AN INEQUALITY FOR EXPECTED VALUES OF SAMPLE QUANTILES [1)]

By W.R. VAN ZWET

University of Leiden

1. INTRODUCTION

Let F be a continuous distribution function on R^1 that is strictly increasing on the (finite or infinite) open interval I where $0 < F < 1$, and let G denote the inverse of F. For $n = 1, 2, \ldots$ and $0 < \lambda < 1$, let

$$(1.1) \quad \gamma_n(\lambda) = \frac{\Gamma(n+1)}{\Gamma(\lambda(n+1))\Gamma((1-\lambda)(n+1))} \int_0^1 G(y)y^{\lambda(n+1)-1} \cdot (1-y)^{(1-\lambda)(n+1)-1}dy.$$

Obviously, if $X_{i:n}$ denotes the i-th order statistic of a sample of size n from the parent distribution F, then

$$\gamma_n(\frac{i}{n+1}) = E\, X_{i:n}\ , \quad i = 1,2,\ldots,n.$$

We shall call $\gamma_n(\lambda)$ the expected value of the λ-quantile of a sample of size n from F, even though this interpretation is meaningless when $\lambda(n+1)$ is not an integer.

We shall assume that for some λ the integral converges for sufficiently large n, which ensures that the same will hold for every $0 < \lambda < 1$. By making minor changes in W. HOEFFDING's proof in [2], one shows that γ_n converges to G on (0,1) for $n \to \infty$.

---

1) Report S 369, Mathematisch Centrum, Amsterdam.

Consider another continuous distribution function F^* that is strictly increasing on the interval I^* where $0 < F^* < 1$, and let G^*, γ_n^* and $X_{i:n}^*$ be defined for F^* analogous to G, γ_n and $X_{i:n}$ for F. Furthermore let

(1.2) $$\phi(x) = G^*F(x) \ , \quad x \in I.$$

In [5] the author studied the following order relations between F and F^*:

(1.3) ϕ is convex on I;

(1.4) F and F^* represent symmetric distributions and ϕ is concave-convex on I.

If x_0 denotes the median of F, relation (1.4) implies that ϕ is antisymmetric about x_0 (i.e. $\phi(x_0+x) + \phi(x_0-x) = 2\phi(x_0)$) and hence that ϕ is concave for $x < x_0$ and convex for $x > x_0$.

Let ϕ_n be the function that maps the expected value of the λ-quantiles of a sample of size n from F on the corresponding quantities for F^*:

(1.5) $$\phi_n(x) = \gamma_n^* \gamma_n^{-1}(x).$$

For $n \to \infty$, ϕ_n will converge to the function ϕ on I that maps the population quantiles of F on those of F^*. It is shown in this note that if relations (1.3) or (1.4) hold, ϕ_n shares the convexity or concave-convexity of ϕ, and the convergence of ϕ_n to ϕ is monotone. The convexity property yields a theorem on the behavior of the ratio of expected values of spacings of consecutive order statistics from F and F^*. Simple applications are given in section 3.

2. THE RESULTS

THEOREM 2.1

If condition (1.3) holds, $\phi_n(x)$ is convex in x for fixed n, and non-increasing in n for fixed x.

PROOF

For each fixed n the densities

$$(2.1)\quad f_\lambda(y)=\frac{\Gamma(n+1)}{\Gamma(\lambda(n+1))\Gamma((1-\lambda)(n+1))}\,y^{\lambda(n+1)-1}(1-y)^{(1-\lambda)(n+1)-1}$$

constitute a one-parameter exponential family for $0 < \lambda, y < 1$, and consequently the family is strictly totally positive of order ∞ in λ and y (cf. [3]). According to a slight elaboration of a result due to S. KARLIN that is given in [4], the convexity of ϕ_n follows from the definition of γ_n and γ_n^*, the total positivity of $f_\lambda(y)$, the monotonicity of F and the convexity of ϕ. Also

$$(2.2)\quad \gamma_n(\lambda) = \lambda\gamma_{n+1}(\lambda + \frac{1-\lambda}{n+2}) + (1-\lambda)\gamma_{n+1}(\lambda - \frac{\lambda}{n+2})$$

and the same holds for γ_n^*. This is easily verified by adding integrands in expression (1.1). Hence, because of the convexity of ϕ_{n+1},

$$\phi_{n+1}\gamma_n(\lambda) = \phi_{n+1}(\lambda\gamma_{n+1}(\lambda + \frac{1-\lambda}{n+2})+(1-\lambda)\gamma_{n+1}(\lambda - \frac{\lambda}{n+2})) \leqq$$

$$(2.3)\quad \leqq \lambda\phi_{n+1}\gamma_{n+1}(\lambda + \frac{1-\lambda}{n+2})+(1-\lambda)\phi_{n+1}\gamma_{n+1}(\lambda - \frac{\lambda}{n+2}) =$$

$$= \lambda\gamma_{n+1}^*(\lambda + \frac{1-\lambda}{n+2})+(1-\lambda)\gamma_{n+1}^*(\lambda - \frac{\lambda}{n+2}) = \gamma_n^*(\lambda),$$

or, replacing $\gamma_n(\lambda)$ by x,

$$\phi_{n+1}(x) \leqq \gamma_n^*\,\gamma_n^{-1}(x) = \phi_n(x).$$

In the same vein we have

THEOREM 2.2

If condition (1.4) holds, $\phi_n(x)$ is antisymmetric concave-convex about x_0 for fixed n, and non-increasing in n for fixed $x > x_0$.

PROOF

Obviously ϕ_n is antisymmetric about x_0. Since ϕ is concave-convex, G^* is a concave-convex function of G and hence

$$h(y) = G^*(y) - a - bG(y)$$

can have at most three changes of sign on (0,1) for any a and b. If it does change sign three times, the signs occur in the order (-, +, -, +) for increasing values of the argument. It follows from the variation diminishing property of totally positive kernels (cf. [3]) that

$$\gamma_n^*(\lambda) - a - b\gamma_n(\lambda) = \int_0^1 h(y)f_\lambda(y)dy$$

changes sign at most three times; if it does have three sign changes, the signs occur in the order (-, +, -, +). Substituting $\gamma_n(\lambda) = x$ we find that

$$\phi_n(x) - a - bx$$

possesses the same property for any a and b. A simple geometrical argument based on the antisymmetry of ϕ_n shows that this implies that ϕ_n is concave-convex about x_0. Since for $\lambda > \frac{1}{2}$

$$(\lambda + \frac{1-\lambda}{n+2}) + (\lambda - \frac{\lambda}{n+2}) > 1,$$

and hence by the antisymmetry of γ_{n+1}

$$\gamma_{n+1}(\lambda + \frac{1-\lambda}{n+2}) + \gamma_{n+1}(\lambda - \frac{\lambda}{n+2}) > 2x_0$$

the inequality of (2.3) remains valid now that ϕ_n is antisymmetric and concave-convex instead of convex. This completes the proof.

We note that in the proofs of theorems 2.1 and 2.2 we have only made use of the total positivity of $f_\lambda(y)$. Exploiting the fact that the total positivity is strict one finds that the convexity (or concave-convexity) in x as well as the monotonicity in n of $\phi_n(x)$ are strict, unless ϕ is linear on I.

The quantities $\gamma_n(\lambda)$ for non-integer $\lambda(n+1)$ were introduced to facilitate the discussion of λ-quantiles for fixed λ and varying n. However, in considering the convexity of ϕ_n for fixed n, we may as well restrict ourselves to the case where $i = \lambda(n+1)$ is an integer. Theorem 2.1 then states that if condition (1.3) holds, i.e. if G^* is a convex function of G, then $EX^*_{i:n}$ is a convex function of $EX_{i:n}$ for varying i and fixed n, i.e.

$$\frac{EX^*_{i+1:n} - EX^*_{i:n}}{EX_{i+1:n} - EX_{i:n}} \tag{2.4}$$

is non-decreasing in i for fixed n. We recall that the proof of this assertion rests solely on the fact that the family (2.1), which for $i = \lambda(n+1)$ becomes

$$f_{i:n}(y) = \frac{n!}{(i-1)!(n-i)!}\, y^{i-1}(1-y)^{n-i}, \tag{2.5}$$

is totally positive of order infinity in i and y for fixed n. However, the family (2.5) is also totally positive of order infinity in n and (1-y) for fixed i. One easily verifies that this implies that $EX^*_{i:n}$ is also a convex function of $EX_{i:n}$ for varying n and fixed i. Since $EX_{i:n}$ is decreasing in n for fixed i, it follows that

$$\frac{EX^*_{i:n} - EX^*_{i:n+1}}{EX_{i:n} - EX_{i:n+1}}$$

is non-increasing in n. Using formula (2.2) for $\lambda(n+1) = i$, i.e.

$$EX_{i:n} = \frac{i}{n+1} EX_{i+1:n+1} + \frac{n+1-i}{n+1} EX_{i:n+1}, \tag{2.6}$$

and the corresponding expression for $EX^*_{i:n}$, we find

$$\frac{EX^*_{i:n} - EX^*_{i:n+1}}{EX_{i:n} - EX_{i:n+1}} = \frac{EX^*_{i+1:n+1} - EX^*_{i:n+1}}{EX_{i+1:n+1} - EX_{i:n+1}} ,$$

and hence (2.4) is non-increasing in n.

By considering the distribution functions $1 - F^*(-x)$ and $1 - F(-x)$ instead of F and F^* one easily shows that

$$\frac{EX^*_{n-i+1:n} - EX^*_{n-i:n}}{EX_{n-i+1:n} - EX_{n-i:n}} \tag{2.7}$$

is non-increasing in i and non-decreasing in n. The former conclusion is of course equivalent to the monotonicity in i of (2.4). We have proved

Theorem 2.3

If condition (1.3) holds, the quantities (2.4) are non-decreasing in i and non-increasing in n, whereas (2.7) is non-decreasing in n.

We note that the last assertion of the theorem may also be proved directly by using the total positivity of (2.5) in i and y for fixed (n-i) and applying (2.6).

It may be of interest to point out the similarity of theorem 2.3 to inequalities that were recently obtained by R.E. BARLOW and F. PROSCHAN [1] for the case where $F(0) = F^{*}(0) = 0$ and ϕ is starshaped (i.e. $\phi(x)/x$ non-decreasing on I). By total positivity arguments similar to those given above they show that

$$\frac{EX^{*}_{i:n}}{EX_{i:n}}$$

is non-decreasing in i and non-increasing in n, whereas

$$\frac{EX^{*}_{n-i:n}}{EX_{n-i:n}}$$

is non-decreasing in n.

3. APPLICATIONS

Let F be the uniform distribution function on (0,1). Then

$$\gamma_n(\lambda) = \lambda \qquad \text{for} \qquad 0 < \lambda < 1,$$

$\phi = G^{*}$ and $\phi_n = \gamma^{*}_n$. If F^{*} is differentiable on I^{*}, it satisfies conditions (1.3) or (1.4) if its density $F^{*\prime}$ is non-increasing on I^{*}, or symmetric and unimodal respectively. Consequently we have:

The expected value of the λ-quantile of a sample of size n from a distribution with non-increasing density is a non-increasing function of n; if the density is symmetric and unimodal the conclusion remains valid for $\lambda > \frac{1}{2}$. Moreover, if $F^{*\prime}$ is non-increasing, $(n+1)(EX^{*}_{i+1:n} - EX^{*}_{i:n})$ is non-decreasing in i and non-increasing in n, whereas $(n+1)(EX^{*}_{n-i+1:n} - EX^{*}_{n-i:n})$ is non-decreasing in n.

As a second example consider the case where F^{*} denotes the exponential distribution function. Then condition (1.3) is satisfied if the distribution F has increasing failure rate

$$q(x) = \frac{F'(x)}{1 - F(x)}$$

(cf. [1] or [5]). We have (cf. similar results in [1]): If F has increasing failure rate, then $(n-i)(EX_{i+1:n} - EX_{i:n})$ is non-increasing in i and non-decreasing in n, whereas $(EX_{n-i+1:n} - EX_{n-i:n})$ is non-increasing in n.

For other cases where relations (1.3) or (1.4) are satisfied and the results of this paper may be applied, the reader is referred to [5].

ACKNOWLEDGMENT

The author is endebted to Professor Richard E. Barlow for a stimulating discussion during which theorem 2.3 was put into shape.

REFERENCES

[1] BARLOW, R.E., and PROSCHAN, F. (1966). Inequalities for linear combinations of order statistics from restricted families. Ann. Math. Statist. 37, 1574-1592.

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

[3] KARLIN, S. (1963). Total positivity and convexity preserving transformations. Convexity, Proc. Symp. Pure Math. 7, 329-347, Amer. Math. Soc., Providence.

[4] MOLENAAR, W., and VAN ZWET, W.R. (1966). On mixtures of distributions. Ann. Math. Statist. 37, 281-283.

[5] VAN ZWET, W.R. (1964). Convex transformations of random variables. Mathematical Centre Tract 7, Mathematisch Centrum, Amsterdam.