PUBLICATIONS DE L'INSTITUT MATHÉMATIQUE
Nouvelle série tome 56 (70), 1994, 111–118

ON SCHUR-CONVEXITY OF SOME DISTRIBUTION FUNCTIONS

Milan Merkle and **Ljiljana Petrović**

Communicated by Zoran Ivković

Abstract. If $X_1, \dots, X_n$ are independent geometric random variables with parameters $p_1, \dots, p_n$ respectivelly, we prove that the function $F(p_1, \dots, p_n; t) = P(X_1 + \dots + X_n \le t)$ is Schur-concave in $(p_1, \dots, p_n)$ for every real t. We also give a new proof for a theorem due to P. Diaconis on Schur-convexity of distribution fuction of linear combination of two exponential random variables.

I. Introduction

Let $X_1, \dots, X_n$ be independent and identically distributed random variables and let

$$F(c_1, \dots, c_n; t) = P(c_1 X_1 + \dots + c_n X_n \le t).$$

By a result of Proschan [**7**], if the common density of $X_1, \dots, X_n$ is symmetric about zero and log-concave, then the function F is Schur-concave in $(c_1, \dots, c_n)$. For nonsymmetric densities, analogous results are known only in several particular cases of Gamma and Weibull distributions (see [**1**], [**2**], [**8**] and a survey in [**6**]).

For discrete distributions, there are Schur-convexity results for Bernoulli random variables in [**3**] and [**4**], and a more general result in [**5**].

Investigation of Schur-convexity is of an interest because Schur-convexity implies certain useful inequalities for tail probabilities.

Let us briefly review concepts of majorization and Schur-convexity (see [**6**] for details). Let $\vec{x}$ and $\vec{y}$ be vectors in $\mathbf{R}^n$, and let $x_{[i]}$, $y_{[i]}$ denote the i-th largest component of $\vec{x}$, $\vec{y}$ respectively. Then we say that $\vec{x} \prec \vec{y}$ ($\vec{x}$ is majorized by $\vec{y}$) if

$$\sum_{i=1}^{k} x_{[i]} \le \sum_{i=1}^{k} y_{[i]}, \ \ k = 1, 2, \dots, n-1, \ \ \sum_{i=1}^{n} x_{[i]} = \sum_{i=1}^{n} y_{[i]}.$$

A function f of n arguments is said to be Schur-convex on a set $A \subset \mathbf{R}^n$ if, for all $\vec{x}, \vec{y} \in A$,

$$\vec{x} \prec \vec{y} \Rightarrow f(\vec{x}) \le f(\vec{y}).$$

AMS Subject Classification (1990): Primary 60 E 15.

A function f is Schur-concave if $-f$ is Schur-convex.

Let f be a Schur-convex function on a convex set A. Then

$$f(x_1, x_2, \ldots, x_n) \geq f(m, m, \ldots, m), \quad m = \frac{x_1 + \ldots + x_n}{n}. \tag{1}$$

In this paper we give our proof of a statement regarding two exponential random variables and discuss a general case, which has not been solved yet. Further, we present a result about sums of geometric random variables, thus adding one more case to the list of particular results.

2. Exponential random variables

In [**1**, p. 377, 12K.3.] the following result of Persi Diaconis is stated without a proof: If X_1 and X_2 are iid random variables with exp(1) distribution, then the function

$$P(c_1 X_1 + c_2 X_2 \leq x), \qquad (c_1, c_2 > 0) \tag{2}$$

is a Schur-concave function on (c_1, c_2) for $x \leq (c_1 + c_2)$ and it is a Schur-convex function on (c_1, c_2) for $x \geq \frac{3}{2}(c_1 + c_2)$. There is a misprint in this statement: the function is, in fact, Schur-convex for $x \leq (c_1 + c_2)$ and Schur-concave for $x \geq \frac{3}{2}(c_1 + c_2)$.

We present here our proof of the correct statement. A more general result for two Gamma random variables can be found in [**1**], but the proof is more involved.

The key tools are inequalities given in the Lemma below, that might be of an independent interest. Note that without loss of generality we may assume $c_1 + c_2 = 1$.

Lemma. *Let $c \in (1/2, 1)$ and let $\varphi(t) = 1 - e^{-t} - \frac{t}{ct+1}$. Then* (i) $\varphi(t) > 0$ *for $0 < t \leq \frac{2c-1}{c(1-c)} \equiv t_1$.* (ii) $\varphi(t) < 0$ *for $t \geq \frac{3}{2} \cdot \frac{2c-1}{c(1-c)} \equiv t_2$.*

Proof. We shall prove both assertions by proving that for all $c \in (1/2, 1)$: (a) The function $t \mapsto \varphi(t)$ has at most one zero in $(0, +\infty)$, (b) $\varphi(t_1) > 0$, (c) $\varphi(t_2) < 0$.

Proof of (a): The equation $\varphi(t) = 0$ is equivalent to

$$\frac{ct + 1 - t}{ct + 1} = e^{-t} \tag{3}$$

The left hand side in (3) is negative if $t > 1/(1-c)$. Therefore, (3) can have real solutions only for $t < 1/(1-c)$. With this restriction, (3) is equivalent to

$$\psi(t) = \log(ct + 1 - t) - \log(ct + 1) + t = 0, \qquad 0 < t < \frac{1}{1-c}. \tag{4}$$

Suppose now that the equation $\psi(t) = 0$ has at least two positive solutions $\beta > \alpha > 0$. Then the function ψ defined by (4) would also have roots α, β and we would have $\psi(0) = \psi(\alpha) = \psi(\beta) = 0$. This would imply (by Rolle's theorem) that the derivative ψ' has at least two positive roots. However, this is not possible because roots of ψ' are 0 and t_1.

Proof of (b): From the proof of (a) we have

$$\psi(0) = 0, \qquad \psi'(t) > 0 (0 < t < t_1), \qquad \psi'(t_1) = 0,$$

which implies $\psi(t) > \psi(0) = 0$, $(0 < t \leq t_1)$, and, in particular, $\psi(t_1) > 0$. The latter inequality is equivalent to $\varphi(t_1) > 0$ as can easily be seen from the proof of (a).

Proof of (c): After subsitution, the inequality $\varphi(t_2) < 0$ reduces to

$$\frac{(1-c)(3-4c)}{c(4c-1)} < \exp\left(-\frac{3}{2}\cdot\frac{2c-1}{c(1-c)}\right), \qquad \frac{1}{2} < c < 1. \tag{5}$$

If $c \geq 3/4$, then (13) trivially holds. For $c < 3/4$, (13) is equivalent to

$$\begin{aligned} h(c) = {} & \log(1-c) + \log(3-4c) \\ & - \log c - \log(4c-1) + \frac{3}{2}\cdot\frac{2c-1}{c(1-c)} < 0, \quad \frac{1}{2} < c < \frac{3}{4}. \end{aligned} \tag{6}$$

It is straightforward to check that $h(1/2) = 0$ and

$$h'(c) = \frac{9(2c-1)^4}{2c^2(4c-1)(4c-3)(c-1)^2} < 0 \text{ for } \frac{1}{2} < c < \frac{3}{4},$$

and (6) is proved.

Theorem 1. *Let X_1 and X_2 be independent random variables with exponential distribution, $EX_1 = EX_2 = 1$. Then the function*

$$F(c_1, c_2, x) = P(c_1X_1 + c_2X_2 \leq x), \quad c_1 + c_2 = 1, \ c_1, c_2 > 0,$$

is Schur convex on (c_1, c_2) if $x \leq 1$ and it is Schur concave on (c_1, c_2) if $x \geq 3/2$.

Proof. It is easy to show that

$$\begin{aligned} F(c_1, c_2, x) &= 1 - \frac{c_1 e^{-x/c_1} - c_2 e^{-x/c_2}}{c_1 - c_2}, \quad c_1 \neq c_2, \\ F\left(\frac{1}{2}, \frac{1}{2}, x\right) &= 1 - e^{-2x}(1+2x) \end{aligned} \tag{7}$$

Since $\lim_{c_1 \to 1/2} F(c_1, 1-c_1, x) = F(1/2, 1/2, x)$, it suffices to show Schur-convexity (Schur-concavity) of F with respect to (c_1, c_2) in the domain $c_1 + c_2 = 1$, $c \neq c_2, c_1, c_2 > 0$.

If we assume that $c_1 > c_2$, then by $c_1 + c_2 = 1$ we have $c_1 > 1/2$ and $c_2 = 1 - c_1$. Therefore,

$$F(c_1, c_2, x) = 1 - f(c_1, x), \tag{8}$$

where

$$f(c,x)=\frac{ce^{-x/c}-(1-c)e^{-x/(1-c)}}{2c-1}.$$

In our case majorization of vectors $\vec{c}=(c_1,c_2),\vec{c}'=(c_1',c_2')$ with $c_1>c_2$ means:

$$\vec{c}\prec\vec{c}' \iff c_1\le c_1'\wedge c_1+c_2=c_1'+c_2'=1,$$

i.e, $\vec{c}\prec\vec{c}' \iff c_1\le c_1'$. Thus, the function F in (8) is Schur convex (Schur concave) if and only if the function f is nonincreasing (nondecreasing) in $c\in(1/2,1]$.

The derivative of the function f with respect to c reads:

$$\frac{\partial f(c,x)}{\partial c}=\frac{1}{(2c-1)^2}\left(\frac{1}{c}e^{-x/c}(2cx-x-c)+\frac{1}{1-c}e^{-x/(1-c)}(2cx-x+1-c)\right).$$

The sign of this expression is the same as the sign of

$$(cx(2c-1)+c(1-c))\left(e^{-x/(1-c)}-e^{-x/c}\right)+x(2c-1)e^{-x/c}.$$

For $1/2<c<1$ and $x>0$, the sign of the latter expression is the same as the sign of

$$A(x,c)=e^{-(x/(1-c)-x/c)}-1+\frac{x(2c-1)}{cx(2c-1)+c(1-c)}. \tag{9}$$

Let

$$t=\frac{x}{1-c}-\frac{x}{c}=\frac{x(2c-1)}{c(1-c)},\quad x=\frac{c(1-c)}{2c-1}. \tag{10}$$

By substitution in (9), we get

$$A(x,c)=-1+e^{-t}+\frac{t}{ct+1}=-\varphi(t),$$

where φ is the function introduced in the Lemma. By (i) of the Lemma and (10), we have that $A(x,c)<0$ for $x\le 1$, and the function f is decreasing on c, which implies Schur-convexity of F. Similarly, from (ii) of the Lemma we conclude that F is a Schur-concave function on (c_1,c_2) for $x\ge 3/2$.

Corollary. *For X_1,X_2,c_1,c_2 as in Theorem 1, the following inequalities hold:*

$$P(c_1X_1+c_2X_2\le x)\ge P\left(\frac{X_1+X_2}{2}\le x\right)\ \text{ if }\ 0\le x\le 1$$

$$P(c_1X_1+c_2X_2\le x)\le P\left(\frac{X_1+X_2}{2}\le x\right)\ \text{ if }\ x\ge\frac{3}{2}.$$

Proof. Immediate by Theorem 1 and the inequality (1).

Remarks and comments. There is a numerical evidence that a statement analogous to the one proved in Theorem 1, holds for more than two random variables. The Schur-convexity of the function

$$F(c_1,\ldots c_n,x)=P(c_1X_1+\ldots c_nX_n\le x)$$

should be investigated on the set

$$\left\{(c_1, \dots, c_n) | \sum_{i=1}^{n} c_1 = 1, c_1, \dots, c_n > 0\right\}.$$

If $X_1, \dots X_n$ are iid exponential random variables with the mean equal to 1, and if all c_j are different and positive, then

$$F(c_1, \dots, c_n, x) = 1 - \sum_{j=1}^{n} \frac{e^{-x/c_j}}{\prod_{k \neq j}(1 - c_k/c_j)}.$$

If $c_j = 1/n$ for all $j = 1, 2, \dots, n$, then

$$F(1/n, \dots, 1/n, x) = 1 - e^{-nx}\left(1 + nx + \dots + \frac{(nx)^{n-1}}{(n-1)!}\right).$$

In intermediate case when some of constants are equal, the corresponding formulas can be derived, but it seems that this can be avoided by observing that the function F has to continuous and so it suffices to look at the set of all different $c_1, \dots c_n$ that sum up to 1.

It seems that the case $n > 2$ requires a different approach. The difference between $n > 2$ and $n = 2$ is also reflected in the following fact. If $n = 2$, then

$$\mathrm{Var}\,(cX_1 + (1-c)X_2) = c^2 + (1-c)^2 = 2c^2 - 2c + 1.$$

For $c > 1/2$, the function $c \mapsto 2c^2 - 2c + 1$ is increasing. So, in this case, F is Schur-convex if and only if

$$\mathrm{Var}\,(Y) \leq \mathrm{Var}\,(Y') \Rightarrow P(Y \leq x) \leq P(Y' \leq x), \tag{11}$$

where $Y = \sum_{i=1}^{n} c_i X_i$, $Y' = \sum_{i=1}^{n} c_i' X_i$, $n = 2$.

Theorem 1 looks very natural in the light of relation (11). If one draws a graph of two unimodal distribution function F_1 and F_2 with the same mean and variances $\sigma_1 < \sigma_2$, then one would expect that $F_1(x) < F_2(x)$ for $x \leq x_0$, and that $F_1(x) > F_2(x)$ for $x > x_0$. So, it seems intuitively clear that the family of unimodal distribution functions $F(c_1, \dots, c_n, x)$ should be increasing with respect to the variance (= Schur-convex) for small values of x, and decreasing (= Schur-concave) for large values of x.

However, for $n > 2$, it can be numerically shown that (11) does not necessarily hold even if F is Schur-convex.

3. Geometric random variables

Let $X_1, \dots, X_n$ be independent geometric random variables with parameters $p_1, \dots, p_n$ respectively:

$$P(X_i = k) = p_i(1-p_i)^k, \qquad k = 1, 2, \dots$$

Let $\mathbf{p} = (p_1, \dots, p_n)$, $\mathbf{q} = (1 - p_1, \dots, 1 - p_n)$.

In this part we will prove the following result:

THEOREM 2. *The function* $F(p_1, \dots, p_n; t) = P(X_1 + \dots + X_n \le t)$ *is Schur-concave with respect to* $\mathbf{p}$*, for every real* t.

If $t \ge n$, then the explicit expression for F reads

$$p_1 \cdots p_n \sum_{k_1 + \cdots + k_n \le [t] - n} (1 - p_1)^{k_1} \cdots (1 - p_n)^{k_n},$$

where $k_i \ge 0$ are integers. The product $p_1 \cdots p_n$ is a Schur-concave function. Second term in (12) is a sum of complete symmetric functions and it is Schur-convex with respect to $\mathbf{q}$ (see [**1**, Ch. 3, F5]). By $\mathbf{p} \prec \mathbf{p}' \iff \mathbf{q} \prec \mathbf{q}'$, it follows that this term is also Schur-convex with respect to $\mathbf{p}$. Therefore, an immediate conclusion about Schur-convexity of (12) is not possible.

Let $S = X_1 + \cdots + X_n$. Then

$$ES = \sum_{i=1}^{n} \frac{1}{p_i}, \qquad \operatorname{Var} S = \sum_{i=1}^{n} \frac{1 - p_i}{p_i^2}$$

and both expressions are convex and symmetric, hence Schur-convex functions of $p_1, \dots, p_n$. By $(\bar{p}, \dots, \bar{p}) \prec (p_1, \dots, p_n)$, $\bar{p} = (p_1 + \cdots p_n)/n$, we conclude that, for $p_1 + \cdots + p_n = c = \text{const.}$, ES and $\operatorname{Var} S$ are minimal when probabilities p_i are equal.

By Theorem 2, we have a stronger conclusion:

$$F(p_1, \dots, p_n, t) \le F(\bar{p}, \dots, \bar{p}, t), \qquad \bar{p} = \frac{p_1 + \cdots + p_n}{n}. \tag{13}$$

Therefore, the value of expression (12) is not greater than

$$\bar{p}^n \sum_{k=0}^{[t]-n} \binom{k + n - 1}{k} (1 - \bar{p})^k.$$

Proof of Theorem 2. By [**6**], it suffices to show that the function

$$x \mapsto F(p_1, c - p_1, p_3, \dots, p_n, t)$$

is decreasing in p_1 when $p_1 \in (c/2, c)$ (and when F is defined), with $p_3, \dots, p_n$ being fixed. Since t is arbitrary and

$$P(X_1 + \cdots + X_n \le t) = E(P(X_1 + X_2 \le t - X_3 - \cdots X_n | X_3, \cdots, X_n)),$$

it suffices to prove the result for $n = 2$. Further, it is clear that it suffices to give a proof for integers $t \ge 2$.

For $n = 2$, (12) reads:

$$F(p_1, p_2, t) = p_1 p_2 \sum_{k=0}^{[t]-2} \sum_{j=0}^{k} (1 - p_1)^j (1 - p_2)^{k-j},$$

or, after summation,

$$F(p_1,p_2,t) = \frac{p_1p_2}{p_1-p_2}\left((1-p_2)\frac{1-(1-p_2)^{m+1}}{p_2} - (1-p_1)\frac{1-(1-p_1)^{m+1}}{p_1}\right), \tag{14}$$

where $m=[t]-2$. Letting $p_1=x$, $p_2=c-x$, we obtain the expression

$$f(x) = F(x,c-x,t) = \frac{x(c-x)}{2x-c}\Bigg((1-c+x)\cdot\frac{1-(1-c+x)^{m+1}}{c-x} - (1-x)\cdot\frac{1-(1-x)^{m+1}}{x}\Bigg),$$

where $m=0,1,2,\dots$.

We have to show that f is decreasing function in $x\in(c/2,c)$. However, the conditions $0\le x\le 1$, $0\le 1-c+x\le 1$, $c/2\le x\le c$, $0\le c\le 2$ are met if and only if

$$\frac{c}{2}\le x\le \min(c,1), \qquad 0\le c\le 2. \tag{15}$$

Therefore, we will show that the function f is decreasing in x, for x and c as specified in (15).

We have the following:

$$\begin{aligned} f'(x) &= -\frac{1}{(2x-c)^2}\Big(\big(2x^2(m+2)-cx(m+3)+c(c-1)\big)(x-c+1)^{m+1} \\ &\qquad - \big(2x^2(m+2)-cx(3m+5)+c(c(m+2)-1)\big)(1-x)^{m+1}\Big) \\ &= -\frac{1}{(2x-c)^2}g(x), \qquad g(c/2)=0, \end{aligned}$$

$$\begin{aligned} g'(x) &= (m+2)(2x-c)\Big(\big(x(m+3)-c(m+1)-2\big)(1-x)^m \\ &\qquad + \big(x(m+3)-2c+2\big)(x-c+1)^m\Big) \\ &= (m+2)(2x-c)\varphi(x). \end{aligned}$$

It follows that the proof will be finished if we show that $\varphi(x)\ge 0$. This inequality is equivalent to

$$\alpha(x)(x-c+1)^m \ge \beta(x)(1-x)^m, \tag{16}$$

where $\alpha(x)=x(m+3)-2c+2$, $\beta(x)=2+c(m+1)-x(m+3)$. Let us show that both α and β are non-negative for x as in (15).

1° We have that

$$\alpha(c/2) = \frac{c}{2}(m-1)+2>0,$$

and since α is increasing, $\alpha(x) > 0$ for all $x \geq c/2$.

2° To show positivity of β, note that $\beta(c/2) = \alpha(c/2) > 0$ and

$$\begin{aligned} \beta(1) &= (c-1)(m+1) \geq 0 \quad \text{for} \quad c \geq 1, \\ \beta(c) &= 2(1-c) \geq 0 \qquad\qquad \text{for} \quad c \leq 1. \end{aligned}$$

Since β is a monotone function, $\beta(x) \geq 0$ for x as indicated.

Further,

$$\alpha(x) - \beta(x) = (2x - c)(m + 3) \geq 0 \tag{17}$$

and, by $x \geq c/2$,

$$(x - c + 1)^m \geq (1 - x)^m. \tag{18}$$

Inequalities (17) and (18) imply (16), which ends the proof.

Remark. There is an analogy between geometric and exponential distributions, both representing a waiting time untill the first event (Bernoulli or Poisson). Therefore, one would expect that a similar result holds for exponential distribution. Indeed, as indicated in [**1**], Tong [**8**] has proved an analogous result for the Gamma distribution.

REFERENCES

[1] M.E. Bock, P. Diaconis, F.W. Huffer, M.D. Perlman, *Inequalities for linear combinations of Gamma random variables*, Canad. J. Statistics **15** (1987), 387–395

[2] P. Diaconis, M.D. Perlman, *Bounds for tail probabilities of linear combinations of independent Gamma random variables*, The Symposium on Dependence in Statistics and Probability, Hidden Valley, Pennsylvania, 1987

[3] L. Gleser, *On the distribution of the number of successes in independent trials*, Ann. Probab. **3** (1975), 713–721

[4] W. Hoeffding, *On the distribution of the number of successes in independent trials*, Ann. Math. Statist. **27** (1956), 713–721

[5] M. Kanter, *Probality inequalities for convex sets and multidimensional concentration functions*, J. Multivariate Anal. **6** (1976), 222–236

[6] A. Marshall, I. Olkin, *Inequalities: Theory of Majorization and its Applications*, Academic Press, New York, 1979

[7] F. Proschan, *Peakedness of distributions of convex combinations*, Ann. Math. Statist. **36** (1965), 1703–1706

[8] Y.L. Tong, *Some inequalities for sums of independent exponential and gamma variables with applications*, unpublished report, Department of Mathematics and Statistics, University of Nebraska, Lincoln, 1980

Elektrotehnički fakultet
Beograd, p.p. 816
Yugoslavia

(Received 14 06 1994)

Prirodno-matematički fakultet
Kragujevac, Radoja Domanovića 12
Yugoslavia