BUREAU OF THE CENSUS

STATISTICAL RESEARCH DIVISION REPORT SERIES

SRD Research Report Number: CENSUS/SRD/RR-84/04

USE OF ONE INSTEAD OF TWO OBSERVATIONS

by

Beverley D. Causey

U.S. Bureau of the Census
Washington, D.C. 20233

Recommended by: Paul P. Biemer

Report completed: November 14, 1983

Report issued: January 19, 1984

USE OF ONE INSTEAD OF TWO OBSERVATIONS

The use of a single sample observation to estimate the center of a (symmetric) distribution can be, in important senses, preferable to the use of the mean/median/midrange of a sample of 2.

Key words: sample mean of two; closeness to true parameter

Beverley D. Causey
Mathematical Statistician
Statistical Research Division
Bureau of the Census
Washington, DC 20233
November 14, 1983

Let x_1, x_2, and x_3 denote independent observations from a continuous distribution symmetric around (for the sake of simplicity) 0; let $\overline{x}_2 = (x_1 + x_2)/2$. For given $\varepsilon>0$ we let P_1 denote $P(|x_1|<\varepsilon)$ and P_2 denote $P(|\overline{x}_2|<\varepsilon)$. Stigler (1980) gives examples of density functions for which $\varepsilon>0$ can be found such that $P_1>P_2$ and thus for which x_1 might in a sense be preferred to $\overline{x}_2$ as an estimator of 0, the mean (i.e., center) of the distribution. Here we continue this review of the perverse circumstances under which x_1 might be preferred to $\overline{x}_2$ and then under which, even when all moments are finite, x_3 might be preferred to $\overline{x}_2$.

Example 1. We begin with "symmetric stable distributions" (Stigler 1980); one seeks a distribution for which the log of the characteristic function (LCF) (of t) is $-|t|^{\alpha}$; thus the LCF of $\overline{x}_2$ is $-|t/2|^{\alpha}\, 2 = -|tR|^{\alpha}$ with $R = 2^{(1-\alpha)/\alpha}$. For $\alpha<1$ we have $R>1$; thus the distribution of $\overline{x}_2$ is that of Rx_1 with $R>1$ (by taking α close to 0 we can make R as large as we like), and we of course have $P_1>P_2\ \forall\ \varepsilon$. The corresponding density function $f(x)$ is $(1/\pi)\int_0^{\infty} \exp(-t^{\alpha}) \cos tx\, dt$, in general not readily computable, with $f(0) = (1/\pi)\ \Gamma(1/\alpha + 1)$.

Example 2. For $\alpha = 1$ and $R = 1$ we have a Cauchy distribution. Let y_1, y_2, and y_3 be independent observations from this distribution, with density function $1/\pi(1 + y^2)$; let $\overline{y}_2 = (y_1 + y_2)/2$; let x_i have the magnitude y_i^2 and the sign of y_i. Thus the density function $f(x)$ is $1/2\pi|x|^{.5}(1 + |x|)$ (with $f(0) = \infty$), and the c.d.f. is $.5 + (\text{sign } x)(1/\pi) \arctan |x|^{.5}$. We now show that $P_1>P_2\ \forall\ \varepsilon$. It is well known (and implied above) that y_1 and $\overline{y}_2$ have identical distributions; thus for any $\delta>0$ we have $P(|y_1|<\delta) = P(|\overline{y}_2|<\delta)$. Let $\varepsilon = \delta^2$; we have $P(|y_1|<\delta) = P_1$, and also $P(|\overline{y}_2|<\delta) = P(|\overline{y}_2^2|<\varepsilon)$. The proof is completed by showing that $|\overline{x}_2|>\overline{y}_2^2$ always (so that $P_2<P(\overline{y}_2^2<\varepsilon)$). Let $t_i = |x_i|$; suppose first that x_1 and x_2 are of the same sign; then $|\overline{x}_2| - \overline{y}_2^2 = (t_1^2 + t_2^2)/2 - (t_1 + t_2)^2/4 = (t_1 - t_2)^2/4>0$. Suppose they are of different sign with $t_1>t_2$ (the case $t_1<t_2$ is of course similar); then $|\overline{x}_2| - \overline{y}_2^2$ $= (t_1^2 - t_2^2)/2 - (t_1 - t_2)^2/4 = (t_1^2 + 2t_1t_2 - 3t_2^2)/4$ $= (t_1 + 3t_2)(t_1 - t_2)/4>0$.

Example 3. Stigler (1980) considers the density function $f(x) = (1 + |x|)^{-c}(c - 1)/2$, $c>1$, with c.d.f. $.5 + .5$ (sign x) $[1 - (1 + |x|)^{1-c}]$. For c a multiple of .5 one may obtain P_2 explicitly; for c = 1.5 and 2 we have found empirically that $P_1>P_2$ (apparently) $\forall$ ε.

Let $P_{12} = P(|x_1|<|\overline{x}_2|)$. In spite of the result $P_1>P_2$ $\forall$ ε in these examples, we have (for any distribution) $P_{12}<.5$, it is easily shown. We compute P_{12} = .445 in Example 2, and .433 (c = 1.5) and .412 (c = 2) in Example 3. Thus more than half the time $\overline{x}_2$ is closer than x_1 to 0; but in these examples, apparently, the difference in closeness is generally greater when x_1 is closer than when $\overline{x}_2$ is closer.

Let $P_{32} = P(|x_3|<|\overline{x}_2|)$. Although we always have $P_{12}<.5$, it is possible to have $P_{32}>.5$, e.g., in Example 3 .860 and .529. These values, like the above values for P_{12}, are obtained by numerical integration: for $0<u<1$ let ε_u be such that $P_1 = u$, and let $h(u) = P_2$ (for ε_u), then $P_{32} = 1 - \int_0^1 h(u)du$. For Example 2 we may show $P_{32}>.5$ based on the fact $P(|y_3|<|\overline{y}_2|) = .5$ and on the above reasoning to show $P_1>P_2$.

In all the above "heavy-tailed" examples no moments of x exist. We now consider the variate z having the distribution of x except truncated at $\pm$ T: that is, for $-T<a<b<T$, $P(a<z<b) = P(a<x<b)/ P(|x|<T)$; all moments of z are finite. By taking T (>0) as large as we like, we can (in analogous notation) make $P(|z_3|<|\overline{z}_2|)$ as close to P_{32} (> .5) as we like. Despite this result we have (along with $Var(\bar{z}_2) = .5\ Var(z_3)$ and the fact that $P(|z_1|<\varepsilon)>P(|\bar{z}_2|<\varepsilon)\ \forall\ \varepsilon$ is impossible) the fact $E(|\bar{z}_2|)<.5[E(|z_1|) + E(|z_2|)] = E(|z_3|)$. Thus in these examples more than half the time z_3 is closer than $\bar{z}_2$ to 0, but apparently the difference in closeness is generally greater when $\bar{z}_2$ is closer than when z_3 is closer (cf. the pattern for P_{12}).

REFERENCES

STIGLER, S.M. (1980), "An Edgeworth Curiosum," Annals of Statistics, 8, 931-934.

WILKS, S.S. (1962), Mathematical Statistics, Wiley, New York.