SOME BOUNDS FOR EXPECTED VALUES OF ORDER STATISTICS

BY MIR M. ALI AND LAI K. CHAN

University of Western Ontario

1. Summary. Let the function $F(x)$ be a distribution function for a continuous symmetric distribution, and let $X_{(i)}$ represent the ith order statistics from a sample of size n. It is shown in this paper that for $i \geqq (n+1)/2$

$$E(X_{(i)}) \geqq G(i/(n+1)) \quad \text{if} \quad F \quad \text{is } \textit{unimodal}$$

and

$$E(X_{(i)}) \leqq G(i/(n+1)) \quad \text{if} \quad F \quad \text{is } \textit{U-shaped},$$

where $x = G(u)$ is the inverse function of $F(x) = u$. The definitions of unimodal and U-shaped distributions are given in Section 3.

The above inequalities are of interest, since it is known (Blom (1958), Chapters 5 and 6) that for sufficiently large n the bound $G(i/(n+1))$ approaches $E(X_{(i)})$.

2. Introduction. Studies of bounds for $E(X_{(i)})$ in terms of i and n have appeared in the literature, for instance, Plackett (1947), Moriguti (1951, 1953) and Hartley and David (1954). Blom (1958) has remarked that if $G(u)$ is convex and continuous then by Jensen's inequality $E(X_{(i)}) \geqq G(i/(n+1))$. An example is provided by the negative exponential distribution. If, however, $G(u)$ is concave and continuous then $E(X_{(i)}) \leqq G(i/(n+1))$. An example of this case is the distribution with probability density function $[(m+1)/a] \cdot [1+(x/a)]^m$, $0 \leqq m \leqq 1$, $-a \leqq x \leqq 0$. For the rectangular distribution, $G(u)$ is both convex and concave and we have $E(X_{(i)}) = G(i/(n+1))$.

3. The bounds.

DEFINITIONS. Following Gnedenko and Kolmogorov (1949, p. 157), F is defined to be *unimodal* if there exists at least one real number c such that $F(x)$ is convex for $x < c$ and concave for $x > c$. On the basis of this definition we generalize somewhat the concept of U-shaped distribution (Kendall and Stuart, 1958, p. 10). F is defined to be *U-shaped* if there exists at least one real number c such that $F(x)$ is concave for $x < c$ and convex for $x > c$.

The convexity and concavity properties in the above definitions are understood to be restricted to the range of the distribution which may be a finite, semi-finite or infinite interval. Furthermore, if $F(x)$ is strictly increasing over the range of the distribution, the definition of a unimodal [U-shaped] distribution is equivalent to the following: F is unimodal [U-shaped] if there exists at least one real number c such that G(u) is concave [convex] for $x < c$ and convex [concave] for $x > c$. This is because, for example, for any $x_1 < c$, $x_2 < c$ and $0 \leqq \alpha \leqq 1$,

Received 27 July 1964; revised 15 January 1965.

$\alpha F(x_1) + (1 - \alpha)F(x_2) \geqq F(\alpha x_1 + (1 - \alpha)x_2)$ if and only if $G(\alpha F(x_1) + (1 - \alpha)F(x_2)) \geqq \alpha x_1 + (1 - \alpha)x_2$.

THE RESULT. Let $F(x)$ be symmetric, continuous and strictly increasing. Then for $i \geqq (n + 1)/2$,

$$\text{(A)} \qquad E(X_{(i)}) \geqq G(i/(n + 1)) \qquad \text{if } F \text{ is unimodal.}$$

PROOF. Without loss of generality we let $G(\frac{1}{2}) = 0$. For $i = (n + 1)/2$, it is trivial that the equality in (A) holds. Hence we assume that $i > (n + 1)/2$. Let $h(u) = n\binom{n-1}{i-1}u^{i-1}(1 - u)^{n-i}$. Then we have

$$E(X_{(i)}) = \int_0^1 G(u)h(u)\,du.$$

Also let $C_i = \int_{\frac{1}{2}}^1 [h(u) - h(1 - u)]\,du$. By straightforward calculations it can be seen that $[h(u) - h(1 - u)] > 0$ for $\frac{1}{2} < u < 1$ and that $0 < C_i < 1$.

The conditions imposed on $F(x)$ imply that $G(u)$ is continuous and convex for $\frac{1}{2} \leqq u \leqq 1$. Hence by Jensen's inequality (Natanson, 1957, p. 46) we have

$$E(X_{(i)})/C_i = \int_{\frac{1}{2}}^1 G(u)\{[h(u) - h(1 - u)]/C_i\}\,du$$
$$\geqq G(\int_{\frac{1}{2}}^1 u\{[h(u) - h(1 - u)]/C_i\}\,du).$$

But $\int_{\frac{1}{2}}^1 u\{[h(u) - h(1 - u)]/C_i\}\,du = \frac{1}{2} + (1/C_i)\{[i/(n + 1)] - \frac{1}{2}\}$. Then we have

$$E(X_{(i)}) \geqq C_i G\{\tfrac{1}{2} + (1/C_i)[(i/(n + 1)) - \tfrac{1}{2}]\}$$
$$= C_i G\{\tfrac{1}{2} + (1/C_i)[(i/(n + 1)) - \tfrac{1}{2}]\} + (1 - C_i)G(\tfrac{1}{2}) \geqq G[i/(n + 1)]$$

since $G(u)$ is convex for $\frac{1}{2} \leqq u \leqq 1$ and $G(\frac{1}{2}) = 0$.

Similar considerations show that for $i \geqq (n + 1)/2$

$$\text{(A}'\text{)} \qquad E(X_{(i)}) \leqq G(i/(n + 1)) \qquad \text{if } F \text{ is } U\text{-shaped.}$$

By symmetry we of course have, for $i < (n + 1)/2$, $E(X_{(i)}) \leqq G(i/(n + 1))$ if F is unimodal and $E(X_{(i)}) \geqq G(i/(n + 1))$ if F is U-shaped.

EXAMPLES. The normal, the logistic, the Student, the Laplace and the Cauchy distributions satisfy (A). For the distribution with probability density function

$$\{\Gamma(m + \tfrac{3}{2})/[a\Gamma(\tfrac{1}{2})\Gamma(m + 1)]\}(1 - (x^2/a^2))^m, \qquad -a \leqq x \leqq a,$$

(A) is satisfied if $m \geqq 0$ while (A$'$) is satisfied when $-1 < m \leqq 0$. When $m = 0$, both (A) and (A$'$) must be satisfied so that $E(X_{(i)}) = G(i/(n + 1))$. Actually in this case the distribution is the uniform distribution.

4. Acknowledgment. The authors express their indebtedness to the referee for suggesting a shorter proof of the inequality (A) and for helpful comments.

REFERENCES

BLOM, G. (1958). *Statistical Estimates and Transformed Beta-Variables.* Wiley, New York.
GNEDENKO, B. V. and KOLMOGOROV, A. N. (1949). *Limit Distributions For Sums of Inde-*

pendent Random Variables. English translation (1954), Addison-Wesley, Cambridge, Mass.

HARTLEY, H. O. and DAVID, H. A. (1954). Universal bounds for mean range and extreme observation. *Ann. Math. Statist.* **25** 85–99.

KENDALL, M. G. and STUART, A. (1958). *The Advanced Theory of Statistics*, **1**. Griffin, London.

MORIGUTI, S. (1951). Extremal properties of extreme value distributions. *Ann. Math. Statist.* **22** 523–536.

MORIGUTI, S. (1953). A modification of Schwarz's inequality with applications to distributions. *Ann. Math. Statist.* **24** 107–113.

NATANSON, I. P. (1957). *Theory of Functions of a Real Variable*, **2**. English translation (1960), Frederick Ungar, New York.

PLACKETT, R. A. (1947). Limits of the ratio of mean range to standard deviation. *Biometrika* **34** 120–122.