Ann. Inst. Statist. Math.
Vol. 40, No. 2, 273-277 (1988)

RECURRENCE RELATIONS FOR ORDER STATISTICS FROM n INDEPENDENT AND NON-IDENTICALLY DISTRIBUTED RANDOM VARIABLES

N. BALAKRISHNAN

Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada

(Received May 25, 1987; revised September 29, 1987)

Abstract. Some well-known recurrence relations for order statistics in the i.i.d. case are generalized to the case when the variables are independent and non-identically distributed. These results could be employed in order to reduce the amount of direct computations involved in evaluating the moments of order statistics from an outlier model.

Key words and phrases: Order statistics, recurrence relation, single moments, product moments, permanent, outliers.

1. Introduction

Let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ denote the order statistics obtained from the realizations of n independent random variables $X_1, X_2, \ldots, X_n$. Let us denote $E(X_{r:n}^k)$ by $\mu_{r:n}^{(k)}$ $(1 \leq r \leq n, k \geq 1)$ and $E(X_{r:n}X_{s:n})$ by $\mu_{r,s:n}$ $(1 \leq r < s \leq n)$. Then, when the X_i's are identically distributed, it is known that

$$(1.1) \qquad r\mu_{r+1:n}^{(k)} + (n-r)\mu_{r:n}^{(k)} = n\mu_{r:n-1}^{(k)}, \quad 1 \leq r \leq n, \quad k \geq 1,$$

and

$$(1.2) \qquad \begin{aligned}(r-1)\mu_{r,s:n} + (s-r)\mu_{r-1,s:n} + (n-s+1)\mu_{r-1,s-1:n} \\ = n\mu_{r-1,s-1:n-1}, \quad 1 \leq r < s \leq n,\end{aligned}$$

(David (1981), pp. 46–49). Relation (1.1) has been derived by Cole (1951) in the continuous case and by Melnick (1964) in the discrete case. Arnold (1977) has given a proof which covers also mixtures of continuous and discrete distributions. Relation (1.2) has been proved by Govindarajulu (1963) and Balakrishnan (1986) for the continuous and discrete cases, respectively; see also Balakrishnan and Malik (1986) for some comments on these two relations. While all these results have been obtained under the assumption of

X_i's being i.i.d., David and Joshi (1968) have demonstrated that Relations (1.1) and (1.2) remain valid even when the order statistics arise from n exchangeable random variables.

In this note, we generalize the above relations to the case when the order statistics are obtained from n independent and non-identically distributed random variables. These relations may be employed in a very simple recursive way in order to reduce the amount of direct computations (which, quite often, are laborious and cumbersome) involved in evaluating the moments of order statistics from an outlier model.

2. Relations

By assuming that $X_1, X_2, \ldots, X_n$ are independent variates with X_i ($i=1, 2, \ldots, n$) having pdf $f_i(x)$ and cdf $F_i(x)$, Vaughan and Venables (1972) have shown that the density function of $X_{r:n}$ ($1 \le r \le n$) can be written down as

(2.1) $h_{r:n}(x)$

$$= \frac{1}{(r-1)!(n-r)!} \; {}^{+}\left| \begin{array}{cccc} F_1(x) & F_2(x) & \cdots & F_n(x) \\ \vdots & \vdots & & \vdots \\ F_1(x) & F_2(x) & \cdots & F_n(x) \\ f_1(x) & f_2(x) & \cdots & f_n(x) \\ 1-F_1(x) & 1-F_2(x) & \cdots & 1-F_n(x) \\ \vdots & \vdots & & \vdots \\ 1-F_1(x) & 1-F_2(x) & \cdots & 1-F_n(x) \end{array} \right|^{+} \begin{array}{l} \left.\begin{array}{l} \\ \\ \\ \end{array}\right\} r-1 \text{ rows} \\ \\ \left.\begin{array}{l} \\ \\ \\ \end{array}\right\} n-r \text{ rows} \end{array},$$

where ${}^{+}|A|^{+}$ denotes the permanent of a square matrix A; the permanent is defined just like the determinant, except that all signs in the expansion are positive.

Let us now use $h_{r:n-m}^{[i_1,\ldots,i_m]}(x)$, $1 \le r \le n-m$, to denote the density function of the r-th order statistic in a sample of size $n-m$ obtained by dropping X_{i_1}, $X_{i_2}, \ldots, X_{i_m}$ from the original set of n variables. We then have the following relation.

RELATION 1. For $1 \le r \le n-1$,

$$rh_{r+1:n}(x) + (n-r)h_{r:n}(x) = \sum_{i=1}^{n} h_{r:n-1}^{[i]}(x) \,. \tag{2.2}$$

PROOF. First, consider the permanent expression of $rh_{r+1:n}(x)$ from equation (2.1). Upon expanding this permanent by its first row, we get

$$rh_{r+1:n}(x) = \sum_{i=1}^{n} F_i(x) h_{r:n-1}^{[i]}(x) \,. \tag{2.3}$$

Next, consider the expression of $(n-r)h_{r:n}(x)$ from equation (2.1). Upon expanding this permanent by its last row, we get

$$(n-r)h_{r:n}(x)=\sum_{i=1}^{n}(1-F_i(x))h_{r:n-1}^{[i]}(x)\,. \tag{2.4}$$

Relation (2.2) follows immediately upon adding equations (2.3) and (2.4).

Let us now denote

$$S_{1:n-m}(x)=\sum_{1\le i_1<i_2<\cdots<i_m\le n}h_{1:n-m}^{[i_1,\ldots,i_m]}(x)$$

and

$$S_{n-m:n-m}(x)=\sum_{1\le i_1<i_2<\cdots<i_m\le n}h_{n-m:n-m}^{[i_1,\ldots,i_m]}(x)\,,$$

with $S_{1:n}(x)\equiv h_{1:n}(x)$ and $S_{n:n}(x)\equiv h_{n:n}(x)$. Then by repeated application of Relation 1, we directly obtain the following relations.

RELATION 2. For $1\le r\le n-1$,

$$h_{r:n}(x)=\sum_{j=r}^{n}(-1)^{j-r}\binom{j-1}{r-1}S_{j:j}(x)\,. \tag{2.5}$$

RELATION 3. For $2\le r\le n$,

$$h_{r:n}(x)=\sum_{j=n-r+1}^{n}(-1)^{j-n+r-1}\binom{j-1}{n-r}S_{1:j}(x)\,. \tag{2.6}$$

Remark 1. For the case when the X_i's are identically distributed, it is easy to see that

$$S_{j:j}(x)=\sum_{1\le i_1<\cdots<i_{n-j}\le n}h_{j:j}^{[i_1,\ldots,i_{n-j}]}(x)=\binom{n}{j}h_{j:j}(x)$$

and

$$S_{1:j}(x)=\sum_{1\le i_1<\cdots<i_{n-j}\le n}h_{1:j}^{[i_1,\ldots,i_{n-j}]}(x)=\binom{n}{j}h_{1:j}(x)\,.$$

As a result, Relations 2 and 3 yield (in terms of moments)

$$\mu_{r:n}^{(k)}=\sum_{j=r}^{n}(-1)^{j-r}\binom{j-1}{r-1}\binom{n}{j}\mu_{j:j}^{(k)}\,,\qquad 1\le r\le n-1\,,$$

and

$$\mu_{r:n}^{(k)} = \sum_{j=n-r+1}^{n} (-1)^{j-n+r-1} \binom{j-1}{n-r} \binom{n}{j} \mu_{1:j}^{(k)}, \quad 2 \le r \le n .$$

These two recurrence relations are quite well known and are due to Srikantan (1962).

Vaughan and Venables (1972) have also shown that the joint density function of $X_{r:n}$ and $X_{s:n}$ $(1 \le r < s \le n)$ can be written down as

$$\text{(2.7)} \quad h_{r,s:n}(x,y) = \frac{1}{(r-1)!(s-r-1)!(n-s)!}$$

$$\times \begin{array}{l} {}^{+}\left|\begin{array}{cccc} F_1(x) & F_2(x) & \cdots & F_n(x) \\ \vdots & \vdots & & \vdots \\ F_1(x) & F_2(x) & \cdots & F_n(x) \\ f_1(x) & f_2(x) & \cdots & f_n(x) \\ F_1(y)-F_1(x) & F_2(y)-F_2(x) & \cdots & F_n(y)-F_n(x) \\ \vdots & \vdots & & \vdots \\ F_1(y)-F_1(x) & F_2(y)-F_2(x) & \cdots & F_n(y)-F_n(x) \\ f_1(y) & f_2(y) & \cdots & f_n(y) \\ 1-F_1(y) & 1-F_2(y) & \cdots & 1-F_n(y) \\ \vdots & \vdots & & \vdots \\ 1-F_1(y) & 1-F_2(y) & \cdots & 1-F_n(y) \end{array}\right|^{+} \end{array} \begin{array}{l} \left.\begin{array}{l} \\ \\ \\ \end{array}\right\} r-1 \text{ rows} \\ \\ \left.\begin{array}{l} \\ \\ \\ \end{array}\right\} s-r-1 \text{ rows} \\ \\ \left.\begin{array}{l} \\ \\ \\ \end{array}\right\} n-s \text{ rows} \end{array}$$

Now using $h_{r,s:n-1}^{[i]}(x, y)$, $1 \le r < s \le n-1$, to denote the joint density function of the r-th and s-th order statistics in a sample of size $n-1$ obtained by dropping X_i from the original set of n variables, we have the following recurrence relation.

RELATION 4. For $2 \le r < s \le n$,

$$\text{(2.8)} \quad (r-1)h_{r,s:n}(x,y) + (s-r)h_{r-1,s:n}(x,y) + (n-s+1)h_{r-1,s-1:n}(x,y) = \sum_{i=1}^{n} h_{r-1,s-1:n-1}^{[i]}(x,y) .$$

PROOF. Expanding the permanent in (2.7) by its first, r-th, and last row respectively, we obtain

$$\text{(2.9)} \qquad (r-1)h_{r,s:n}(x,y) = \sum_{i=1}^{n} F_i(x) h_{r-1,s-1:n-1}^{[i]}(x,y) ,$$

$$\text{(2.10)} \qquad (s-r)h_{r-1,s:n}(x,y) = \sum_{i=1}^{n} (F_i(y) - F_i(x)) h_{r-1,s-1:n-1}^{[i]}(x,y) ,$$

$$\text{(2.11)} \qquad (n-s+1)h_{r-1,s-1:n}(x,y) = \sum_{i=1}^{n} (1 - F_i(y)) h_{r-1,s-1:n-1}^{[i]}(x,y) .$$

Relation (2.8) follows upon adding equations (2.9), (2.10) and (2.11).

Remark 2. For the p-outlier model, that is, $F_1=F_2=\cdots=F_{n-p}=F$ and $F_{n-p+1}=\cdots=F_n=G$, Relations 1 and 4, respectively, yield

$$rh_{r+1:n}(x)+(n-r)h_{r:n}(x)=(n-p)h^{[F]}_{r:n-1}(x)+ph^{[G]}_{r:n-1}(x)$$

and

$$\begin{aligned}(r-1)h_{r,s:n}(x,y)&+(s-r)h_{r-1,s:n}(x,y)+(n-s+1)h_{r-1,s-1:n}(x,y)\\&=(n-p)h^{[F]}_{r-1,s-1:n-1}(x,y)+ph^{[G]}_{r-1,s-1:n-1}(x,y)\ ,\end{aligned}$$

where $h^{[F]}_{r:n-1}(x)$ and $h^{[G]}_{r:n-1}(x)$ are the density function of the r-th order statistic in a sample of size $n-1$ from the p-outlier model and the $(p-1)$-outlier model, respectively.

Acknowledgement

Thanks are due to Mrs. Janet Leach for the excellent typing of the manuscript.

REFERENCES

Arnold, B. C. (1977). Recurrence relations between expectations of functions of order statistics, *Scand. Actuar. J.*, 169–174.

Balakrishnan, N. (1986). Order statistics from discrete distributions, *Comm. Statist. A—Theory Methods*, **15**(3), 657–675.

Balakrishnan, N. and Malik, H. J. (1986). A note on moments of order statistics, *Amer. Statist.*, **40**, 147–148.

Cole, R. H. (1951). Relations between moments of order statistics, *Ann. Math. Statist.*, **22**, 308–310.

David, H. A. (1981). *Order Statistics*, 2nd ed., Wiley, New York.

David, H. A. and Joshi, P. C. (1968). Recurrence relations between moments of order statistics for exchangeable variates, *Ann. Math. Statist.*, **39**, 272–274.

Govindarajulu, Z. (1963). On moments of order statistics and quasi-ranges from normal populations, *Ann. Math. Statist.*, **34**, 633–651.

Melnick, E. L. (1964). Moments of ranked Poisson variates, M. S. Thesis, Virginia Polytechnic Institute.

Srikantan, K. S. (1962). Recurrence relations between the PDF's of order statistics, and some applications, *Ann. Math. Statist.*, **33**, 169–177.

Vaughan, R. J. and Venables, W. N. (1972). Permanent expressions for order statistics densities, *J. Roy. Statist. Soc. Ser. B*, **34**, 308–310.