A MOMENT PROBLEM FOR ORDER STATISTICS

By Joseph B. Kadane

Center for Naval Analyses and Carnegie–Mellon University

Necessary and sufficient conditions are given for a triangular array of numbers to be expectations of order statistics of some nonnegative random variable. Using well-known recurrence relations, the expectations of all order statistics of the largest sample size, n, in the triangular array, or the expectations of the smallest of every sample size up to and including n are sufficient to determine the whole array. The former are reduced to a Stieltjes moment problem, the latter to a Hausdorff moment problem. These results are applied to show that for every sample size, there is a positive random variable with geometrically increasing expectations of order statistics with arbitrary ratio and expectation of smallest order statistic. However, only the degenerate distributions have geometrically increasing expectations of order statistics for more than one sample size, even when the ratio and mean of the smallest order statistic can depend on the sample size. These results were required for a study of participation in discussion groups.

1. Introduction. Consider a triangular array of nonnegative numbers

$$
\begin{array}{ccccc}
 & & a_{1,1} & & \\
 & a_{1,2} & & a_{2,2} & \\
 & a_{1,3} & a_{2\ 3} & a_{3,3} & \\
\end{array}
$$

(1)

$$
a_{1,n} \qquad \cdots \qquad \cdots \qquad \cdots \qquad a_{n,n}
$$

Could such an array be expectations of order statistics from some positive random variable? That is, is there a distribution function F, with $F(0-) = 0$, such that if $X_{1,k} \leqq X_{2,k} \leqq \cdots \leqq X_{k,k}$ are order statistics of a sample of size k from a population with distribution function F, then

(2) $$E(X_{i,k}) = a_{i,k} \qquad\qquad 1 \leqq i \leqq k \leqq n \quad ?$$

Section 2 reviews the linear recurrence relations known to apply between expectations and distribution functions of order statistics. Theorems 1 and 2, in Section 3, give necessary and sufficient conditions for (2) to hold. In Section 4, Theorem 2 is applied to show the existence of a nonnegative distribution with geometrically increasing expectations of order statistics at a sample size n, where the distribution can depend on the ratio and n (Theorem 3). Finally Theorem 4 shows that only degenerate distributions have geometrically increasing expectations of order statistics for more than one sample size. Theorem 5, due to Kemperman, gives an elegant inequality which proves Theorem 4.

Mallows [4], considers whether the triangular array (1) could be expectations of order statistics on $[\alpha, \beta]$, $-\infty \leqq \alpha < \beta \leqq \infty$. His conditions are different from those here, however.

2. Recurrence relations. Linear recurrence relations have been established for distribution functions, and hence integrals, of order statistics for arbitrary exchangeable random variables, discrete or continuous. The following are two ways of expressing them (for means):

$$(3) \qquad a_{i,k} = i\binom{k}{i} \sum_{j=0}^{i-1} (-1)^j \binom{i-1}{j} \frac{1}{k-i+j+1} a_{1,k-i+j+1} \qquad 1 \leqq i \leqq k \leqq n$$

$$(4) \qquad\qquad a_{i,k} = \sum_{j=0}^{n-k} \frac{\binom{n-j-i}{k-i}\binom{i+j-1}{i-1}}{\binom{n}{k}} a_{i+j,n} \qquad 1 \leqq i \leqq k \leqq n.$$

Formula (3) expresses an arbitrary element of the array (1) as a linear combination of expectations of smallest order statistics from various smaller sample sizes (See Young [8]). Thus the array (1) is a vector space of dimension at most n; $\{a_{1,k}, 1 \leqq k \leqq n\}$ spans this vector space, as (3) shows. Furthermore, if there is a distribution function F for which $\{a_{1,k}, 1 \leqq k \leqq n\}$ are expectations of smallest order statistics, and the array (1) satisfies (3), then (1) represents expectations of order statistics from F.

Formula (4) expresses an arbitrary element of the array (1) as a linear combination of expectations of order statistics of sample size n (see McCool [5] and Sillitto [7]). Thus $\{a_{k,n}, 1 \leqq k \leqq n\}$ also spans the vector space of array (1), as shown by (4). Again, if there is a distribution function F for which $\{a_{k,n}, 1 \leqq k \leqq n\}$ are expectations of order statistics of sample size n, and if the array (1) satisfies (4), then (1) represents expectations of order statistics from F.

3. The reduced problem. The results cited above allow reduction of the search for necessary and sufficient conditions to the following two questions:

(i) What sets of numbers $\{a_{1,k}, 1 \leqq$ like $k \leqq n\}$ can be expectations of smallest order statistics of various sample sizes from some distribution F?

(ii) What sets of numbers $\{a_{k,n}, 1 \leqq k \leqq n\}$ can be expectations of order statistics of the sample size n from some distribution F?

If necessary and sufficient conditions for (i) [or (ii)] can be found, then those conditions and (3) [or (4)] give necessary' and sufficient conditions for the array (1) to be expectations of order statistics from some distribution.

For the remainder of this section, the possible random variables are restricted to be nonnegative; that is, $F(0-) = 0$. Some element $a_{i,k}$ of (1) equals zero if and only if $X_{i,k}$ is zero with probability one, which occurs if and only if $F(0) = 1$. Therefore without loss of generality, take all $a_{i,k}$'s to be positive and assume $F(0) < 1$.

To begin question (i), consider $F_{1,k}(x)$, the probability that the smallest of k is less than or equal to x. This happens except when all k are larger than x. That is, $F_{1,k}(x) = 1 - (1 - F(x))^k$. Hence.

$$(5) \qquad a_{1,k} = \int_0^\infty (1 - F(x))^k \, dx.$$

The form of (5) is reminiscent of a moment problem, except that the unknown function is involved in the power, and the measure is fixed. Thus a change of variable is suggested. Proceeding formally, let $y = 1 - F(x)$, so $x = F^{-1}(1-y)$. Then

$$(6) \qquad a_{1,k} = \int_0^1 y^k \, d\{-F^{-1}(1-y)\}.$$

If F is monotone increasing, $F^{-1}(1-y)$ is well defined on $(0, 1]$. More generally, let $T(y) = -\inf_{x \geq 0}\{x \mid F(x) \geq 1-y\}$. T is monotone non-decreasing and right continuous. Also

$$(7) \qquad a_{1,k} = \int_0^\infty (1 - F(x))^k \, dx = \int_{0+}^1 y^k \, dT(y), \qquad 1 \leq k \leq n.$$

Notice that the mapping from possible F's satisfying (5) to possible measures dT satisfying (7) and $T(1) = 0$ is $1-1$ and onto. Therefore there is a dT satisfying (7) and $T(1) = 0$ if and only if there is an F satisfying (5).

Formula (7) is in the form of the classical Hausdorff moment problem except that dT need not be a probability measure. Therefore define

$$d\mathscr{B}(y) = \frac{y \, dT(y)}{a_{1,1}} \qquad \text{on} \quad (0, 1].$$

Now $d\mathscr{B}$ is a probability measure satisfying

$$(8) \qquad \frac{a_{1,k}}{a_{1,1}} = \int_{0+}^1 y^{k-1} \, d\mathscr{B}(y) \qquad 2 \leq k \leq n.$$

Again notice that the mapping form dT satisfying (7) to $d\mathscr{B}$ satisfying (8) is $1-1$ and onto. This proves

THEOREM 1. *A necessary and sufficient condition for* (1) *to represent expected values of order statistics from some nonnegative distribution is that the array* (1) *satisfy* (3) *and that* $m_k = a_{1,k+1}/a_{1,1}$ *be* kth *moment* $(1 \leq k \leq n-1)$ *of a probability distribution on* (0, 1].

In the above treatment, there is no reason why n cannot be taken to be infinity.

From some theorems in Krein [3] (see also Mallows [4] and Karlin and Studden [2] page 106 ff) the following can be derived:

Let $\mu_0 = 1, \mu_1, \mu_2, \cdots$ be a sequence of numbers and consider the following four determinants:

$$\Delta_{2K} = |\mu_{i+j}| \qquad\qquad i, j = 0, \cdots, K \quad K = 0, 1, \cdots$$

$$\Delta_{2K+1} = |\mu_{i+j+1}| \qquad\qquad i, j = 0, \cdots, K \quad K = 0, 1, \cdots$$

$$\Gamma_{2K} = |\mu_{i+j-1} - \mu_{i+j}| \qquad\qquad i, j = 1, \cdots, K \quad K = 1, 2, \cdots$$

$$\Gamma_{2K+1} = |\mu_{i+j} - \mu_{i+j+1}| \qquad\qquad i, j = 0, 1, \cdots, K \quad K = 0, 1, \cdots.$$

Then a necessary and sufficient condition for the existence of a measure $d\mathscr{B}$ satisfying

(*) $\mu_i = \int_{0^+}^1 t^i \, d\mathscr{B}(t)$ $i = 1, \cdots, n < \infty$

is that one of the following holds:

(a) $\Gamma_0, \Delta_0, \Gamma_1, \cdots, \Delta_{k-1}, \Gamma_k > 0$, $\Delta_k = \Gamma_{k+1} = \cdots = 0$ and k is odd iff there is a $d\mathscr{B}$ satisfying (*) and having exactly $(k+1)/2$ points of rise, one of which is one. In this case, there is only one $d\mathscr{B}$ satisfying (*).

(b) $\Delta_0, \Gamma_0, \cdots, \Gamma_{k-1}, \Delta_k > 0$, $\Gamma_k = \Delta_{k+1} = \cdots = 0$ and k is even iff there is a $d\mathscr{B}$ satisfying (*) and having exactly $k/2$ points of rise, none of which is one. In this case also there is only one $d\mathscr{B}$ satisfying (*).

(c) $\Delta_0, \Gamma_0, \cdots, \Delta_n, \Gamma_n > 0$. Then there are many $d\mathscr{B}$'s satisfying (*).

To begin examination of question (ii) above, consider $F_{k,n}(x)$, the probability that k or more of the nX's are less than or equal to x. Then

$$F_{k,n}(x) = \sum_{i=k}^n \binom{n}{i} F^i(x)(1 - F(x))^{n-i} = 1 - \sum_{i=0}^{k-1} \binom{n}{i} F^i(x)(1 - F(x))^{n-i}.$$

Using the same integration by parts,

$$a_{k,n} = \int_0^\infty \sum_{i=0}^{k-1} \binom{n}{i} F^i(x)(1 - F(x))^{n-i} \, dx,$$
$$= \sum_{i=0}^{k-1} \binom{n}{i} \int_{0^+}^1 (1 - y)^i y^{n-i} \, dT(y) \qquad\qquad 1 \leqq k \leqq n,$$

using the argument preceding (7).

Then

(9) $a_{k,n} - a_{k-1,n} = \binom{n}{k-1} \int_{0^+}^1 (1 - y)^{k-1} y^{n-k+1} \, dT(y)$, $2 \leqq k \leqq n$
$$a_{1,n} = \int_{0^+}^1 y^n \, dT(y).$$

Rewriting (9),

$$\frac{a_{k,n} - a_{k-1,n}}{\binom{n}{k-1}} = \int_{0^+}^1 \left(\frac{1-y}{y}\right)^{k-1} y^n \, dT(y) \qquad\qquad 2 \leqq k \leqq n.$$

Let $d\gamma(y) = y^n dT(y)/a_{1,n}$.

Then

$$\frac{a_{k,n} - a_{k-1,n}}{a_{1,n}\binom{n}{k-1}} = \int_{0^+}^1 \left(\frac{1-y}{y}\right)^{k-1} d\gamma(y) \qquad\qquad 2 \leqq k \leqq n$$

and $d\gamma(y)$ is a probability measure. Finally let $z = ((1 - y)/y)$. Then

$$\frac{a_{k,n} - a_{k-1,n}}{a_{1,n}\binom{n}{k-1}} = \int_0^\infty z^{k-1} \, d\mathscr{B}(z) \qquad\qquad 2 \leqq k \leqq n$$

and $d\mathscr{B}(z)$ is a probability measure. This proves

THEOREM 2. *A necessary and sufficient condition for* (1) *to represent expected values of order statistics from some nonnegative random variable is that the array* (1) *satisfy* (4) *and that*

$$m_k = \frac{a_{k+1,n} - a_{k,n}}{a_{1,n}\binom{n}{k}}$$

be kth moments $(1 \leq k \leq n-1)$ *of a probability distribution on* $[0, \infty)$.

Recalling the definition of Δ_{2K} and Δ_{2K+1}, the following can be derived from Krein [3] or Shohat and Tamarkin ([6] page 6):

A necessary and sufficient condition for the existence of a measure $d\Psi$ satisfying

$$(*) \qquad\qquad \mu_i = \int_0^\infty t^i d\Psi(t) \qquad\qquad i = 0, 1, \cdots, n < \infty$$

is that, for some k, $0 \leq k \leq n$,

$$\Delta_0 > 0, \quad \Delta_1 > 0, \cdots, \Delta_k > 0, \quad \Delta_{k+1} = \cdots = \Delta_n = 0.$$

The interpretation of k is as follows:

$k < n$ is odd iff there is a measure $d\Psi$ satisfying (*) and having exactly $(k+1)/2$ points of rise, none of which is zero. In this case $d\Psi$ is the only measure satisfying (*).

$k < n$ is even iff there is a measure $d\Psi$ satisfying (*) and having exactly $(k+2)/2$ points of rise, one of which is zero. In this case $d\Psi$ is the only measure satisfying (*).

If $k = n$, there are many measures satisfying (*).

4. An application. In a study of participation rates in small groups, Kadane and Lewis [1] encountered the following problems

(i) For what values of n, f, and s are there nonnegative distributions such that $a_{k,n} = fs^{k-1}$ $1 \leq k \leq n$?

(ii) For what values of f_n and s_n are there distributions such that $a_{k,n} = f_n s_n^{k-1}$ for all n and $1 \leq k \leq n$? The first question is in the form of Theorem 2. It makes sense only for $f > 0$ and $s \geq 1$. When $s = 1$, the distribution is degenerate, placing all its mass at f. Thus the only interesting case is $s > 1$.

To apply Theorem 2, consider

$$m_k = \frac{fs^k - fs^{k-1}}{f\binom{n}{k}} = s^k \left(\frac{s-1}{s\binom{n}{k}}\right) \qquad\qquad 1 \leq k \leq n-1.$$

Does there exist a probability distribution $d\mathscr{B}(z)$ such that

$$\int_0^\infty z^i d\mathscr{B}(z) = s^i \left(\frac{s-1}{s}\right)\frac{1}{\binom{n}{i}} \qquad\qquad 1 \leq i \leq n-1?$$

Let $y = z/s$. Then we wish to find a distribution γ satisfying $\int_0^\infty y^i d\gamma(y)$

$= (s-1/s) \, 1/\binom{n}{i}$, or a measure σ satisfying $\int_0^\infty y^i \, d\sigma(y) = 1/\binom{n}{i} \, 1 \leq i \leq n-1$ and $\int_0^\infty d\sigma(y) = s/(s-1) = 1 + 1/(s-1)$.

Consider the measure

$$d\mu(y) = \frac{(n+1) \, dy}{(1+y)^{n+2}}.$$

As is well known, $\int d\mu(y) = 1$ and $\int y^i \, d\mu(y) = 1/\binom{n}{i}$. Adding a jump of size $1/(s-1)$ at zero does not change any of the moments, but does increase the total measure to $1 + 1/(s-1) = s/(s-1)$, as desired. This proves

THEOREM 3. *For every* $f \geq 0$, $s \geq 1$ *and* $n \geq 1$ *there is a nonnegative distribution such that* $a_{k,n} = fs^{k-1} \quad 1 \leq k \leq n$.

The second question above is answered in a strong way by the following theorem:

Let X be a nonnegative nondegenerate random variable, and let $X_{1k} \leq X_{2k} \leq \cdots \leq X_{kk}$ be the order statistics for X of order k. Assume that $a_{ik} = E(X_{ik}) < \infty$.

THEOREM 4. (a) *If for some* $n > 3$

(10) $a_{i,n} = fs^{i-1}$ $i = 1, \cdots, n$

and for some m *satisfying* $n-3 \geq m > 0$, $a_{i,n-m} = td^{i-1} \quad i = 1, \cdots, n-m$, *then the distribution F is degenerate.*

(b) *If* (10) *holds and F is nondegenerate,*

(11) $a_{i-1,k} \, a_{i+1,k} < a_{i,k}^2 \qquad 3 \leq k \leq n-1 \text{ and } i = 2, \cdots, k-1$.

REMARK. Let $A_{f,s,n}$ be the set of all nonnegative distributions satisfying (10). If $n \geq 2$, $s = 1$ iff the distribution is degenerate with all its mass at f. For $f \geq 0$, $s \geq 1$ and $n \geq 1$, Theorem 3 shows that $A_{f,s,n}$ is non-empty. For $f > 0$, $s > 1$ and $n \geq 3$, Theorem 4 shows that any two distinct A's are disjoint.

PROOF OF THEOREM 4. Since (a) is implied by (b), only (b) need be proved. My somewhat cumbersome proof of (b) can be replaced by the following result of Kemperman, which he has kindly allowed me to include.

THEOREM 5 (*Kemperman*). (11) *is implied by*

(12) $a_{i-1,n} a_{i+1,n} \leq a_{i,n}^2$ $(i = 2, \cdots, n-1)$.

PROOF. It suffices to show that (12) implies

(13) $a_{i-1,n-1} \, a_{i+1,n-1} < a_{i,n-1}^2$ $(i = 2, \cdots, n-2)$.

Removing a random member from a sample of size n one obtains a sample of size $n-1$, so that $X_{i,n-1}$ equals $X_{i,n}$ with probability $1 - i/n$, $X_{i+1,n}$ with probability i/n. In this way we have the recurrence relation.

$$a_{i,n-1} = (1 - i/n) a_{i,n} + (i/n) a_{i+1,n} \qquad (i-1, \cdots, n-1)$$

which is a special case of (4).

Using that $a_{i-1,n} \leqq a_{i,n}^2/a_{i+1,n}$ and $a_{i+2,n} \leqq a_{i+1,n}^2/a_{i,n}$, we obtain that

$$a_{i-1,n-1} a_{i+1,n-1} \leqq \left[\left(1 - \frac{i-1}{n}\right) a_{i,n}^2/a_{i+1,n} + \frac{i-1}{n} a_{i,n}\right]$$

$$\cdot \left[\left(1 - \frac{i+1}{n}\right) a_{i+1,n} + \frac{i+1}{n} a_{i+1,n}^2/a_{i,n}\right]$$

$$= \left[\left(1 - \frac{i-1}{n}\right) a_{i,n} + \frac{i-1}{n} a_{i+1,n}\right] \left[\left(1 - \frac{i+1}{n}\right) a_{i,n} + \frac{i+1}{n} a_{i+1,n}\right]$$

$$< [a_{i,n-1}]^2,$$

since $xy < ((x+y)/2)^2$ for $x \neq y$. $\square$

Professor Kemperman remarks that a sufficient condition for (12) is that

$$a_{j+1,n} = \frac{1}{j!} \int_0^\infty u^j e^{-g(u)} du \qquad\qquad j = 0, 1, \cdots$$

where $g(u)$ is convex. A boundary case is when $g(u)$ is linear, which happens when (10) holds.

Acknowledgment. The following gave helpful comments on aspects of this paper: Donald Berry, Joseph Bram, Paul Feder, David Grether, J. H. B. Kemperman, Gordon Lewis, Colin Mallows, Barry Margolin, Richard Olshen, John Ramage, Daniel Relles and I. R. Savage.

I am especially grateful to Professor Kemperman for his permission to report iiis Theorem 5.

REFERENCES

[1] KADANE, J. B. and LEWIS, G. (1969). The distribution of participation in group discussions: an empirical and theoretical reappraisal. *Amer. Sociological Rev.* **34** 710–723.
[2] KARLIN, S. J. and STUDDEN, W. J. (1966). *Tchebycheff Systems: with Applications in Analysis and Statistics.* Wiley, New York.
[3] KREIN, M. G. (1951). The ideas of P. L. Chebysev and A. A. Markov in the theory of limiting values of integrals and their further developments. *Amer. Math. Soc. Transl. Ser.* 2 **12** 1–122.
[4] MALLOWS, C. L. (1969). Bounds on distribution functions in terms of expectations of order statistics. Unpublished.
[5] McCOOL, J. I. (1965). The construction of good linear unbiased estimates from the best linear estimates for a smaller sample size. *Technometrics* **7** 543–552.
[6] SHOHAT, J. A. and TAMARKIN, J. D. (1943). *The Problem of Moments, Mathematical Surveys,* No. 1. *Amer. Math. Soc.,* New York.
[7] SILLITTO, G. P. (1964). Some relations between expectations of order statistics in samples of different sizes. *Biometrika* **51** 259–262.
[8] YOUNG, D. H. (1967). Recurrence relations between the p.d.f's of order statistics of dependent variables, and some applications. *Biometrika* **54** 283–292.