ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XIX, 2 (1987), pp. 239–240

R. ZIELIŃSKI (Warszawa)

ROBUSTNESS OF SAMPLE MEAN AND SAMPLE MEDIAN UNDER RESTRICTIONS ON OUTLIERS

Summary. Under restrictions on outliers, the sample mean may be a more robust estimator of the location that the median.

The problem is to estimate the location θ of the distribution $F_\theta(x) = F(x-\theta)$, where F is assumed to be symmetric ($F(x) = 1-F(-x)$) and unimodal (mode = 0). Suppose that the observations are ε-contaminated and their true distribution is $G_\theta(x) = G(x-\theta)$ such that $G = (1-\varepsilon)F+\varepsilon H$, where H is any distribution. We consider as estimators the statistics $T_n = T(G_n)$ derived from a translation invariant functional T; here G_n is the empirical distribution function. We are interested in finding such a T which minimizes the maximum asymptotic bias $\sup|T(G)-T(F)|$, where the supremum is taken over all H (see [1], p. 11). The well-known solution is the sample median (see [1], Section 4.2).

It is interesting to observe that the solution may be quite different if H is known to belong to a smaller class $\mathscr{H}$ of distributions.

THEOREM. *Let $\mathscr{H}$ be a class of distributions such that*

(i) $\operatorname{supp} H \subset [F^{-1}(\frac{1}{2}(1-\varepsilon)), \infty)$,

(ii) *the expected value μ of the outlier exists and satisfies the inequality* $\varepsilon\mu \leqslant F^{-1}(\frac{1}{2}(1-\varepsilon))$.

Then, for every $\varepsilon \in (0, \frac{1}{2})$ and $H \in \mathscr{H}$, the bias of the sample mean $T_1(G)$ is smaller than that of the sample median $T_2(G)$.

The proof is trivial: $T_1(G) = \varepsilon\mu$ and $T_2(G) = F^{-1}(\frac{1}{2}(1-\varepsilon))$.

Comment. Conditions (i) and (ii) seem to be realistic. Condition (i) says that a "very small outlier is no outlier". For instance, if F is $N(0, \sigma^2)$ and $\varepsilon = 0.05$, then the lower bound for an outlier is 0.07σ. Condition (ii) ensures that large outliers can occur with small probability.

Reference

[1] P. J. Huber, *Robust statistics*, J. Wiley, New York 1981.

MATHEMATICAL INSTITUTE
POLISH ACADEMY OF SCIENCES
00-656 WARSZAWA

Received on 1985.12.23