CHAPTER II

INFLUENCE FUNCTION

2.0 INTRODUCTION

In this chapter we define the influence function of an estimator (statistical functional) and explain the use of it for the study of robustness.

Some properties of influence function (IF) are also given. The IF is essentially the first derivative of an estimator (viewed as functional) at some distribution (for details see definition 2.1.6). A number of examples of robust and non-robust estimators are given. IF of r^{th} raw and central moments at distribution F are obtained. The relation between the IF of S.D. and that of variance is obtained. The influence functions for the s^{th} quantile and the q-quantile range have been obtained. The IF of MLE is also obtained. Also influence functions for coefficient of skewness and coefficient of kurtosis are calculated. Some theorem and lemmas, related to IF, are given.

2.1 INFLUENCE FUNCTION AND RELATED MEASURES OF ROBUSTNESS

Hample (1968) introduced the idea of the IF. The IF was originally called the "influence curve", however, the more general name "influence function", in view of the generalization to higher dimensions. The IF describes the effect of an additional observation x, on a statistic (functional T(.)), given a (large) sample with distribution F and it is denoted by the notation IF(x;T,F).

Before going to definition of IF, let us introduce the definition of derivative that has been used to investigate the properties of estimators. This derivative is known as Gateaux derivative.

Gateaux derivative [Differentiation of functionals T(.)] : Given two points F and H in the space $\mathcal{F}$ of all distribution functions, the "line segment" in $\mathcal{F}$ joining F and H consists of the set of distribution functions $\{(1-t)F+tH,\ 0 \le t \le 1\}$ [also written as $\{F+t(H-F),\ 0 \le t \le 1\}$]. Consider a functional T defined on $F+t(H-F)$ for all sufficiently small t. If the limit

$$d_1T(F;\ H-F) = \lim_{t \downarrow 0} \frac{T(F+t(H-F)) - T(F)}{t} \qquad (2.1.1)$$

exists, it is called the Gateaux differential of T at F in the direction of H.

Note that $d_1T(F;\ H-F)$ is simply the ordinary right-hand derivative, at $t = 0$, of the function

$Q(t) = T[F + t(H-F)]$ of the real variable t. That is

$$d_1T(F;\ H-F) = \frac{\partial}{\partial t} T[F + t(H-F)] \Big|_{t=0}, \qquad (2.1.2)$$

provided the limit exists.

Let us illustrate one example of Gateaux derivative. The r^{th} central moment of a distribution F may be expressed as a functional

$$T(F) = \mu_r = \int (x - \mu_F)^r \, dF(x)$$

where $\mu_F = \int x \, dF(x)$

If distribution G is such that $G = (1-t)F+tH = F+t(H-F)$, $o \le t \le 1$, then

$$\mu_G = \int x \, dG(x) = \mu_F + t\,(\mu_H - \mu_F)$$

Therefore,

$$T(G) = \int (x-\mu_G)^r \, dG(x)$$

$$= \int (x-\mu_G)^r \, dF(x)$$

$$+ t \int (x-\mu_G)^r \, d\left[H(x)-F(x)\right] \qquad (2.1.3)$$

In order to obtain Gateaux derivative $d_1T(F;H-F)$, it is enough to find $\frac{\partial}{\partial t} T(\,F + t(\,H - F\,))\,\Big|_{t=0}$. Hence, by differentiating (2.1.3) w.r.t. t, we get

$$\frac{\partial}{\partial t} T(G) = -r(\mu_H - \mu_F) \int (x-\mu_G)^{r-1}\, dF(x) + \int (x-\mu_G)^r\, d[H(x)-F(x)] - tr(\mu_H-\mu_F)\int (x-\mu_G)^{r-1}\, d[H(x)-F(x)]$$

$$= -r(\mu_H - \mu_F')\int (x-\mu_G)^{r-1}\, dG(x) + \int (x-\mu_G)^r\, d[H(x)-F(x)].$$

Therefore,

$$\frac{\partial}{\partial t} T(G)\Big|_{t=0} = \int [(x-\mu)^r - r\mu_{r-1}x]\, d[H(x)-F(x)]$$

Thus, Gateaux derivative is given by

$$d_1T(F,\ H-F) = (x-\mu)^r - r\mu_{r-1}x - E_F[(x-\mu)^r - r\mu_{r-1}x]. \quad \blacksquare$$

Now in the following, we define the IF. Let $\mathbb{R}$ be the real line, let T be a real-valued functional defined on some subset of the set of all probability measures on R, and let F denote a probability measure on $\mathbb{R}$ for which T is defined. Denote Δ_x the probability measure determined by the point mass one in any given point $x \in \mathbb{R}$. Mixture of F and some Δ_x are written as $(1-t)F + t\Delta_x$, for $o<t<1$. Then the influence function of T at F is defined as

$$IF(x;T,F) = \lim_{t \downarrow o} \left[\frac{T[(1-t)\ F + t\Delta_x] - T(F)}{t} \right], -\infty<x<\infty. \quad (2.1.5)$$

if the limit defined for every point $x \in \mathbb{R}$.

It may also be written as

$$IF(x;T,F) = \frac{\partial}{\partial t} T\left[(1-t)F + t\Delta_x\right] \Big|_{t=o}, \quad -\infty<x<\infty. \qquad (2.1.6)$$

provided the limit exists.

Note that the right member of (2.1.5) is the directional derivative of T at F, in the direction of Δ_x.

From the above definition of IF, we observe that the IF describes the influence of an additional observation x on the estimate T {This statistic T depends on a large number of i.i.d. observations drawn from F and is consistent for functional T(F)}. The IF(x;T,F) is the first derivative of functional T at an underlying distribution F. The IF is a collection of directional derivatives in the directions of the point masses Δ_x, and is usually evaluated at the model distribution F. It is a very useful tool to define some important robustness measures such as the gross-error sensitivity, local-shift sensitivity and rejection point which will be defined in subsequent section. The influence function measures the effects of infinitesimal perturbations and its purpose is to measure the differential effect of a point mass at x that has on the functional of interest. ∎

Example (2.1.1). For exponential distribution with parameter $\theta > 0$ having p.d.f.

$$f(x) = \theta e^{-\theta x} \,.\quad x \geq 0$$
$$= 0 \quad , \text{Otherwise}$$

Let $T(F)$ = mean of the exponential distribution

$$= \frac{1}{\theta}$$

Therefore,

$$T\left[(1-t)F + t\Delta_n\right] = \text{Mean of } \left[(1-t)F + t\Delta_x\right]$$
$$= (1-t)\frac{1}{\theta} + tx,$$

where Δ_x is the distribution function having point mass of one at x ($x \geq o$).

By definition (2.1.5), we have

$$IF(x;T,F) = \lim_{t \downarrow o} \left[\frac{T[(1-t)F + t\Delta_x] - T(F)}{t} \right]$$
$$= \lim_{t \downarrow o} \left[\frac{\frac{(1-t)}{\theta} + tx - \frac{1}{\theta}}{t} \right]$$
$$= x - \frac{1}{\theta}, \qquad x \geq 0$$

and

$$V(T,F) = \int IF(x;T,F)^2 \, dF(x)$$
$$= \int_o^{\infty} \left[x - \frac{1}{\theta}\right]^2 \theta e^{-\theta x} \, dx$$
$$= \frac{1}{\theta^2}.$$

Where V(T,F) is an asymptotic varianee

<u>Remark (2.1.1)</u> : From the remark (iv) on page 222 of Serfling (1980), note that there exists a function $T_1[F;x]$, $x \in \mathbb{R}$, such that

$$d_1T(F;\Delta_x-F) = T_1[F;x] - \int T_1[F;x]\, dF(x)$$

In the following we prove some mathematical properties of IF. The proof of <u>Lemma (2.1.1)</u> basically depends on remark (2.1.1), which states that the expected value of IF is always zero. <u>Lemma (2.1.2)</u> is concerned with the asymptotic variance of statistical functional T(F). We note that, if T(F) = K (constant) for all F then IF(x;T,F) = o.

<u>Lemma (2.1.1)</u> : If F is a distribution function and T(F) is a statistical functional at F, then

$$\int IF(x;T,F)\, dF(x) = o$$

proof : consider

$$\begin{aligned}
L.H.S &= \int IF(x;T,F)\, dF(x) \\
&= \int d_1T(F;\Delta_x-F)\, dF(x) \quad \text{By using (2.1.2) and (2.1.6)} \\
&= \int \Big[T_1[F:x] - \int T_1[F;x]\, dF(x)\Big]dF(x) \quad \text{by remark (2.1.1)} \\
&= \int T_1[F;x]\, dF(x) - \int T_1[F;x]\, dF(x) \\
&= 0 \\
&= R.H.S.
\end{aligned}$$

That is the expected value of IF is always zero. ∎

9925

Let $x_1, x_2, \ldots, x_n$ be i.i.d. from F and F_n be the emperical distribution function. Let T(F) be the functional of interest and define the statistic $T_n = T(x_1, x_2 \ldots x_n) = T(F_n)$.

By asymptotic variance of T(F) we mean the asymptotic variance of $\sqrt{n}\ \left[T(F_n) - T(F)\right]$.

Lemma (2.1.2) :

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

Proof : Let the observations $x_i (i=1,2,\ldots)$ be i.i.d. with common distribution F and T(F) be the statistical functional at F. If some distribution H is near F, then the leading terms of first-order Von Mises expansion of T at F which is derived from a Taylor expansion [Hample(1986),P.85], evaluated in H is given by

$$T(H) = T(F) + \int IF(x;T,F) \, d(H-F)(x) + \text{remainder},$$

that is

$$T(H) - T(F) = \int IF(x;T,F) \, dH(x) + \text{remainder} \qquad (2.1.7)$$

since, by lemma (2.1.1), $\int IF(x;T,F) \, dF(x) = 0$

If the observations x_i are i.i.d. with common continuous distribution F, then by the Glivenko-Cantelli theorem (Rohatgi, (1976),P.300), the emperical distribution F_n will tend to F. In order to obtain the required relation, we substitute the emperical distribution F_n for H in the above expression (2.1.7), we obtain

$$\sqrt{n}\,\left[T(F_n) - T(F)\right] = \sqrt{n}\int IF(x;T,F)\,dF_n(x) + \text{remainder}$$

$$\sqrt{n}\,\left[T_n - T(F)\right] = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} IF(x_i;T,F) + \text{remainder} \quad (2.1.8)$$

$$= \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i + \text{remainder}$$

$$= \sqrt{n}\;\bar{Z}$$

Here, the remainder is asymptotically negligible. Since the observations x_i are independent with common distribution F, then by CLT, we have

$$\sqrt{n}\;\bar{Z} \sim N\left[\sqrt{n}\;E(Z_i),\; V(Z_i)\right]$$

where

$$E(Z_i) = \int IF(x;T,F)\,dF(x) = 0$$

and

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

Hence, the leading term on the right-hand side of (2.1.8) is asymptotically normal with mean zero and variance

$$\int IF(x;T,F)^2\,dF(x)$$

Thus,

$$\sqrt{n}\,\left[T_n - T(F)\right] \sim N\left[0, V(T,F)\right]$$

where, the asymptotic variance equals

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

Hence the required. ∎

Following are some important measures based on IF, which measure the robustness properties.

1) Gross-Error Sensitivity [GES] : The gross-error sensitivity of T at F is the supremum of the absolute value of

the IF. That is

$$\text{Gross-Error Sensitivity} = r^* = \sup_x |IF(x;T,F)|, \qquad (2.1.9)$$

where $IF(x;T,F)$ exists for some T and F, and the supremum being taken over all x. Gross-error sensitivity r^* may also be denoted as $r^*(T,F)$.

Gross-error sensitivity is the central local robustness measure, measuring the maximum bias caused by infinitesimal contamination. It measures the worst possible influence which a small amount of contamination of fixed size can have on the value of the estimator. Therefore, it may be regarded as an upper bound on the asymptotic bias of the estimator.

From the value of $r^*(T,F)$ we can conclude the following.

i) Since $r^*(T,F)$ is the supremum of the absolute value of IF, so the influence of any outlier cannot exceed $r^*(T,F)$.

ii) If $r^*(T,F)$ is finite, then we can say that T is B-robust at F. Here B is used to mean the boundedness of gross-error.

iii) If $r^*(T,F)$ is positive minimum for Fisher-consistent estimators, then T is the most B-robust at F. That is an estimator minimizing GES is known as the most B-robust.

iv) If the gross-error sensitivity of T at F is infinite, then we can say that T is not robust.

11) Local-Shift Sensitivity : The second summary value of the IF, which is also important for robustness considerations, is the local-shift sensitivity. When some values of observations are changed slightly (as happens in rounding and grouping and due to some local inaccuracies), this has a certain measurable effect on estimate. Intuitively, the effect of shifting an observation slightly from the point x to some neighboring point y can be measured by means of IF(y;T,F) - IF(x;T,F), because an observation is added at y and another one is removed at x. Therefore, the effect of "Wiggling" around x is approximatly described by a normalized defference or simply the slope of IF in that point. A measure for the worst effect of "Wiggling" the observations is therefore provided by the local-shift sensitivity, which is denoted as λ^* and is defined as

$$\lambda^* = \sup_{x \neq y} \left| IF(y;T,F) - IF(x;T,F) \right| \; . \; |y-x|^{-1} \qquad (2.1.10)$$

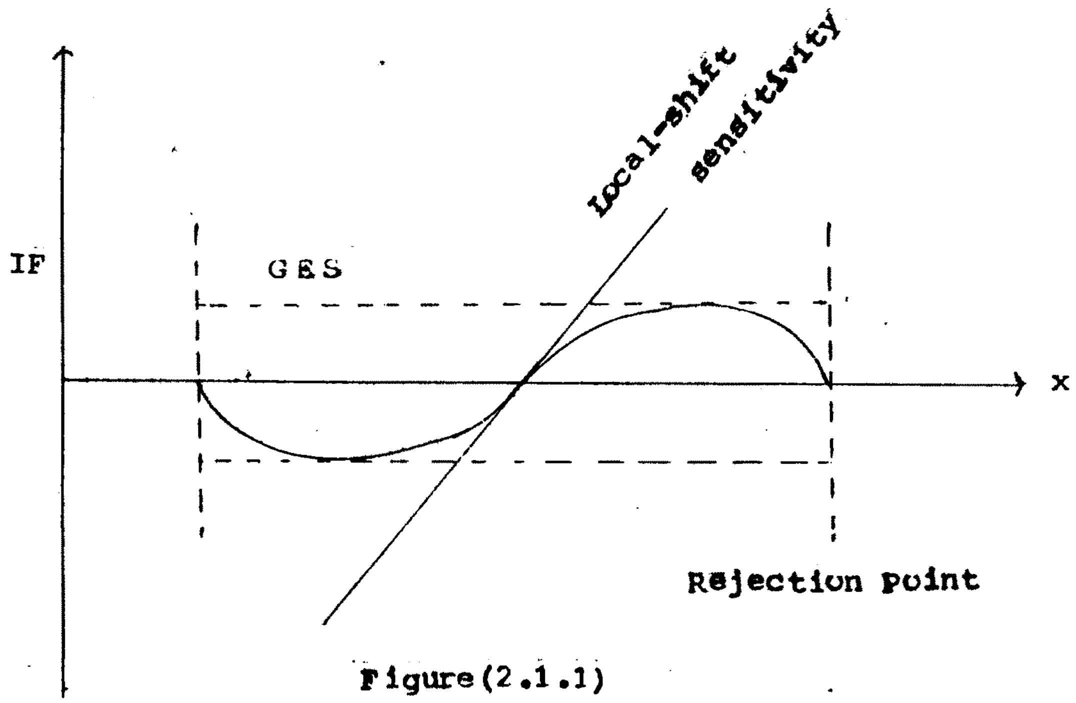
For the proper interpretation of λ^*, however, one has to keep in mind that it refers only to local changes of the value of the estimator, so that even an infinite value of λ^* may refer only to a very limited actual changes.

111) Rejection point : It is an old robustness idea to reject extreme outliers entirely. It is aften of interest to know whether an estimator rejects outliers and, if so, at what

distance?. In the language of the IF, this means that the IF vanishes outside a certain area. Indeed, there may be a region outside of which the influence function is identically zero. The contamination in those points does not have any influence at all. The distance from the centre of symmetry of a distribution to the point at which the influence function becomes identically zero, is called the "rejection point". It is denoted by the notation ρ^* and is defined as

$$\rho^* = \inf\{r>0; IF(x;T,F) = 0 \text{ when } |x|>r\}. \qquad (2.1.11)$$

All observations farther away than the rejection point are rejected completely. Therefore, it is a desirable feature, if ρ^* is finite. Note that, if there exists no such r, then $\rho^* = \infty$ by definition of the infimum.



Figure(2.1.1)

The sketch of various properties of an influence function

Let us work-out an example to illustrate the above.

Example(2.1.2) : Let us compute the IF of mean for poisson distribution where the sample space $\mathcal{X}$ equals the set of nonnegative integers $\{0,1,2,...\}$. Let $x_1, x_2, ..., x_n$ are i.i.d. with respect to poisson distribution $F_\theta(x)$, where the unknown parameter θ belongs to the parameter space $\Theta = (0,\infty)$. The density function of F_θ is

$$f_\theta(x) = \frac{e^{-\theta}\theta^x}{x!}, \quad x = 0,1,2,...$$

Let the estimator $T_n = \frac{1}{n} \sum_{i=1}^{n} x_i$, and the corresponding functional is $T(F) = \sum_{x=0}^{\infty} xf(x)$ with existing first moment about origin (mean) for any distribution F on $\mathcal{X} = \{0,1,2,...\}$. This functional is clearly Fisher consistent, because

$$T(F_\theta) = \sum_{x=0}^{\infty} xf_\theta(x) = \sum_{x=0}^{\infty} x \frac{e^{-\theta}\theta^x}{x!} = \theta e^{-\theta} \sum_{x=1}^{\infty} \frac{\theta^{x-1}}{(x-1)!} = \theta,$$

for all θ in $\Theta = (0,\infty)$.

Now,

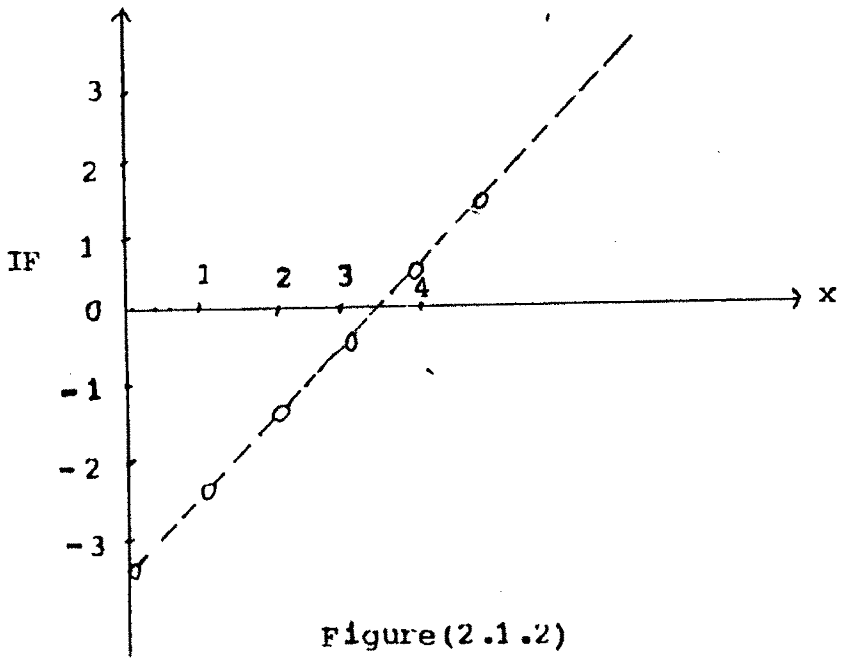
$$T\left[(1-t)F_\theta + t\Delta_x\right] = \text{Mean of } \left[(1-t)F_\theta + t\Delta_x\right]$$
$$= (1-t)\theta + tx$$

Therefore, by definition (2.1.5), we have

$$\begin{aligned} IF(x;T,F) &= \lim_{t \downarrow 0} \left[\frac{T[(1-t)F_\theta + t\Delta_x] - T(F_\theta)}{t} \right] \\ &= \lim_{t \downarrow 0} \left[\frac{(1-t)\theta + tx - \theta}{t} \right] \\ &= x - \theta \end{aligned}$$

Here the contaminating point masses may only occur in points $X \in \mathcal{X}$, so the IF(x;T,F) can only be calculated at integer values of x. Thus, the IF of T at poission distribution is a discrete set of points.

The pictorial representation of IF of the mean at poisson distribution F with a specific value of $\theta = 3.5$ is shown in the Figure(2.1.2). The IF is only defined for integer arguments.



Figure(2.1.2)

Influence function of the mean at the poisson distribution with $\theta = 3.5$

The gross-error sensitivity is given by

$$r^*(T,F) = \sup_x |IF(x;T,F)|$$

$$= \sup_x |x-\theta|$$

$$= \infty$$

since IF is unbounded so that the mean of the poisson distribution is not robust.

Remark(2.1.4) : We can generalise the above result for any distribution F. If $T(F) = \int x dF(x)$ with existing mean θ at F, then IF of the mean at F is given by

$$IF(x;T,F) = x-\theta, \quad -\infty<x<\infty \tag{2.1.12}$$

and GES is

$$r^{*}(T,F) = \infty , \quad \text{when } x\to\infty \tag{2.1.13}$$

Comment : The IF of the mean for any distribution is equal to $x-\theta$, $-\infty<x<\infty$, where the mean of the distribution exists and is equal to θ and the IF of variance at any distribution is $(x-\theta)^2 - \sigma^2$, $-\infty<x<\infty$, (proved in particular case of Lemma (2.2.2)), whenever the second moment about mean exists. In case of poisson distribution although the mean and the variance are equal, their influence functions are not equal. ∎

In the following sections we compute the IF of various functionals like,

i) $T^r(F) = \int x^r dF(x) = \int_0^1 \left[F^{-1}(t)\right]^r dt = \mu^r$

ii) $T_r(F) = \int (x-\mu)^r dF(x) = r^{th}$ central moment of distribution $F=\mu_r$, where μ is the mean of distribution F

iii) $Q_s(F) = s^{th}$ quantile of distribution F

iv) $S_q(F) =$ q-quantile range of distribution F

v) $T_s(F)$ = skewness of distribution F

vi) $T_{cs}(F)$ = coefficient of skewness of distribution F

vii) $T_{ck}(F)$ = coefficient of Kurtosis of distribution F.

The existence of all these functionals is assumed.

In further discussion, the distribution G is used to denote the mixture of distributions F and Δ_x such that

$$G = (1-t)F + t\Delta_x, \quad 0 \le t \le 1.$$

2.2 IF OF THE r^{th} MOMENT :

Lemma(2.2.1) : If the r^{th} moment $T^r(F) = \int x^r dF(x)$ at F with existing r^{th} moment μ^r about the origin, then

$$IF(x;T^r,F) = x^r - \mu^r, \qquad x \in \mathbb{R}.$$

Proof : We are given the r^{th} moment $T^r(F) = \int x^r dF(x)$ at F with existing r^{th} moment μ^r about the origin.

Therefore,

$$\begin{aligned} T^r(G) &= r^{th} \text{ moment about origin of distribution } G \\ &= \int y^r \, dG(y) \\ &= \int y^r \, d\left[(1-t)F + \Delta_x\right](y) \\ &= (1-t)\mu^r + x^r \end{aligned}$$

Thus, by definition (2.1.5), we have

$$\begin{aligned} IF(x;T^r,F) &= \lim_{t \downarrow 0} \left\{ \frac{T^r(G) - T^r(F)}{t} \right\} \\ &= \lim_{t \downarrow 0} \left\{ \frac{(1-t)\mu^r + tx^r - \mu^r}{t} \right\} \\ &= x^r - \mu^r \qquad x \in \mathbb{R} \qquad (2.2.1) \end{aligned}$$

Hence the proof.

In partcular, if r=1, we get the IF of the mean which is

$$IF(x;T^1,F) = x - \mu.$$

Lemma(2.2.2): If for a given distribution F,

$T_r(F) = \int (x-\mu)^r \, dF(x)$ with existing the r^{th} order central moment μ_r, then $IF(x;T_r,F)$ exists and is given by

$$IF(x;T_r,F) = (x-\mu)^r - r\mu_{r-1}(x-\mu) - \mu_r, \quad x \in \mathbb{R}.$$

where μ is the mean of distribution F which is known.

Proof: Here, we have

$T_r(F) = \int (x-\mu)^r \, dF(x)$ at F with existing the r^{th} order central moment μ_r.

Therefore,

$$T_r(G) = r^{th} \text{ order central moment of distribution } G.$$

$$= \int \left[y - \{(1-t)\mu + tx\}\right]^r dG(y)$$

where $(1-t)\mu + tx$ is the mean of distribution G.

That is

$$T_r(G) = \int \left[(y-\mu) - t(x-\mu)\right]^r d\left[(1-t)F + t\Delta_x\right](y)$$

$$= \int \Big[(y-\mu)^r - rt(y-\mu)^{r-1}(x-\mu) + \frac{r(r-1)}{2!} t^2(y-\mu)^{r-2}(x-\mu)^2$$

$$- \ldots + (-1)^r t^r (x-\mu)^r\Big] \, d\left[(1-t)F + t\Delta_x\right](y)$$

$$= (1-t)\ \left[\mu_r - rt(x-\mu)\mu_{r-1} + \frac{r(r-1)}{2!}\, t^2(x-\mu)^2\mu_{r-2} - \ldots + (-1)^r t^r (x-\mu)^r\right] + t(1-t)^r(x-\mu)^r$$

Hence

$$IF(x;T_r,F) = \lim_{t\downarrow 0}\left\{\frac{T_r\{(1-t)F + t\Delta_x\} - T_r(F)}{t}\right\}$$

$$= (x-\mu)^r - r(x-\mu)\mu_{r-1} - \mu_r \qquad (2.2.2)$$

Hence the required.

In partcular, if $r = 2$, we get the IF of variance at F which is

$$IF(x;T_2,F) = (x-\mu)^2 - \sigma^2$$

where $T_2(F) = \int (x-\mu)^2\, dF(x)$ with existing the variance σ^2.

Lemma(2.2.3): If the standard deviation (S.D)

$T'(F) = \sqrt{\int (x-\mu)^2\, dF(x)}$ at F with existing S.D. σ and known mean μ of F. Then

$$IF(x;T',F) = \frac{(x-\mu)^2 - \sigma^2}{2\sigma}\ , \quad x \in \mathbb{R},\ \sigma>0$$

Proof: We are given the S.D. $T'(F) = \sqrt{\int (x-\mu)^2\, dF(x)}$ which is defined for all probability measures with existing S.D. σ.

Therefore,

$T'(G)$ = S.D. of distribution G.

$$= \sqrt{\int \left[y-[(1-t)\mu+tx]\right]^2 \, dG(y)}$$

$$= \sqrt{\int \left[(y-\mu) - t(x-\mu)\right]^2 \, d\left[(1-t)F+t\Delta_x\right](y)}$$

$$= \sqrt{\int \left[(y-\mu)^2 + t^2(x-\mu)^2\right] \, d\left[(1-t)F+t\Delta_x\right](y)}$$

$$= \sqrt{\int (1-t)\left[\sigma^2 + t^2(x-\mu)^2\right] + t(1+t^2)(x-\mu)^2}$$

$$= \sqrt{B} \quad \text{(say)}$$

Hence

$$IF(x;T',F) = \lim_{t \downarrow 0} \left\{ \frac{T'(G) - T'(F)}{t} \right\}$$

$$= \lim_{t \downarrow 0} \left\{ \frac{\sqrt{B} - \sqrt{A}}{t} \right\} \text{ since } T'(F) = \sqrt{\sigma^2} = \sqrt{A}$$

$$= \lim_{t \downarrow 0} \left\{ \frac{B - A}{t(\sqrt{B} + \sqrt{A})} \right\}$$

$$= \lim_{t \downarrow 0} \left\{ \frac{t(1-t)(x-\mu)^2 - \sigma^2 + (1+t^2)(x-\mu)^2}{\sqrt{B} + \sqrt{A}} \right\}$$

$$= \frac{(x-\mu)^2 - \sigma^2}{2\sigma} \qquad (2.2.3)$$

Hence the proof.

Now we obtain the relation between the IF of S.D. and that of variance. From Lemma(2.2.2), the IF of variance is given by

$$IF(x;T_2,F) = (x-\mu)^2 - \sigma^2 \qquad (2.2.4)$$

and from Lemma(2.2.3), the IF of S.D. is given by

$$IF(x;T',F) = \frac{(x-\mu)^2 - \sigma^2}{2\sigma} \tag{2.2.5}$$

from (2.2.4) and (2.2.5), we have

$$IF(x;T_2,F) = 2\sigma\ IF(x;T',F) \tag{2.2.6}$$

which is required relation between IF of S.D. and that of variance.

2.3 THE INFLUENCE FUNCTIONS OF s^{th} QUANTILE, q-QUANTILE RANGE AND MLE :

For the distribution H, let $Q_s(H)$ be the s^{th} quantile of H. That is, $Q_s(H)$ is such that $H(Q_s(H)) = H(H^{-1}(s)) = s$, $0<s<1$.

IF for the s^{th} quantile : Let $Q_s(F)$ and $Q_s(G)$ respectively be the s^{th} quantiles of F and G. Where $G = (1-t)F + t\Delta_x$. Here existence of quantile is assumed. We know that from (2.1.5)

$$IF(x;Q_s,F) = \lim_{t \downarrow 0} \left\{ \frac{Q_s(G)-Q_s(F)}{t} \right\} = \frac{\partial}{\partial t} \left[Q_s(G)\right]\Big|_{t=0} \tag{2.3.1}$$

In the following we shall find the R.H.S of (2.3.1). We need to consider the following three cases.

Case I : If $x > Q_s(F) = F^{-1}(s)$

If $x > Q_s(F)$, then

$$G\left[Q_s(G)\right] = (1-t)\ F(Q_s(G)) + 0.t$$

$$s = (1-t)\ F(Q_s(G))$$

$$F\left[Q_s(G)\right] = \frac{s}{1-t}$$

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

$$\frac{\partial}{\partial t}\left[Q_s(G)\right] = \frac{s}{(1-t)^2}\ \frac{1}{f(Q_s(G))}$$

$$\frac{\partial}{\partial t}\left[Q_s(G)\right]\Big|_{t=0} = \frac{s}{f(F^{-1}(s))} \qquad (2.3.2)$$

Therefore, from (2.3.1), we have

$$IF(x;Q_s,F) = \frac{s}{f(F^{-1}(s))}, \qquad x>F^{-1}(s)$$

Case II : If $x< Q_s(F) = F^{-1}(s)$

If $x< Q_s(F)$, then

$$G\left[Q_s(G)\right] = (1-t)\ F(Q_s(G)) + t$$

$$s = (1-t)\ F(Q_s(G)) + t$$

$$F\left[Q_s(G)\right] = \frac{s-t}{1-t}$$

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

$$\frac{\partial}{\partial t}\left[Q_s(G)\right] = \frac{s-1}{(1-t)^2}\ \frac{1}{f(Q_s(G))}$$

$$\frac{\partial}{\partial t}\left[Q_s^s(G)\right]\Big|_{t=0} = \frac{s-1}{f(F^{-1}(s))}$$

Therefore,

$$IF(x;Q_s,F) = \frac{s-1}{f(F^{-1}(s))}, \qquad x<F^{-1}(s)$$

Thus, the IF for s^{th} quantile of F is

$$\begin{aligned} IF(x;Q_s,F) &= \frac{s}{f(F^{-1}(s))}, && \text{if } x>F^{-1}(s) \\ &= \frac{s-1}{f(F^{-1}(s))}, && \text{if } x<F^{-1}(s) \qquad (2.3.3) \\ &= 0, && \text{if } x = F^{-1}(s) \end{aligned}$$

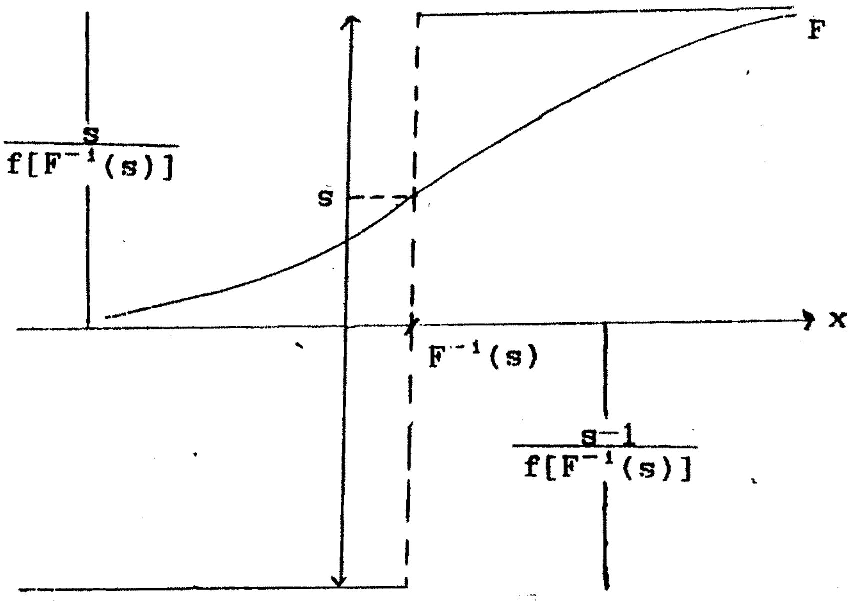
and GES is

$$r^*(Q_s,F) = \frac{s}{f(F^{-1}(s))}$$

$$= \text{finite}$$

This implies that the s^{th} quantile (if it is exists) is B-robust at any distribution.

Graphical representation of the influence function for the s^{th} quantile at F is shown in the following figure (2.3.1).



Figure(2.3.1)
Influence function of the s^{th} quantile

<u>Particular cases</u> :

i) If $s = \frac{1}{2}$, then $Q_{1/2}(F)$ and $Q_{1/2}(G)$ become the medians of distributions F and G respectively.

Therefore,

$$IF\left[x;Q_{1/2},F\right] = \frac{1}{2f(F^{-1}(1/2))}, \quad \text{if } x>F^{-1}(1/2)$$

$$= \frac{-1}{2f(F^{-1}(1/2))}, \quad \text{if } x<F^{-1}(1/2)$$

$$= 0 \quad , \quad \text{if } x = F^{-1}(1/2).$$

This can also be written as

$$IF(x;Q_{1/2},F) = \frac{\text{sign}(x-F^{-1}(1/2))}{2f(F^{-1}(1/2))} \qquad (\text{sign } 0 = 0)$$

$$= \frac{\text{sign}(F(x)-1/2)}{2f(F^{-1}(1/2))}$$

and

$$GES = \frac{1}{2f(F^{-1}(1/2))} = \text{finite}.$$

ii) If $s = \frac{1}{2}$ and $F = \Phi$ (the standard normal distribution), then $Q_{1/2}(\Phi) = 0$

Therefore,

$$F^{-1}(s) = \Phi^{-1}(1/2) = 0$$

and

$$f\left[F^{-1}(s)\right] = \phi(0) = \frac{1}{\sqrt{2\pi}}$$

Thus, the IF of the median at the standard normal distribution (S.N.D) is

$$IF\left[x;Q_{1/2},\Phi\right] = \frac{1}{2\phi(0)}, \quad \text{if } x>0$$

$$= \frac{-1}{2\phi(0)}, \quad \text{if } x<0$$

$$= 0 \quad , \quad \text{if } x = 0$$

It can also be written as

$$IF\left[x;Q_{1/2},\Phi\right] = \frac{\text{sign}(x)}{2\phi(0)} = \text{sign}(x)\ \left[\left(\frac{\pi}{2}\right)^{1/2}\right],$$

where sign(x) = 1,0,−1 as x>0, x = 0 or x<0.

and

$$GES = \frac{1}{2\phi(0)} = \left(\frac{\pi}{2}\right)^{1/2} = 1.253,$$ which is minimal value, so the median is most B-robust. ∎

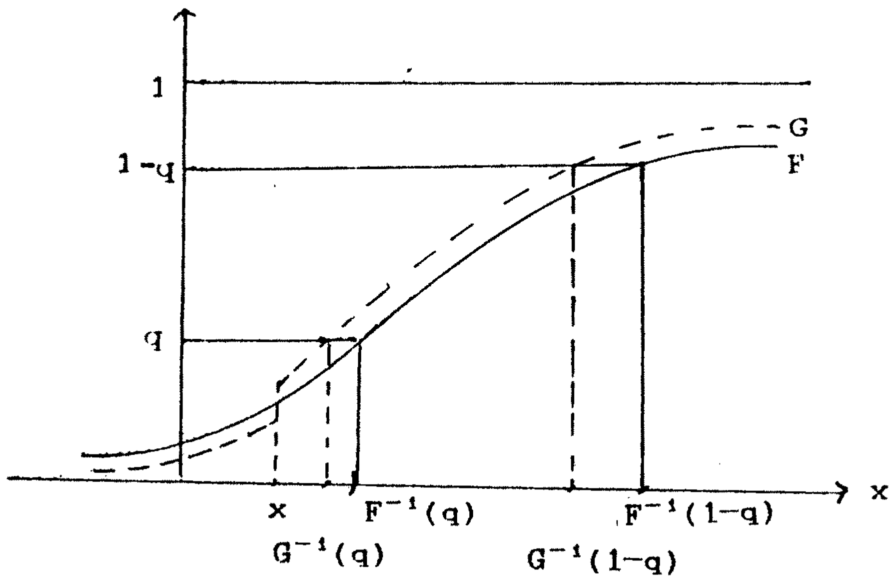
For distribution F, let $S_q(F)$ be the q-quantile range of F, defined by $S_q(F) = F^{-1}(1-q) - F^{-1}(q)$, $0<q<1/2$.

<u>IF for the q-quantile range</u> :Let $S_q(F)$ and $S_q(G)$ respectively be the q-quantile ranges of F and G. Where distribution G is, as usual, the mixture of distributions F and Δ_x. Here the existence of q-quantile range corresponding to distributions F and G is assumed. We know that

$$IF(x;S_q,F) = \frac{\partial}{\partial t}\left[S_q(G)\right]\Big|_{t=0} . \qquad (2.3.4)$$

So that, in order to obtain the IF for the q-quantile range at F, it is enough to find $\frac{\partial}{\partial t}\left[S_q(G)\right]\Big|_{t=0}$. We need to consider the following three cases.

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



Figure(2.3.2)

The q-quantile range of G is

$$S_q(G) = G^{-1}(1-q) - G^{-1}(q)$$

$$= F^{-1}\left[\frac{1-q-t}{1-t}\right] - F^{-1}\left[\frac{q-t}{1-t}\right], \qquad (2.3.5)$$

where

$$G\left[G^{-1}(q)\right] = (1-t)\ F(G^{-1}(q))+t$$

$$q = (1-t)\ F(G^{-1}(q))+t$$

Therefore,

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

and

$$G\left[G^{-1}(1-q)\right] = (1-t)\ F(G^{-1}(1-q))+t$$

Therefore,

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

Differentiating (2.3.5) w.r.t. t, we get

$$\frac{\partial}{\partial t}\left[S_q(G)\right] = \frac{1}{f[F^{-1}(\frac{1-q-t}{1-t})]}\ \frac{\partial}{\partial t}\left[\frac{1-q-t}{1-t}\right] - \frac{1}{f[F^{-1}(\frac{q-t}{1-t})]}\ \frac{\partial}{\partial t}\left[\frac{q-t}{1-t}\right]$$

$$= \frac{1}{f[F^{-1}(\frac{1-q-t}{1-t})]}\ \frac{-q}{(1-t)^2} - \frac{1}{f[F^{-1}(\frac{q-t}{1-t})]}\ \frac{q-1}{(1-t)^2}$$

$$\frac{\partial}{\partial t}\left[S_q(G)\right]\Big|_{t=0} = \frac{-q}{f(F^{-1}(1-q))} - \frac{q-1}{f(F^{-1}(q))}$$

Therefore, from(2.3.4), we have

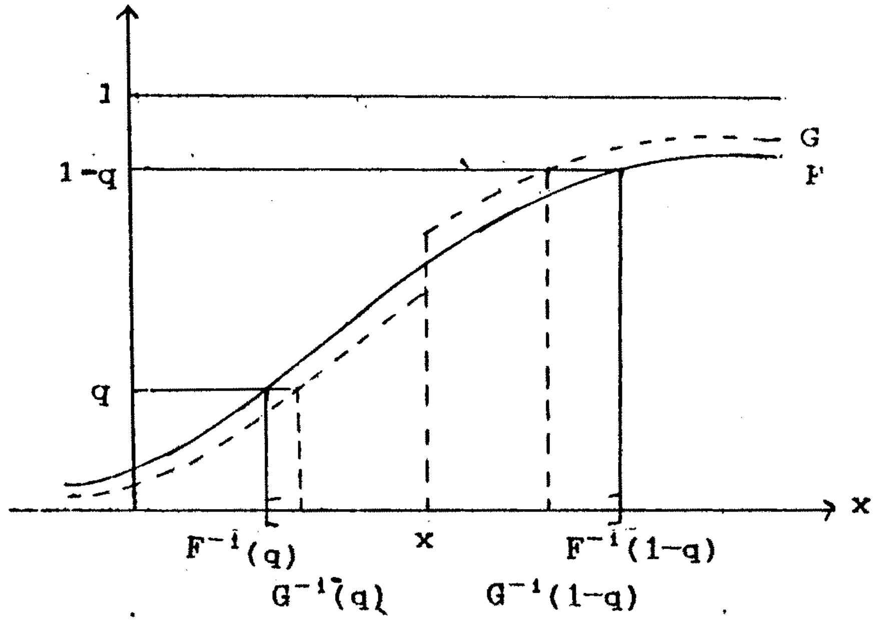
$$IF(x;S_q,F) = \frac{1}{f(F^{-1}(q))} - q\left[\frac{1}{f(F^{-1}(q))} + \frac{1}{f(F^{-1}(1-q))}\right.$$

$$= \frac{1}{f(F^{-1}(q))} - C(F),\ \text{if } x<F^{-1}(q), \qquad (2.3.6)$$

where

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

Case II : If $F^{-1}(q)<x<F^{-1}(1-q)$



Figure(2.3.3)

Here the q-quantile range of G is

$$S_q(G) = G^{-1}(1-q) - G^{-1}(q)$$

$$= F^{-1}\left[\frac{1-q-t}{1-t}\right] - F^{-1}\left[\frac{q}{1-t}\right], \qquad (2.3.7)$$

where

$$G\left[G^{-1}(q)\right] = (1-t)\ F(G^{-1}(q))$$

Therefore,

$$G^{-1}(q) = F^{-1}\left(\frac{q}{1-t}\right)$$

and

$$G\left[G^{-1}(q)\right] = (1-t)\ F^{-1}(G^{-1}(1-q))+t$$

Therefore,

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

Differentiating (2.3.7) w.r.t. t, we get

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

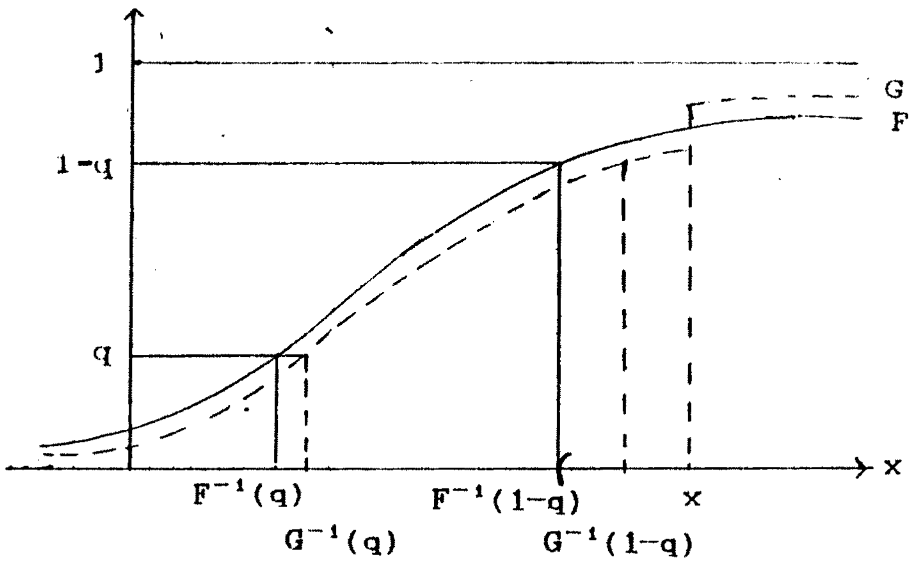
$$\frac{\partial}{\partial t}\left[S_q(G)\right]\Big|_{t=0} = \frac{-q}{f(F^{-1}(1-q))} - \frac{q}{f(F^{-1}(q))}$$

Therefore,

$$IF(x;S_q,F) = -q\left[\frac{1}{f(F^{-1}(1-q))} + \frac{1}{f(F^{-1}(q))}\right]$$

$$= -C(F), \text{ if } F^{-1}(q)<x<F^{-1}(1-q) \qquad (2.3.8)$$

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



Figure(2.3.4)

In this case the q-quantile range of G is

$$S_q(G) = G^{-1}(1-q) - G^{-1}(q)$$

$$= F^{-1}\left[\frac{1-q}{1-t}\right] - F^{-1}\left[\frac{q}{1-t}\right], \tag{2.3.9}$$

where

$$G\left[G^{-1}(q)\right] = (1-t)\ F(G^{-1}(q))$$

Therefore,

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

and

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

Differentiating (2.3.9) w.r.t. t and putting t = 0, we get

$$\frac{\partial}{\partial t}\left[S_q(G)\right]\Big|_{t=0} = \frac{1-q}{f(F^{-1}(1-q))} - \frac{q}{f(F^{-1}(q))}$$

Therefore,

$$IF(x;S_q,F) = \frac{1}{f(F^{-1}(1-q))} - q\left[\frac{1}{f(F^{-1}(q))} + \frac{1}{f(F^{-1}(1-q))}\right]$$

$$= \frac{1}{f(F^{-1}(1-q))} - C(F),\ \text{if } x>F^{-1}(1-q) \tag{2.3.10}$$

Thus, from(2.3.6), (2.3.8) and (2.3.10), the IF for the q-quantile range will be

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

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

$$= \frac{1}{f(F^{-1}(1-q))} - C(F), \quad \text{for } x>F^{-1}(1-q), \tag{2.3.11}$$

where

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

and GES is given by

$$r^*(S_q, F) = \frac{1}{f(F^{-1}(1-q))} - C(F)$$

$$= \text{Finite}$$

This implies that the q-quantile range is robust at any distribution.

In particular, (i) if $q = \frac{1}{4}$ then $S_{1/4}$ (F) becomes interquartile range of F which is

$$S_{1/4}(F) = F^{-1}(3/4) - F^{-1}(1/4) \qquad (2.3.12)$$

and its corresponding IF of interquartile range at F is

$$\begin{aligned} IF(x;S_{1/4},F) &= \frac{1}{f(F^{-1}(1/4))} - C(F), && \text{if } x<F^{-1}(1/4) \\ &= - C(F), && \text{if } F^{-1}(1/4)<x<F^{-1}(3/4) \\ &= \frac{1}{f(F^{-1}(3/4))} - C(F), && \text{if } x>F^{-1}(3/4), \qquad (2.3.13) \end{aligned}$$

where

$$C(F) = \frac{1}{4}\left[\frac{1}{f(F^{-1}(1/4))} + \frac{1}{f(F^{-1}(3/4))}\right]$$

ii) if $q = \frac{1}{4}$ and if $S^*_{1/4}(F)$ denotes the Quartile Deviation of F, then

$$S^*_{1/4}(F) = \frac{1}{2}\left[F^{-1}(3/4) - F^{-1}(1/4)\right]$$

and its corresponding $IF(x;S^*_{1/4},F)$ will be

$$IF(x;S^*_{1/4},F) = \frac{1}{2f(F^{-1}(1/4))} - C(F), \quad \text{if } x<F^{-1}(1/4)$$
$$= -C(F), \quad \text{if } F^{-1}(1/4)<x<F^{-1}(3/4)$$
$$= \frac{1}{2f(F^{-1}(3/4))} - C(F), \quad \text{if } x>F^{-1}(3/4), \qquad (2.3.14)$$

where

$$C(F) = \frac{1}{8}\left[\frac{1}{f(F^{-1}(1/4))} + \frac{1}{f(F^{-1}(3/4))}\right]$$

and (iii) if F is symmetric about zero, then these formulae (2.3.11), (2.3.13) and (2.3.14) respectively reduce to

$$IF(x;S_q,F) = \frac{1-2q}{f(F^{-1}(q))}, \quad \text{if } x<F^{-1}(q) \text{ or } x>F^{-1}(1-q)$$
$$= \frac{-2q}{f(F^{-1}(q))}, \quad \text{if } F^{-1}(q)<x<F^{-1}(1-q),$$
$$IF(x;S_{1/4},F) = \frac{1}{2f(F^{-1}(1/4))}, \quad \text{if } x<F^{-1}(1/4) \text{ or } x>F^{-1}(3/4)$$
$$= \frac{-1}{2f(F^{-1}(1/4))}, \quad \text{if } F^{-1}(1/4)<x<F^{-1}(3/4)$$

and

$$IF(x;S^*_{1/4},F) = \frac{1}{4f(F^{-1}(1/4))}, \quad \text{if } x<F^{-1}(1/4) \text{ or } x>F^{-1}(3/4)$$
$$= \frac{1}{4f(F^{-1}(1/4))}, \quad \text{if } F^{-1}(1/4)<x<F^{-1}(3/4).$$

<u>IF of maximum likelihood estimator</u> : Under regularity conditions (given in section 4.4.2, P.152 of Serfling,(1980)) on the family of distributions $\{F(x;\theta),\theta\in\Theta\}$ under consideration, the maximum likelihood estimate of θ is the

solution of

$$\int \frac{\partial}{\partial\theta} \ln f(x;\theta)\, dF_n(x) = 0$$

That is, the maximum likelihood estimate is $\theta(F_n)$, where $\theta(F)$ is the functional defined as the solution of

$$\int \frac{\partial}{\partial\theta} \ln f(x;\theta)\, dF(x) = 0$$

Let $\theta(G) = \theta(F+t(\Delta_x-F))$ denote the functional of maximum likelihood estimate defined as the solution of

$$\int \frac{\partial}{\partial\theta} \ln g(x;\theta)\, dG(x) = 0$$

That is solution of $H(\theta(G),\ t) = 0$

In order to obtain $IF(x;\theta(F),F)$, it is enough to find

$$\frac{\partial}{\partial t}\, \theta(G)\Big|_{t=0} = \frac{\partial}{\partial t}\, \theta(F+t(\Delta_x-F))\Big|_{t=0}$$

by implicit differentiation through the equation

$$H(\theta(G),t) = 0,$$

where

$$G = F+t(\Delta_x-F) \text{ and } H(\theta,t) = \int \frac{\partial}{\partial\theta} \ln g(x;\theta)\, dG(x).$$

we have

$$\frac{\partial H}{\partial\theta}\Big|_{\theta=\theta(F)} \cdot \frac{\partial}{\partial t}\, \theta(G)\Big|_{t=0} + \frac{\partial H}{\partial t}\Big|_{t=0} = 0$$

that is

$$\frac{\partial}{\partial t}\, \theta(G)\Big|_{t=0} = -\frac{\partial H}{\partial t}\Big|_{t=0} \cdot \left\{ \frac{\partial H}{\partial t}\Big|_{\theta=\theta(F)} \right\}^{-1} \qquad (2.3.15)$$

Here

$$\frac{\partial H}{\partial t}\Big|_{t=0} = \frac{\partial}{\partial\theta} \ln f(x;\theta) \qquad (2.3.16)$$

and

$$\frac{\partial H}{\partial \theta}\Big|_{\theta=\theta(F)} = \int \frac{\partial^2}{\partial \theta^2} \ln f(x;\theta)\, dF(x)$$

$$= -\int \left[\frac{\partial}{\partial \theta} \ln f(x;\theta)\right]^2 dF(x) \qquad (2.3.17)$$

using (2.3.16) and (2.3.17) in (2.3.15), we get

$$\frac{\partial}{\partial t}\, \theta(F+t(\Delta_x - F))\Big|_{t=0} = \frac{\partial}{\partial \theta} \ln f(x:\theta).\ \frac{1}{\int \left[\frac{\partial}{\partial \theta} \ln f(x;\theta)\right]^2 dF(x)}$$

thus,

$$IF(x;\theta(F),F) = \frac{\frac{\partial}{\partial \theta} \ln f(x;\theta)}{\int \left[\frac{\partial}{\partial \theta} \ln f(x;\theta)\right]^2 dF(x)}, \qquad (2.3.18)$$

which is IF of the MLE

2.4 INFLUENCE FUNCTIONS FOR THE COEFFICIENT OF SKEWNESS AND COEFFICIENT OF KURTOSIS :

IF for the coefficient of skewness : Let the functionals $T_{cs}(F)$ and $T_r(F)$ be the coefficient of skewness and the r^{th} central moment of distribution F respectively and let, corresponding to distribution G, these functionals be $T_{cs}(G)$ and $T_r(G)$ respectively.

Therefore,

$$T_{cs}(F) = \frac{[T_3(F)]^2}{[T_2(F)]^3}$$

and

$$T_{cs}(G) = \frac{[T_3(G)]^2}{[T_2(G)]^3}$$

where

$$T_2(G) = \text{second central moment at } G$$

$$= \int \left[y-[(1-t)\mu+tx]\right]^2 dG(y)$$

$$= (1-t)\mu_2+t(1-t)(x-\mu)^2 \qquad (2.4.1)$$

and

$$T_3(G) = \text{third central moment at } G$$

$$= (1-t)\left[\mu_3-t^3(x-\mu)^3-3t(x-\mu)\mu_2\right]+t(1-t)^3(x-\mu)^3$$

Here, the existence of all these functionals at F and G is assumed.

Therefore,

$$IF(x;T_{cs},F) = \lim_{t \downarrow 0} \left[\frac{T_{cs}(G)-T_{cs}(F)}{t} \right]$$

$$= \lim_{t \downarrow 0} \frac{1}{t} \left\{ \frac{[T_3(G)]^2}{[T_2(G)]^3} - \frac{[T_3(F)]^2}{[T_2(F)]^3} \right\}$$

$$= \lim_{t \downarrow 0} \frac{1}{t} \left\{ \frac{\left\{(1-t)\left[\mu_3-t^3(x-\mu)^3-3t(x-\mu)\mu_2\right]+t(1-t)^3(x-\mu)^3\right\}^2}{\left[(1-t)\mu_2+t(1-t)(x-\mu)^2\right]^2} - \frac{\mu_3^2}{\mu_2^3} \right\}$$

$$= \lim_{t \downarrow 0} \left[\frac{\left\{ \mu_2^3 (1-t)^2 \left[\mu_3 - t^3 (x-\mu)^3 - 3t(x-\mu)\mu_2\right]^2 + \mu_2^3 \left[t(1-t)^3 (x-\mu)\right]^2 + 2\mu_2^3 t(1-t)^4 (x-\mu)^3 \left[\mu_3 - t^3 (x-\mu)^3 - 3t(x-\mu)\mu_2\right] - \mu_3^2 \left[(1-t)\mu_2 + t(1-t)(x-\mu)^2\right]^3 \right\}}{t\mu_2^3 \left[(1-t)\mu_2 + t(1-t)(x-\mu)^2\right]^3} \right]$$

$$= \lim_{t \downarrow 0} \left[\frac{\left\{ \mu_2^3\ \mu_3^2 - 6\ \mu_2^4\ \mu_3 (x-\mu) + 2(1-t)^4\ \mu_2^3\ \mu_3 (x-\mu)^3 - 3(1-t)^3 \mu_2^2\ \mu_3^2 (x-\mu)^2 + \text{terms containing } t \text{ with powers more than one} \right\}}{\mu_2^3 \left[(1-t)\mu_2 + t(1-t)(x-\mu)^2\right]^3} \right]$$

$$= \frac{\mu_3 \left[\mu_2\ \mu_3 - 6\mu_2^2 (x-\mu) - 3\mu_3 (x-\mu)^2 + 2\mu_2 (x-\mu)^3\right]}{\mu_2^4}, \quad \mu_2 \neq 0 \qquad (2.4.2)$$

Remark(2.4.1): For symmetric distribution, the IF of the coefficient of skewness is zero.

Remark(2.4.2): The resut (2.4.2) can also be abtained by defferentiating $\left\{ \frac{[T_3(G)]^2}{[T_2(G)]^3} \right\}$ partially w.r.t. t at t = 0.

IF for the coefficient of kurtosis: Let the functionals $T_{ck}(F)$ and $T_{ck}(G)$ be the coefficients of Kurtosis of distribution F and G respectively and let the functionals $T_r(F)$ and $T_r(G)$ be the r^{th} central moments of F and G respectively.

Therefore,

$$T_{ck}(F) = \frac{T_4(F)}{\left[T_2(F)\right]^2}$$

and

$$T_{ck}(G) = \frac{T_4(G)}{\left[T_2(G)\right]^2}$$

where

$$T_2(G) = (1-t)\mu_2+t(1-t)(x-\mu)^2 \quad \text{from} \quad (2.4.3)$$

and

$$T_4(G) = (1-t)\left[\mu_4-4t\mu_3(x-\mu)+6t^2\mu_2(x-\mu)^2 +t^4(x-\mu)^4\right]+t(1-t)^4(x-\mu)^4.$$

Here we assume the existence of mean, second and fourth central moments at F and G.

Therefore,

$$IF(x;T_{ck},F) = \lim_{t\downarrow 0}\left[\frac{T_{ck}(G) - T_{ck}(F)}{t}\right]$$

$$= \lim_{t\downarrow 0}\frac{1}{t}\left\{\frac{T_4(G)}{\left[T_2(G)\right]^2} - \frac{T_4(F)}{\left[T_2(F)\right]^2}\right\}$$

$$= \lim_{t\downarrow 0}\frac{1}{t}\left\{\frac{(1-t)\left[\mu_4-4t\mu_3(x-\mu)+6t^2\mu_2(x-\mu)^2+t^4(x-\mu)^4\right]+t(1-t)^4(x-\mu)^4}{\left[(1-t)\mu_2+t(1-t)(x-\mu)^2\right]^2} - \frac{\mu_4}{\mu_2^2}\right\}$$

$$= \lim_{t\downarrow 0}\left\{\frac{\mu_4\mu_2^2-4(1-t)\mu_3(x-\mu)\mu_2^2+(1-t)^4(x-\mu)^4\mu_2^2-\mu_2\mu_4(x-\mu)^2 + \text{terms containing t with powers han one}}{\mu_2^2\left[(1-t)\mu_2+t(1-t)(x-\mu)^2\right]^2}\right\}$$

$$= \frac{\mu_2\left[\mu_4-4\mu_3(x-\mu)+(x-\mu)^4\right] - 2\mu_4(x-\mu)^2}{\mu_2^3}, \quad \mu_2 \neq 0 \qquad (2.4.4)$$

2.5 SOME RESULTS RELATED TO INFLUENCE FUNCTION:

In the following we are giving some theorems which are related to IF.

Theorem (2.5.1): Let $T_1, T_2, \ldots, T_k$ be functionals and

$$T(F) = \sum_{i=1}^{k} a_i T_i(F). \text{ Then}$$

$$IF(x;T,F) = \sum_{i=1}^{k} a_i IF(x;T_i,F)$$

provided the influence functions of $T_i(i=1,2,\ldots,k)$ exist.

Proof: By definition (2.1.5), we have

$$IF(x;T,F) = \lim_{t \downarrow 0} \left\{ \frac{T\left[(1-t)F+t\Delta_x\right]-T(F)}{t} \right\}$$

$$= \lim_{t \downarrow 0} \left\{ \frac{\sum_{i=1}^{k} a_i T_i\left[(1-t)F+t\Delta_x\right] - \sum_{i=1}^{k} a_i T_i(F)}{t} \right\}$$

$$= \sum_{i=1}^{k} a_i \left\{ \lim_{t \downarrow 0} \frac{T_i\left[(1-t)F+t\Delta_x\right]-T_i(F)}{t} \right\}$$

$$= \sum_{i=1}^{k} a_i IF(x;T_i,F)$$

Hence the proof.

symilarly, we can prove the following

Corollary(2.5.1): If T_1 and T_2 be the two functionals corresponding to the distribution F and $T = aT_1 - bT_2$, then

$$IF(x;T,F) = aIF(x;T_1,F) - bIF(x;T_2,F)$$

where a and b are the constants.

For example, i) if we put $s = \frac{1}{4}$ in (2.3.3), then the IF of the first quartile at F will be

$$IF(x:Q_{1/4},F) = -\frac{3}{4}\,\frac{1}{f(F^{-1}(1/4))}, \quad \text{for } x<F^{-1}(1/4)$$
$$= \frac{1}{4}\,\frac{1}{f(F^{-1}(1/4))}, \quad \text{for } x>F^{-1}(1/4)$$

and if we put $s = \frac{3}{4}$ in (2.3.3), then the IF of the third quartile at F will be

$$IF(x;Q_{3/4},F) = -\frac{1}{4}\,\frac{1}{f(F^{-1}(3/4))}, \quad \text{for } x<F^{-1}(3/4)$$
$$= \frac{3}{4}\,\frac{1}{f(F^{-1}(3/4))}, \quad \text{for } x>F^{-1}(3/4)$$

thus,

$$IF(x;Q_{3/4})F) - IF(x,Q_{1/4},F)$$

$$= \frac{3}{4}\,\frac{1}{f(F^{-1}(1/4))} - \frac{1}{4}\,\frac{1}{f(F^{-1}(3/4))}, \quad \text{for } x<F^{-1}(1/4)$$
$$= -\frac{1}{4}\,\frac{1}{f(F^{-1}(3/4))} - \frac{1}{4}\,\frac{1}{f(F^{-1}(1/4))}; \quad \text{for } F^{-1}(1/4)<x<F^{-1}(3/4)$$
$$= \frac{3}{4}\,\frac{1}{f(F^{-1}(3/4))} - \frac{1}{4}\,\frac{1}{f(F^{-1}(1/4))}, \quad \text{for } x>F^{-1}(3/4)$$

That is,

$$IF(x;Q_{3/4},F)-IF(x;Q_{1/4},F) = \frac{1}{f(F^{-1}(1/4))} - C(F), \quad \text{for } x<F^{-1}(1/4)$$
$$= -C(F), \quad \text{for } F^{-1}(1/4)<x<F^{-1}(3/4)$$
$$= \frac{1}{f(F^{-1}(3/4))} - C(F), \quad \text{for } x>F^{-1}(3/4),$$

which is equivalent to $IF(x;S_{1/4},F)$,

where

$$C(F) = \frac{1}{4}\left\{ \frac{1}{f(F^{-1}(1/4))} + \frac{1}{f(F^{-1}(3/4))} \right\}.$$

This implies that, if $S_{1/4} = Q_{3/4} - Q_{1/4}$, then

$$IF(x;S_{1/4},F) = IF(x;Q_{3/4},F) - IF(x;Q_{1/4},F).$$

ii) we know that, skewness = 3 (mean-median).

that is, $T_S = 3(T_1-T_2)$.

Therefore, by Corollary (2.5.1), we can show that

$$IF(x;T_S,F) = 3IF(x;T_1,F) - 3IF(x;T_2,F).$$ ∎

In the following, we state and prove some results regarding asymptotic information inequality for a sequence of Fisher consistent estimators, which are related to IF.

Theorem(2.5.2): Asymptotic Information Inequality

[Asymptotic Cramer-Rao Inequality]: If the sequence of estimators $\{T_n;n\geq 1\}$ for which the corresponding functional T of distribution F is Fisher consistent, then

$$\int IF(x,T,F_*)^2\, dF_*(x) \geq \frac{1}{J(F_*)}$$

where $J(F_*)$ is Fisher information

Proof: Let density of F_θ be f_θ and put $F_{\theta*} = F_*$, where θ_* is some fixed member of Θ.

The functional T is Fisher consistent, therefore,

$$T(F_\theta) = \theta, \text{ for all } \theta \text{ in } \Theta. \qquad (2.5.1)$$

By the difinition of Fisher information, we have

$$J(F_*) = \int \left[\frac{\partial}{\partial\theta} [\ln f_\theta(x)]_{\theta_*} \right]^2 dF(x) \qquad (2.5.2)$$

where $0<J(F_*)<\infty$.

If some distribution H is "near" F, then the first-order Von-Mises expansion of T at F (which is derived from a Taylor series) evaluated in H is given by

$$T(H) = T(F)+\int IF(x;T,F)d(H-F)(x) + \text{remainder}$$

In the above expression if we put $H = F_\theta$ and $F = F_*$, we get

$$T(F_\theta) = T(F_*)+\int IF(x;T,F_*)d(F_\theta-F_*)(x) + \text{remainder}. \qquad (2.5.3)$$

We note that the remainder is asymptotically negligible, hence the expression(2.5.3) becomes

$$T(F_\theta) = T(F_*)+\int IF(x;T,F_*)d(F_\theta-F_*)(x),$$

which reduces to

$$T(F_\theta) = T(F_*)+\int IF(x;T,F_*)\ dF_\theta(x),$$

due to the fact that

$$\int IF(x;T,F_*)\ dF_*(x) = 0. \qquad \text{by lemma}$$

Therefore,

$$\frac{\partial}{\partial\theta} \left[T(F_\theta)\right] = \frac{\partial}{\partial\theta}\left[\int IF(x;T,F_*)dF_\theta\right].$$

That is

$$\frac{\partial}{\partial\theta} \left[\int IF(x;T,F_*)\ dF_\theta\right]_{\theta*} = \frac{\partial}{\partial\theta} \left[T(F_\theta)\right]_{\theta*}$$

$$= \left\{\frac{\partial\theta}{\partial\theta}\right\}_{\theta*} \qquad \text{since from (2.5.1)}$$

$$= 1$$

The L.H.S of (2.5.4) can also be written, by changing the order of defferentiation and integration, as

$$1 = \int IF(x;T,F_*)\frac{\partial}{\partial\theta}\left[f_\theta(x)\right]_{\theta*}dx$$

$$= \int IF(x;T,F_*)\frac{\partial}{\partial\theta}\left[\ln f_\theta(x)\right]_{\theta*}dF_*(x).$$

using the cauchy-schwarz inequality, we have

$$1 \le \left[\int IF(x;T,F_*)^2 dF_*(x)\right]\left[\int\left\{\frac{\partial}{\partial\theta}\left[\ln f_\theta(x)\right]_{\theta*}\right\}^2 dF_*(x)\right]$$

Therefore,

$$\int IF(x;T,F_*)^2 dF_*(x) \ge \frac{1}{J(F_*)} \qquad (2.5.5)$$

Hence the proof.

Remark(2.5.1): The equality holds in the asymptotic information inequality(2.5.5) iff

$IF(x;T,F_*)$ is proportional to $\frac{\partial}{\partial\theta}\left[\ln f_\theta(x)\right]_{\theta*}$,

that is, the estimator is asymptotically efficient iff

$$IF(x;T,F_*) = J^{-1}(F_*)\frac{\partial}{\partial\theta}\left[\ln f_\theta(x)\right]_{\theta*}$$

Remark(2.5.2): based on (2.5.5) the (absolute) asymptotic efficiency of an estimator is given by

$$e = \left[V(T,F_*)J(F_*)\right]^{-1}$$

Theorem(2.5.3): If the estimator $T(F_n)$ for which the corresponding functional at F is $T(F) = E(x)$ and x has the p.d.f $f_\theta(x) = h(x)\ C(\theta)\ \exp\left[xQ(\theta)\right]$, then to prove that the estimator is asymptotically efficient, provided

$$Q'(\theta) \neq 0, \ \forall \ \theta \in \Theta, \text{ where } Q'(\theta) = \frac{\partial}{\partial\theta} Q(\theta)$$

Proof: Here we have to prove the asymptotic efficiency of the estimator $T(F_n)$. That is, from remark (2.5.1), it is enough to show that

$$IF(x;T,F) = J^{-1}(F) \frac{\partial}{\partial T(F)} \left[\ln f_\theta(x)\right] \tag{2.5.6}$$

If T(F) is the mean of the distribution F (= $m(\theta)$ say), then from Remark(2.1.4) the L.H.S of (2.5.6) will be

$$IF(x;T,F) = x-m(\theta) \tag{2.5.7}$$

Now, we have

$$f_\theta(x) = h(x)C(\theta) \exp\{x.Q(\theta)\}$$

$$\ln f_\theta(x) = \ln h(x)+\ln C(\theta) + x.Q(\theta)$$

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

$$\frac{\partial}{\partial\theta} \left[\ln f_\theta(x)\right] = \frac{C'(\theta)}{C(\theta)} + xQ'(\theta), \text{ where } C'(\theta) = \frac{\partial}{\partial\theta} C(\theta)$$

Also, we have

$$m(\theta) = \int xh(x) \ C(\theta) \ \exp\{x.Q(\theta)\}dx$$

Differentiating this partially w.r.t. θ under integral sign, we get

$$\frac{\partial}{\partial\theta} m(\theta) = \int xh(x) \left[C'(\theta)\exp[x.Q(\theta)]+C(\theta)\exp[x.Q(\theta)]xQ'(\theta)\right]dx$$

$$= \frac{m(\theta)}{C(\theta)} C'(\theta) + E(X^2) \ Q'(\theta)$$

Therefore,

$$\frac{\partial}{\partial m(\theta)}\left[\ln f_\theta(x)\right] = \frac{\partial}{\partial\theta}\left[\ln f_\theta(x)\right]. \frac{\partial}{\partial m(\theta)}$$

$$= \left[\frac{C'(\theta)}{C(\theta)} + xQ'(\theta)\right]\left[\frac{m(\theta)}{C(\theta)} C'(\theta) + E(X^2)Q'(\theta)\right]^{-1}$$

$$= \left[\frac{C'(\theta)}{C(\theta)Q'(\theta)} + x\right]\left[\frac{C'(\theta)}{C(\theta)} . \frac{m(\theta)}{Q'(\theta)} + E(X^2)\right]^{-1} \quad (2.5.8)$$

We know that,

$$\int h(x)C(\theta) \exp\{x.Q(\theta)\}dx = 1$$

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

$$\int h(x) \left[C(\theta) \exp\{x.Q(\theta)\}. \ x.Q'(\theta) + C'(\theta) \exp\{x.Q(\theta)\}\right] dx = 0$$

That is

$$m(\theta)Q'(\theta) + \frac{C'(\theta)}{C(\theta)} = 0$$

Therefore,

$$m(\theta) = -\frac{C'(\theta)}{C(\theta)Q'(\theta)}$$

Using this result in (2.5.8), we get

$$\frac{\partial}{\partial m(\theta)}\left[\ln f_\theta(x)\right] = \left[X - m(\theta)\right]\left[V(x)\right]^{-1}$$

Thus,

$$J(F) = \int \left[\frac{\partial}{\partial m(\theta)}\left[\ln f_\theta(x)\right]\right]^2 dF(x)$$

$$= \int \left[x - m(\theta)\right]^2 \left[V(x)\right]^{-2} dF(x)$$

$$= \left[V(x)\right]^{-1}$$

Hence, the R.H.S of (2.5.6) becomes

$$J^{-1}(F) \frac{\partial}{\partial m}\left[\ln f_\theta(x)\right] = x - m(\theta) \quad (2.5.9)$$

From(2.5.7) and (2.5.9) the result follows. ∎

<u>Properties of the IF</u>: Following are the some properties of IF.

i) An appealing heuristic interpretation.

ii) An indicator via GES of maximum bias due to infinitesimal contamination.

iii) To compare the influence of individual observation on estimator.

iv) To study the robustness properties of an estimator

v) From the IF we can obtain GES and local-shift sensitivity

vi) IF measures the effects of infinitesimal contamination at the point x on the estimate. ∎