexity and Applications*, Academic Press, Boston.

Eurolink Overseas Ltd
F-2 Block B-1
Mohan Co-operative Ind. Est
New Delhi 110 044.