Convexity of the Bounds induced by Markov's Inequality

by

Marcel F. Neuts and David B. Wolfson (*)

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series No. 289

July, 1972

(*) Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. AFOSR-72-2331. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

Abstract

X is a nonnegative random variable such that $EX^t < \infty$ for $0 \leq t < \lambda \leq \infty$. The $(1-\varepsilon)$-quantile of the distribution of X is bounded above by $[\varepsilon^{-1} EX^t]^{1/t}$. We show that there exist positive $\varepsilon_1 \geq \varepsilon_2$ such that for all $0 < \varepsilon \leq \varepsilon_1$, the function $g(t) = [\varepsilon^{-1} EX^t]^{1/t}$ is log-convex in $[0,c]$ and such that for all $0 < \varepsilon \leq \varepsilon_2$, $\log g(t)$ is nonincreasing in $[0,c]$.

$$P\left(X \geq \left(\frac{EX^t}{\varepsilon}\right)^{1/t}\right) \leq \varepsilon$$

$$P\left(X^t \geq \frac{EX^t}{\varepsilon}\right) \leq \varepsilon$$

Convexity of the Bounds induced by Markov's Inequality

by

Marcel F. Neuts and David B. Wolfson (*)

Purdue University

Introduction

Let X be a nonnegative random variable such that $EX^t < \infty$, for $0 \leq t < \lambda \leq \infty$. Throughout this discussion, c is a number such that $0 < c < \lambda$, and X is not identically zero. The classical Markov inequality states that for all $a > 0$, $P\{X \geq a\} \leq a^{-t} EX^t$, for $0 \leq t < \lambda$. The upper bound $a^{-t} EX^t$ does not exceed a given $\varepsilon > 0$, provided a is not less than $[\varepsilon^{-1} EX^t]^{1/t}$. The object of this note is to examine the behavior of this bound on the $(1-\varepsilon)$-quantile of the distribution $F(\cdot)$ of X, as a function of the order t of the moment used.

We shall consider the functions

$$(1) \qquad f(t) = \log EX^t,$$

$$g(t) = [A\ EX^t]^{1/t},$$

$$h(t) = \log g(t) = t^{-1}[f(t) + \log A],$$

where A is a positive constant. We recall the well-known results that $f(t)$ is a convex function of t in $(0,\lambda)$ and that for $A = 1$, $g(t)$ is nondecreasing in $(0,\lambda)$, [1]. Furthermore, a function is termed log-convex in an interval, if and only if its logarithm is defined and convex in that interval.

(*) Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. AFOSR-72-2331. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

Main Results

Lemma 1

The function $f(t)$ has continuous derivatives of all orders in any interval $(0,c]$.

Proof

The Mellin transform EX^t is strictly positive for $0 \leq t < \lambda$ and may be written as

$$EX^t = \int_{-\infty}^{+\infty} e^{-tu} \, d\, G(u) \quad , \tag{2}$$

where $G(u) = 1-F(e^{-u})$. Since this bilateral Laplace-Stieltjes transform converges for $0 \leq t < \lambda$, it is analytic and hence continuously differentiable of all orders in $(0,c]$. [3].

Theorem 1

For every c, with $0 < c < \lambda$, there exist positive constants $A_1 \leq A_2$ such that $g(t)$ is log-convex in $[0,c]$ for all $A \geq A_1$ and such that in addition, for all $A \geq A_2$ the function $h(t)$ is nonincreasing in $[0,c]$.

Proof

The first two derivatives of $h(t)$ are given by

$$\begin{aligned} h'(t) &= t^{-2} \, [t\, f'(t) - f(t) - \log A], \\ h''(t) &= t^{-3}[t^2 \, f''(t) - 2t\, f'(t) + 2f(t) + 2 \log A] \\ &= t^{-3}[t^2 f''(t) - 2t^2 \, h'(t)] \quad . \end{aligned} \tag{3}$$

We define A_1 and A_2 by

$$(4) \qquad \log A_1 = \max_{0\le t\le c} \left[t\, f'(t) - f(t) - \frac{t^2}{2} f''(t)\right] ,$$

$$\log A_2 = \max_{0\le t\le c} \left[t\, f'(t) - f(t)\right] .$$

Since $f''(t) \ge 0$, it follows that $A_1 \le A_2$. Moreover the derivative of $t\, f'(t) - f(t)$ is $t\, f''(t)$, which is nonnegative for $0 \le t \le c$. It follows that

$$(5) \qquad A_2 = \exp[c\, f'(c) - f(c)] .$$

The derivative of $t\, f'(t) - f(t) - \frac{t^2}{2} f''(t)$ is given by $-\frac{t^2}{2} f'''(t)$, so that the value of A_1 may be determined by examination of the third derivative of $f(t)$. Clearly for $A \ge A_1$, $h(t)$ is convex and for $A \ge A_2$, $h(t)$ is in addition nonincreasing.

<u>Remark</u>

By replacing A by ε^{-1}, we immediately obtain the appropriate statement for the bounds induced by Markov's inequality. The following weaker result can be proved without relying on differentiability properties.

<u>Theorem 2</u>

Let $k(t)$ be a bounded positive, log-convex function defined on the interval $[0,c]$, then there exists a positive constant A_0 such that for all $A \ge A_0$, the function $g(t)$ defined by

$$(6) \qquad g(t) = [A\, k(t)]^{1/t} ,$$

is log-convex on $[0,c]$.

<u>Proof</u>

Set $\log k(t) = f(t)$, then since $f(t)$ is convex, there exists for every

t_0 with $0 \leq t_0 \leq c$, a linear function $a(t_0)t + b(t_0)$ such that

(7) $$a(t_0)t_0 + b(t_0) = f(t_0) \quad ,$$

and

$$a(t_0)t + b(t_0) \leq f(t), \qquad \text{for } 0 \leq t \leq c \quad .$$

We may add a constant $\log A$ to both sides of the inequalities in (7) and divide by t, to obtain

(8) $$a(t_0) + \frac{b(t_0) + \log A}{t} \leq \frac{f(t) + \log A}{t} = \log\,[A\;k(t)]^{1/t} \quad ,$$

for all $0 < t \leq c$, with equality holding for $t = t_0$. Since the function $f(t)$ is convex, the intercept $b(t_0)$ of the supporting line $a(t_0)t + b(t_0)$ is nonincreasing in $[0,c]$. We choose A_0 such that $\log A_0 + b(c) \geq 0$, then for any $A \geq A_0$, the function $a(t_0) + [b(t_0) + \log A]t^{-1}$ is convex in $(0,c]$. It follows from equation (8) that for $A \geq A_0$, the function $\log\,[A\;k(t)]^{1/t}$ has a supporting family of convex functions for all $0 < t \leq c$, and is therefore itself convex. We note that if $f(t)$ is differentiable at $t = c$, then the smallest A_0 which validates theorem 2 is equal to A_2 determined in theorem 1.

<u>Corollary 1</u>

Let $S(Q,\mu)$ be a finite measure space and ψ a function from S to R^n, which does not vanish almost everywhere with respect to μ. Let $k(t)$ be defined by

(9) $$k(t) = \int_S \Big[\sum_{j=1}^{n} \psi_j^2\Big]^{\frac{t}{2}} \, d\mu \ ,$$

for $0 \leq t \leq c$ and assume that $k(c) < \infty$. It is then well-known that $\log k(t)$ is convex in $[0,c]$. Theorem 2 implies that for all A sufficiently large, $\log [A^{1/t}||\psi||_t]$ is convex in $[0,c]$, where $||\psi||_t$ denotes the usual L_t-norm of ψ.

Corollary 2.

Let $S(Q,\mu), \psi$, and $k(t)$ be defined as in corollary 1 with $f(t) = \log k(t)$. In addition let $k(t)$ be continuously differentiable for $0 \leq t \leq c$. Then $\log [A^{1/t}||\psi||_t]$ is convex and nonincreasing if and only if $A \geq A_2$ where $A_2 = cf'(c) - f(c)$.

Proof.

The definition of A_2 and corollary 1 together yield corollary 2. We need only observe that $\log A^{1/t}||\psi||_t$ is nonincreasing if and only if $\frac{d}{dt} \log A^{1/t}||\psi||_t \leq 0$ for all $t \in [0,c]$. cf. equation (3).

Applications

The Markov bounds on the tail probabilities of a random variable X are generally crude. However in some applications, it is possible to compute a large number of moments of probability distributions of interest, while the computation of the distribution itself may remain a laborious task. An example of this is the computation of the distribution of the busy period in queues of $M|G|1$ type. Klimko and Neuts [2] have shown that moments of order up to order forty may be readily computed by using the functional equation for the moment generating function. The computation of the busy period distribution itself by iterative procedures requires the knowledge of an upper bound on the $(1-\varepsilon)$-quantile of the distribution. The bounds induced by the Markov inequality are appealing because of their easy computability in terms of the moments. For the values of ε which occur in practice, one is usually in the case where the logarithm of the Markov bound is convex, but not yet monotone decreasing. This implies that in such cases there is a smallest Markov bound for the $(1-\varepsilon)$-quantile, but this bound need not correspond to the highest available moment.

Bibliography

[1] Feller, W. (1966). An Introduction to Probability Theory and its Applications, Vol. II. J. Wiley and Sons, New York, N.Y.

[2] Klimko, E. and Neuts, M.F. (1972). The Single Server Queue in Discrete Time-Numerical Analysis II. Purdue Mimeo Series No. 276, February, 1972, Department of Statistics, Purdue University, West Lafayette, IN.

[3] Widder, D. V. (1946). The Laplace Transform. Princeton University Press, Princeton, N.J.

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Probability theory						
Markov inequality						
Convexity						