ELSEVIER

Statistics & Probability Letters 28 (1996) 191–194

STATISTICS & PROBABILITY LETTERS

A generalization of variance bounds

N. Papadatos, V. Papathanasiou*

Department of Mathematics, University of Athens, 157 84 Panepistemiopolis Athens, Greece

Received December 1993; revised April 1995

Abstract

Upper and lower bounds for the variance of a random variable are obtained in terms of the density-quantile function. Some applications of these bounds to order statistics are given.

Keywords: Variance bounds; Density-quantile function; Order statistics

1. Introduction

Arnold and Brockett (1988) using Polya's inequality derived several variance bounds of interest in statistics. Their bounds include the density-quantile function $h_X(u) = f(F^{-1}(u))$, where f and F are, respectively, the density and the distribution function of some absolutely continuous random variable X. They also obtained a characterization of such functions.

Cacoullos and Papathanasiou (1985, 1989) obtained upper and lower bounds for the variance of a function of a random variable Y in terms of a w-function associated with its density g. This function is defined for every y in the interval support of g by the following relation:

$$\sigma^2 g(y) w(y) = \int_{-\infty}^{y} (\mu - t) g(t)\, dt, \tag{1.1}$$

provided that the mean μ and the variance σ^2 of Y exist. Papathanasiou (1990) gave similar upper bounds for the variance of order statistics.

In the present note, by using the results of Cacoullos and Papathanasiou, some generalization of Arnold and Brockett's bounds are given. The achievement of these bounds characterizes the corresponding distributions. Moreover, the asymptotic behaviour of the variance for the ith-order statistic is investigated.

* Corresponding author.

SSDI 0167-7152(95)00117-4

2. Variance bounds with applications to order statistics

Consider an absolutely continuous random variable Y having distribution function G, density g with support and interval (not necessarily finite) and finite variance. Let $w(y)$ be the w-function associated with the density g (see (1.1)). Let also X be an absolutely continuous random variable with distribution function F, density f and finite variance. The inequality of Cacoullos and Papathanasiou (1985) is stated as follows: For any absolutely continuous function h defined on the support of g,

$$\operatorname{Var}[h(Y)] \leqslant \operatorname{Var}[Y]E[w(Y)(h'(Y))^2], \tag{2.1}$$

with equality iff (if and only if) h is linear.

For the special choice $h = F^{-1} \circ G$ we have $X =_{\mathrm{d}} h(Y)$ ($X =_{\mathrm{d}} Y$ means: X and Y have the same distribution) and from (2.1) we conclude that the following inequality holds:

$$\operatorname{Var}[X] \leqslant \operatorname{Var}[Y] \int_0^1 w(G^{-1}(u))[h_Y(u)/h_X(u)]^2 \,\mathrm{d}u, \tag{2.2}$$

where $h_Y = g \circ G^{-1}$, $h_X = f \circ F^{-1}$ are the density-quantile functions of Y and X, respectively. Equality holds in (2.2) iff $X =_{\mathrm{d}} \alpha Y + b$, $\alpha > 0$.

Remark. Arnold and Brockett (1988) proved (Theorem 3.1) that if the random variable $0 \leqslant Y \leqslant 1$ has distribution function G, then for any differentiable function h defined on $(0, 1)$,

$$\operatorname{Var}[h(Y)] \leqslant E[Y] \int_0^1 [G(u) - G^{(1)}(u)](h'(\mathrm{u}))^2 \,\mathrm{d}u, \tag{2.3}$$

where $G^{(1)}$ is the first moment distribution of Y. Observe that in the special case where the density g of Y exists and is positive in some interval, (2.1) and (2.3) are equivalent, since

$$\operatorname{Var}[Y]w(u)g(u) = \int_0^u (E[Y] - t)g(t)\,\mathrm{d}t = E[Y]\cdot[G(u) - G^{(1)}(u)].$$

Therefore, (2.1) may be regarded as an extension of (2.3), since Y has an arbitrary interval support.

By using the lower bound for the variance given by Cacoullos and Papathanasiou (1989), namely,

$$\operatorname{Var}[h(Y)] \geqslant \operatorname{Var}[Y]E^2[w(Y)h'(Y)], \tag{2.4}$$

we can easily establish the following inequality:

$$\operatorname{Var}[X] \geqslant \operatorname{Var}[Y]\left(\int_0^1 w(G^{-1}(u))[h_Y(u)/h_X(u)]\,\mathrm{d}u\right)^2,$$

where equality holds iff $X =_{\mathrm{d}} \alpha Y + b$, $\alpha > 0$.

Consider now $Y_{(1)} < Y_{(2)} < \cdots < Y_{(n)}$ the order statistics from a sample of size n from the population Y and $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ the order statistics from a population X. Let G and F be the absolutely continuous distributions of Y and X, respectively. Suppose that the density g of Y has an interval support. Then, we have the following:

Theorem 2.1. *If* $\operatorname{Var}[X_{(i)}] < \infty$, $\operatorname{Var}[Y_{(i)}] < \infty$, *then*

$$\operatorname{Var}[X_{(i)}] \leqslant \frac{n!\operatorname{Var}[Y_{(i)}]}{(i-1)!(n-i)!} \int_0^1 w_i(G^{-1}(u)) \left[\frac{h_Y(u)}{h_X(u)}\right]^2 u^{i-1}(1-u)^{n-i}\,\mathrm{d}u, \tag{2.5}$$

where w_i is the w-function associated with $Y_{(i)}$. The equality holds iff $X =_{\mathrm{d}} \alpha Y + b$, $\alpha > 0$.

Proof. Since $X_{(i)} =_{\mathrm{d}} F^{-1}(G(Y_{(i)}))$, (2.1) gives

$$\mathrm{Var}[X_{(i)}] = \mathrm{Var}[h(Y_{(i)})] \leqslant \mathrm{Var}[Y_{(i)}] \int_{-\infty}^{+\infty} w_i(y) g_i(y) (h'(y))^2 \,\mathrm{d}y, \tag{2.6}$$

where $h = F^{-1} \circ G$ and g_i is the density of $Y_{(i)}$. Changing the variable $G(y) = u$ in (2.6) we obtain (2.5). □

By using inequality (2.4) and the same arguments as in Theorem 2.1 we give without proof the following theorem.

Theorem 2.2. *Under the conditions of Theorem* 2.1 *we have*

$$\mathrm{Var}[X_{(i)}] \geqslant \left\{ \frac{n!\sqrt{\mathrm{Var}[Y_{(i)}]}}{(i-1)!(n-1)!} \int_0^1 w_i(G^{-1}(u)) \frac{h_Y(u)}{h_X(u)} u^{i-1}(1-u)^{n-i} \,\mathrm{d}u \right\}^2, \tag{2.7}$$

with equality iff $X =_{\mathrm{d}} \alpha Y + b$, $\alpha > 0$.

We give now some applications to specific distributions. Set $G(x) = x$, $0 \leqslant x \leqslant 1$ in (2.7). Then we have

$$\mathrm{Var}[X_{(i)}] \geqslant \frac{n+2}{i(n+1-i)} \left[\frac{n!}{(i-1)!\,(n-i)!} \int_0^1 u^i (1-u)^{n+1-i} (h_X(u))^{-1} \,\mathrm{d}u \right]^2, \tag{2.8}$$

with equality iff X is uniform over some interval. Setting $G(x) = \mathrm{e}^x$, $x \leqslant 0$ and $i = n$ in (2.5) and (2.7), one has

$$n^2 \left\{ \int_0^1 (-\ln u) u^n (h_X(u))^{-1} \,\mathrm{d}u \right\}^2 \leqslant \mathrm{Var}[X_{(n)}] \leqslant \int_0^1 (-\ln u) u^{n+1} (h_X(u))^{-2} \,\mathrm{d}u,$$

with equality iff $f(x) = \lambda \mathrm{e}^{\lambda(x-c)}$, $x < c$.

As an example, consider the case where X has the sine distribution

$$F(x) = [1 + \sin(x-c)]/2, \quad c - \pi/2 < x < c + \pi/2. \tag{2.9}$$

Burrows (1986) showed a complicated series expansion for the variance of $X_{(n)}$ from this distribution and proved that $\mathrm{Var}[X_{(n)}]$ behaves like $(4-\pi)/n$ for large n. Arnold and Brockett proved that $\mathrm{Var}[X_{(n)}] \leqslant 1/(n+1)$. Here, by using (2.8) (in this case $h_X(u) = \sqrt{u(1-u)}$), we have

$$\mathrm{Var}[X_{(n)}] \geqslant \frac{\pi}{4} n(n+2) \left[\frac{\Gamma(n+1/2)}{\Gamma(n+2)} \right]^2 \approx \frac{\pi}{4n}$$

for large n, which implies that $\mathrm{Var}[X_{(n)}] = \mathrm{O}(1/n)$.

3. Asymptotic behaviour of the variance

It is known that if $i/n \to p$ as $n \to \infty$, then $\sqrt{n}(X_{(i)} - \xi_p)$ has a limiting normal distribution with mean 0 and variance $p(1-p)/f^2(\xi_p)$, where $\xi_p = F^{-1}(p)$, provided that f is positive and continuous at ξ_p.

In this section, by using the previous upper and lower bounds, we investigate the asymptotic behaviour of $\mathrm{Var}[X_{(i)}]$.

Theorem 3.1. *Let* X *be a r.v. with density* f, *distribution function* F *and density-quantile function* $h_X(u) = f(F^{-1}(u))$. *Suppose also that the function* $g(u) = [u(1-u)]^{s+1}[h_X(u)]^{-2}$, $0 < u < 1$, *is bounded for* *some* $s \geqslant -1$. *If* $n \to \infty$ *such that* $i/n \to p$ $(0 < p < 1)$, *then*

$$n \operatorname{Var}[X_{(i)}] \to p(1-p)/h_X^2(p) \tag{3.1}$$

provided that $h_X(p)$ *is continuous and positive at* p.

Proof. Theorem 2.1 with $G(x) = x$, $0 \leqslant x \leqslant 1$ (or from Papathanasiou's, 1990, Theorem 3.1) and inequality (2.8) may be rewritten as

$$\mathrm{LB}(n) \leqslant n \operatorname{Var}[X_{(i)}] \leqslant \mathrm{UB}(n),$$

where

$$\mathrm{UB}(n) = \frac{nB(i-s, n+1-i-s)}{(n+1)B(i, n+1-i)} E[g(U_n)],$$

$$\mathrm{LB}(n) = \frac{n(n+2)B^2(i-s, n+1-i-s)}{i(n+1-i)B^2(i,\ n+1-i)} E^2[g(U_n)h_X(U_n)]$$

and U_n is a beta random variable with parameters $i-s$ and $n+1-i-s$. Obviously, $g(U_n) \to_w g(p)$ (weakly) as $n \to \infty$. Since g is bounded, the r.v.'s $g(U_n)$, $n \geqslant 1$ are uniformly integrable and thus (see Billingsley, 1968, p. 32)

$$\lim_{n \to \infty} \mathrm{UB}(n) = [p(1-p)]^{-s} g(p) = p(1-p)/h_X^2(p). \tag{3.3}$$

On the other hand, the r.v.'s $g(U_n)h_X(U_n)$ are also uniformly integrable. Using similar arguments we have

$$\lim_{n \to \infty} \mathrm{LB}(n) = p(1-p)/h_X^2(p) \tag{3.4}$$

and by virtue of (3.3) and (3.4) the proof is complete. □

As an example, consider again the case where F is given by (2.9). The exact value of $\operatorname{Var}[X_{(i)}]$ is difficult to evaluate, while Theorem 3.1. simply gives (in this case $g(u) \equiv 1$ (bounded) for $s = 0$)

$$n \operatorname{Var}[X_{(i)}] \to 1$$

whenever $i/n \to p$.

References

Arnold, B.C. and P.L. Brockett (1988), Variance bounds using a theorem of Polya, *Statist. Probab. Lett.* **6**, 321–326.

Billingsley, P. (1968), *Convergence of Probability Measures* (Wiley, New York).

Burrows, P.M. (1986), Extreme statistics from the sine distribution, *Amer. Statist.* **40**, 216–217.

Cacoullos, T. and V. Papathanasiou (1985), On upper bounds for the variance of functions of random variables, *Statist. Probab. Lett.* **3**, 175–184.

Cacoullos, T. and V. Papathanasiou (1989), Characterizations of distributions by variance bounds, *Statist. Probab. Lett.* **7**, 351–356.

Papathanasiou, V. (1990), Some characterizations of distributions based on order statistics, *Statist. Probab. Lett.* **9**, 145–147.