J. Austral. Math. Soc. Ser. B 40(1998), 80–85

JENSEN'S INEQUALITY FOR DISTRIBUTIONS POSSESSING HIGHER MOMENTS, WITH APPLICATION TO SHARP BOUNDS FOR LAPLACE-STIELTJES TRANSFORMS

B. GULJAŠ[1], C. E. M. PEARCE[2] and J. PEČARIĆ[3]

(Received 7 May 1996; revised 23 October 1996)

Abstract

A new version of Jensen's inequality is established for probability distributions on the non-negative real numbers which are characterized by moments higher than the first. We deduce some new sharp bounds for Laplace-Stieltjes transforms of such distribution functions.

1. Introduction

In a previous article [4] we established the following variant of Jensen's inequality. For an earlier discussion of this theme and examples of applications see Pittenger [6].

THEOREM A. *Suppose $f(x)/x$ is a positive, convex function on $(0, \infty)$ and σ a probability measure on $[0, \infty)$, not consisting entirely of an atom at the origin, whose second moment exists. Then*

$$\int f(x)\,d\sigma \geq \frac{\left(\int x\,d\sigma\right)^2}{\int x^2\,d\sigma} f\left(\frac{\int x^2\,d\sigma}{\int x\,d\sigma}\right).$$

If $f(x)/x$ is strictly convex then strict inequality applies unless the support of σ intersects $(0, \infty)$ in a single point.

This result may be put to use to give a transparent derivation of the following well-known inequality in the teletraffic literature relating to the $G/M/n$ queue (see, for example, Rolski [8]).

[1]Mathematics Department, University of Zagreb, Bijenička Cesta 30, 41000 Zagreb, Croatia
[2]Department of Applied Mathematics, The University of Adelaide, Adelaide SA 5005, Australia.
[3]Faculty of Textile Technology, University of Zagreb, Pierottijeva 6, 41000 Zagreb, Croatia.

THEOREM B. *Let σ be a probability measure with nonnegative support and positive moments $m_i = \int t^i \, d\sigma(t)$ $(i = 1, 2)$. Then the functional $\int e^{-st} \, d\sigma$ achieves its supremum uniquely at $\sigma = \lambda_2$, where the measure λ_2 is given by*

$$d\lambda_2(t) = (1 - m_1^2/m_2)\delta(t)\,dt + (m_1^2/m_2)\delta(t - m_2/m_1)\,dt$$

and where as usual $\delta(\cdot)$ represents the Dirac delta.

A systematic provision of candidates for applications of Theorem A emerges from the notion of n-convexity ($n \geq 2$). See Popoviciu [7], Aumann and Haupt [1], Bullen [2] and Pečarić, Proschan and Tong [5] for a discussion of n-convex functions. We note in particular that this useful class of functions can be characterized by the property that, for $n \geq 2$, f is n-convex if and only if $f^{(n-2)}$ exists and is convex (see [1, p. 286]). Thus 2-convexity is just ordinary convexity. We have the following theorem.

THEOREM C. *Suppose f is an n-convex function on $(0, \infty)$ with $f^{(i)}(0) = 0$ $(0 \leq i < n - 2)$. Then the map: $x \to f(x)/x^{n-2}$ is convex on $(0, \infty)$.*

Thus we have that $f^{(3)} \geq 0$ on $(0, \infty)$ implies that the map: $x \to [f(x) - f(0)]/x$ is convex. In our earlier paper use was made of the homely particular case $f(x) = -e^{-sx}$.

In this note we pursue the foregoing approach to derive some new sharp bounds for the Laplace-Stieltjes transform of a probability distribution on $[0, \infty)$ characterized by higher moments. In Section 2 we present a more general version of Theorem A for higher moments.

The requirements on the derivatives of f in Theorem C are rather restrictive from the viewpoint of some probabilistic applications and it turns out to be preferable to proceed directly from the results of elementary calculus. These are codified in Section 3 as Proposition 1. This agrees with Theorem C for $n = 3$ but offers further scope for applications when $n > 3$. In Section 4 we marry the results of Sections 2 and 3 to engender a generalization of Theorem B. Finally, in Section 5, we illustrate by an example based on Section 4 the advantages that Proposition 1 can offer over Theorem C.

2. Jensen's inequality

THEOREM 1. *Suppose that r is nonnegative, that the map: $x \to G(x) = f(x)/x^r$ is positive and convex on $(0, \infty)$ and that σ is a probability measure on $[0, \infty)$ possessing an $(r + 1)$-st moment and not consisting simply of an atom at the origin. Then*

$$\int f(x)\,d\sigma \geq \frac{\left(\int x^r \, d\sigma\right)^{r+1}}{\left(\int x^{r+1} \, d\sigma\right)^r} f\left(\frac{\int x^{r+1} \, d\sigma}{\int x^r \, d\sigma}\right). \tag{2.1}$$

If G is strictly convex then the inequality is strict unless the support of σ intersects $(0, \infty)$ in a single point.

PROOF. Let X be a random variable with probability measure ν given by

$$d\nu(t) = \frac{t^r d\sigma(t)}{\int x^r\, d\sigma(x)},$$

so that

$$E(X) = \int x\, d\nu = \frac{\int x^{r+1}\, d\sigma}{\int x^r\, d\sigma}.$$

By convexity, Jensen's inequality yields

$$E[G(X)] \geq G(E(X)),$$

or

$$\frac{\int G(x) x^r\, d\sigma}{\int x^r\, d\sigma} \geq G\left[\frac{\int x^{r+1}\, d\sigma}{\int x^r\, d\sigma}\right],$$

whence we have (2.1). The statement on strict inequality is inherited from the corresponding result for Jensen's inequality.

3. An analogue for Theorem C

PROPOSITION 1. *Suppose that f is a function on $(0, \infty)$ with a second derivative and that r is a positive integer. A necessary and sufficient condition that the map: $t \to f(t)/t^r$ be convex is that*

$$h(t) \equiv r(r+1) f(t) - 2rtf'(t) + t^2 f''(t) \geq 0.$$

PROOF. The result is immediate from

$$d^2/dt^2\,[f(t)/t^r] = t^{-r-2} h(t).$$

COROLLARY 1. *If h is differentiable, its nonnegativity is guaranteed by the conditions $h(0) \geq 0$ and $h' \geq 0$ on $[0, \infty]$. Now*

$$h'(t) = r(r-1) f'(t) - 2(r-1) t f''(t) + t^2 f'''(t).$$

For $r = 1$, our conditions reduce to $f(0) \geq 0$ and $f''' \geq 0$, which hold automatically for any nonnegative 3-convex function. Similarly the conditions are satisfied trivially for $r > 1$ by any function f with $f(0) \geq 0$ for which f' and f''' are nonnegative and f'' is nonpositive.

4. Bounds for Laplace-Stieltjes transforms

We now proceed to a generalization of Theorem B.

THEOREM 2. *Let σ be a probability measure with nonnegative support not consisting purely of an atom at the origin and with given positive moments $m_j = \int t^i \, d\sigma(t)$ $(j = r, r+1)$. Then the functional $\phi(s) = \int e^{-st} \, d\sigma$ $(s \geq 0)$ achieves its supremum uniquely at $\sigma = \lambda_{r+1}$, where the measure λ_{r+1} is given by*

$$d\lambda_{r+1}(t) = \left[1 - \frac{m_r^{r+1}}{m_{r+1}^r}\right] \delta(t)\, dt + \frac{m_r^{r+1}}{m_{r+1}^r} \delta(t - m_{r+1}/m_r)\, dt.$$

PROOF. By Corollary 1, Proposition 1 applies for $f(x) = 1 - e^{-sx}$ $(s > 0)$ for all positive integral r. Hence this choice of f satisfies the conditions of Theorem 1 and from (2.1) we have

$$1 - \int e^{-sx} \, d\sigma \geq \frac{m_r^{r+1}}{m_{r+1}^r} \left[1 - \exp(-s m_{r+1}/m_r)\right]. \tag{4.1}$$

The fundamental inequality for L_p norms gives

$$\left(\int |x|^p d\sigma\right)^{1/p} \leq \left(\int |x|^q \, d\sigma\right)^{1/q} \quad \text{if} \quad 0 < p < q$$

for a probability measure σ, with strict inequality if σ does not consist of a single atom. Therefore $m_r^{r+1} < m_{r+1}^r$, so that λ_{r+1} is a proper two-point probability measure and (4.1) may be cast as

$$\phi(s) \leq \int e^{-sx} d\lambda_{r+1}(x).$$

A simple calculation shows that λ_{r+1} has r-th and $(r+1)$-st moments m_r and m_{r+1} respectively. This gives the main part of the enunciation. Uniqueness follows from the final statement in Theorem 1, since there is a unique measure on $[0, \infty)$ with the two given moments whose support intersects $(0, \infty)$ in a single point.

This result appears to be new for $r > 1$. For $r = 2$ it takes m_2 and m_3 as given and provides an interesting complement to a result of Eckberg [3]. Eckberg showed that if all three moments m_1, m_2, m_3 are given, then

$$\phi(s) \leq \left(1 - \frac{m_2}{m_1^2}\right) + \frac{m_2}{m_1^2} e^{-s m_2/m_1},$$

that is, the same upper bound applies as when only m_1, m_2 are given. Eckberg remarked that the upper bound needs an "infinitesimal mass at ∞" to achieve the correct third moment.

This last result extends to our general context.

COROLLARY 2. *The upper bound given in Theorem 2 applies if* m_{r+2} *is also given.*

PROOF. The probability measure λ_{r+1} has $(r+2)$-nd moment

$$\tilde{m}_{r+2} = \frac{m_r^{r+1}}{m_{r+1}^r} \cdot \left(\frac{m_{r+1}}{m_r}\right)^{r+2} = \frac{m_{r+1}^2}{m_r}.$$

By Cauchy's theorem $m_{r+2} \geq m_{r+1}^2/m_r$ for any probability measure σ, so that $m_{r+2} \geq \tilde{m}_{r+2}$. There is nothing to prove if equality holds, so suppose $m_{r+2} > \tilde{m}_{r+2}$. This enables us to construct (for some positive integer K) a sequence $(\mu_k)_{k\geq K}$ of probability measures whose moments of orders $r, r+1, r+2$ are respectively m_r, m_{r+1}, m_{r+2} with μ_k converging weakly to λ_{r+1} as $k \to \infty$. Since the Laplace-Stieltjes transform of μ_k converges to that of λ_{r+1} we shall then have the desired result. The construction may be implemented as follows. Set $m = m_{r+1}/m_r$ and define

$$\begin{aligned}
\epsilon_{2,k} &= (m_{r+2} - \tilde{m}_{r+2}) \,/\, [(k^\alpha + m)^r (k^\alpha + 1/k) k^\alpha]\,,\\
\epsilon_{3,k} &= (m_{r+2} - \tilde{m}_{r+2})\, k / \,[(m - 1/k)^r (k^\alpha + 1/k)]\,,\\
\epsilon_{1,k} &= (m_{r+2} - \tilde{m}_{r+2})\, k / \,[m^r k^\alpha]\,,\\
\epsilon_{0,k} &= \epsilon_{2,k} + \epsilon_{3,k} - \epsilon_{1,k},
\end{aligned}$$

for all $k \geq K$. Here K is chosen sufficiently large that $m - 1/K > 0$, $\epsilon_{3,K} > \epsilon_{1,K}$, $\epsilon_{0,K} < 1 - m_r^{r+1}/m_{r+1}^r$ and $\epsilon_{1,K} < m_r^{r+1}/m_{r+1}^r$ and α is chosen sufficiently large that $[1 - 1/(mK)]^{-r} > 1 + K^{-1-\alpha}$. We readily verify that, for $k \geq K$, the measure μ_k given by

$$\begin{aligned}
d\mu_k(t) = &\left[1 - \frac{m_r^{r+1}}{m_{r+1}^r} - \epsilon_{0,k}\right]\delta(t)\,dt + \epsilon_{3,k}\delta(t - (m - 1/k))\,dt\\
&+ \left[\frac{m_r^{r+1}}{m_{r+1}^r} - \epsilon_{1,k}\right]\delta(t - m)\,dt + \epsilon_{2,k}\delta(t - (k^\alpha + m))\,dt
\end{aligned}$$

is a probability measure with moments as stated that converges weakly to λ_{r+1}.

5. Theorem C and Proposition 1

It is interesting to compare the analysis of Theorem 2, based on Proposition 1, for $r = 2$ with a parallel development using Theorem C with $n = 4$. The corresponding natural choice with the latter is then $f(x) = e^{-sx} - 1 + sx$, the last term being forced on us by the requirement that $f'(0)$ be zero. We have at once that $f(x)/x^2$ is convex. Theorem 1 leads to

$$\phi(s) - 1 + sm_1 \geq \frac{m_2^3}{m_3^2}\left[e^{-sm_3/m_2} - 1 + sm_3/m_2\right]$$

or

$$\phi(s) \geq \left[1 - \frac{m_2^3}{m_3^2}\right] + \frac{m_2^3}{m_3^2} e^{-sm_3/m_2} - s\left[m_1 - m_2^2/m_3\right].$$

By Cauchy's inequality, the last term in brackets in nonnegative. If it is strictly positive, as must happen for $\sigma \neq \lambda_3$, then we obtain a very poor lower bound for ϕ, since the last term on the right is unbounded for $s \to \infty$. Moreover, we appear to lack an appropriate probabilistic interpretation for this result.

References

[1] G. Aumann and O. Haupt, *Einführung in die Reelle Analysis, Band II* (Walter de Gruyter, Berlin, 1938).

[2] P. S. Bullen, "A criterion for n-convexity", *Pacific J. Math.* **36** (1971) 81–98.

[3] A. E. Eckberg, "Sharp bounds on Laplace-Stieltjes transforms, with applications to various queueing problems", *Math. of Oper. Res.* **2** (1977) 135–142.

[4] C. E. M. Pearce and J. E. Pečarić, "An integral inequality for convex functions, with application to teletraffic congestion problems", *Math. of Oper. Res.* **20** (1995) 526–528.

[5] J. E. Pečarić, F. Proschan and Y. L. Tong, *Convex functions, partial orderings and statistical applications*, Mathematics in Science and Engineering, Vol. 187 (Academic Press, San Diego, 1992).

[6] A. O. Pittenger, "Sharp mean-variance bounds for Jensen-type inequalities", *Probability and Statistics Letters* **10** (1990) 91–94.

[7] T. Popoviciu, *Les fonctions convexes* (Hermann, Paris, 1945).

[8] T. Rolski, "Some inequalities for $GI/M/n$ queues", *Zastos. Mat.* **13** (1972) 43–47.