The Annals of Probability
1981, Vol. 9, No. 6, 1041–1042

THE NATURAL MEDIAN

DIETER LANDERS AND LOTHAR ROGGE

University of Cologne and University of Duisburg

In general the median of a random variable is not uniquely determined. This leads to difficulties in estimation and test theory. This paper suggests a limit process for choosing exactly one median. The choice uses the whole distribution of the random variable in a natural and reasonable way.

Consider a random variable X on a probability space. Denote by P and E the probability and expectation, respectively. Assume $X \in L_\varepsilon(P)$ for some $\varepsilon > 1$. For each $s \in (1, \varepsilon]$ let m_s be the s-mean of X, i.e., the unique number closest to X in $(L_s(P), \| \|_s)$. Denote by $[\underline{m}, \bar{m}]$ the interval of all medians of X and define

$$\rho(y) = E[\text{sign}(y - X) \log |y - X|] \qquad \text{for } y \in (\underline{m}, \bar{m}).$$

THEOREM. *If* $\varepsilon > 1$ *and* $X \in L_\varepsilon(P)$, *there is an* m_1 *in the interval* $[\underline{m}, \bar{m}]$ *of all medians of* X *such that* $m_s \to m_1$ *as* $s \downarrow 1$. *Moreover either* (i) $\rho(y) = 0$ *for exactly one* y *or* (ii) $\rho(y) > 0$ *for all* y *or* (iii) $\rho(y) < 0$ *for all* y *and* m_1 *satisfies* $\rho(m_1) = 0$ *in case* (i), $m_1 = \underline{m}$ *in case* (ii), $m_1 = \bar{m}$ *in case* (iii).

REMARK. The theorem suggests to define a "natural median", in a unique way, as the median m_1, described above. This m_1 is not only a median but it is also close to the s-mean m_s for s close to 1.

PROOF OF THE THEOREM. Write $a^{\underline{\delta}}$ for $|a|^\delta \text{ sign } a$ and define $\phi(y, \delta) = E(y - X)^{\underline{\delta}}$, $y \in \mathbb{R}$ and $0 \le \delta \le \varepsilon$. It is well known (see [1]) that the s-mean m_s fulfills

$$\phi(m_s, s - 1) = 0 \qquad 1 < s \le \epsilon. \tag{1}$$

Let $m_0 \in \bar{\mathbb{R}}$ be an accumulation point of m_s for $s \downarrow 1$. Then there exist $s_n \downarrow 1$ with $m_{s_n} \to m_0$. Write $\delta_n = s_n - 1$. If $m_0 \in [-\infty, \underline{m} - \eta)$ for some $\eta > 0$, then we obtain from (1) and the monotonicity of $\phi(\cdot, \delta)$ that

$$0 = \phi(m_{s_n}, \delta_n) \le \liminf_n \phi(\underline{m} - \eta, \delta_n) = P\{X < \underline{m} - \eta\} - P\{X > \underline{m} - \eta\}$$

contradicting the meaning of $\underline{m}$. Thus $m_0 \ge \underline{m}$; by symmetry $m_0 \le \bar{m}$. Hence $m_{s_n} \to \underline{m}$ if $\underline{m} = \bar{m}$, whence $\lim_{s\downarrow 1} m_s = \underline{m}$ in this case.

Now assume $\underline{m} < \bar{m}$: For each $y \in (\underline{m}, \bar{m})$ the function $[0, \infty) \ni \delta \to (y - X)^{\underline{\delta}}$ is twice differentiable P-a.e.

The first and second derivatives of this function are for $\delta \in [0, \epsilon/2]$ bounded by a fixed P-integrable function. Hence, as $E(y - X)^{\underline{0}} = 0$, we obtain from the Taylor expansion of $\delta \to \phi(y, \delta)$

$$\phi(y, \delta) = \delta\rho(y) + \tfrac{1}{2}\delta^2\phi''(y, \xi(\delta)) \tag{2}$$

with $0 < \xi(\delta) < \delta$ and

$$\sup_{0 \le \delta \le \varepsilon/2} |\phi''(y, \xi(\delta))| < \infty. \tag{3}$$

Received November 1, 1979; revised October 17, 1980.
AMS 1970 *subject classifications.* Primary 60G25; secondary 62G05.
Key words and phrases. Median, approximation.

Let $y \in (\underline{m}, \bar{m})$ with $\rho(y) > 0$. Then according to (2) and (3) eventually $\phi(y, \delta_n) > 0$. Since $0 = \phi(m_{s_n}, \delta_n) \geq \phi(y, \delta_n)$ if $m_{s_n} \geq y$, we obtain $m_{s_n} < y$ eventually and hence $m_0 \leq y$. Similarly $\rho(y) < 0$ implies $y \leq m_0$.

Since ρ is an increasing and continuous function, one of the cases (i), (ii), (iii) holds, and $m_0 = m_1$ as specified in the assertion. Since this holds for each accumulation point m_0, we have $\lim_{s\downarrow 1} m_s = m_1$.

In [2] the concept of s-means was extended from $L_s(P)$ to $L_{s-1}(P)$. The s-mean m_s of $X \in L_{s-1}(P)$ is the unique number fulfilling (1), i.e., $E(m_s - X)^{\underline{s-1}} = 0$. Using this generalization of an s-mean the assumption $\varepsilon > 1$ in the theorem above can be weakened to $\varepsilon > 0$.

Hence a natural median can be defined for each $X \in \cup_{\varepsilon>0} L_\varepsilon(P)$.

Acknowledgement. We thank the referee for improving the presentation.

REFERENCES

[1] Ando, T. and Amemiya, L. (1965). Almost everywhere convergence of prediction sequences in $L_p (1 < p < \infty)$. *Z. Wahrsch. verw. Gebiete* **4** 113–120.

[2] Landers, D. and Rogge, L. (1981). Isotonic approximation in L_s. *Journal of Approximation Theory.* **31** 199–223.

Mathematisches Institut der Universität Köln
Weyertal 86–90
D-5000 Köln 41
WEST GERMANY

Universität Duisburg
Fachbereich 11 - Mathematik
Lotharstr. 65
D-4100 Duisburg 1
WEST GERMANY