Ann. Inst. Statist. Math.
Vol. 41, No. 2, 323-329 (1989)

RECURRENCE RELATIONS AMONG MOMENTS OF ORDER STATISTICS FROM TWO RELATED SETS OF INDEPENDENT AND NON-IDENTICALLY DISTRIBUTED RANDOM VARIABLES

N. BALAKRISHNAN

Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada L8S 4K1

(Received April 25, 1988; revised August 26, 1988)

Abstract. Some recurrence relations among moments of order statistics from two related sets of variables are quite well-known in the i.i.d. case and are due to Govindarajulu (1963*a*, *Technometrics*, **5**, 514–518 and 1966, *J. Amer. Statist. Assoc.*, **61**, 248–258). In this paper, we generalize these results to the case when the order statistics arise from two related sets of independent and non-identically distributed random variables. These relations can be employed to simplify the evaluation of the moments of order statistics in an outlier model for symmetrically distributed random variables.

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

1. Introduction

Let $X_{1:n} \le X_{2:n} \le \cdots \le X_{n:n}$ denote the order statistics obtained from n independent absolutely continuous random variables X_i $(i = 1, 2, \dots, n)$, with X_i having pdf $f_i(x)$ and cdf $F_i(x)$. Then, Vaughan and Venables (1972) have shown that the density function of $X_{r:n}$ $(1 \le r \le n)$ can be written as

$$(1.1) \qquad h_{r:n}(x) = \frac{1}{(r-1)!(n-r)!} \times {}^{+}\left|\begin{array}{cccc} F_1(x) & F_2(x) & \cdots & F_n(x) \\ f_1(x) & f_2(x) & \cdots & f_n(x) \\ 1-F_1(x) & 1-F_2(x) & \cdots & 1-F_n(x) \end{array}\right|^{+} \begin{array}{l} r-1 \text{ rows} \\ \\ n-r \text{ rows} \end{array},$$

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 denote the single moments of order statistics by $\mu_{r:n}^{(k)}$, $1 \le r \le n$, where

$$(1.2) \qquad \mu_{r:n}^{(k)} = E(X_{r:n}^k) = \int_{-\infty}^{\infty} x^k h_{r:n}(x) dx .$$

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

$$(1.3) \qquad h_{r,s:n}(x,y) = \frac{1}{(r-1)!(s-r-1)!(n-s)!}$$

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

Let us denote the product moments of order statistics by $\mu_{r,s:n}$, $1 \le r < s \le n$, where

$$(1.4) \qquad \mu_{r,s:n} = \iint_{-\infty<x<y<\infty} xy h_{r,s:n}(x,y) dy dx .$$

Let us now assume that the density functions $f_i(x)$ are all symmetric about 0. Then for $x > 0$, let

$$(1.5) \qquad G_i(x) = 2F_i(x) - 1, \quad g_i(x) = 2f_i(x) .$$

That is, the density functions $g_i(x)$, $i = 1, 2, \ldots, n$, are obtained by folding the density functions $f_i(x)$ at zero. Now, let $Y_{1:n} \le Y_{2:n} \le \cdots \le Y_{n:n}$ denote the order statistics obtained from n independent, absolutely continuous random variables Y_i ($i = 1, 2, \ldots, n$), with Y_i having pdf $g_i(x)$ and cdf $G_i(x)$. Let us denote $\nu_{r:n-l}^{(k)[i_1,\ldots,i_l]}$ for the k-th single moment of $Y_{r:n-l}^{[i_1,\ldots,i_l]}$ and $\nu_{r,s:n-l}^{[i_1,\ldots,i_l]}$ for the product moment of $Y_{r:n-l}^{[i_1,\ldots,i_l]}$ and $Y_{s:n-l}^{[i_1,\ldots,i_l]}$, where $Y_{r:n-l}^{[i_1,\ldots,i_l]}$ denotes the r-th order statistic in a sample of size $n-l$ obtained by dropping $Y_{i_1}, Y_{i_2}, \ldots, Y_{i_l}$ from the original set of n variables $Y_1, Y_2, \ldots, Y_n$.

For the i.i.d. case, that is, $F_1 = F_2 = \cdots = F_n = F$, Govindarajulu (1963*a*) has derived some relationships among these two sets of moments of order statistics. He (1966) also employed these relations successfully to compute the moments of order statistics from a double exponential distribution. These results have been recently generalized by Balakrishnan

(1988*a*) to the case when the order statistics arise from a sample containing a single outlier. Furthermore, Balakrishnan and Ambagaspitiya (1988) have made use of these relations in studying the robustness properties of various estimators of the location and scale parameters of the double exponential distribution in the presence of a single outlier. Interested readers may refer to the recent monograph by Arnold and Balakrishnan (1989) which presents a compendium of these and many other relations.

In this paper, we generalize all these results to the case when the order statistics arise from two related sets of independent and non-identically distributed random variables. These results are established by using some properties of permanents. Similar work has been done recently by Balakrishnan (1988*b*) while generalizing the results of Govindarajulu (1963*b*), David (1981), Balakrishnan (1986) and Balakrishnan and Malik (1986) to the case when the order statistics arise from n independent and non-identically distributed random variables.

2. Recurrence relations

In this section, we derive some relations which express the moments $\mu_{r:n}^{(k)}$ $(1 \le r \le n)$ and $\mu_{r,s:n}$ $(1 \le r < s \le n)$ in terms of the moments $\nu_{r:n}^{(k)}$ and $\nu_{r,s:n}$.

Relation A. For $1 \le r \le n$ and $k = 1, 2, \ldots,$

$$\mu_{r:n}^{(k)} = 2^{-n}\left\{\sum_{l=0}^{r-1} \sum_{1 \le i_1 < \cdots < i_l \le n} \nu_{r-l:n-l}^{(k)[i_1,\ldots,i_l]} + (-1)^k \sum_{l=r}^{n} \sum_{1 \le i_1 < \cdots < i_{n-l} \le n} \nu_{l-r+1:l}^{(k)[i_1,\ldots,i_{n-l}]}\right\}. \tag{2.1}$$

PROOF. From equations (1.2) and (1.5), we have

$$\mu_{r:n}^{(k)} = \frac{2^{-n}}{(r-1)!(n-r)!} \times \left\{\int_0^\infty x^k I_{r-1,n-r}(x)dx + (-1)^k \int_0^\infty x^k I_{n-r,r-1}(x)dx\right\}, \tag{2.2}$$

where

$$I_{r-1,n-r}(x) = \begin{array}{|ccccc|l} 1+G_1(x) & 1+G_2(x) & \cdots & 1+G_n(x) & & {}^{+}\; r-1 \text{ rows} \\ g_1(x) & g_2(x) & \cdots & g_n(x) & & \\ 1-G_1(x) & 1-G_2(x) & \cdots & 1-G_n(x) & & n-r \text{ rows} \end{array}^{+},$$

and a similar expression for $I_{n-r,r-1}(x)$. By using the Cauchy expansion of permanents (Aitken (1944), p. 74), we have

$$I_{r-1,n-r}(x) = {}^{+}\left|\begin{array}{cccc} 1+G_2(x) & 1+G_3(x) & \cdots & 1+G_n(x) \\ g_2(x) & g_3(x) & \cdots & g_n(x) \\ 1-G_2(x) & 1-G_3(x) & \cdots & 1-G_n(x) \end{array}\right|^{+} \begin{array}{l} r-2 \text{ rows} \\ \\ n-r \text{ rows} \end{array}$$

$$+ {}^{+}\left|\begin{array}{cccc} G_1(x) & 1+G_2(x) & \cdots & 1+G_n(x) \\ 1+G_1(x) & 1+G_2(x) & \cdots & 1+G_n(x) \\ g_1(x) & g_2(x) & \cdots & g_n(x) \\ 1-G_1(x) & 1-G_2(x) & \cdots & 1-G_n(x) \end{array}\right|^{+} \begin{array}{l} \\ r-2 \text{ rows} \\ \\ n-r \text{ rows} \end{array} .$$

By repeated application, we get

$$I_{r-1,n-r}(x) = \sum_{i_1=1}^{n} J_{0,r-2,n-r}^{[i_1]}(x) + J_{1,r-2,n-r}(x) ,$$

where $J_{0,r-2,n-r}^{[i_1]}(x)$ is the permanent obtained from $I_{r-1,n-r}(x)$ by dropping the first row and i_1-th column, and $J_{1,r-2,n-r}(x)$ is the permanent obtained from $I_{r-1,n-r}(x)$ by replacing the first row by $G_1(x), G_2(x), \ldots, G_n(x)$. Proceeding in a similar way, we obtain

$$I_{r-1,n-r}(x) = \sum_{l=0}^{r-1} (r-1-l)! \binom{r-1}{l} \sum_{1 \le i_1 < \cdots < i_{r-1-l} \le n} J_{l,0,n-r}^{[i_1,\ldots,i_{r-1-l}]}(x) ,$$

where

$$J_{l,0,n-r}^{[i_1,\ldots,i_{r-1-l}]}(x) = {}^{+}\left|\begin{array}{cccc} G_1(x) & G_2(x) & \cdots & G_n(x) \\ g_1(x) & g_2(x) & \cdots & g_n(x) \\ 1-G_1(x) & 1-G_2(x) & \cdots & 1-G_n(x) \end{array}\right|^{+} \begin{array}{l} l \text{ rows} \\ \\ n-r \text{ rows} \end{array} ,$$

with columns $(i_1, i_2, \ldots, i_{r-1-l})$ having been dropped. By realizing that

$$\int_0^\infty x^k J_{l,0,n-r}^{[i_1,\ldots,i_{r-1-l}]}(x) dx = l!(n-r)!\, \nu_{l+1:n-r+l+1}^{(k)[i_1,\ldots,i_{r-1-l}]}$$

and

$$\int_0^\infty x^k J_{r-1,0,n-r}(x) dx = (r-1)!(n-r)!\, \nu_{r:n}^{(k)} ,$$

we immediately obtain that

$$\frac{1}{(r-1)!(n-r)!}\int_0^\infty x^k I_{r-1,n-r}(x)dx = \sum_{l=0}^{r-1} \sum_{1\le i_1<\cdots<i_l\le n} \nu_{r-l:n-l}^{(k)[i_1,\ldots,i_l]} .$$

Proceeding exactly on the same lines, we also obtain that

$$\frac{1}{(r-1)!(n-r)!}\int_0^\infty x^k I_{n-r,r-1}(x)dx = \sum_{l=r}^{n} \sum_{1\le i_1<\cdots<i_{n-l}\le n} \nu_{l-r+1:l}^{(k)[i_1,\ldots,i_{n-l}]} .$$

Making use of these expressions on the RHS of (2.2), we derive the required relation.

Remark 1. If we set $F_1 = F_2 = \cdots = F_n = F$ and $f_1 = f_2 = \cdots = f_n = f$, Relation A then reduces to

$$\mu_{r:n}^{(k)} = 2^{-n}\left\{\sum_{l=0}^{r-1}\binom{n}{l}\nu_{r-l:n-l}^{(k)} + (-1)^k\sum_{l=r}^{n}\binom{n}{l}\nu_{l-r+1:l}^{(k)}\right\},$$

a relation that has been derived by Govindarajulu (1963*a*).

Remark 2. If we set $F_1 = F_2 = \cdots = F_{n-1} = F$ and $f_1 = f_2 = \cdots = f_{n-1} = f$ (that is, a single outlier model), Relation A then reduces to

$$\mu_{r:n}^{(k)} = 2^{-n}\left\{\sum_{l=0}^{r-1}\binom{n-1}{l}\nu_{r-l:n-l}^{(k)} + (-1)^k\sum_{l=r}^{n}\binom{n-1}{l-1}\nu_{l-r+1:l}^{(k)}\right.$$
$$\left. + \sum_{l=1}^{r-1}\binom{n-1}{l-1}\nu_{r-l:n-l}^{*(k)} + (-1)^k\sum_{l=r}^{n-1}\binom{n-1}{l}\nu_{l-r+1:l}^{*(k)}\right\},$$

a relation that has been established recently by Balakrishnan (1988*a*) and has been used by Balakrishnan and Ambagaspitiya (1988) in studying the robustness properties of various estimators of the location and scale parameters of the double exponential distribution in the presence of a single scale-outlier; here, $\nu^{*(k)}$ denotes the k-th moment in the non-outlier case.

In the following, we present a relation which expresses the moments $\mu_{r,s:n}$ $(1 \le r < s \le n)$ in terms of the moments $\nu_{r:n}$ and $\nu_{r,s:n}$.

Relation B. For $1 \le r < s \le n$,

$$\mu_{r,s:n} = 2^{-n}\left\{ \sum_{l=0}^{r-1} \sum_{1\le i_1<\cdots<i_l\le n} \nu_{r-l,\,s-l:n-l}^{[i_1,\ldots,i_l]} + \sum_{l=s}^{n} \sum_{1\le i_1<\cdots<i_{n-l}\le n} \nu_{l-s+1,\,l-r+1:l}^{[i_1,\ldots,i_{n-l}]} - \sum_{l=r}^{s-1} \sum_{1\le i_1<\cdots<i_l\le n} \nu_{s-l:n-l}^{[i_1,\ldots,i_l]}\, \nu_{l-r+1:l}^{[i_{l+1},\ldots,i_n]} \right\}. \tag{2.3}$$

The above relation may be proved by following exactly the same lines as in the proof of Relation A.

Remark 3. If we set $F_1 = F_2 = \cdots = F_n = F$ and $f_1 = f_2 = \cdots = f_n = f$, then Relation B reduces to the corresponding result for the product moments that has been derived by Govindarajulu (1963*a*).

Remark 4. If we set $F_1 = F_2 = \cdots = F_{n-1} = F$ and $f_1 = f_2 = \cdots = f_{n-1} = f$ (that is, a single outlier model), then Relation B reduces to the corresponding result for the product moments that has been established recently by Balakrishnan (1988*a*), and has also been applied by Balakrishnan and Ambagaspitiya (1988) in robustness studies.

Acknowledgements

Thanks are due to Mrs. Edna Pathmanathan for the excellent typing of the manuscript and to the Natural Sciences and Engineering Research Council of Canada for funding this project.

References

Aitken, A. C. (1944). *Determinants and Matrices*, Oliver and Boyd, London.

Arnold, B. C. and Balakrishnan, N. (1989). *Relations, Bounds and Approximations for Order Statistics*, Lecture Notes in Statistics, No. 53, Springer, New York.

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

Balakrishnan, N. (1988*a*). Recurrence relations among moments of order statistics from two related outlier models, *Biomet. J.*, **30**, 741–746.

Balakrishnan, N. (1988*b*). Recurrence relations for order statistics from n independent and non-identically distributed random variables, *Ann. Inst. Statist. Math.*, **40**, 273–277.

Balakrishnan, N. and Ambagaspitiya, R. S. (1988). Relationships among moments of order statistics from two related outlier models and some applications, *Comm. Statist. A—Theory Methods*, **17**, 2327–2341.

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

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

Govindarajulu, Z. (1963*a*). Relationships among moments of order statistics in samples from two related populations, *Technometrics*, **5**, 514–518.

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

Govindarajulu, Z. (1966). Best linear estimates under symmetric censoring of the parameters of a double exponential population, *J. Amer. Statist. Assoc.*, **61**, 248–258.

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