PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 131, Number 5, Pages 1601–1606
S 0002-9939(02)06655-8
Article electronically published on September 19, 2002

ON A CHARACTERIZATION OF MEASURES OF DISPERSION

A. S. FAINLEIB

(Communicated by Claudia M. Neuhauser)

ABSTRACT. Measures of dispersion are characterized by the set of all bounded random variables whose dispersion is minimized when taken around the origin.

1. INTRODUCTION

Let φ be a real valued function on $\mathbf{R}$, X a bounded random variable (b.r.v.), and a a real number. The functional $E\varphi(X-a)$ may be used as a measure of dispersion of X around a. The base of the measure is the set of all b.r.v. X such that

$$\min_a E\varphi(X-a) = E\varphi(X). \tag{1}$$

For example, the base of the first absolute moment $E|X-a|$ is the set of all b.r.v. with zero median; the base of the second moment $E(X-a)^2$ is the set of all b.r.v. with zero mean value.

In this paper, we consider a characterization of the measures of dispersion by their bases. Kagan and Shepp [2] proved that if φ is continuous and the base of the measure $E\varphi(X-a)$ contains all b.r.v. with $EX=0$, then $\varphi(x)=\alpha x^2+\varphi(0)$ with some $\alpha \geq 0$, and they also obtained a multivariate version of the result.

In what follows all the functions are real valued; f is a non-negative continuous function on $\mathbf{R}$ with $f(0)=0$; B_φ denotes the base of the measure $E\varphi(X-a)$ (so B_0 is the set of all b.r.v.).

Theorem 1. *Let f satisfy the following conditions:*

$$f(x) \text{ does not vanish identically on } (-\infty,0) \text{ or on } (0,\infty) \tag{2}$$

and

$$y\int_0^z \{f(x+y)-f(x)-f(y)\}\,dx \geq 0 \quad \text{for any } y,z\in \boldsymbol{R}. \tag{3}$$

If

$$\varphi \text{ is continuous on } \boldsymbol{R} \text{ and } B_f \subseteq B_\varphi, \tag{4}$$

then

$$\varphi(x) = \alpha f(x) + \varphi(0) \tag{5}$$

with some $\alpha \geq 0$.

Received by the editors July 5, 2001 and, in revised form, December 12, 2001.
2000 *Mathematics Subject Classification.* Primary 60E15.

In particular, if f is convex on $\mathbf{R}$, then (2) is equivalent to $f(\pm\infty) = \infty$. Moreover, in this case, the difference $f(x+y) - f(x)$ is an increasing function of x for any fixed $y > 0$ (see, for example, [1, 3.18]). Therefore, (3) is fulfilled and we have the following:

Corollary. *If f is convex on $\boldsymbol{R}$ and $f(\pm\infty) = \infty$, then* (4) *implies* (5).

The bases of convex measures are described in the last section. Note that convexity of f on $\mathbf{R}$ is not necessary for (3). For example, the function

$$f(x) = x^2(x^2 - 3x + 3)$$

satisfies (2) and (3) but is not convex on $\mathbf{R}$.

Theorem 2. *Let f be absolutely continuous on each finite interval and satisfy* (2). *Moreover, let g be defined on $\boldsymbol{R}$, bounded on each finite interval, $g(0) = 0$ and $g(x) = f'(x)$ at all the points of differentiability of f (hence, almost everywhere). If φ is continuous on $\boldsymbol{R}$ and B_φ contains all $X \in B_0$ with $Eg(X) = 0$, then* (5) *holds.*

Condition (2) is essential. The functions

$$f(x) = (x + |x|)^2, g(x) = 4(x + |x|) \quad \text{and } \varphi(x) = (x + |x|)^3$$

satisfy all the conditions of Theorems 1 and 2 except (2). Moreover,

$$B_f = B_\varphi = \{X \in B_0 : P(X > 0) = 0\},$$

and $Eg(X) = 0$ is equivalent to $X \in B_\varphi$. However, (5) is obviously not valid in this case.

The functions $f(x) = |x|$ and $g(x) = \operatorname{sign} x$ satisfy all the conditions of Theorem 2. It follows from $E \operatorname{sign} X = 0$ that X has zero median. So if B_φ contains all b.r.v. with zero median, then we have (5) with $f(x) = |x|$ (this also follows from the Corollary). The result holds under more general conditions (in particular, the function φ may be a priori discontinuous).

Theorem 3. *Let φ be a function on $\boldsymbol{R}$ bounded from either above or below on some interval and let $0 < p < 1$. If B_φ contains all binary r.v. X with $\min X \le 0 \le \max X$ and $P(X = \min X) = p$, then* (5) *holds with*

$$f(x) = |x| + (2p - 1)x. \tag{6}$$

Note that in this case $B_f = \{X \in B_0 : P(X < 0) \le p \le P(X \le 0)\}$ (so that B_f consists of all bounded r.v. with zero quantile of order p).

2. Proof of Theorems 1 and 2

Let Y_w denote an r.v. equal to w with probability 1,

$$M = \{x \in \mathbf{R} : f(x) > 0\}$$

and $[M]$ is the closure of M. Set, moreover, for $u, v \in M$ and $u < 0 < v$ (there exist the such u and v in view of (2))

$$\lambda = \lambda(u, v) = \{vf(u) - uf(v)\}^{-1}.$$

Let $Y = Y(u, v)$ be an r.v. with the distribution function $F(x) = F(x, u, v)$ and

$$F(x) = \begin{cases} \lambda f(v)(x - u) & \text{for} \quad x \in [u, 0], \\ \lambda\{f(u)x - uf(v)\} & \text{for} \quad x \in [0, v]. \end{cases}$$

Lemma 1. *Let f satisfy* (2). *If B_φ contains Y_0, Y_w for $w \notin [M]$ and $Y(u,v)$ for $u, v \in M, u < 0 < v$, then* (5) *holds with some $\alpha \geq 0$.*

Proof. It follows from $Y_w \in B_\varphi$ that

$$\varphi(w) = E\varphi(Y_w) = \min_a E\varphi(Y_w - a) = \min_t \varphi(t).$$

Therefore,

$$\varphi(0) = \varphi(w) = \min_t \varphi(t) \tag{7}$$

for all $w \notin [M]$ and we obtain (5) for all $x \notin [M]$. Now let $u, v \in M, u < 0 < v$. Putting for any integrable function r

$$E_r(z) = Er(Y+z) = \lambda\{f(v)\int_{z+u}^{z} r(x)\,dx + f(u)\int_{z}^{z+v} r(x)\,dx\}, \tag{8}$$

and taking into account that $Y(u,v) \in B_\varphi$, we get $E'_\varphi(0) = 0$, since φ is continuous so $E_\varphi(z)$ is differentiable. Hence

$$s(u) = s(v) \quad \text{for } u, v \in M, u < 0 < v,$$

where

$$s(x) = \frac{\varphi(x) - \varphi(0)}{f(x)}.$$

It follows that $s(x)$ has the same value α for all $x \in M$, so we have (5) for all such x. Since f and φ are continuous, it implies (5) for all $x \in [M]$ and thus for all real x. It follows from (5) and (7) that

$$\min_x \alpha f(x) = 0$$

so $\alpha \geq 0$. □

To prove Theorem 1, it is enough now to show that

$$Y_0, Y_w, Y(u,v) \in B_f \quad \text{for any } u, v \in M, u < 0 < v, \quad \text{and any } w \notin [M].$$

Since

$$f(w) = f(0) = 0 = \min_t f(t),$$

we have $Y_0, Y_w \in B_f$. It follows from (3) and (8) that

$$\int_0^z \frac{f(x+u) - f(x)}{f(u)}\,dx \leq z \leq \int_0^z \frac{f(x+v) - f(x)}{f(v)}\,dx$$

and

$$E_f(z) \geq E_f(0) \quad \text{for all } z \in \mathbf{R},$$

so $Y(u,v) \in B_f$ for $u, v \in M, u < 0 < v$. □

Similarly, to prove Theorem 2, it is enough to show that $Eg(X) = 0$ for

$$X = Y_0, Y_w, Y(u,v), \quad \text{where } u, v \in M, u < 0 < v, \quad \text{and } w \notin [M].$$

Indeed, $Eg(Y_0) = g(0) = 0$. If $w \notin [M]$, then $f(x) = 0$ in some open interval containing w; therefore, also in this interval, $g(x) = f'(x) = 0$, so

$$Eg(Y_w) = g(w) = 0.$$

Moreover, it follows from (8) that

$$Eg\{Y(u,v)\} = E_g(0) = E_f'(0) = 0$$

because f is absolutely continuous and so $f(x) = \int_0^x g(t)\,dt$ (see, for example, [4, 11.7]). $\square$

3. Proof of Theorem 3

Continuity of φ is essential for the proof of Theorems 1 and 2. Therefore, we now use another approach.

Let $U = U_p(u,v)$ $(u < v)$ be a binary r.v. defined by

$$P(U = u) = p, \quad P(U = v) = q = 1 - p. \tag{9}$$

Let $x > 0$ and $u \in [0,x]$. Then the r.v. $U_p(0,x)$ and $U_p(-u, x-u)$ satisfy the conditions of Theorem 3. It follows from (1) and (9) that

$$p\varphi(-u) + q\varphi(x-u) \geq p\varphi(0) + q\varphi(x)$$

and

$$p\varphi(0) + q\varphi(x) \geq p\varphi(-u) + q\varphi(x-u),$$

whence

$$p\{\varphi(-u) - \varphi(0)\} = q\{\varphi(x) - \varphi(x-u)\}. \tag{10}$$

In particular, we have by setting $x = u$ that

$$p\{\varphi(-u) - \varphi(0)\} = q\{\varphi(u) - \varphi(0)\}. \tag{11}$$

It follows from (10) and (11) that

$$\varphi(x) - \varphi(x-u) = \varphi(u) - \varphi(0) \quad \text{for } x \geq 0, u \in [0,x]$$

and (replacing x by $u+v$)

$$\psi(u+v) = \psi(u) + \psi(v) \quad \text{for any } u, v \geq 0, \tag{12}$$

where $\psi(x) = \varphi(x) - \varphi(0)$. So both the functions ψ and $-\psi$ are convex on $[0,\infty)$ [1, 3.20]. Since one of them is bounded from above on some interval, they are continuous [1, 3.18] and therefore linear [1, 3.19]. Thus $\psi(x) = \beta x$, where β is a constant, and

$$\psi(x) = \frac{\beta}{2p}\{|x| + (2p-1)x\} \quad \text{for } x \geq 0.$$

In view of (11), the last equality is also valid for $x < 0$. Setting $\alpha = \beta/2p$, we obtain (5) with f defined by (6). Finally, it follows from (5) and (1) for $X = U_p(0,1)$ that $\alpha \geq 0$. $\square$

Remark. According to the known Blumberg-Sierpinski theorem [3], every measurable convex function is continuous. So the proof shows that the condition on φ in Theorem 3 may be replaced by measurability of φ.

4. Convex measures of dispersion

A convex measure of dispersion is a measure $E\varphi(X-a)$ generated by a convex continuous function φ. The bases of the such measures may be described as follows.

Theorem 4. *If φ is convex and continuous on $\boldsymbol{R}$, then*

$$B_\varphi = \{X \in B_0 : E\varphi'_-(X) \le 0 \le E\varphi'_+(X)\}, \tag{13}$$

where φ'_- and φ'_+ denote the left and right derivatives of φ, respectively.

In particular, if φ is convex and differentiable on $\mathbf{R}$, then

$$B_\varphi = \{X \in B_0 : E\varphi'(X) = 0\}.$$

Proof. The proof of Theorem 4 is based on the following lemmas.

Lemma 2. *Let functions $\psi_n(x)$ $(n = 1, 2, \dots)$ and their variations be uniformly bounded on an interval $[a, b]$ and let*

$$\lim_{n\to\infty} \psi_n(x) = \psi(x) \quad \text{for each } x \in [a, b].$$

If $K(x)$ is a function of bounded variation on $[a, b]$, then

$$\lim_{n\to\infty} \int_a^b \psi_n(x)\, dK(x) = \int_a^b \psi(x)\, dK(x).$$

It is enough to prove it for the cases in which $\psi(x) \equiv 0$ and $K(x)$ is either continuous or discrete on $[a, b]$. In the first case, it follows from the known Helly's theorem by integration by parts. In the second case,

$$I_n = \int_a^b \psi_n(x)\, dK(x) = \sum_m \psi_n(x_m) h_m,$$

where $m = 1, 2, \dots,$ x_m runs over all the points of discontinuity of $K(x)$ on $[a, b]$ and h_m are the corresponding jumps, so that

$$\sum_m |h_m| < \infty.$$

Let $A > 0$, $|\psi_n(x)| \le A$ for all $x \in [a, b], n = 1, 2, \dots,$ and let $\varepsilon > 0$ and

$$\sum_{m>N} |h_m| \le \varepsilon / A,$$

where $N = N(\varepsilon)$. Then

$$|I_n| \le \sum_{m\le N} |\psi_n(x_m) h_m| + \varepsilon,$$

whence it follows that

$$\limsup_{n\to\infty} |I_n| \le \varepsilon, \quad \text{so } \lim_{n\to\infty} I_n = 0,$$

because $\psi_n(x) \to 0$ and $\varepsilon > 0$ is arbitrary. □

Lemma 3. *Let the functions $\psi_n(x)$ increase on $\boldsymbol{R}$ and be uniformly bounded on each finite interval. If*

$$\lim_{n\to\infty} \psi_n(x) = \psi(x) \quad \textit{for all real } x,$$

then

$$\lim_{n\to\infty} E\psi_n(X) = E\psi(X) \quad \textit{for all } X \in B_0.$$

It follows immediately from the previous lemma.

Lemma 4. *Let $\tau(x)$ be a convex continuous function on $\boldsymbol{R}$. Then:*

(i) *the ratio*

$$\frac{\tau(x+h)-\tau(x)}{h} \quad (h \neq 0)$$

is an increasing function of x and h, bounded for bounded x and h;

(ii) *the equality*

$$\tau(x_0) = \min_x \tau(x)$$

is equivalent to

$$\tau'_-(x_0) \leq 0 \leq \tau'_+(x_0).$$

It follows from known properties of convex functions [1, 3.18].

To prove Theorem 4, note that the function $\mu(x) = E\varphi(X+x)$ is also convex and continuous on $\mathbf{R}$ for any fixed $X \in B_0$. By Lemmas 3 and 4,

$$\mu'_\pm(0) = \lim_{h\to\pm 0} E\frac{\varphi(X+h)-\varphi(X)}{h} = E\varphi'_\pm(X). \tag{14}$$

By Lemma 4, $X \in B_\varphi$ if and only if $\mu'_-(0) \leq 0 \leq \mu'_+(0)$. Taking (14) into account, we obtain (13). □

References

[1] G.H. Hardy, J.E. Littlewood, and G. Polya, *Inequalities.* University Press, Cambridge(1934).
[2] A. Kagan and L.A. Shepp, *Why the variance?*, Statist. Probab. Lett. 38(1998), 329-333. MR **99c:**60031
[3] W. Sierpinski, *Sur les fonctions convexes mesurable*, Fundamenta Math. 1(1920), 125-129.
[4] E.C. Titchmarsh, *The Theory of Functions.* University Press, Oxford(1939).

9-4990 Ed. Montpetit, Montreal, Quebec, Canada H3W 1P9
E-mail address: a_fainleib@hotmail.com