AD614788

D1-82-0385

BOEING SCIENTIFIC RESEARCH LABORATORIES

Peakedness of Distributions of Convex Combinations

Frank Proschan

Mathematics Research

November 1964

D1-82-0385

PEAKEDNESS OF DISTRIBUTIONS OF CONVEX COMBINATIONS

by

Frank Proschan

Mathematical Note No. 374

Mathematics Research Laboratory

BOEING SCIENTIFIC RESEARCH LABORATORIES

November 1964

This report represents a continuation of work begun at the Boeing Scientific Research Laboratories and completed during the author's sojourn at the University of California. It is being issued simultaneously there as Operations Research Center Report No. ORC 64-26(RR).

The author's work at the University of California has been partially supported by the Office of Naval Research under Contract Nonr-3656(18) with the University of California.

ABSTRACT

This paper presents simple sufficient conditions under which the distribution of the sample average shows increasing peakedness with increasing sample size. The results are actually more general, permitting a comparison of the peakedness of distributions of various convex combinations of sample observations.

BLANK PAGE

I. INTRODUCTION

Roughly speaking, the law of large numbers states that under mild restrictions the average of a random sample has small probability of deviating from the population mean if the sample size n is taken large enough. However, nothing is said about the probability of a given size deviation decreasing monotonically as n increases. In this paper we develop conditions under which such monotonicity can be established. Another way of stating this is that under appropriate conditions the "peakedness" of the distribution of the average of n increases with n. We use the definition of peakedness given by Birnbaum (1948).

Definition   Let X_1 and X_2 be real random variables and a_1 and a_2 real constants. We say X_1 is more peaked about a_1 than X_2 about a_2 if

$$P[|X_1 - a_1| \geq t] \leq P[|X_2 - a_2| \geq t]$$

for all $t \geq 0$. In the case $a_1 = 0 = a_2$, we shall simply say X_1 is more peaked than X_2.

II. PEAKEDNESS COMPARISONS FOR SYMMETRIC PÓLYA FREQUENCY FUNCTIONS OF ORDER 2

Lemma 2.1   Let f be a Pólya frequency function of order 2 (PF_2), $f(u) = f(-u)$ for all u, $f(u) > 0$ for $-a < u < a$, and 0 elsewhere, $0 < a \leq \infty$, X_1 and X_2 independently distributed with density f, and $G_2(p,t) = P[pX_1 + qX_2 \leq t]$, where $0 \leq p \leq q$, $p + q = 1$. Then for $0 < t < a$, $G_2(p,t)$ is strictly increasing in p, $0 \le p \le \frac{1}{2}$.

<u>Proof</u> For $0 < p < \frac{1}{2}$, write $G_2(p,t) = \int_{-\infty}^{\infty} F\left(\frac{t-qu}{p}\right) f(u)du$.

Then $p^2 \frac{\partial G_2}{\partial p} = \int_{-\infty}^{\infty} f\left(\frac{t-qu}{p}\right) f(u)(u - t)du$; differentiation under the integral sign is permissible since $\left|f\left(\frac{t-qu}{p}\right) f(u)(u - t)\right| \le Mf(u)|u - t|$ and $\int_{-\infty}^{\infty} Mf(u)(u - t)du < \infty$, where M is the modal ordinate of f.

Rewrite

$$p^2 \frac{\partial G_2}{\partial p} = \int_{-\infty}^{t} f\left(\frac{t-qu}{p}\right) f(u)(u - t)du + \int_{t}^{\infty} f\left(\frac{t-qu}{p}\right) f(u)(u - t)du.$$

Let $v = t - u$ in the first integral and $v = u - t$ in the second integral. We get

$$(1) \qquad p^2 \frac{\partial G_2}{\partial p} = \int_0^{\infty} v\{f(t + v)f(t - \tfrac{qv}{p}) - f(t - v)f(t + \tfrac{qv}{p})\}dv$$

$$= \int_0^{\infty} v\left\{\frac{f(t + v)}{f(t + qv/p)} - \frac{f(t - v)}{f(t - qv/p)}\right\} f(t + \tfrac{qv}{p})f(t - \tfrac{qv}{p})dv,$$

with the understanding that whenever a denominator is 0 we use the integrand in (1).

Next note that when f is PF_2, $\frac{f(t + v)}{f(t)}$ is decreasing in t for fixed $v \ge 0$. Since $t > 0$, $\frac{q}{p} > 1$, we have

$$(2) \qquad p^2 \frac{\partial G_2}{\partial p} \ge \int_0^{\infty} v\left\{\frac{f(v)}{f(qv/p)} - \frac{f(-v)}{f(-qv/p)}\right\} f(t + \tfrac{qv}{p})f(t - \tfrac{qv}{p})dv.$$

No difficulty arises in case $f(\frac{qv}{p}) = 0$, since in this case

$f(t + \frac{qv}{p}) = 0$. Thus the inequality remains valid if we interpret the integrand in (2) as 0 whenever $f(\frac{qv}{p}) = 0$. By symmetry of f, we conclude $p^2 \frac{\partial G_2}{\partial p} \geq 0$.

Now suppose $p^2 \frac{\partial G_2}{\partial p} = 0$. Then for all v, except for at most two points (a PF_2 density is continuous except for at most two points),

$$0 \equiv f(t+v)f(t - \frac{qv}{p}) - f(t-v)f(t + \frac{qv}{p}) = \left\{ \frac{f(t+v)}{f(t-v)} - \frac{f(t+qv/p)}{f(t-qv/p)} \right\} f(t-v)f(t-qv/p).$$

For $0 < v \leq \frac{tp}{q}$, $f(t - v)f(t - \frac{qv}{p}) > 0$. But since f is symmetric PF_2, f has a mode at 0. Thus for $0 < v \leq \frac{tp}{q}$, $\frac{f(t + v)}{f(t - v)} \geq \frac{f(t + qv/p)}{f(t - qv/p)}$, with equality possible only if $f(u)$ is constant for $t - tp/q \leq u \leq t + tp/q$. This last however implies $f(u)$ is constant for $-a < u < a$, that is, f is the uniform density. But a direct examination of the uniform shows $p^2 \frac{\partial G_2}{\partial p} > 0$.

Finally note that at $p = 0$, $G_2(p,t)$ is continuous by Cramér (1946), p. 254. ||

<u>Lemma 2.2</u> Let f be PF_2, $f(t) = f(-t)$ for all t, $f(u) > 0$ for $-a < u < a$ and 0 elsewhere, $0 < a \leq \infty$, $X_1, \ldots, X_n$ independently distributed with density f, and $G_n(p_1,t) = P\left[\sum_{i=1}^{n} p_i X_i \leq t\right]$, where $0 \leq p_i \leq 1$, $i = 1, \ldots, n$, $\sum_{i=1}^{n} p_i = 1$, $p_1 + p_2 = b$, a positive constant. Then for $0 < t < a$, $G_n(p_1,t)$ is strictly increasing in p_1 for $0 \leq p_1 \leq b/2$.

<u>Proof</u> Write $G_n(p_1,t) = P[pX_1 + qX_2 \leq t' - Y]$ where

$p = p_1/b$, $q = p_2/b$, $t' = t/b$, and $Y = \sum_{i=3}^{n} \frac{p_i}{b} X_i$. Suppose Y has density h. Since f is symmetric unimodal, then so is h, Wintner (1938).

Write $G_n(p_1,t) = \int_{-\infty}^{\infty} G_2(p,t' - y)h(y)dy$, so that

$$b \frac{\partial G_n}{\partial p_1} = \int_{-\infty}^{\infty} \frac{\partial G_2(p,t' - y)}{\partial p} h(y)dy \; .$$

Differentiation under the integral sign is permissible as before. Rewrite

$$b \frac{\partial G_n}{\partial p_1} = \int_{-\infty}^{t'} \frac{\partial G_2(p,t' - y)}{\partial p} h(y)dy + \int_{t'}^{\infty} \frac{\partial G_2(p,t' - y)}{\partial p} h(y)dy \; .$$

Let $v = t' - y$ in the first integral and $v = y - t'$ in the second. Using the fact that $\frac{\partial G(p,v)}{\partial p} = - \frac{\partial G_2(p, -v)}{\partial p}$ we get

$$b \frac{\partial G_n}{\partial p_1} = \int_{0}^{\infty} \frac{\partial G_2(p,v)}{\partial p} \{h(t' - v) - h(t' + v)\}dv.$$

Now $\frac{\partial G_2(p,v)}{\partial p} > 0$ for $0 < v < a$. Also since h is symmetric unimodal $h(t' - v) \geq h(t' + v)$. Thus $\frac{\partial G_n}{\partial p_1} > 0$.

If $\frac{\partial G_n}{\partial p_1} = 0$, then $h(t' - v) \equiv h(t' + v)$ for all v in $(0,a)$ except for at most two points, i.e., h is uniform. But this implies $a < \infty$. Thus $h(t' + u) \equiv 0$ for $u > a - t'$ so that h is identically 0, an obvious contradiction. Thus $\frac{\partial G_n}{\partial p_1} > 0$. ||

In Hardy, Littlewood, Pólya (1952), Chapter II, the concept of majorization is defined, and its equivalence to several related concepts is shown. A vector $\underline{\alpha} = (\alpha_1, \cdots, \alpha_n)$ is said to <u>majorize</u> a vector $\underline{\alpha}' = (\alpha_1', \cdots, \alpha_n')$ (written $\underline{\alpha} > \underline{\alpha}'$) if the components can be arranged so that

(i) $$\sum_{i=1}^{n} \alpha_i = \sum_{i=1}^{n} \alpha_i' ;$$

(ii) $$\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n \geq 0 \qquad \alpha_1' \geq \alpha_2' \geq \cdots \geq \alpha_n' \geq 0;$$

(iii) $$\sum_{j=1}^{i} \alpha_j \geq \sum_{j=1}^{i} \alpha_j' \quad \text{for} \quad 1 \leq i \leq n.$$

A linear transformation T of $\underline{\alpha}$ is defined as follows. Consider a pair α_k, α_ℓ with $\alpha_k > \alpha_\ell$; write

$$\alpha_k = \rho + \tau, \ \alpha_\ell = \rho - \tau \ (0 < \tau \leq \rho).$$

If $0 \leq \delta < \tau$, then T is defined by

$$\alpha_k' = \frac{\tau + \delta}{2\tau} \alpha_k + \frac{\tau - \delta}{2\tau} \alpha_\ell = \rho + \delta$$

$$\alpha_\ell' = \frac{\tau - \delta}{2\tau} \alpha_k + \frac{\tau + \delta}{2\tau} \alpha_\ell = \rho - \delta$$

$$\alpha_j' = \alpha_j (j \neq k, \ j \neq \ell).$$

We write $\underline{\alpha}' = T\underline{\alpha}$.

Hardy, Littlewood, Pólya show that a necessary and sufficient condition that $\underline{\alpha} > \underline{\alpha}'$ but $\underline{\alpha}$ is not identical with $\underline{\alpha}'$ is that $\underline{\alpha}'$ can be derived from $\underline{\alpha}$ by a finite number of transformations T.

A third concept is that of forming averages. We say $\underline{\alpha}'$ is an <u>average</u> of $\underline{\alpha}$ if there exist $p_{ij} \geq 0$, $i = 1,\cdots,n$, $j = 1,\cdots,n$, such that

$$\sum_{i=1}^{n} p_{ij} = 1,\ j = 1,\cdots,n;\ \sum_{j=1}^{n} p_{ij} = 1,\ i = 1,\cdots,n;$$

$$\text{and}\quad \alpha_i' = \sum_{j=1}^{n} p_{ij}\alpha_j,\ i = 1,\cdots,n.$$

Hardy, Littlewood, Pólya show that a necessary and sufficient condition that $\underline{\alpha}'$ should be an average of $\underline{\alpha}$ is that $\underline{\alpha} > \underline{\alpha}'$.

<u>Theorem 2.1</u> Let f be PF_2, $f(t) = f(-t)$ for all t, $f(t) > 0$ for $-a < t < a$ and 0 elsewhere, $0 < a \leq \infty$, $X_1,\cdots,X_n$ independently distributed with density f, $H(t) = P\left[\sum_{i=1}^{n} p_i X_i \leq t\right]$, $H'(t) = P\left[\sum_{i=1}^{n} p_i' X_i \leq t\right]$, $\underline{p} > \underline{p}'$, $\underline{p}$, $\underline{p}'$ not identical, $\sum_{i=1}^{n} p_i = 1 = \sum_{i=1}^{n} p_i'$. Then for $0 < t < a$, $H'(t) > H(t)$.

<u>Proof</u> $\underline{p}'$ can be obtained from $\underline{p}$ by a finite number of T transformations. Applying Lemma 2 as many times, we obtain the desired conclusion. ||

We may define a continuous version of majorization. Let

(i) $\int_a^b h(t)dt = \int_a^b h'(t)dt$, where a and/or b may be infinite;

(ii) $h(t)$, $h'(t)$ be nonnegative and decreasing for $a < t < b$;

(iii) $\int_a^x h(t)dt \geq \int_a^x h'(t)dt$ for $a < x < b$.

Then we say that h majorizes h' on $[a,b]$.

In Theorem 2.1 if we let h and h' be the densities corresponding to the distribution functions H and H', then we see that the conclusion of the theorem states that h' majorizes h on $[0,\infty)$. Thus majorization in the weights of the convex combination of random variables carries over into majorization in the resulting densities (though in the reverse direction).

<u>Corollary 2.1</u> Let f be PF_2, $f(t) = f(-t)$ for all t, $f(t) > 0$ for $-a < t < a$ and 0 elsewhere, $0 < a \leq \infty$, $X_1, X_2, \cdots$ independently distributed with density f. Then for $0 < t < a$, $F^{(n)}(nt) = P\left[\frac{1}{n}\sum_1^n X_i \leq t\right]$ is strictly increasing in $n = 1, 2, \cdots$.

<u>Proof</u> Note that $\underline{p} = \left(\frac{1}{n}, \frac{1}{n}, \cdots, \frac{1}{n}, 0\right) > \underline{p}' = \left(\frac{1}{n+1}, \frac{1}{n+1}, \cdots, \frac{1}{n+1}, \frac{1}{n+1}\right)$, where each vector contains $n+1$ components. Thus Theorem 2.1 immediately yields the desired conclusion. ||

In words, for symmetric PF_2 densities peakedness of means increases with sample size.

We can extend the class of densities for which the conclusion of Theorem 2.1 and consequently that of Corollary 2.1 applies as follows. First we prove

<u>Lemma 2.3</u> Let $f_i(t) = f_i(-t)$ for all t, $f_i(t) > 0$, and decreasing for $0 < t < a$ and 0 elsewhere, $0 < a \leq \infty$, $i = 1,2$. Let $X_1,\cdots,X_n$ be independently distributed with density f_1, $Y_1,\cdots,Y_n$ be independently distributed with density f_2. Suppose $\underline{p} > \underline{p}'$ implies $\sum_{i=1}^n p_i X_i$ more peaked than $\sum_{i=1}^n p_i' X_i$ and $\sum_{i=1}^n p_i Y_i$ more peaked than $\sum_{i=1}^n p_i' Y_i$. Then $\underline{p} > \underline{p}'$ implies $\sum_{i=1}^n p_i (X_i + Y_i)$ more peaked than $\sum_{i=1}^n p_i'(X_i + Y_i)$.

<u>Proof</u> $\sum_{i=1}^n p_i X_i$, $\sum_{i=1}^n p_i Y_i$, $\sum_{i=1}^n p_i' X_i$, and $\sum_{i=1}^n p_i' Y_i$ are symmetric unimodal random variables, Wintner (1938). Hence by the lemma of Birnbaum (1948) the result follows. ||

Note that if $X_1,\cdots,X_n$ are independently distributed with Cauchy density $g_a(x) = \dfrac{a}{\pi(1 + a^2x^2)}$, $a > 0$, then $\sum_{i=1}^n p_i X_i (0 \leq p_i < 1$, $\sum_{i+1}^n p_i = 1)$ is distributed with the same density. Note too that if X_1 and X_2 are independent Cauchy variates with corresponding densities g_{a_1} and g_{a_2}, then $X_1 + X_2$ is also a Cauchy variate with density g_a for appropriate a. We may now state

<u>Theorem 2.2</u> Let f be PF_2, with $f(t) = f(-t)$, $f(t) > 0$ for $-a < t < a$ and 0 elsewhere, $0 < a \leq \infty$. Let $X_1,\cdots,X_n$ be independently distributed with density $f * g_a$,

$$H(t) = P\left[\sum_{i=1}^n p_i X_i \leq t\right],\quad H'(t) = P\left[\sum_{i=1}^n p_i' X_i \leq t\right],\quad \underline{p} > \underline{p}',\ \underline{p},\ \underline{p}'$$

not identical, $\sum_{i=1}^{n} p_i = 1 = \sum_{i=1}^{n} p_i'$. Then for $0 < t < a$, $H'(t) > H(t)$.

Proof From Lemma 2.3 it follows that $H'(t) \geq H(t)$. Strictness follows from the fact that corresponding strictness holds for the PF_2 component of the convolution. ||

Thus Theorem 2.1 holds when the underlying density is the convolution of a symmetric PF_2 density and a Cauchy density.

It is worthwhile to consider symmetric distributions for which the conclusions of Theorem 2.1 do not hold. One such is the Cauchy distribution with density $f(x) = \frac{a}{\pi(1 + a^2x^2)}$; it represents a "boundary" distribution in that if $X_1, \cdots, X_n$ are independently distributed with density f, then $\sum_{i=1}^{n} p_i X_i$, $p_i \geq 0$, $i = 1, \cdots, n$, $\sum_{i=1}^{n} p_i = 1$, also has Cauchy density f. We can actually produce a distribution G such that if Y_1, Y_2 are independently distributed with distribution G, then for $0 < t < \infty$

$$G^{(2)}(2t) = P[\tfrac{1}{2}Y_1 + \tfrac{1}{2}Y_2 \leq t] < P[Y_1 \leq t] = G(t).$$

Lemma 2.4 Let X_1, X_2 be independently distributed with density $f(x) = \frac{a}{\pi(1 + a^2x^2)}$. Let $\phi(x)$ be strictly convex and increasing for $0 \leq x < \infty$ and $\phi(x) = -\phi(-x)$ for all x. Define $Y_i = \phi(X_i)$, $i = 1,2$. Then for $t > 0$ $P[\tfrac{1}{2}Y_1 + \tfrac{1}{2}Y_2 \leq t] < P[Y_1 \leq t]$.

Proof For $X_1, X_2 \geq 0$ but not both 0, $\phi(\tfrac{1}{2}X_1 + \tfrac{1}{2}X_2) < \tfrac{1}{2}\phi(X_1) + \tfrac{1}{2}\phi(X_2)$. By symmetry for $X_1, X_2 \leq 0$ but not both 0,

$|\phi(\tfrac{1}{2}X_1 + \tfrac{1}{2}X_2)| < |\tfrac{1}{2}\phi(X_1) + \tfrac{1}{2}\phi(X_2)|$. For $X_1 \leq 0$, $X_2 > 0$, $|X_1| < X_2$, we have $\phi(\tfrac{1}{2}X_1 + \tfrac{1}{2}X_2) = \phi(\tfrac{1}{2}(X_2 - |X_1|)) < \tfrac{1}{2}\phi(X_2 - |X_1|) \leq \tfrac{1}{2}\phi(X_2) - \tfrac{1}{2}\phi(|X_1|) = \tfrac{1}{2}\phi(X_1) + \tfrac{1}{2}\phi(X_2)$. By symmetry, for $X_1 < 0$, $X_2 \geq 0$, $|X_1| > X_2$, $|\phi(\tfrac{1}{2}X_1 + \tfrac{1}{2}X_2)| < |\tfrac{1}{2}\phi(X_1) + \tfrac{1}{2}\phi(X_2)|$. Thus for all X_1, X_2 for which $X_1 + X_2 \neq 0$, $|\phi(\tfrac{1}{2}X_1 + \tfrac{1}{2}X_2)| < |\tfrac{1}{2}\phi(X_1) + \tfrac{1}{2}\phi(X_2)|$.

But $\tfrac{1}{2}X_1 + \tfrac{1}{2}X_2$ has the same distribution as X_1. Thus $|Y_1|$ is strictly stochastically smaller than $|\tfrac{1}{2}Y_1 + \tfrac{1}{2}Y_2|$ by Lemma 1, page 73, of Lehmann (1959). The result follows. ||

Thus the distribution of the mean of two is actually less peaked than that of a single random variable. In analogous fashion we may show

Lemma 2.5 Let X_1, X_2 be independently distributed with density $f(x) = \dfrac{a}{\pi(1 + a^2x^2)}$. Let $\phi(x)$ be strictly concave and increasing for $0 \leq x < \infty$ and $\phi(x) = -\phi(-x)$ for all x. Define $Y_i = \phi(X_i)$, $i = 1,2$. Then for $t > 0$, $P[\tfrac{1}{2}Y_1 + \tfrac{1}{2}Y_2 \leq t] > P[Y_1 \leq t]$.

Note that a very strong form of stochastic comparison is involved, since for each sample outcome in Lemma 2.3, (2.4), $|Y| <(>) |\tfrac{1}{2}Y_1 + \tfrac{1}{2}Y_2|$. It does not seem possible to use the same method to obtain stochastic comparisons between averages of n and $n + 1$ variables for $n > 1$. However, using Birnbaum's lemma we can obtain stochastic comparisons between averages of 2^n and 2^{n+1} variables, $n = 1,2,\cdots$.

BIBLIOGRAPHY

E. F. Beckenbach and R. Bellman (1961), "Inequalities," Springer-Verlag, Berlin.

Z. W. Birnbaum (1948), "On Random Variables with Comparable Peakedness," Annals of Mathematical Statistics, Vol. 19, No. 1, pp. 76-81.

H. Cramér (1946), "Mathematical Methods of Statistics," Princeton University Press, Princeton.

G. H. Hardy, J. E. Littlewood, and G. Pólya (1952), "Inequalities," Cambridge University Press.

E. L. Lehmann (1959), "Testing Statistical Hypotheses," John Wiley and Sons, Inc., New York.

A. Wintner (1938), "Asymptotic Distributions and Infinite Convolutions," Edwards Brothers, Ann Arbor.