CHAPTER III

INFLUENCE FUNCTION OF CERTAIN FUNCTIONALS

3.0 INTRODUCTION : In this chapter we are considered various types of functionals viz.

i) α-trimmed mean

ii) α-trimmed variance

iii) α-Winsorized mean

iv) α-Winsorized variance

v) M-estimator.

The influence function of each of them is obtained. In section 3.3 of this Chapter the finite-sample versions of influence function is considered and some simple theoretical and numerical examples on empirical influence function and sensitivity curve are given.

3.1 INFLUENCE FUNCTIONS OF α-TRIMMED MEAN AND α-TRIMMED VARIANCE :

In this section we shall define i) α-trimmed mean and ii) α-trimmed variance and obtain their influence functions. Ofcourse, the final result of IF of α-trimmed mean has been given directly in the paper (Hample,(1974)). The α-trimmed variance is defined and its IF is obtained independently.

The α-trimmed mean : One class of estimators with very good robustness properties are the α-trimmed means($0<\alpha<1/2$) which have been proposed by Hample(1968) as substitutes for the mean.

The α-trimmed mean($0<\alpha<1/2$) can be obtained as follows : Let $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ be the ordered sample of size n and α($0<\alpha<1/2$) be the given number. Define h to be $n\alpha$(or $[n\alpha]+\ell$) if $n\alpha$ is (or is not) an integer. Then the sample mean computed, by deleting h smallest and h largest observations, is called the sample α-trimmed mean and is denoted by $T_{n,\alpha}$. That is,

$$T_{n,\alpha} = \frac{1}{n-2h} \sum_{i=h+1}^{n-h} x_{(i)} \tag{3.1.1}$$

Here $[n\alpha]$ denote the integer part of $n\alpha$.

Further, corresponding to a distribution function F the α-trimmed mean of F is given by

$$T_\alpha(F) = \frac{1}{1-2\alpha} \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x dF(x), \qquad 0<\alpha<1/2$$

$$= \frac{1}{1-2\alpha} \int_{\alpha}^{1-\alpha} F^{-1}(P) dP, \qquad \alpha<P<1-\alpha. \tag{3.1.2}$$

The α-trimmed mean represents a compromise between the mean and the median, which represent the limiting cases as $\alpha \to 0$ and $\alpha \to 1/2$, respectively.

<u>IF of α-trimmed mean</u> : Let F be a continuous distribution function. In order to obtain $IF(x;T_\alpha,F)$, it is enough to find $\frac{\partial}{\partial t}(T_\alpha(G))\Big|_{t=0}$, where $T_\alpha(G)$ is the α-trimmed mean of

the distribution G. Here we need to consider the following three cases.

<u>Case I</u>: <u>If $x<F^{-1}(\alpha)$</u>

The α-trimmed mean of the distribution G is

$$T_\alpha(G) = \int_{G^{-1}(\alpha)}^{G^{-1}(1-\alpha)} y\,dG(y)$$

$$= \int_{G^{-1}(\alpha)}^{G^{-1}(1-\alpha)} y g(y)\,dy$$

$$= \frac{1-t}{1-t\alpha} \int_{F^{-1}(\frac{\alpha-t}{1-t})}^{F^{-1}(\frac{1-\alpha-t}{1-t})} y f(y)\,dy \qquad (3.1.3)$$

Differentiating (3.1.3) partially w.r.t. t, we get

$$\frac{\partial}{\partial t}\left[T_\alpha(G)\right] = \frac{1-t}{1-2\alpha}\Big\{F^{-1}\left(\frac{1-\alpha-t}{1-t}\right) f\left[F^{-1}\left(\frac{1-\alpha-t}{1-t}\right)\right]\frac{\partial}{\partial t}\left[F^{-1}\left(\frac{1-\alpha-t}{1-t}\right)\right]$$

$$- F^{-1}\left(\frac{\alpha-t}{1-t}\right) f\left[F^{-1}\left(\frac{\alpha-t}{1-t}\right)\right]\frac{\partial}{\partial t}\left[F^{-1}\left(\frac{\alpha-t}{1-t}\right)\right]\Big\} - \frac{1}{1-2\alpha}\int_{F^{-1}(\frac{\alpha-t}{1-t})}^{F^{-1}(\frac{1-\alpha-t}{1-t})} y\,dF(y)$$

$$= \frac{1-t}{1-2\alpha}\Big\{F^{-1}\left(\frac{1-\alpha-t}{1-t}\right)\frac{\partial}{\partial t}\left[\frac{1-\alpha-t}{1-t}\right] - F^{-1}\left(\frac{\alpha-t}{1-t}\right)\frac{\partial}{\partial t}\left[\frac{\alpha-t}{1-t}\right]\Big\}$$

$$- \frac{1}{1-2\alpha}\int_{F^{-1}(\frac{\alpha-t}{1-t})}^{F^{-1}(\frac{1-\alpha-t}{1-t})} y\,dF(y)$$

$$= \frac{1-t}{1-2\alpha}\left\{F^{-1}\left(\frac{1-\alpha-t}{1-t}\right)\left[\frac{-\alpha}{(1-t)^2}\right] - F^{-1}\left(\frac{\alpha-1}{1-t}\right)\left[\frac{\alpha-1}{(1-t)^2}\right]\right\}$$

$$- \frac{1}{1-2\alpha}\int_{F^{-1}\left(\frac{\alpha-t}{1-t}\right)}^{F^{-1}\left(\frac{1-\alpha-t}{1-t}\right)} y\,dF(y)$$

$$\frac{\partial}{\partial t}\left[T_\alpha(G)\right]\Big|_{t=0} = \frac{1}{1-2\alpha}\left[F^{-1}(1-\alpha)(-\alpha)-F^{-1}(\alpha)(\alpha-1)\right] - \frac{1}{1-2\alpha}\int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} y\,dF(y)$$

Therefore,

$$IF(x;T_\alpha,F) = \frac{1}{1-2\alpha}\left[F^{-1}(\alpha)-\int_{\alpha}^{1-\alpha}F^{-1}(P)dP-\alpha F^{-1}(\alpha)-\alpha F^{-1}(1-\alpha)\right] \quad \alpha<P<1-\alpha$$

$$= \frac{1}{1-2\alpha}\left[F^{-1}(\alpha)-T_W(F)\right] \ , \quad x<F^{-1}(\alpha) \ , \qquad (3.1.4)$$

Where $T_W(F)$ is the α-Winsorized mean (defined in subsequent section)

<u>Case II</u> : <u>If $F^{-1}(\alpha)\le x\le F^{-1}(1-\alpha)$</u>

In this case $T_\alpha(G)$ is

$$T_\alpha(G) = \int_{G^{-1}(\alpha)}^{G^{-1}(1-\alpha)} y\,dG(y)$$

$$= \int_{G^{-1}(\alpha)}^{G^{-1}(1-\alpha)} y\,g(y)dy + \frac{t}{1-2\alpha}\,x$$

$$= \frac{1-t}{1-2\alpha} \int_{F^{-1}(\frac{\alpha}{1-t})}^{F^{-1}(\frac{1-\alpha-t}{1-t})} yf(y)dy + \frac{t}{1-2\alpha}\, x \qquad (3.1.5)$$

Differentiating (3.1.5) partially w.r.t. t, we get

$$\frac{\partial}{\partial t}\left[T_\alpha(G)\right] = \frac{1-t}{1-2\alpha}\left\{F^{-1}\left(\frac{1-\alpha-t}{1-t}\right) f\left[F^{-1}\left(\frac{1-\alpha-t}{1-t}\right)\right]\frac{\partial}{\partial t}\left[F^{-1}\left(\frac{1-\alpha-t}{1-t}\right)\right]\right.$$

$$\left. - F^{-1}\left(\frac{\alpha}{1-t}\right) f\left[F^{-1}\left(\frac{\alpha}{1-t}\right)\right]\frac{\partial}{\partial t}\left[F^{-1}\left(\frac{\alpha}{1-t}\right)\right]\right\}$$

$$- \frac{1}{1-2\alpha}\int_{F^{-1}(\frac{\alpha}{1-t})}^{F^{-1}(\frac{1-\alpha-t}{1-t})} ydF(y) + \frac{1}{1-2\alpha}\, x$$

$$= \frac{1-t}{1-2\alpha}\left\{F^{-1}\left(\frac{1-\alpha-t}{1-t}\right)\left[\frac{-\alpha}{(1-t)^2}\right] - F^{-1}\left(\frac{\alpha}{1-t}\right)\left[\frac{\alpha}{(1-t)^2}\right]\right\}$$

$$- \frac{1}{1-2\alpha}\int_{F^{-1}(\frac{\alpha}{1-t})}^{F^{-1}(\frac{1-\alpha-t}{1-t})} ydF(y) + \frac{1}{1-2\alpha}$$

$$\frac{\partial}{\partial t}\left[T_\alpha(G)\right]\Big|_{t=0} = \frac{1}{1-2\alpha}\left[F^{-1}(1-\alpha)(-\alpha) - F^{-1}(\alpha)(\alpha)\right]$$

$$- \frac{1}{1-2\alpha}\int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} ydF(y) + \frac{x}{1-2\alpha}$$

Therefore,

$$IF(x;T_\alpha,F) = \frac{1}{1-2\alpha}\left[x - \int_\alpha^{1-\alpha} F^{-1}(P)dP - \alpha F^{-1}(\alpha) - \alpha F^{-1}(1-\alpha)\right], \quad \alpha<P<1-\alpha$$

$$= \frac{1}{1-2\alpha}\left[x - T_W(F)\right], \quad F^{-1}(\alpha)\le x\le F^{-1}(1-\alpha) \qquad (3.1.6)$$

<u>Case III : If $x>F^{-1}(1-\alpha)$</u>

Here $T_\alpha(G)$ is

$$T_\alpha(G) = \int_{G^{-1}(\alpha)}^{G^{-1}(1-\alpha)} y\,dG(y)$$

$$= \int_{G^{-1}(\alpha)}^{G^{-1}(1-\alpha)} y\,g(y)\,dy$$

$$= \frac{1-t}{1-2\alpha} \int_{F^{-1}(\frac{\alpha}{1-t})}^{F^{-1}(\frac{1-\alpha}{1-t})} y\,f(y)\,dy \qquad (3.1.7)$$

Differentiating (3.1.7) partially w.r.t. t, we get

$$\frac{\partial}{\partial t}\left[T(G)\right] = \frac{1-t}{1-2\alpha}\left\{F^{-1}\left[\frac{1-\alpha}{(1-t)}\right]\left[\frac{1-\alpha}{(1-t)^2}\right] - F^{-1}\left(\frac{\alpha}{1-t}\right)\left[\frac{\alpha}{(1-t)^2}\right]\right\}$$

$$- \frac{1}{1-2\alpha}\int_{F^{-1}(\frac{\alpha}{1-t})}^{F^{-1}(\frac{1-\alpha}{1-t})} y\,dF(y)$$

$$\frac{\partial}{\partial t}\left[T(G)\right]\Big|_{t=0} = \frac{1}{1-2\alpha}\left[F^{-1}(1-\alpha)(1-\alpha) - F^{-1}(\alpha)(\alpha)\right] - \frac{1}{1-2\alpha}\int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} y\,dF(y)$$

Therefore,

$$IF(x;T_\alpha,F) = \frac{1}{1-2\alpha}\left[F^{-1}(1-\alpha) - \int_{\alpha}^{1-\alpha} F^{-1}(P)\,dP - \alpha F^{-1}(\alpha) - \alpha F^{-1}(1-\alpha)\right],$$

$$\alpha<P<1-\alpha$$

$$= \frac{1}{1-2\alpha}\left[F^{-1}(1-\alpha) - T_W(F)\right], \quad x>F^{-1}(1-\alpha) \qquad (3.1.8)$$

Thus, from (3.1.4), (3.1.6) and (3.1.8), the IF of the α-trimmed mean at F is given by

$$IF(x;T_\alpha,F) = \frac{1}{1-2\alpha}\left[F^{-1}(\alpha)-T_W(F)\right] , \quad \text{for} \quad x<F^{-1}(\alpha)$$

$$= \frac{1}{1-2\alpha}\left[x-T_W(F)\right] , \quad \text{for} \quad F^{-1}(\alpha)\le x\le F^{-1}(1-\alpha)$$

$$= \frac{1}{1-2\alpha}\left[F^{-1}(1-\alpha)-T_W(F)\right] , \quad \text{for} \quad x>F^{-1}(1-\alpha)$$

and GES is given by (3.1.9)

$$r^*(T_\alpha,F) = \frac{1}{1-2\alpha}\left[F^{-1}(1-\alpha)-T_W(F)\right]$$

$$= \text{finite} \tag{3.1.10}$$

This implies that the α-trimmed mean is B-robust at F.

In particular, if F is symmetric about zero, then the IF of the α-trimmed mean reduces to,

$$IF(x;T_\alpha,F) = \frac{1}{1-2\alpha}\, F^{-1}(\alpha) , \quad \text{if } x<F^{-1}(\alpha)$$

$$= \frac{1}{1-2\alpha}\, x , \quad \text{if } F^{-1}(\alpha)\le x\le F^{-1}(1-\alpha)$$

$$= \frac{1}{1-2\alpha}\, F^{-1}(1-\alpha) , \quad \text{if } x>F^{-1}(1-\alpha).$$

The same can be written as,

$$IF(x;T_\alpha,F) = \frac{x}{1-2\alpha} , \quad \text{if } |x| \le F^{-1}(1-\alpha)$$

$$= F^{-1}(1-\alpha)\,\frac{\text{sign}(x)}{1-2\alpha} , \quad \text{if } |x|> F^{-1}(1-\alpha)$$

and GES will be

$$r^*(T_\alpha, F) = \frac{1}{1-2\alpha} F^{-1}(1-\alpha)$$

$$= \text{Finite.}$$

The corresponding asymptotic variance of estimator is given by

$$V(T_\alpha, F) = \int IF^2(x; T_\alpha, F)dF(x)$$

$$= \frac{1}{(1-2\alpha)^2} \int_{-\infty}^{F^{-1}(\alpha)} \left[F^{-1}(\alpha)\right]^2 dF(x) + \frac{1}{(1-2\alpha)^2} \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x^2 dF(x)$$

$$+ \frac{1}{(1-2\alpha)^2} \int_{F^{-1}(1-\alpha)}^{\infty} \left[F^{-1}(1-\alpha)\right]^2 dF(x)$$

$$= \frac{\left[F^{-1}(x)\right]^2}{(1-2\alpha)^2} \left[\left[F(x)\right]_{-\infty}^{F^{-1}(\alpha)} + \left[F(x)\right]_{F^{-1}(1-\alpha)}^{\infty} \right]$$

$$+ \frac{1}{(1-2\alpha)^2} \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x^2 dF(x)$$

$$= \frac{1}{(1-2\alpha)^2} \left\{ \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x^2 dF(x) + 2\alpha(F^{-1}(\alpha))^2 \right\}$$

$$= \frac{1}{(1-2\alpha)^2} \left\{ \int_{\alpha}^{1-\alpha} \left[F^{-1}(P)\right]^2 dP + 2\alpha(F^{-1}(\alpha))^2 \right\}$$

<u>Remark(3.1.1)</u> : Comparing the results (2.1.13) and (3.1.10), we conclude that, the α-trimmed mean for $0<\alpha<1/2$, provides a correction for the deficiency of the result (2.1.13)

<u>The α-trimmed variance</u> : Let $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ denote the order Statistic and $\alpha (0<\alpha<1/2)$ be the given number. Define $h = n\alpha$ (or $h = [n\alpha]+\ell$) if $n\alpha$ is (or is not) an integer. Let the sample variance computed by deleting h observations from each end of the sample be called the sample α-trimmed variance and be denoted by $V_{n,\alpha}$.

That is,

$$V_{n,\alpha} = \frac{1}{n-2h} \sum_{i=h+1}^{n-h} (x_{(i)}-T_{n,\alpha})^2 , \qquad (3.1.11)$$

Where $T_{n,\alpha}$ is the mean of the trimmed sample.

Furthermore, for a distribution function F, let the α-trimmed variance of F be given by

$$V_\alpha(F) = \frac{1}{1-2\alpha} \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} (x-T_\alpha(F))^2 dF(x) , \qquad 0<\alpha<1/2 \qquad (3.1.12)$$

Where $T_\alpha(F)$ is the α-trimmed mean at F.

If F is symmetric about zero, then

$$V_\alpha(F) = \frac{1}{1-2\alpha} \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x^2 dF(x) , \qquad 0<\alpha<1/2$$

$$= \frac{1}{1-2\alpha} \int_{\alpha}^{1-\alpha} [F^{-1}(P)]^2 dP , \qquad \alpha<P<1-\alpha \qquad (3.1.13)$$

<u>IF of α-trimmed variance</u> : Consider a continuous distribution function F, symmetric about zero and the corresponding α-trimmed variance $V_\alpha(F)$. Here, as before, we have to find $\frac{\partial}{\partial t}\left[V_\alpha(F)\right]\Big|_{t=0}$ in order to obtain $IF(x;V_\alpha,F)$. We need to consider the following three cases.

<u>Case I</u> : <u>If $x<F^{-1}(\alpha)$</u>

The α-trimmed variance of the distribution G is

$$V_\alpha(G) = \int_{G^{-1}(\alpha)}^{G^{-1}(1-\alpha)} y^2 dG(y) - \left[T_\alpha(G)\right]^2$$

$$= \frac{1-t}{1-2\alpha}\int_{F^{-1}(\frac{\alpha-t}{1-t})}^{F^{-1}(\frac{1-\alpha-t}{1-t})} y^2 f(y)dy - \left[T_\alpha(G)\right]^2 \qquad (3.1.14)$$

Differentiating (3.1.14) partially w.r.t. t, we get

$$= \frac{1-t}{1-2\alpha}\left\{\left[F^{-1}(\frac{1-\alpha-t}{1-t})\right]^2 f\left[F^{-1}(\frac{1-\alpha-t}{1-t})\right]\frac{\partial}{\partial t}\left[F^{-1}(\frac{1-\alpha-t}{1-t})\right]\right.$$

$$\left. - \left[F^{-1}(\frac{\alpha-t}{1-t})\right]^2 f\left[F^{-1}(\frac{\alpha-t}{1-t})\right]\frac{\partial}{\partial t}\left[F^{-1}(\frac{\alpha-t}{1-t})\right]\right\} - \frac{1}{1-2\alpha}\int_{F^{-1}(\frac{\alpha-t}{1-t})}^{F^{-1}(\frac{1-\alpha-t}{1-t})} y^2 dF(y)$$

$$- 2\ T_\alpha(G)\frac{\partial}{\partial t}\left[T_\alpha(G)\right]$$

$$= \frac{1-t}{1-2\alpha}\left\{\left[F^{-1}\left(\frac{1-\alpha-t}{1-t}\right)\right]^2\left[\frac{-\alpha}{(1-t)^2}\right] - \left[F^{-1}\left(\frac{\alpha-t}{1-t}\right)\right]^2\left[\frac{\alpha-1}{(1-t)^2}\right]\right\}$$

$$- \frac{1}{1-2\alpha}\int_{F^{-1}\left(\frac{\alpha-t}{1-t}\right)}^{F^{-1}\left(\frac{1-\alpha-t}{1-t}\right)} y^2 dF(y) - 2T_\alpha(G)\frac{\partial}{\partial t}\left[T_\alpha(G)\right]$$

$$\frac{\partial}{\partial t}\left[V_\alpha(G)\right]\Big|_{t=0} = \frac{1}{1-2\alpha}\left[(F^{-1}(1-\alpha))^2(-\alpha)-(F^{-1}(\alpha))^2(\alpha-1)\right]$$

$$- \frac{1}{1-2\alpha}\int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} y^2 dF(y)$$

Therefore,

$$IF(x;V_\alpha,F) = \frac{1}{1-2\alpha}\left[(F^{-1}(\alpha))^2 - \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} y^2 dF(y) - 2\alpha(F^{-1}(\alpha))^2\right]$$

$$= \frac{1}{1-2\alpha}\left[(F^{-1}(\alpha))^2 - V_W(F)\right], \qquad (3.1.15)$$

Where $V_W(F) = \int_{\alpha}^{1-\alpha}\left[F^{-1}(P)\right]^2 dP + 2\alpha\left[F^{-1}(\alpha)\right]^2, \quad \alpha<P<1-\alpha,$

is a α-Winsorized variance of distribution F, which is symmetric about zero (defined in subsequent section).

<u>Case II</u> : <u>If $F^{-1}(\alpha)<x<F^{-1}(1-\alpha)$</u>

The α-trimmed variance of the distribution G is

$$V_\alpha(G) = \int_{G^{-1}(\alpha)}^{G^{-1}(1-\alpha)} y^2 dG(y) - \left[T_\alpha(G)\right]^2$$

SHIVAJI UNIVERSITY LIBRARY KOLHAPUR

$$= \int_{G^{-1}(\alpha)}^{G^{-1}(1-\alpha)} y^2 g(y)dy + \frac{t}{1-2\alpha} x^2 - \left[T_\alpha(G)\right]^2$$

$$= \frac{1-t}{1-2\alpha} \int_{F^{-1}(\frac{\alpha}{1-t})}^{F^{-1}(\frac{1-\alpha-t}{1-t})} y^2 f(y)dy + \frac{t}{1-2\alpha} x^2 - \left[T_\alpha(G)\right]^2$$

Differentiating this partially w.r.t. t and substituting $t = 0$, we get

$$\frac{\partial}{\partial t}\left[V_\alpha(G)\right]\Big|_{t=0} = \frac{1}{1-2\alpha}\left\{\left[F^{-1}(1-\alpha)\right]^2(-\alpha)-\left[F^{-1}(\alpha)\right]^2(\alpha) - \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} y^2 dF(y) + x^2\right\}$$

Therefore,

$$IF(x;V_\alpha,F) = \frac{1}{1-2\alpha}\left\{x^2-2\alpha\left[F^{-1}(\alpha)\right]^2 - \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} y^2 dF(y)\right\}$$

$$= \frac{1}{1-2\alpha}\left[x^2-V_w(F)\right] \qquad (3.1.16)$$

<u>Case III</u> : <u>$x>F^{-1}(1-\alpha)$</u>

In this case $V_\alpha(G)$ is

$$V_\alpha(G) = \int_{G^{-1}(\alpha)}^{G^{-1}(1-\alpha)} y^2 dG(y) - \left[T_\alpha(G)\right]^2$$

$$= \frac{1-t}{1-2\alpha} \int_{F^{-1}(-\frac{\alpha}{1-t})}^{F^{-1}(-\frac{1-\alpha}{1-t})} y^2 f(\ y)dy - \left[T_\alpha(G)\right]^2$$

Differentiating this partially w.r.t. t and substituting $t = 0$, we get

$$IF(x;V_\alpha,F) = \frac{1}{1-2\alpha}\left[(F^{-1}(\alpha))^2 - V_W(F)\right] \qquad (3.1.17)$$

Thus, from (3.1.15),(3.1.16) and (3.1.17), the IF of α-trimmed variance at a symmetric distribution F is given by

$$IF(x;V_\alpha,F) = \frac{1}{1-2\alpha}\left[(F^{-1}(\alpha))^2-V_W(F)\right], \text{for } x<F^{-1}(\alpha) \text{ or } x>F^{-1}(1-\alpha)$$

$$= \frac{1}{1-2\alpha}\left[x^2-V_W(F)\right], \qquad \text{for } F^{-1}(\alpha)<x<F^{-1}(1-\alpha). \quad \blacksquare$$

3.2 INFLUENCE FUNCTIONS OF α-WINSORIZED MEAN AND α-WINSORIZED VARIANCE :

In this section we shall define i) α-Winsorized mean and α-Winsorized variance and the influence functions for both are obtained. Ofcourse, the final result of IF of α-Winsorized mean has been given directly in the paper (Hample (1974)). The α-Winsorized variance is defined and its IF is obtained independently.

<u>α-Winsorized Mean</u> : Another class of the order Statistics are the α-Winsorized means ($0<\alpha<1/2$) which satisfy the robustness properties. In case of the α-trimmed means the extreme observations are not considered (given zero weight). But for α-Winsorized means, in a sense weightes are given to the extreme observations. The purpose in doing is that we do not want to "throw away" the $n\alpha$ leftmost and $n\alpha$ rightmost observations as in the α-trimmed mean.

BARR. BALASAHEB KHARDEKAR LIBRARY
SHIVAJI UNIVERSITY, KOLHAPUR.

The α-Winsorized mean can be obtained as follows : Let $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ be the ordered sample of size n and $\alpha (0<\alpha<1/2)$ be the given number. Define $h=n\alpha$, if $n\alpha$ is an integer and $h=[n\alpha]+\ell$, if $n\alpha$ is not an integer. Replace the values of the h leftmost and h rightmost observations by that of $x_{(h+1)}$ and $x_{(n-h)}$ respectively. Count these values (h+1) times. The mean obtained from this modified sample is known as the α-Winsorized mean and is denoted by $W_{n,\alpha}$. That is,

$$W_{n,\alpha} = \frac{hx_{(h+1)}+x_{(h+1)}+x_{(h+2)}+\ldots+x_{(n-h)}+hx_{(n-h)}}{n}$$

$$= n^{-1}\left\{(h+1)(x_{(h+1)}+x_{(n-h)})+\sum_{i=h+2}^{n-h-1} x_{(i)}\right\}, \quad h<\frac{n-1}{2} \qquad (3.2.1)$$

Furthermore, an asymptotically equivalent definition, which holds for all distributions F, is given by

$$T_w(F) = \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x dF(x)+ \alpha\left[F^{-1}(\alpha)+F^{-1}(1-\alpha)\right], \qquad 0<\alpha<1/2$$

$$= \int_{\alpha}^{1-\alpha} F^{-1}(P)dP+ \alpha\left[F^{-1}(\alpha)+F^{-1}(1-\alpha)\right], \qquad 0<P<1-\alpha$$

$$= (1-2\alpha)T_\alpha(F)+ \alpha\left[F^{-1}(\alpha)+F^{-1}(1-\alpha)\right], \qquad (3.2.2)$$

Where $T_\alpha(F)$ is the α-trimmed mean at the distribution F.

<u>Remark(3.2.1)</u> : There will be trouble if the corner points $F^{-1}(\alpha)$ and $F^{-1}(1-\alpha)$ are not uniquely determined that is if F^{-1} has jumps there.

<u>Remark(3.2.2)</u> : i) If F is symmetric about zero, then $T_\alpha(F) = T_W(F) = 0$

ii) If F is symmetric about θ, $(\theta>0)$, then $T_\alpha(F)<T_W(F)$

iii) If F is symmetric about θ, $(\theta<0)$, then $T_\alpha(F)>T_W(F)$

iv) For a random sample, $T_{n,\alpha} >_{\leq} W_{n,\alpha}$. ∎

<u>IF of α-Winsorized Mean</u> :

Consider a continuous distribution function F. In order to obtain $IF(x;T_W,F)$, it is enough to find $\frac{\partial}{\partial t}\left[T_W(G)\right]\Big|_{t=0}$, where $T_W(G)$ is the α-Winsorized mean of distribution G. We need to consider the following three cases.

<u>Case I</u> : <u>If $x<F^{-1}(\alpha)$</u>

Here $T_W(G)$ is

$$T_W(G) = \int_{G^{-1}(\alpha)}^{G^{-1}(1-\alpha)} yg(y)dy + \alpha\left[G^{-1}(\alpha) + G^{-1}(1-\alpha)\right]$$

$$= (1-t)\int_{F^{-1}(\frac{\alpha-t}{1-t})}^{F^{-1}(\frac{1-\alpha-t}{1-t})} yf(y)dy + \alpha\left[F^{-1}(\frac{\alpha-t}{1-t}) + F^{-1}(\frac{1-\alpha-t}{1-t})\right]$$

Differentiating this partially w.r.t. t, we get

$$\frac{\partial}{\partial t}\left[T_w(G)\right] = (1-t)\left\{F^{-1}\left(\frac{1-\alpha-t}{1-t}\right)\frac{\partial}{\partial t}\left[\frac{1-\alpha-t}{1-t}\right] - F^{-1}\left(\frac{\alpha-t}{1-t}\right)\frac{\partial}{\partial t}\left[\frac{\alpha-t}{1-t}\right]\right\}$$

$$- \int_{F^{-1}\left(\frac{\alpha-t}{1-t}\right)}^{F^{-1}\left(\frac{1-\alpha-t}{1-t}\right)} yf(y)dy$$

$$+ \alpha\left\{\frac{1}{f[F^{-1}(\frac{\alpha-t}{1-t})]}\frac{\partial}{\partial t}\left[\frac{\alpha-t}{1-t}\right] + \frac{1}{f[F^{-1}(\frac{1-\alpha-t}{1-t})]}\frac{\partial}{\partial t}\left[\frac{1-\alpha-t}{1-t}\right]\right\}$$

$$= (1-t)\left\{F^{-1}\left(\frac{1-\alpha-t}{1-t}\right)\left[\frac{-\alpha}{(1-t)^2}\right] - F^{1}\left(\frac{\alpha-t}{1-t}\right)\left[\frac{\alpha-1}{(1-t)^2}\right]\right\}$$

$$- \int_{F^{-1}\left(\frac{\alpha-t}{1-t}\right)}^{F^{-1}\left(\frac{1-\alpha-t}{1-t}\right)} yf(y)dy + \alpha\left\{\frac{\alpha-1}{(1-t)^2}\frac{1}{f[F^{-1}(\frac{\alpha-t}{1-t})]} + \frac{(-\alpha)}{(1-t)^2}\frac{1}{f[F^{-1}(\frac{1-\alpha-t}{1-t})]}\right\}$$

$$\frac{\partial}{\partial t}\left[T_w(F)\right]\Big|_{t=0} = F^{-1}(1-\alpha)(-\alpha) - F^{-1}(\alpha)(\alpha-1) - \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} ydF(y)$$

$$+ \alpha\left\{\frac{\alpha-1}{f[F^{-1}(\alpha)]} + \frac{(-\alpha)}{f[F^{-1}(1-\alpha)]}\right\}$$

Therefore,

$$IF(x;T_w,F) = F^{-1}(\alpha) - \frac{\alpha}{f(F^{-1}(\alpha)} - \int_{\alpha}^{1-\alpha} F^{-1}(P)dP - \alpha F^{-1}(\alpha) - \alpha F^{-1}(1-\alpha)$$

$$+ \frac{\alpha^2}{f[F^{-1}(\alpha)]} - \frac{\alpha^2}{f[F^{-1}(1-\alpha)]}$$

$$= F^{-1}(\alpha) - \frac{\alpha}{f[F^{-1}(\alpha)]} - T_w(F) - \alpha^2\left\{\frac{1}{f[F^{-1}(1-\alpha)]} - \frac{1}{f[F^{-1}(\alpha)]}\right\}$$

$$= F^{-1}(\alpha) - \frac{\alpha}{f[F^{-1}(\alpha)]} - U(F) \ , \qquad x<F^{-1}(\alpha), \qquad (3.2.3)$$

Where $U(F) = T_w(F) + \alpha^2\left\{\frac{1}{f[F^{-1}(1-\alpha)]} - \frac{1}{f[F^{-1}(\alpha)]}\right\}$

<u>Case II</u> : <u>If $F^{-1}(\alpha)<x<F^{-1}(1-\alpha)$</u>

In this case $T_w(G)$ is

$$T_w(G) = \int_{G^{-1}(\alpha)}^{G^{-1}(1-\alpha)} yg(y)dy + t.x + \alpha\left[G^{-1}(\alpha)+G^{-1}(1-\alpha)\right]$$

$$= (1-t)\int_{F^{-1}(\frac{\alpha}{1-t})}^{F^{-1}(\frac{1-\alpha-t}{1-t})} yf(y)dy+t.x + \alpha\left[F^{-1}(\frac{\alpha}{1-t})+F^{-1}(\frac{1-\alpha-t}{1-t})\right]$$

Differentiating this partially w.r.t. t, we get

$$\frac{\partial}{\partial t}\left[T_w(G)\right] = (1-t)\left\{F^{-1}(\frac{1-\alpha-t}{1-t})\left[\frac{-\alpha}{(1-t)^2}\right] - F^{-1}(\frac{\alpha}{1-t})\left[\frac{\alpha}{(1-t)^2}\right]\right\}$$

$$- \int_{F^{-1}(\frac{\alpha}{1-t})}^{F^{-1}(\frac{1-\alpha-t}{1-t})} yf(y)dy$$

$$+ \alpha\left\{\frac{1}{f[F^{-1}(\frac{\alpha}{1-t})]}\left[\frac{\alpha}{(1-t)^2}\right] + \frac{1}{f[F^{-1}(\frac{1-\alpha-t}{1-t})]}\left[\frac{-\alpha}{(1-t)^2}\right]\right\} + x$$

$$\frac{\partial}{\partial t}\left[T_w(G)\right]\Big|_{t=0} = F^{-1}(1-\alpha)(-\alpha) - F^{-1}(\alpha)(\alpha) - \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} ydF(y)$$

$$+ \alpha\left\{\frac{\alpha}{f[F^{-1}(\alpha)]} - \frac{\alpha}{f[F^{-1}(1-\alpha)]}\right\} + x$$

Therefore,

$$IF(x;T_w,F) = x - \int_{\alpha}^{1-\alpha} F^{-1}(P)dP - \alpha\, F^{-1}(\alpha) - \alpha F^{-1}(1-\alpha)$$

$$-\alpha^2 \left\{ \frac{1}{f[F^{-1}(1-\alpha)]} - \frac{1}{f[F^{-1}(\alpha)]} \right\}$$

$$= x-U(F), \qquad F^{-1}(\alpha)<x<F^{-1}(1-\alpha) \tag{3.2.4}$$

<u>Case III</u> : <u>If $x>F^{-1}(1-\alpha)$</u>

Here $T_w(G)$ is

$$T_w(G) = \int_{G^{-1}(\alpha)}^{G^{-1}(1-\alpha)} yg(y)dy + \alpha \left[G^{-1}(\alpha)+G^{-1}(1-\alpha)\right]$$

$$= (1-t) \int_{F^{-1}(\frac{\alpha}{1-t})}^{F^{-1}(\frac{1-\alpha}{1-t})} yf(y)dy + \alpha \left[F^{-1}(\frac{\alpha}{1-t}) + F^{-1}(\frac{1-\alpha}{1-t})\right]$$

Differentiating this partially w.r.t. t, we get

$$\frac{\partial}{\partial t}\left[T_w(G)\right] = (1-t)\left\{F^{-1}(\frac{1-\alpha}{1-t})\left[\frac{1-\alpha}{(1-t)^2}\right] - F^{-1}(\frac{\alpha}{1-t})\left[\frac{\alpha}{(1-t)^2}\right]\right\}$$

$$- \int_{F^{-1}(\frac{\alpha}{1-t})}^{F^{-1}(\frac{1-\alpha}{1-t})} yf(y)dy$$

$$+ \alpha\left\{\frac{1}{f[F^{-1}(\frac{\alpha}{1-t})]}\left[\frac{\alpha}{(1-t)^2}\right] + \frac{1}{f[F^{-1}(\frac{1-\alpha}{1-t})]}\left[\frac{1-\alpha}{(1-t)^2}\right]\right\}$$

$$\frac{\partial}{\partial t}\left[T_w(G)\right]\Big|_{t=0} = F^{-1}(1-\alpha)(1-\alpha) - F^{-1}(\alpha)(\alpha) - \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} ydF(y)$$

$$+\alpha\left\{\frac{\alpha}{f[F^{-1}(\alpha)]} + \frac{1-\alpha}{f[F^{-1}(1-\alpha)]}\right\}$$

Therefore,

$$IF(x;T_w,F) = F^{-1}(1-\alpha) + \frac{\alpha}{f[F^{-1}(1-\alpha)]} - \alpha F^{-1}(\alpha) - \alpha F^{-1}(1-\alpha)$$

$$- \int_{\alpha}^{1-\alpha} F^{-1}(P)dP - \alpha^2\left\{\frac{1}{f[F^{-1}(1-\alpha)]} - \frac{1}{f[F^{-1}(\alpha)]}\right\}$$

$$= F^{-1}(1-\alpha) + \frac{\alpha}{f[F^{-1}(1-\alpha)]} - U(F), \quad x>F^{-1}(1-\alpha) \quad (3.2.5)$$

Thus, from (3.2.3),(3.2.4) and (3.2.5), the IF of the α-Winsorized mean at F is given by

$$IF(x;T_w,F) = F^{-1}(\alpha) - \frac{\alpha}{f(F^{-1}(\alpha))} - U(F), \quad \text{if } x<F^{-1}(\alpha)$$

$$= x - U(F) \qquad\qquad , \text{ if } F^{-1}(\alpha)<x<F^{-1}(1-\alpha)$$

$$= F^{-1}(1-\alpha)+\frac{\alpha}{f[F^{-1}(1-\alpha)]} - U(F), \quad \text{if } x>F^{-1}(1-\alpha) \quad (3.2.6)$$

The IF of the α-Winsorized mean has jumps at $F^{-1}(\alpha)$ and $F^{-1}(1-\alpha)$.

The GES is given by

$$r^*(T_w,F) = F^{-1}(1-\alpha)+ \frac{\alpha}{f[F^{-1}(1-\alpha)]} - U(F)$$

$$= \text{Finite}$$

This implies that the α-Winsorized mean is B-robust at any distribution

In particular, if F is symmetric about zero, then the $IF(x;T_w,F)$ reduces to,

$$IF(x;T_w,F) = F^{-1}(\alpha) - \frac{\alpha}{f(F^{-1}(1-\alpha))} \quad , \quad \text{for } x<F^{-1}(\alpha)$$

$$= x \qquad\qquad , \quad \text{for } F^{-1}(\alpha)<x<F^{-1}(1-\alpha)$$

$$= F^{-1}(1-\alpha)+\frac{\alpha}{f(F^{-1}(1-\alpha))} \quad , \quad \text{for } x>F^{-1}(1-\alpha).$$

It may also be written as

$$IF(x;T_w,F) = x \qquad , \quad \text{for } |x|<F^{-1}(1-\alpha)$$
$$= F^{-1}(1-\alpha)+\frac{\alpha}{f[F^{-1}(1-\alpha)]}\,\text{Sign}(x), \quad \text{for}|x|>F^{-1}(1-\alpha)$$

and GES is given by

$$GES = F^{-1}(1-\alpha) + \frac{\alpha}{f[F^{-1}(1-\alpha)]}$$
$$= \text{Finite}$$

The asymptotic variance of estimator is

$$V(T_w,F) = \int IF^2(x;T_w,F)dF(x)$$
$$= \left\{F^{-1}(\alpha) - \frac{\alpha}{f[F^{-1}(\alpha)]}\right\}^2 \int_{-\infty}^{F^{-1}(\alpha)} dF(x) + \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x^2 dF(x)$$
$$+ \left\{F^{-1}(1-\alpha) + \frac{\alpha}{f[F^{-1}(1-\alpha)]}\right\}^2 \int_{F^{-1}(1-\alpha)}^{\infty} dF(x)$$
$$= \left\{F^{-1}(1-\alpha) + \frac{\alpha}{f[F^{-1}(1-\alpha)]}\right\}^2 \left[F(x)\right]_{-\infty}^{F^{-1}(\alpha)} + \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x^2 dF(x)$$
$$+ \left\{F^{-1}(1-\alpha) + \frac{\alpha}{f[F^{-1}(1-\alpha)]}\right\}^2 \left[F(x)\right]_{F^{-1}(1-\alpha)}^{\infty}$$
$$= \left\{F^{-1}(1-\alpha) + \frac{\alpha}{f[F^{-1}(1-\alpha)]}\right\}^2 \left[(\alpha-0)+(1-(1-\alpha))\right] + \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x^2 dF(x)$$

$$= \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x^2 dF(x) + 2\alpha\left\{F^{-1}(1-\alpha)+\frac{\alpha}{F^{-1}(1-\alpha)}\right\}^2 \qquad \blacksquare$$

<u>α-Winsorized Variance</u> : Let $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ be the ordered sample of size n and $\alpha(0<\alpha<1/2)$ be the given number. Define h to be $n\alpha$(or $[n\alpha]+1$) if $n\alpha$ is (or is not) an integer. Replace the h extreme observations on either side of sample by $x_{(h+1)}$ and $x_{(n-h)}$ respectively. Count these values (h+1) times. Let the variance obtained from this modified sample be called the sample α-Winsorized variance and be denoted by $V_{w,n}$. That is,

$$V_{w,n} = n^{-1}(h+1)\left[(x_{(h+1)}-W_{n,\alpha})^2 + (x_{n-h}-W_{n,\alpha})^2\right] + n^{-1}\sum_{i=h+2}^{n-h-1}(x_{(i)}-W_{n,\alpha})^2 , \qquad (3.2.7)$$

Where $W_{n,\alpha}$ is α-Winsorized mean.

Furthermore, if F is distribution function, then α-Winsorized variance of F is given by

$$V_w(F) = \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)}(x-T_w(F))^2 dF(x)+\alpha\left[(F^{-1}(\alpha)-T_w(F))^2+(F^{-1}(1-\alpha)-T_w(F))^2\right], \quad 0<\alpha<1/2 \qquad (3.2.8)$$

Where $T_w(F)$ is α-Winsorized mean at F.

In particular, if F is symmetric about zero, then α-Winsorized variance is given by

$$V_w(F) = \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x^2 dF(x)+ \alpha\left[(F^{-1}(\alpha))^2+(F^{-1}(1-\alpha))^2\right]$$

$$= \int_{\alpha}^{1-\alpha} \left[F^{-1}(P)\right]^2 dP + 2\alpha \left[F^{-1}(\alpha)\right]^2, \qquad \alpha<P<1-\alpha \qquad (3.2.9)$$

We note that α-Winsorized variance at any distribution is always greater than that of the α-trimmed variance.

<u>IF of α-Winsorized Variance</u> : Let F be a continuous distribution function which is symmetric about zero and $V_w(F)$ be the α-Winsorized variance of F. $IF(x;V_w,F)$ is obtained by $\frac{\partial}{\partial t}\left[V_w(G)\right]\Big|_{t=0}$. Here also, we will consider the following three cases.

<u>Case I</u> : <u>If $x<F^{-1}(\alpha)$</u>

The α-Winsorized variance of the distribution G is

$$V_w(G) = \int_{G^{-1}(\alpha)}^{G^{-1}(1-\alpha)} y^2 g(y)dy + \alpha\left[(G^{-1}(\alpha))^2 + (G^{-1}(1-\alpha))^2\right] - (T_w(G))^2$$

$$= (1-t)\int_{F^{-1}(\frac{\alpha-t}{1-t})}^{F^{-1}(\frac{1-\alpha-t}{1-t})} y^2 f(y)dy + \alpha\left[(F^{-1}(\frac{\alpha-t}{1-t}))^2 + (F^{-1}(\frac{1-\alpha-t}{1-t}))^2\right] - (T_w(G))^2$$

Differentiating this partially w.r.t. t and substituting t=0, we get

$$\frac{\partial}{\partial t}\left[V_w(G)\right]\Big|_{t=0} = \left[F^{-1}(1-\alpha)\right]^2(-\alpha) - \left[F^{-1}(\alpha)\right]^2(\alpha-1) - \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} y^2 dF(y)$$

$$+ 2\alpha\left\{\frac{(-\alpha)F^{-1}(1-\alpha)}{f(F^{-1}(1-\alpha))} + \frac{(\alpha-1)F^{-1}(\alpha)}{f(F^{-1}(\alpha))}\right\}$$

Therefore,

$$IF(x;V_w,F) = \left[F^{-1}(\alpha)\right]^2 - \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} y^2 dF(y) - 2\alpha\left[F^{-1}(\alpha)\right]^2 + 2\alpha \frac{F^{-1}(\alpha)}{f(F^{-1}(\alpha))}(2\alpha-1)$$

$$= \left[F^{-1}(\alpha)\right]^2 + 2\alpha(2\alpha-1)\frac{F^{-1}(\alpha)}{f(F^{-1}(\alpha))} - V_w(F) \qquad (3.2.10)$$

Where $V_w(F)$ is the α-Winsorized variance of distribution F which is symmetric about zero

<u>Case II</u> : <u>If $F^{-1}(\alpha)<x<F^{-1}(1-\alpha)$</u>

Here, α-Winsorized variance of the distribution G is

$$V_w(G) = \int_{G^{-1}(\alpha)}^{G^{-1}(1-\alpha)} y^2 g(y)dy + tx^2 + \alpha\left[(G^{-1}(\alpha))^2 + (G^{-1}(1-\alpha))^2\right] - \left[T_w(G)\right]^2$$

$$= (1-t)\int_{F^{-1}(-\frac{\alpha}{1-t})}^{F^{-1}(\frac{1-\alpha-t}{1-t})} y^2 f(y)dy + tx^2 + \alpha\left[(F^{-1}(\frac{1-\alpha-t}{1-t}))^2 + (F^{-1}(\frac{\alpha}{1-t}))^2\right] - \left[T_w(G)\right]^2$$

Differentiating this partially w.r.t. t and substituting t=0, we get

$$\frac{\partial}{\partial t}\left[V_w(G)\right]\Big|_{t=0} = \left[(F^{-1}(1-\alpha))^2(-\alpha) - (F^{-1}(\alpha))^2(\alpha)\right] - \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} y^2 f(y)dy + x^2$$

$$+ 2\alpha\left\{\frac{F^{-1}(1-\alpha)}{f(F^{-1}(1-\alpha))}(-\alpha) + \frac{F^{-1}(\alpha)}{f(F^{-1}(\alpha))}(\alpha)\right\}$$

Therefore,

$$IF(x;V_w,F) = x^2 + 4\alpha^2 \frac{F^{-1}(\alpha)}{f(F^{-1}(\alpha))} - 2\left[F^{-1}(\alpha)\right]^2 - \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} y^2 f(y)dy$$

$$= x^2 + 4\alpha^2 \frac{F^{-1}(\alpha)}{f(F^{-1}(\alpha))} - V_w(F) \qquad (3.2.11)$$

<u>Case III</u> : <u>$x>F^{-1}(1-\alpha)$</u>

For this case $V_w(G)$ is

$$V_w(G) = \int_{G^{-1}(\alpha)}^{G^{-1}(1-\alpha)} y^2 g(y)dy + \alpha\left[(G^{-1}(\alpha))^2 + (G^{-1}(1-\alpha))^2\right] - (T_w(G))^2$$

$$= (1-t) \int_{F^{-1}(\frac{\alpha}{1-t})}^{F^{-1}(\frac{1-\alpha}{1-t})} y^2 f(y)dy + \alpha\left[(F^{-1}(\frac{1-\alpha}{1-t}))^2 + (F^{-1}(\frac{\alpha}{1-t}))^2\right] - (T_w(G))^2$$

Differentiating this partially w.r.t. t and substituting t=0, we get

$$\frac{\partial}{\partial t}\left[V_w(G)\right]\Big|_{t=0} = \left[(F^{-1}(1-\alpha))^2(1-\alpha) - (F^{-1}(\alpha))^2(\alpha)\right] - \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} y^2 f(y)dy$$

$$+ 2\alpha \left\{\frac{(1-\alpha)F^{-1}(1-\alpha)}{f[F^{-1}(1-\alpha)]} + \frac{\alpha F^{-1}(\alpha)}{f[F^{-1}(\alpha)]}\right\}$$

Therefore,

$$IF(x;V_w,F) = (F^{-1}(\alpha))^2 - 2\alpha(F^{-1}(\alpha))^2 - \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} y^2 f(y)dy$$

$$+ \frac{2\alpha F^{-1}(\alpha)}{f[F^{-1}(1-\alpha)]}(2\alpha-1)$$

$$= (F^{-1}(\alpha))^2 + 2\alpha(2\alpha-1)\frac{F^{-1}(\alpha)}{f(F^{-1}(1-\alpha))} - V_w(F) \qquad (3.2.12)$$

From (3.2.10), (3.2.11) and (3.2.12), the IF of the α-Winsorized variance at F, which is symmetric about zero, is given by

$$IF(x;V_W,F) = (F^{-1}(\alpha))^2+2\alpha(2\alpha-1)\frac{F^{-1}(\alpha)}{f(F^{-1}(\alpha))} - V_W(F),$$

for $x>F^{-1}(1-\alpha)$ or $x<F^{-1}(\alpha)$

$$= x^2 + 4\alpha^2\frac{F^{-1}(\alpha)}{f(F^{-1}(\alpha))} - V_W(F),$$

for $F^{-1}(\alpha)<x<F^{-1}(1-\alpha)$. ∎

3.3 FINITE-SAMPLE VERSIONS OF INFLUENCE FUNCTION :

In this section we shall consider the finite-sample versions of IF. The IF defined (2.1.5) is concerned with the underlying distribution F. This IF can also be viewed as an asymptotic functional of a sequence of suitable estimators of F. However, there exist also some simple finite-sample versions of IF, which can easily be computed and they are :

i) Empirical Influence Function

ii) Sensitivity Curve.

The graphical representation of these versions of influence function for various estimators is given in this section.

Now, we shall discuss these versions in detail.

1) <u>Empirical Influence Function</u> : Suppose we have a sample $(x_1, x_2, \ldots, x_n)$ of n observations and an estimator $\{T_n; n \geq 1\}$. Let an observation in the sample (say x_n) be replaced by an arbitrary value x. Then the empirical influence function (EIF) of the estimator at that sample is a function $T_n(x_1, x_2, \ldots, x_{n-1}, x)$ (as a function of x). This simplest idea is suggested by Hample (1982). Let us consider some special cases of EIF.

i) If n=10, α=.1, then $[n\alpha]=1$ and suppose $x_{(10)}$ is replaced by x. Therefore, the EIF of the 10%- trimmed mean, which is a function of x, can be obtained by deleting the smallest and the largest items from the modified sample. That is,

$$\text{EIF of 10\% trimmed mean} = \frac{x_{(1)}+x_{(2)}+\ldots+x_{(8)}}{8}, \text{ if } x<x_{(1)}$$

$$= \frac{x+x_{(2)}+x_{(3)}+\ldots+x_{(8)}}{8}, \text{if } x_{(1)} \leq x \leq x_{(9)}$$

$$= \frac{x_{(2)}+x_{(3)}+\ldots+x_{(9)}}{8}, \text{if } x>x_{(9)}$$

This EIF of the 10%-trimmed mean is constant when $x<x_{(1)}$, is continuous nondecreasing $x_{(1)} \leq x \leq x_{(9)}$ and is again constant when $x>x_{(9)}$.

ii) If n=10, α=.1, then $[n\alpha]=1$ and suppose $x_{(10)}$ is replaced by x. Therefore, the EIF of the 10%-Winsorized mean, which is a function of x, can be obtained as follows :

EIF of 10%-Winsorized mean

$$= \frac{2x_{(1)}+x_{(2)}+\ldots+x_{(7)}+2x_{(8)}}{10} \quad , \text{ if } x<x_{(1)}$$

$$= \frac{2x+x_{(2)}+x_{(3)}+\ldots+x_{(7)}+2x_{(8)}}{10} \quad , \text{ if } x_{(1)}\le x<x_{(2)}$$

$$= \frac{x+2x_{(2)}+x_{(3)}+\ldots x_{(7)}+2x_{(8)}}{10} \quad , \text{ if } x_{(2)}\le x<x_{(8)}$$

$$= \frac{2x_{(2)}+x_{(3)}+\ldots+x_{(8)}+2x}{10} \quad , \text{ if } x_{(8)}\le x<x_{(9)}$$

$$= \frac{2x_{(2)}+x_{(3)}+\ldots+x_{(8)}+2x_{(9)}}{10} \quad , \text{ if } x\ge x_{(9)}$$

This EIF of the 10%-Winsorized mean is bounded for both $x\to+\infty$ and $x\to-\infty$. Symilarly, the EIF of α-trimmed mean and that of α-Winsorized mean also be obtained and note that these empirical influence functions are always bounded for finite values of $x_{(i)}$, $(i=1,2,\ldots,n)$.

Now, we shall consider a numerical example.

<u>Example(3.3.1)</u> : For the Cushny and Peebles data (Ref.Hample (1986), page 79), we have n=10, x_n=4.6. Using this data the empirical influence functions of the simple arithmetic mean, the median (50%), the 10% and 20%-trimmed means, 10% and 20%-Winsorized means have been calculated in the following table(3.3.1), where the last observation 4.6 in the data has been replaced by an arbitrary value x.

Table No (3.3.1)

EIF of various estimators based on Cushny and Peebles data

x	EIF of Mean	EIF of 10%-trimmed mean	EIF of 20%-trimmed mean	EIF of 10%-Winsorized mean	EIF of 20%-Winsorized mean	EIF of Median (50%)
-3.0	.82	1.10	1.17	1.06	1.14	1.25
-2.0	.92	1.10	1.17	1.06	1.14	1.25
-1.0	1.02	1.10	1.17	1.06	1.14	1.25
0.0	1.12	1.10	1.17	1.06	1.14	1.25
0.1	1.13	1.11	1.17	1.08	1.14	1.25
0.7	1.19	1.19	1.17	1.18	1.14	1.25
0.8	1.20	1.20	1.17	1.22	1.14	1.25
0.9	1.21	1.21	1.18	1.23	1.17	1.25
1.0	1.22	1.23	1.20	1.24	1.20	1.25
1.2	1.24	1.25	1.23	1.26	1.22	1.25
1.3	1.25	1.26	1.25	1.27	1.23	1.30
1.8	1.30	1.33	1.33	1.32	1.36	1.30
1.9	1.31	1.34	1.33	1.34	1.36	1.30
2.0	1.32	1.35	1.33	1.36	1.36	1.30
2.4	1.36	1.40	1.33	1.44	1.36	1.30
2.5	1.37	1.40	1.33	1.44	1.36	1.30
3.0	1.42	1.40	1.33	1.44	1.36	1.30
4.0	1.52	1.40	1.33	1.44	1.36	1.30

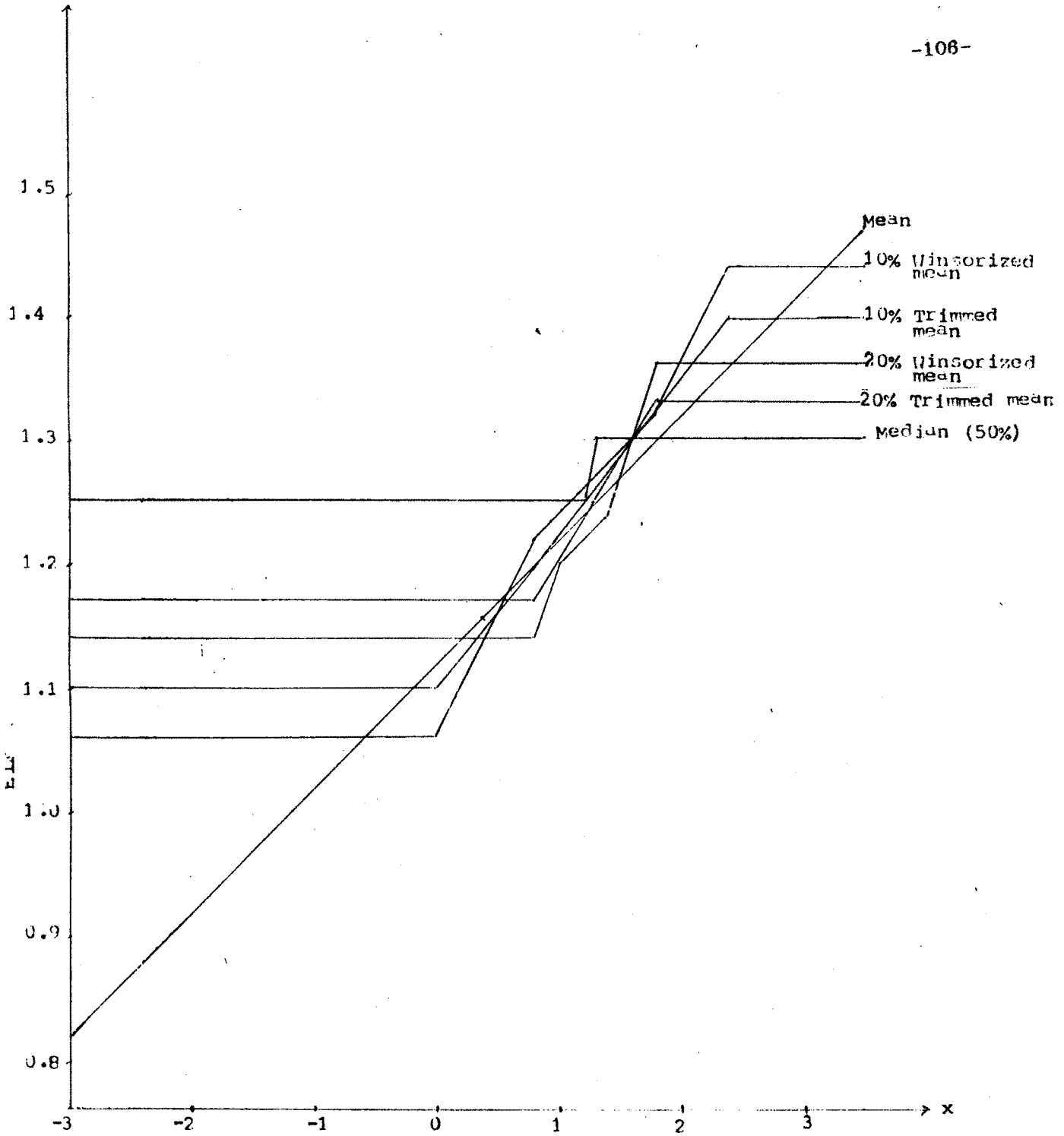

Figure (3.3.1)

Empirical influence functions of the mean, the median, the 10% and 20% - trimmed means and the 10% and 20% - Winsorized means for Cushny and Peebles data.

Figure(3.3.1) shows the empirical influence functions for the various estimators viz. the arithmetic mean, the 10% and 20%-trimmed means, the 10% and 20%-Winsorized means and the median(50%). It simply shows the values of various estimators for other various values of last observation. The EIF of the mean becomes unbounded for both $x \to +\infty$ and $x \to -\infty$, where as for the other estimators it is bounded. ∎

2) <u>Sensitivity Curve</u> : Tukey (1970) suggested another finite-sample version of IF which is known as sensitivity curve and is denoted as $Sc_n(x)$. This can be obtain, if we replace F by F_{n-1} and t by $\frac{1}{n}$ in (2.1.5). That is,

$$Sc_n(x) = \left[T\left(\frac{n-1}{n} F_{n-1} + \frac{1}{n} \Delta_x\right) - T(F_{n-1})\right] . n^{-1} \qquad (3.3.1)$$

where F_{n-1} is the empirical distribution function corresponding to $(x_1, x_2, \ldots, x_{n-1})$.

The above formula (3.3.1) is equivalent to

$$Sc_n(x) = n\left[T_n(x_1, \ldots, x_{n-1}, x) - T_{n-1}(x_1, \ldots, x_{n-1})\right], \qquad (3.3.2)$$

Where $T_{n-1}(x_1, \ldots x_{n-1})$ can be obtained from the sample $(x_1, \ldots, x_{n-1})$ of (n-1) observations and $T_n(x_1, \ldots, x_{n-1}, x)$ can be calculated by introducing an observation x in the sample $(x_1, \ldots, x_{n-1})$.

Equivalently for n fixed, one can consider (3.3.2) as

$$Sc_n(x) = T_n(x_1, \ldots, x_{n-1}, x) - T_{n-1}(x_1, \ldots, x_{n-1})$$

Now, we shall consider some special cases of sensitivity curve

i) If $n=5$ and $\alpha=.2$, then $[n\alpha]=1$. Therefore, the 20%-trimmed mean $T_{n,\alpha}$ would be the average of the remaining items after trimming the smallest and the largest items. That is,

$$T_{n,\alpha} = \left[x_{(2)}+x_{(3)}+x_{(4)}\right] (3)^{-1}$$

An additional value x is introduced in the sample and the 20%-trimmed mean $T_{n+1,\alpha}$ is obtained. The difference $T_{n+1,\alpha}-T_{n,\alpha}$, which is a function of x, is called a Sensitivity curve. That is,

$$\begin{aligned} T_{n+1,\alpha}-T_{n,\alpha} &= \frac{x_{(1)}+x_{(2)}+x_{(3)}+x_{(4)}}{4} - T_{n,\alpha}, && \text{if } x<x_{(1)} \\ &= \frac{x+x_{(2)}+x_{(3)}+x_{(4)}}{4} - T_{n,\alpha}, && \text{if } x_{(1)}\le x\le x_{(5)} \\ &= \frac{x_{(2)}+x_{(3)}+x_{(4)}+x_{(5)}}{4} - T_{n,\alpha} && \text{if } x>x_{(5)}. \end{aligned}$$

This sensitivity curve is continuous when $x_{(1)}\le x\le x_{(5)}$ and is equal to constant when $x<x_{(1)}$ or $x>x_{(5)}$. A sensitivity curve for the 10%-trimmed mean can also be obtained in a similar way. Note that these sensitivity curves are bounded, where as the sensitivity curve for arithmetic mean is not bounded.

ii) If $n=10$, $\alpha=.1$, then $[n\alpha]=1$. Therefore, the 10%-Winsorized mean will be

$$W_{n,\alpha} = \frac{2x_{(2)}+x_{(3)}+\ldots+x_{(8)}+2x_{(9)}}{10} .$$

An additional value x is introduced in the sample and the 10%-Winsorized mean $W_{n+1,\alpha}$ can be obtained. The difference, $W_{n+1,\alpha}-W_{n,\alpha}$, is a function of x and is called a sensitivity curve. That is,

$Sc_n(x)$ for 10%-Winsorized mean

$$= W_{n+1,\alpha}-W_{n,\alpha}$$

$$= \frac{2x_{(1)}+x_{(2)}\cdots+x_{(8)}+2x_{(9)}}{11} - W_{n,\alpha}, \quad \text{if } x<x_{(1)}$$

$$= \frac{2x+x_{(2)}+x_{(3)}+\cdots+x_{(8)}+2x_{(9)}}{11} - W_{n,\alpha}, \text{if } x_{(1)}\le x<x_{(2)}$$

$$= \frac{x+2x_{(2)}+x_{(3)}+\cdots+x_{(8)}+2x_{(9)}}{11} - W_{n,\alpha}, \text{if } x_{(2)}\le x<x_{(9)}$$

$$= \frac{2x_{(2)}+x_{(3)}+\cdots+x_{(9)}+2x}{11} - W_{n,\alpha}, \quad \text{if } x_{(9)}\le x<x_{(10)}$$

$$= \frac{2x_{(2)}+x_{(3)}+\cdots+x_{(9)}+2x_{(10)}}{11} - W_{n,\alpha}, \quad \text{if } x\ge x_{(10)} .$$

This sensitivity curve is bounded for all values of x. The sensitivity curves corresponding to α-trimmed mean and α-Winsorized mean are always bounded for finite values of $x_{(i)}$, $(i=1,2,\ldots,n)$.

In the following Table(3.3.2), the values of 10% and 20%-trimmed means, 10% and 20%-Winsorized means and simple arithmetic mean after introducing an additional observation x in the Cushny and Peebles data are given in column numbers 2,4,6,8 and 10 respectively. The values of sensitivity curves corresponding to these various estimators are given in column numbers 3,5,7,9 and 11 respectively.

Table (3.3.2)

Sensitivity Curve of Various Estimators based on Cushyny and Peebles data

2	3	4	5	6	7	8	9	10	11
10%-trim-med mean	$SC_n(x)$ for 10%-trim-med mean	20%-trim-med mean	$SC_n(x)$ for 20%-trim-med mean	10%-Win-sorized mean	$SC_n(x)$ for 10%-winso-rized mean	20%-Win-sorized mean	$SC_n(x)$ for 20%-Winso-rized mean	Arithme-tic mean	$SC_n(x)$ for Arithmetic mean
$= T_{n+1,\alpha}$	$= T_{n+1,\alpha} - T_{n,\alpha}$	$= \bar{T}_{n+1,\alpha}$	$= \bar{T}_{n+1,\alpha} - \bar{T}_{n,\alpha}$	$= W_{n+1,\alpha}$	$= W_{n+1,\alpha} - W_{n,\alpha}$	$= \bar{W}_{n+1,\alpha}$	$= \bar{W}_{n+1,\alpha} - \bar{W}_{n,\alpha}$	$= \bar{x}_{n+1}$	$= \bar{x}_{n+1} - \bar{x}_n$
) 1.24	-0.16	1.26	-0.07	1.24	-0.20	1.27	-0.09	1.25	-0.33
) 1.24	-0.16	1.26	-0.07	1.24	-0.20	1.27	-0.09	1.35	-0.23
5 1.24	-0.16	1.26	-0.07	1.24	-0.20	1.27	-0.09	1.39	-0.19
) 1.24	-0.16	1.26	-0.07	1.24	-0.20	1.27	-0.09	1.44	-0.14
5 1.30	-0.10	1.26	-0.07	1.33	-0.11	1.27	-0.09	1.48	-0.10
) 1.36	-0.04	1.29	-0.04	1.40	-0.04	1.33	-0.03	1.53	-0.05
5 1.41	0.01	1.36	0.03	1.44	0.00	1.37	0.01	1.56	-0.02
) 1.47	0.07	1.43	0.10	1.49	0.05	1.45	0.09	1.62	0.04
5 1.52	0.12	1.49	0.16	1.55	0.11	1.56	0.20	1.66	0.08
) 1.58	0.18	1.49	0.16	1.67	0.23	1.56	0.20	1.71	0.13
5 1.63	0.23	1.49	0.16	1.73	0.29	1.56	0.20	1.75	0.17
) 1.69	0.29	1.49	0.16	1.82	0.38	1.56	0.20	1.80	0.22
) 1.76	0.36	1.49	0.16	1.93	0.49	1.56	0.20	1.89	0.31
5 1.76	0.36	1.49	0.16	1.93	0.49	1.56	0.20	1.94	0.36
) 1.76	0.36	1.49	0.16	1.93	0.49	1.56	0.20	1.98	0.40
5 1.76	0.36	1.49	0.16	1.93	0.49	1.56	0.20	2.03	0.45

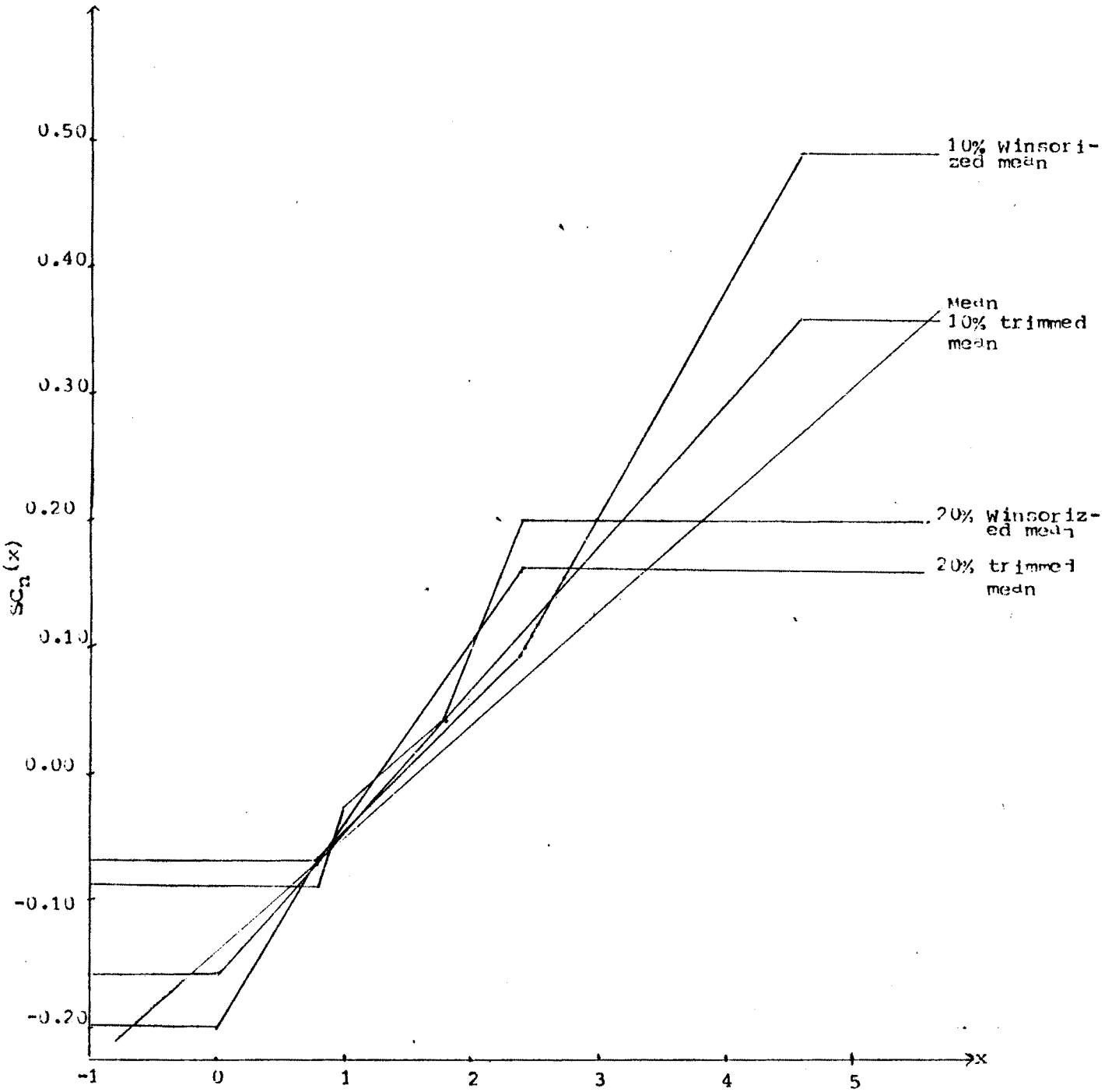

Figure (3.3.2)

Sensitivity Curves of the mean, the 10% and 20% - trimmed means, the 10% and 20% - Winsorized means for Cushny and Peebles data.

Here, for n=10, α=.1, the 10%-trimmed mean $T_{n,\alpha}$ equals 1.40 and the 10%-Winsorized mean $W_{n,\alpha}$ equals 1.44. And for n=10, α=.2, the 20%-trimmed mean $T_{n,\alpha}$ is equal to 1.33 and 20%-Winsorized mean $W_{n,\alpha}$ is equal to 1.36. The simple arithmetic mean $\bar{x}_n$ for n=10, 1.58. The figure (3.3.2) shows the sensitivity curves of the simple arithmetic mean, the 10% and 20%-trimmed means the 10% and 20%-Winsorized means, conputed for the Cushny and Peebles data, where the arbitrary value x is taken as an additional observation. From the graph, we conclude that the sensitivity curve for arithmetic mean is unbounded, where as for other estimators it is bounded. If we compare the graph of sensitivity curve for the arithmetic mean with the corresponding graphs of sensitivity curves for the other estimators, it shows a higher "influence" of the additional observation x in the central part of the sample and no influence if x falls into the left or right tail. Also we find that both the trimmed mean and Winsorized mean restrict the influence of outliers, but in different ways.

3.4 INFLUENCE FUNCTION OF M-ESTIMATORS :

Main aim of this section is to obtain IF of M-estimator. In this section first we shall discuss the Huber's approach for M-estimators and then we define M-estimator and in particular for location parameter. The IF of M-estimators is obtained. Some examples are also given.

<u>Huber's approach for M-estimators</u> : According to Hample(1986), "Huber's(1964) paper on <u>robust estimation</u> formed the first basis for a theory of robust estimation". In that paper he introduced a flexible class of estimators, called "M-estimators", which become a very useful tool in the theory of robust statistics. These estimators are just a slight generalization of MLE's, so it is also known as "generalized maximum likelihood type estimators".

Let $x_1, x_2, \ldots, x_n$ be a random sample from the distribution with density $f(x,\theta)$. To obtain MLE based on a random sample Of size n, we need to minimize $-\sum_{i=1}^{n} \ln f(x_i,\theta)$ or equivalently we need to solve the equation $\sum_{i=1}^{n} \frac{f'(x_i,\theta)}{f(x_i,\theta)} = 0$ (under suitable conditions). Where θ is a parameter.

To find M-estimator, Huber (1964) has considered the problem of minimization of $\sum_{i=1}^{n} \rho(x_i, T_n)$. Sometimes this can be done by solving the equation $\sum_{i=1}^{n} \varphi(x_i, T_n) = 0$, as a function of T, for a suitable function $\rho(\cdot)$. If $\rho = -\ln f$ or $\varphi = -f'/f$, then Huber's problem reduces to that of finding the MLE for θ. ∎

<u>M-estimator</u> : Let $x_1, x_2, \ldots, x_n$ be a random sample that arises from a distribution with density $f(x,\theta)$, where the unknown parameter θ belongs to some parameter space Θ. We consider estimators of θ which are functionals (real-valued statistics) $T_n = T_n(x_1, x_2, \ldots, x_n)$ based on sample. The estimators of type M are solutions of the more general structure, proposed by Huber (1964), is the one that minimizes

$$\sum_{i=1}^{n} \rho(x_i, T_n), \tag{3.4.1}$$

where ρ is same function on $\mathcal{X} \times \Theta$, $\mathcal{X}$ being the sample space. Suppose that ρ has a derivative $\varphi(x,\theta) = \frac{\partial}{\partial\theta} \rho(x,\theta)$, so the estimate T_n satisfies the implicit equation

$$\sum_{i=1}^{n} \varphi(x_i, T_n) = 0. \tag{3.4.2}$$

Any estimator defined by (3.4.1) or (3.4.2) is called an M-estimator.

<u>Remark (3.4.1)</u> : If H_n is the empirical distribution function generated by the sample, then the solution T_n of (3.4.2) can also be written as $T(H_n)$, where T is functional given by

$$\int \varphi(x, T(H))dH(x) = 0, \tag{3.4.3}$$

for all distributions H for which the integral is defined (Ref. Hample (1986), Page 101). ∎

<u>IF of M-estimators</u> : In order to obtain the IF of M-estimator at any distribution, we have to consider contamination

(mixture of distributions). For that purpose, we replace H by $(1-t)F+t\Delta_x$, $0<t<1$

Therefore,

$$\int \varphi(x, T[(1-t)F+t\Delta_x])d[(1-t)F+t\Delta_x](x) = 0$$

By differentiating this w.r.t. t at t=0, we get

$$\int \frac{\partial}{\partial\theta}\left[\varphi(x,\theta)\right]_{T(F)} \frac{\partial}{\partial t}\left\{T[(1-t)F+t\Delta_x]\right\}_{t=0} dF(x)$$

$$+ \int \varphi(x, T(F))d(\Delta_x - F)(x) = 0$$

Making use of (3.4.3) and using definition of IF, we obtain

$$IF(x;\varphi,F) = \frac{\varphi(x,T(F))}{-\int \frac{\partial}{\partial\theta}\left[\varphi(x,\theta)\right]_{T(F)} dF(y)} \qquad (3.4.4)$$

where $\frac{\partial}{\partial t}T\left[(1-t)F+t\Delta_x\right]\Big|_{t=0} = IF(x;\varphi,F)$.

In general (3.4.4) can be written as

$$IF(x;\varphi,F) = -\frac{\varphi}{\int \varphi' \, dF} \quad ,$$

under the assumption that the denominator is nonzero.

From (3.4.4) we can say that the IF of an M-estimate is proportional to φ. The φ is B-robust at F iff $\varphi(\cdot, T(F))$ is bounded.

The GES of M-estimator is given by

$$r^*(\varphi,F) = \sup_x \left|\varphi \cdot \left[\int \varphi' \, dF\right]^{-1}\right| \qquad (3.4.5)$$

and the asymptotic variance of M-estimator is given by

$$V(\varphi,F) = \int IF(x;\varphi,F)^2 dF(x)$$

$$= \frac{\int \varphi^2(x;T(F))\ dF(x)}{\left[\int \frac{\partial}{\partial\theta}[\varphi(x,\theta)]_{T(F)} dF(x)\right]^2} . \qquad (3.4.6) \blacksquare$$

<u>M-estimators for location</u> ; Let $x_1, x_2, \ldots, x_n$ be a random sample that arises from a distribution with density $f(x-\theta)$ of the continuous type, where θ is a location parameter. The logarithm of the likelihood function is

$$\ln L(\theta) = \sum_{i=1}^{n} \ln f(x_i-\theta) = -\sum_{i=1}^{n} \rho(x_i-\theta)$$

where $\rho(x) = -\ln f(x)$.

In maximum likelihood we wish to maximize $\ln L(\theta)$ or, in terms of the ρ function, minimize $K(\theta)=\sum_{i=1}^{n} \rho(x_i-\theta)$. Suppose that this minimization can be achieved by differentiating and solving $K'(\theta)=0$, that is, finding the appropriate θ that satisfies

$$\sum_{i=1}^{n} \varphi(x_i-\theta) = 0,$$

where $\quad \varphi(x) = \rho'(x) = -\dfrac{f'(x)}{f(x)}$.

The solution of this equation that minimizes $K(\theta)$ is called the maximum likelihood or M-estimator of θ and is denoted as T_n.

If the functional $T(\cdot)$ is the M-functional corresponding to φ, where it is assumed that

$$\int \varphi dF(x) = 0,$$

in order that $T(\cdot)$ be Fisher consistent, then, as before, $IF(x;\varphi,F)$ can be obtained and is given by

$$IF(x;\varphi,F) = \frac{\varphi(x-T(F))}{\int \varphi'(x-T(F))dF(x)} , \qquad (3.4.7)$$

with the assumption that the denominator is nonzero. ∎

Let us illustrate (3.4.7) through an example. If F is normal distribution function with mean θ, then using (3.4.7), we have

$$IF(x;\varphi,F) = \frac{x-\theta}{\int \frac{\partial}{\partial\theta}(x-\theta)dF(x)}$$

$$= \frac{x-\theta}{-\int dF(x)}$$

$$= \theta - x , \qquad -\infty < x < \infty.$$

In the following we obtain M-estimators for location parameter corresponding to various $\varphi(\cdot)$ functions. ∎

<u>Example (3.4.1)</u> : Let $\varphi(x) = x$. Therefore, the corresponding M-estimator T_n satisfies the equation

$$\sum_{i=1}^{n} \varphi(x_i - T_n) = 0,$$

that is,

$$\sum_{i=1}^{n} (x_i - T_n) = 0,$$

that is,

$$T_n = \bar{x} .$$

Equivalently, when $\rho(x) = \frac{x^2}{2} + C$, the M-estimator for location parameter will be the sample mean. ∎

R E F E R E N C E S

1. Cox D.R. and Hinkley D.V. (1974).
   Theoritical Statistics. Chapman and Hall, London.

2. Hample F.R. (1968).
   Contribution to the theory of robust estimation.
   Ph.D. Thesis. University of Colifornia, Berkely.

3. Hample F.R. (1971).
   A general qualitative definition of robustness. The Annals of Mathematical Statistics.
   Vol.42, No.6, P.1887-1896.

4. Hample F.R. (1974).
   The influence curve and its role in robust estimation.
   J.Am.Statist.Assoc.69, P.383-398.

5. Hample F.R. et. al (1986).
   Robust Statistics. The approach based on influence functions. John Wiley and Sons.

6. Hodges J.L. and Lehman E.L. (1963).
   Estimates of location based on rank tests.
   Ann.Math.Statist.34, P.598-611.

7. Huber P.J. (1964).
   Robust estimation of location parameter.
   Ann.Math.Statist.35, P.73-101.

8. Huber P.J. (1972).
   Robust Statistics : A review.
   Ann.Math.Statist.43, P.1041-1067.

9. Huber P.J. (1981).
   Robust Statistics.
   Wiley, New York.

10. Kendall M.G. and Stuart A. (1960).
    The Advanced Theory of Statistics.
    Vol-2, Griffin London.

11. Rao C.R. (1966).
    Linear Statistical Inference and its Applications.
    John Wiley and Sons, Inc., New Delhi.

12. Rey W.J.J. (1983).
    Introduction to Robust and Quasi-Robust Statistical Methods. Universitext, Springer, Heidelberg.

13. Rohatgi V.K. (1976).
An Introduction to Probability Theory and Mathematical Statistics. John Wiley and Sons., New York.

14. Serfling R.J. (1980).
Approximation Theorems of Mathematical Statistics. Wiley, New York.

15. Shoemaker L.H. (1984).
Robustness Properties for a class of scale estimators. Commun.Statist.A 13,P.15-28.

16. Tukey J.W. (1960).
A survey of sampling from contaminated distributions, in: Contributions to Probability and Statistics. I.Olkin,Ed.,Stanford University Press, Stanford, Calif. ∎