PEAKEDNESS OF DISTRIBUTIONS OF CONVEX COMBINATIONS[1]

By Frank Proschan

Boeing Scientific Research Laboratories and University of California, Berkeley

1. 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

$$P[|X_1 - a_1| \geqq t] \leqq P[|X_2 - a_2| \geqq t] \tag{1.1}$$

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

If the inequality between the two probabilities in (1.1) is strict whenever the two probabilities are not both 0, we say X_1 is strictly more peaked about a_1 than X_2 about a_2.

2. 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, X_1 and X_2 independently distributed with density f. Then $pX_1 + qX_2$ is strictly increasing in peakedness as p increases from* 0 *to* $\frac{1}{2}$, *with* $p + q = 1$.

Proof. For $0 < p < \frac{1}{2}$, define

$$G_2(p, t) = P[pX_1 + qX_2 \leqq t] = \int_{-\infty}^{\infty} F((t - qu)/p)f(u)\,du.$$

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

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

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

$$p^2(\partial G_2/\partial p) = \int_0^{\infty} v\{f(t + v)f(t - (qv/p)) - f(t - v)f(t + (qv/p))\}\,dv. \tag{1}$$

Received 16 November 1964; revised 20 May 1965.

[1] This work was partially supported by the Office of Naval Research under Contract Nonr-3656(18).

By symmetry of f,

$$f(t + v)f(t - (qv/p)) - f(t - v)f(t + (qv/p))$$
$$= f(t + v)f(-t + (qv/p)) - f(-t + v)f(t + (qv/p)) \geqq 0,$$

since f is PF_2, $t > 0$, and $q/p > 1$. Thus $p^2(\partial G_2/\partial p) \geqq 0$, so that $\partial G_2/\partial p \geqq 0$.

Now suppose $\partial G_2/\partial p = 0$. Then for all $v \geqq 0$ except for at most two points (a PF_2 density is continuous except for at most two points), from (1) we have $f(t + v)f(t - (qv/p)) - f(t - v)f(t + (qv/p)) \equiv 0$. Since f is a symmetric PF_2, f has a mode at 0. Thus $f(u)$ must be constant on its interval of support, that is, f is the uniform density on $(-a, a)$. However, for $(p/q)(a - t) < v < \min \{(p/q)(a + t), a - t\}$, $f(t + v)f(t - (qv/p)) - f(t - v)f(t + (qv/p)) > 0$. From this contradiction it follows that $\partial G_2/\partial p > 0$.

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

LEMMA 2.2. *Let f be PF_2, $f(t) = f(-t)$ for all t, $X_1, \cdots, X_n$ independently distributed with density f. Then $\sum_{i=1}^n p_iX_i$ is strictly increasing in peakedness as p_1 increases from 0 to $\frac{1}{2}b$, with $p_1 + p_2 = b$, $0 < b \leqq 1$, $p_i \geqq 0$, $i = 1, \cdots, n$, and $\sum_{i=1}^n p_i = 1$.*

PROOF. First note that $\sum_{i=1}^2 p_iX_i$ and $\sum_{i=3}^n p_iX_i$ are each symmetric unimodal random variables since each X_i is. (See Wintner (1938.)) Suppose $p_1 < p_1'$, $p_1 < p_2$, $p_1' < p_2'$, $p_1 + p_2 = b = p_1' + p_2'$. Then by Lemma 2.1, $p_1X_1 + p_2X_2$ is less peaked than $p_1'X_1 + p_2'X_2$. By the lemma of Birnbaum (1948), it follows that $\sum_{i=1}^n p_iX_i$ is less peaked than $\sum_{i=1}^2 p_i'X_i + \sum_{i=3}^n p_iX_i$. Finally the strictness in the conclusion of Lemma 2.2 follows from the corresponding strictness in Lemma 2.1. ||

To state the main result, we discuss majorization. A vector $\mathbf{b} = (b_1, \cdots, b_n)$ is said to be *majorized* by a vector $\mathbf{a} = (a_1, \cdots, a_n)$, written $\mathbf{a} \rangle \mathbf{b}$, if the components can be arranged so that $a_1 \geqq \cdots \geqq a_n$, $b_1 \geqq \cdots \geqq b_n$, $\sum_{i=1}^k a_i \geqq \sum_{i=1}^k b_i$, $k = 1, 2, \cdots, n - 1$, and $\sum_{i=1}^n a_i = \sum_{i=1}^n b_i$. If $\mathbf{a} \rangle \mathbf{b}$, then $\mathbf{b}$ can be derived from $\mathbf{a}$ by a finite number of transformations T of the form

$$T(\mathbf{a}) = \alpha(a_1, \cdots, a_n)$$
$$+ (1 - \alpha)(a_1, \cdots, a_j, a_k, a_{j+1}, \cdots, a_{k-1}, a_j, a_{k+1}, \cdots, a_n), \quad 0 \leqq \alpha \leqq 1.$$

(See Hardy, Littlewood, Pólya (1952), p. 47.) We may now obtain

THEOREM 2.3. *Let f be PF_2, $f(t) = f(-t)$ for all t, $X_1, \cdots, X_n$ independently distributed with density f, $\mathbf{p} \rangle \mathbf{p}'$, $\mathbf{p}$, $\mathbf{p}'$ not identical, $\sum_{i=1}^n p_i = 1 = \sum_{i=1}^n p_i'$. Then $\sum_{i=1}^n p_i'X_i$ is strictly more peaked than $\sum_{i=1}^n p_iX_i$.*

PROOF. $\mathbf{p}'$ can be obtained from $\mathbf{p}$ by a finite number of T transformations. Applying Lemma 2.2 in each case, we obtain the desired conclusion. ||

An application of statistical interest is

COROLLARY 2.4. *Let f be PF_2, $f(t) = f(-t)$ for all t, $X_1, X_2, \cdots$ independently distributed with density f. Then $(1/n) \sum_{i=1}^n X_i$ is strictly increasing in peakedness as n increases over the positive integers.*

PROOF. Note that $\mathbf{p} = (1/n, 1/n, \cdots, 1/n, 0) \rangle \mathbf{p}' = (1/(n + 1), 1/(n + 1),$

$\cdots, 1/(n+1), 1/(n+1))$, where each vector contains $n+1$ components. The result follows immediately from Theorem 2.3. ||

We can extend the class of densities for which the conclusion of Theorem 2.3 and consequently that of Corollary 2.4 applies. First we prove

LEMMA 2.5. *Let $f_i(t) = f_i(-t)$ for all t, $f_i(t)$ decreasing for $0 < t < \infty$, $i = 1, 2$. 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 $\mathbf{p} \rangle \mathbf{p}'$ implies $\sum_{i=1}^n p_i' X_i$ is more peaked than $\sum_{i=1}^n p_i X_i$ and $\sum_{i=1}^n p_i' Y_i$ is more peaked than $\sum_{i=1}^n p_i Y_i$. Then $\mathbf{p} \rangle \mathbf{p}'$ implies $\sum_{i=1}^n p_i' (X_i + Y_i)$ is more peaked than $\sum_{i=1}^n p_i(X_i + Y_i)$.*

PROOF. $\sum_{i=1}^n p_i X_i$, $\sum_{i=1}^n p_i Y_i$, $\sum_{i=1}^n p_i' X_i$, $\sum_{i=1}^n p_i' Y_i$ are symmetric unimodal random variables. See 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,

$$g_a(x) = a/\pi(1 + a^2 x^2), \qquad a > 0, \tag{2.1}$$

then $\sum_{i=1}^n p_i X_i (0 \leqq p_i \leqq 1, \sum_{i=1}^n p_i = 1)$ 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

THEOREM 2.6. *Let f be PF_2, with $f(t) = f(-t)$, $X_1, \cdots, X_n$ be independently distributed with density $f * g_a$, where g_a is defined in (2.1), $\mathbf{p} \rangle \mathbf{p}'$, $\mathbf{p}$, $\mathbf{p}'$ not identical, and $\sum_{i=1}^n p_i = 1 = \sum_{i=1}^n p_i'$. Then $\sum_{i=1}^n p_i' X_i$ is strictly more peaked than $\sum_{i=1}^n p_i X_i$.*

PROOF. From Lemma 2.5 it follows that $\sum_{i=1}^n p_i' X_i$ is more peaked than $\sum_{i=1}^n p_i X_i$. The strictness follows from the fact that corresponding strictness holds for the PF_2 component of the convolution. ||

Thus Theorem 2.3 and Corollary 2.4 hold when the underlying density is the convolution of a symmetric PF_2 density and a Cauchy density.

It is of interest of consider symmetric distributions for which the conclusions of Theorem 2.3 do not hold. One such is the Cauchy with density g_a. Actually we can produce a distribution G such that if Y_1 and Y_2 are independently distributed according to G, then $\frac{1}{2}Y_1 + \frac{1}{2}Y_2$ is strictly less peaked than Y_1.

LEMMA 2.7. *Let X_1 and X_2 be independently distributed with density g_a defined in (2.1). Let $\phi(x)$ be strictly convex and increasing for $0 \leqq x < \infty$ and $\phi(x) = -\phi(-x)$ for all x. Define $Y_i = \phi(X_i)$, $i = 1, 2$. Then $\frac{1}{2}Y_1 + \frac{1}{2}Y_2$ is strictly less peaked than Y_1.*

PROOF. For $X_1, X_2 \geqq 0$ but not both 0, $\phi(\frac{1}{2}X_1 + \frac{1}{2}X_2) < \frac{1}{2}\phi(X_1) + \frac{1}{2}\phi(X_2)$. By symmetry for $X_1, X_2 \leqq 0$ 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 $X_1 \leqq 0$, $X_2 > 0$, $|X_1| < X_2$, we have

$$\begin{aligned}\phi(\tfrac{1}{2}X_1 + \tfrac{1}{2}X_2) &= \phi(\tfrac{1}{2}(X_2 - |X_1|)) < \tfrac{1}{2}\phi(X_2 - |X_1|) \\ &\leqq \tfrac{1}{2}\phi(X_2) - \tfrac{1}{2}\phi(|X_1|) = \tfrac{1}{2}\phi(X_1) + \tfrac{1}{2}\phi(X_2).\end{aligned}$$

By symmetry, for $X_1 < 0$, $X_2 \geqq 0$, $|X_1| > X_2$,

$$|\phi(\tfrac{1}{2}X_1 + \tfrac{1}{2}X_2)| < |\tfrac{1}{2}\phi(X_1) + \tfrac{1}{2}\phi(X_2)|.$$

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 $\frac{1}{2}X_1 + \frac{1}{2}X_2$ has the same distribution as X_1. Thus $|Y_1|$ is strictly stochastically smaller than $|\frac{1}{2}Y_1 + \frac{1}{2}Y_2|$ by Lemma 1, p. 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 2.8. *Let X_1, X_2 be independently distributed with density $g_a(x) = a/\pi(1 + a^2x^2)$. Let $\phi(x)$ be strictly concave and increasing for* $0 \leqq x < \infty$ and *$\phi(x) = -\phi(-x)$ for all x. Define $Y_i = \phi(X_i)$, $i = 1, 2$. Then for $t > 0$, $P[\frac{1}{2}Y_1 + \frac{1}{2}Y_2 \leqq t] > P[Y_1 \leqq t]$.*

Note that a very strong form of stochastic comparison is involved, since for each sample outcome in Lemma 2.7, (2.8), $|Y| < (>) |\frac{1}{2}Y_1 + \frac{1}{2}Y_2|$. 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, $n = 1, 2, \cdots$.

REFERENCES

BECKENBACH, E. F. and BELLMAN, R. (1961). *Inequalities*. Springer-Verlag, Berlin.

BIRNBAUM, Z. W. (1948). On random variables with comparable peakedness. *Ann. Math. Statist.* **19** 76–81.

CRAMÉR, H. (1946). *Mathematical Methods of Statistics*. Princeton Univ. Press, Princeton.

HARDY, G. H., LITTLEWOOD, J. E., and PÓLYA, G. (1952). *Inequalities*. Cambridge Univ. Press.

LEHMANN, E. L. (1959). *Testing Statistical Hypotheses*. Wiley, New York.

WINTNER, A. (1938). *Asymptotic Distributions and Infinite Convolutions*. Edwards Brothers, Ann Arbor.