Ann. Inst. Statist. Math.
36 (1984), Part A, 451-453

A NOTE ON EQUAL DISTRIBUTIONS

GWO DONG LIN

(Received May 18, 1983; revised Sept. 13, 1983)

Summary

It is known that the set

$$\{E(X_{k_n,n}) \mid n=1,2,\cdots\}, \qquad \text{where } 1 \leqq k_n \leqq n,$$

of expectations of order statistics of samples from a distribution F which has a finite expectation determines F. In this note, we show that each of the sets

$$\{E(X_{k_j,n_j}) \mid j=1,2,\cdots\}, \quad \text{where } \{(k_j/n_j) \mid j=1,2,\cdots\} \text{ is dense in } [0,1],$$

and

$$\{E(X_{1,1})\} \cup \{E(X_{k_j,2j+1}) \mid j=1,2,\cdots\} \cup \{E(X_{k'_j,2j+1}) \mid j=1,2,\cdots\}, \quad \text{where } 1 \leqq k_j < k'_j \leqq 2j+1,$$

also determines F.

The object of this note is to give two characterizations of distributions of random variables whose expectations exist and are finite.

Let $X_{1,n} \leqq X_{2,n} \leqq \cdots \leqq X_{n,n}$ be the order statistics of n i.i.d. random variables $X_1, X_2, \cdots, X_n$ obeying an arbitrary distribution F with finite expectation $E(X)$. Hoeffding [2] showed that the set

$$(1) \qquad \{E(X_{k,n}) \mid 1 \leqq k \leqq n,\ n=1,2,\cdots\}$$

determines F. His proof was based on the fact that

$$(2) \qquad \lim_{n\to\infty} F_n(x) = F(x), \qquad \text{for each } x \text{ at which } F \text{ is continuous},$$

where F_n is the distribution having mass function

$$f(E(X_{k,n})) = 1/n, \qquad k=1,2,\cdots,n.$$

AMS 1980 subject classifications: Primary 62E10; Secondary 30B60.
Key words and phrases: Distribution, determine, dense, completeness and order statistics.

In fact, for any subsequence $\{n_1, n_2, \cdots, n_j, \cdots\}$ of $\{1, 2, \cdots\}$ we also have

$$\lim_{j\to\infty} F_{n_j}(x)=F(x)\ , \qquad \text{for each } x \text{ at which } F \text{ is continuous}\ ,$$

and hence the set

$$(3)\qquad \{\mathrm{E}\,(X_{k,n_j})\,|\,k=1, 2,\cdots, n_j,\ j=1, 2,\cdots\}\ , \qquad \text{where } \lim_{j\to\infty} n_j=\infty\ ,$$

also determines F. This result was obtained by Hwang [4] by a complete polynomials in summable space $L_1(0, 1)$, and by Arnold and Meeden [1] by a combinatorial approach. Pollak [5] improved (1) based on Hoeffding's paper and got that the set

$$(4)\qquad \{\mathrm{E}\,(X_{k_n,n})\,|\,n=1, 2,\cdots\}\ , \qquad \text{where } 1\leqq k_n\leqq n\ ,$$

also determines F. He first proved that the set (4) determines the set (1), and then proved that the set (1) determines the distribution F. In fact, his result can be seen easily by the triangular equality (6). And a slight modification of the proof of his second lemma yields the following result which can also be obtained from a lemma of Hoeffding ([2], Lemma 5).

THEOREM 1. *Let X, Y be two random variables whose expectations exist and are finite, and let F, G denote their respective distributions. Then $F=G$ if and only if there exists a set $\{(k_j, n_j)\,|\,j=1, 2,\cdots\}$, where $1\leqq k_j\leqq n_j$, such that $\{(k_j/n_j)\,|\,j=1, 2,\cdots\}$ is dense in* $[0, 1]$ *and*

$$\mathrm{E}\,(X_{k_j,n_j})=\mathrm{E}\,(Y_{k_j,n_j})\ , \qquad j=1, 2,\cdots\ .$$

PROOF. Take $\rho_n=k_j$ and $n=n_j$, respectively, in the proof of Pollak's second lemma [5].

Notice that the result (3) is a corollary of Theorem 1. Now, let us define

$$F^{-1}(t)=\inf\{x\,|\,F(x)\geqq t\}\ , \qquad t\in(0, 1)\ ,$$

then it is well-known that (see, for example, Huang and Hwang [3])

$$\mathrm{E}\,(X_{k,n})=k\binom{n}{k}\int_0^1 F^{-1}(t)t^{k-1}(1-t)^{n-k}dt\ ,$$
$$k=1, 2,\cdots, n,\ n=1, 2,\cdots\ .$$

From the identity

$$(5)\qquad x^{k-1}(1-x)^{n-k}+x^k(1-x)^{n-k-1}=x^{k-1}(1-x)^{n-k-1}\ ,$$

we have

$$(6)\qquad (n-k)\,\mathrm{E}(X_{k,n})+k\,\mathrm{E}(X_{k+1,n})=n\,\mathrm{E}(X_{k,n-1})\,.$$

The above equality says that within certain triples of expected values of order statistics, knowledge of any two determines the third. Eliminating the term $x^{k-1}(1-x)^{n-k-1}$ from both sides of (5), we have

$$(1-x)+x=1\,,$$

that is, 1 is a linear combination of $1-x$ and x. In the same spirit, we can see that 1 is a linear combination of $x, x^2, \cdots, x^{n-2}$, $x^{k-1}(1-x)^{n-k}$ and $x^{k'-1}(1-x)^{n-k'}$, where $1\leqq k<k'\leqq n$, and hence knowledge of $\mathrm{E}(X_{1,1})$, $\mathrm{E}(X_{2,2}), \cdots, \mathrm{E}(X_{n-2,n-2})$, $\mathrm{E}(X_{k,n})$ and $\mathrm{E}(X_{k',n})$ also determines $\mathrm{E}(X_{n-1,n-1})$. This is the motivation to get the following

THEOREM 2. *Let X, Y be two random variables whose expectations exist and are finite, and let F, G denote their respective distributions. Then $F=G$ if and only if there exists a set*

$$\{(k_j, 2j+1)\,|\,j=1, 2, \cdots\}\cup\{(k'_j, 2j+1)\,|\,j=1, 2, \cdots\}\,,\quad \textit{where } 1\leqq k_j<k'_j\leqq 2j+1\,,$$

such that $\mathrm{E}(X)=\mathrm{E}(Y)$ *and*

$$\mathrm{E}(X_{k_j,2j+1})=\mathrm{E}(Y_{k_j,2j+1})\,,\qquad \mathrm{E}(X_{k'_j,2j+1})=\mathrm{E}(Y_{k'_j,2j+1})\,,\qquad j=1, 2, \cdots\,.$$

PROOF. Necessity is clear while sufficiency follows inductively by the above discussion and the triangular equality (6).

Acknowledgement

The author is grateful to the referee for the simplification of the proof of Theorem 2.

INSTITUTE OF STATISTICS, ACADEMIA SINICA, TAIWAN

REFERENCES

[1] Arnold, B. C. and Meeden, G. (1975). Characterization of distributions by sets of moments of order statistics, *Ann. Statist.*, **3**, 754-758.

[2] Hoeffding, W. (1953). On the distribution of the expected values of the order statistics, *Ann. Math. Statist.*, **24**, 93-100.

[3] Huang, J. S. and Hwang, J. S. (1975). L_1-completeness of a class of Beta distributions, *Statistical Distributions in Scientific Work*, **3**, 137-141.

[4] Hwang, J. S. (1978). A note on Bernstein and Müntz-Szász Theorems with applications to the order statistics, *Ann. Inst. Statist. Math.*, **30**, A, 167-176.

[5] Pollak, M. (1973). On equal distributions, *Ann. Statist.*, **1**, 180-182.