Qualitative Robustness in Estimation

Mohammed Nasser
Department of Statistics
Rajshahi University, Rajshahi-6205, Bangladesh
mnasser.ru@gmail.com

Nor Aishah Hamzah
Institute of Mathematical Sciences University of Malaya
Kuala Lumpur, Malaysia
naishah.hamzah@gmail.com

Md. Ashad Alam
Department of Statistics
Hajee Mohammad Danesh Science and Technology University
Dinajpur 5200, Bangladesh
ashad.hstu@gmail.com

Abstract

Qualitative robustness, influence function, and breakdown point are three main concepts to judge an estimator from the viewpoint of robust estimation. It is important as well as interesting to study relation among them. This article attempts to present the concept of qualitative robustness as forwarded by first proponents and its later development. It illustrates intricacies of qualitative robustness and its relation with consistency, and also tries to remove commonly believed misunderstandings about relation between influence function and qualitative robustness citing some examples from literature and providing a new counter-example. At the end it places a useful finite and a simulated version of qualitative robustness index (QRI). In order to assess the performance of the proposed measures, we have compared fifteen estimators of correlation coefficient using simulated as well as real data sets.

Keywords: Breakdown Point, Influence Function, Qualitative Robustness and Qualitative Robustness Index.

1. Introduction

Hampel in his Ph.D. thesis (1968) developed three concepts: qualitative robustness (also Π-robustness), breakdown point and influence function to assess robustness in estimation and thus raised rigorousness in robust estimation to a satisfactory level. He developed qualitative robustness to uphold qualitative side of robustness gauging distributional robustness, Π-robustness; a form of qualitative robustness suitable for dependent observations, breakdown point to quantify global side of robustness and influence function to quantify infinitesimal side of robustness. In Huber's (1972) words, "Hampel (1968) recognized and sorted out the stability aspect of robustness, in close analogy to stability of a mechanical structure (say of a bridge): (i) the qualitative aspect: a small perturbation should have small effects; (ii) the breakdown aspect: how big can the

perturbation be before everything breaks down; (iii) the infinitesimal aspect: the effects of infinitesimal perturbations." In Robust Statistics parameters are considered as functionals; $\theta = T(F)$ where domain is $S(X)$, set of finite signed measures defined on the sample space or a subset of it (such as $S_p(X)$, set of probability measures) and co-domain is R^k or R^∞ or a function space or set of sets of a metric space. Fundamental concept of robustness is directly or indirectly related to continuity or differentiability of T. A natural sample estimator of $T(F)$ is provided by statistical function $T(F_n)$ based on the sample d.f. F_n. It is intuitively obvious that when T is continuous at F and F_n is near F, $T(F_n)$ is near $T(F)$. Equicontinuity w.r.t. *n* is more desirable. " Continuity"- one of the fundamental concepts of Classical Analysis, which was generalized to spaces which include Euclidean spaces during first two decades (Frechet, 1906 (metric space); Hausdorff, 1914 (topological space)) of last century is the backbone of the concept; "Qualitative Robustness" of estimators. If T is differentiable at F, one can find the differential of T and thereby can measure the nearness of $T(F_n)$ to $T(F)$ due to infinitesimal change in F through F_n. Now arises the question how one can define and check continuity and differentiability of *T*. Continuity of *T* is related to qualitative robustness, while differentiability to quantitative and infinitesimal robustness. It is important as well as interesting to study relation among them, specially the relation between qualitative robustness and influence function because both deal effect of small perturbations. Hampel discussed and elaborated this concept at the outset of his thesis, breakdown point and influence function in the latter part. His seminal article on qualitative robustness (1971) was published three years before his mostly quoted article on influence function (1974). Breakdown point attracted wide range of researchers only after the development of finite version of breakdown point by Donoho (1982) and Donoho and Huber (1983). Qualitative robustness, though no less important than the other two concepts from the viewpoint of robustness has gained less popularity. Most probably its mathematical complications and absence of finite versions have acted behind this present but undesirable unpopularity.

The concepts of qualitative robustness and $\Pi-$robustness (more restrictive concept than qualitative robustness) introduced by Hampel (1968 and1971) were extended in different directions in last eighties. Huber (1977 and 1981) modified Hampel's definition suggesting asymptotic equicontinuity of sampling distribution of the estimators with respect to *n* on the ground that nonrobustness gets worse for large *n*. Rieder (1982) and Lambert (1982) introduced qualitative robustness in hypothesis testing, Boente et al. (1987) following Papatoni-Kazakos and Gray (1979) and Cox (1981) generalized qualitative robustness for stochastic processes. Cuevas (1987 and 1988) adjusted some results of Hampel (1971) and Huber (1981) in the context of abstract inference. He showed incompatibility of consistency and qualitative robustness in the case of kernel density estimators. Cuevas and Romo (1993) and Nasser (2000) applied this concept in nonparametric bootstrapping and Basu et al. (1998) in Bayesian inference. In case of M-estimators Clarke (1983) extended Huber's results forwarding sufficient conditions for not only weak-continuity (qualitative robustness) but also Fréchet differentiability at a particular parametric model. In 2001 he enriched his previous results showing global weak continuity of some well-known M-functionals in neighbiurhood of a pramatic

model. Fasano et al. (2012) advocated for a novel form of weak differentiability to prove consistency, asymptotic normality and qualitative robustness of M-estimates under more general conditions than those required in standard approaches. Daouia and Ruiz-Gazen (2004) etc studied qualitative robustness of nonparametric frontier estimator. Hable and Christmann (2011) showed weak continuity of support vector machines, hence its qualitative robustness and thus combining with the existence and uniqueness of support vector machines, they can be treated as the solutions of a well-posed mathematical problem in Hadamard's sense. Mizera (2010) and Krätschmer et al.(2012) tried to develop basic concept of qualitative robustness in two different directions - Mizera (2010) presented connection between weak continuity and qualitative robustness in full generality and under minimal assumptions taking Prokhorov metric on both set of models and set of sampling distributions while Krätschmer et al.(2012) using different metric proposed index of qualitative robustness to order statistical procedures on the scale 0 to ∞ in place of qualitative division of robust and non-robust procedures.

In this article we shall try to examine concept of qualitative robustness and its relation with influence function and thereby to alleviate some related and common misunderstandings. In Section 2 we provide and discuss model-based definitions of qualitative robustness as forwarded by Hampel (1968 and 1971), its later development by Huber (1977 and 1981), their results and their complications through some new propositions. In Section 3 we discuss works of Mizera (2010) and Krätschmer et al.(2012) more elaborately while in Section 4 we put forward definition of influence function and some fact regarding its relation with qualitative robustness with a new counter example, and place a finite-version and an index of qualitative robustness in Section 5. Finally, we conclude the work in Section 6.

2. Qualitative robustness

2.1. Gist of Hampel's paper (1971) and Huber's results (1981)

Hampel (1968 and 1971) gave definition of qualitative robustness (Bahadur and Savage (1956) fore-shadowed this idea) and continuity of T_n when sample space is X, a polish space with metric d and parametric space is R^k. Both spaces are endowed with Borel $\sigma -$ algebra to make them measurable. He deduced two main theorems, three lemmas and two corollaries to show relation between concept of qualitative robustness and continuity of T_n in two cases – i) The general case $\{T_n = T_n(F_n)\}$ and ii) The particular case $\{T_n = T(F_n)\}$. Staudte (1980) and Staudte and Sheather (1990) followed the definition and theorems of Hampel (1971). Both Hampel (1968 and 1971) and Huber (1977 and 1981) cited the result of Strassen (1965) to show Prokhorov metric is intuitive to catch up (i) rounding and grouping errors (small errors occurring with large probability) and (ii) gross error (large errors occurring with low probability). Huber quoted and proved two results due to Prokhorov (1956) – i) The Prokhorov metric metrizes the weak topology on $S_p(X)$, the set of probability measures on Xi,e it encompasses weak convergence that is the case of idealized approximation of the underlying chance mechanism and ii) $S_p(X)$ with the topology is a polish space.

Huber's generalization of Hampel's definition. Huber generalized the definition of Hampel on the ground that for "non robust" statistics the modulus of continuity typically gets worse for increasing n. His definition – $\{T_n\}$ is qualitative robust at F_0

$$\Leftrightarrow (\forall \varepsilon > 0 \, \exists \delta > 0 \wedge n_0 s.t. (\forall n > n_0) \wedge (\forall F \in S_p(X))$$

$$\wedge d_1(F_0, F) < \delta \Rightarrow d_2(L_{F_0}(T_n), L_F(T_n)) < \varepsilon)$$

i.e. $h_n : \langle S_p(X), d_1 \rangle \rightarrow \langle S_p(R^k), d_2 \rangle$ with $h_n(F) = L_F(T_n)$ is asymptotically equicontinuous w.r.t. n at F_0. d_i are metrics that induce weak topology. In Hampel's definition $\{h_n\}$ is equicontinuous w.r.t. for all n. d_i are Prokhorov metrics.

Comment 1. It is clear if $\{T_n\}$ is qualitative robustness at F in Hampel's sense, it is qualitative robustness at F in Huber's sense, but converse is not necessarily true.

2.1.1. Three main results of Hampel (1971 and 1986)

Definition 2. A $\{T_n : n \geq 1\}$ is continuous at $F \Leftrightarrow \forall \in > 0 \quad \exists \delta > 0$ $\exists \eta_0 \forall n, m \geq \eta_0, \forall F_n, F_m; \{F_n \in S_{F_n}(X) \wedge F_m \in S_{F_m}(X) \wedge \Pi(F, F_n) < \delta \wedge \Pi(F, F_n) < \delta\} \Rightarrow$ $|T_n(F_n) - T_m(F_m)| < \in$ (where F_j, a discrete probability measures whose atoms have probability equal to $\dfrac{1}{j}$ or multiple of $)\dfrac{1}{j}$.

Let $T_n = T(F_n)$, then T is weakly continuous at $F \Rightarrow T_n$ is continuous at F. Converse is not true.

Theorem 1. (Also theorem 1 in Hampel's paper) If i) $\{T_n\}$ is continuous at F and ii) T_n is continuous as point function on X^n, $\forall$ n, then $\{T_n\}$ is robust at F.

Comment 2. Let $T_n = T(F_n)$ and T be continuous at F, and, $T(F_n)$ continuous point function on $X^n \forall$ n. Then $T_n = T(F_n)$, is robust at F. Since $T_n = T(F_n)$ is nearly always continuous point function on $X^n \forall$ n, weak-continuity of T at $F_0 \Rightarrow$ qualitative robustness of $T_n = T(F_n)$ at F_0. For $X = R$ and $T_n = T(F_n)$, Huber (1981) proved, condition ii in theorem is not required if Huber's definition is adopted (discussed below). Even for general c.s.m.s X, the statement is true is shown in the following subsection.

Theorem 2. (Lemma 3 in Hampel's (1971) let $\{T_n\}$ be robust at F_o and consistent at all G in nbdof F_o. Then $T_\infty(G)$ [where $T_n(G) \overset{p}{\longrightarrow} T_\infty(G)$] is continuous at F_o.

Comment 3. Let $T_n = T(F_n)$ be robust at F_o and consistent at all G in a nbd of F_o (with $T_\infty(G) = T(G)$). Then T is continuous at F_o.

Theorem 3. (Theorem 2 in Hampel's (1971), Let $T_n = T(F_n)$. Then T is continuous at all $F \Leftrightarrow \{T_n\}$ is robust and consistent, tending to $T(F) \forall F \Leftrightarrow \forall K \subset S_p(X)$, relatively compact,

$$\forall F \in S_p(X) \; \forall G \in S_p(X), F \in K \wedge \Pi(F,G) < \delta \Rightarrow |T(F) - T(G)| < \varepsilon$$

The theorem is mathematically very nice but looks very strict from practical viewpoint.

2.1.2. Huber's results

Theorem 4.(Proposition 6.2 (Huber, 1981)). Assume that $\{T_n = T(F_n)\}$ is consistent in anbd of F. Then T is continuous at $F \Leftrightarrow T_n$ is robust at F.

Comment 4. a) He proved it taking $X = R$ and $d_1 =$ Le'vy metric, $d_2 =$ Prokhorov metric. He used the result $d_2(\delta_x, \delta_y) = \Pi(\delta_x, d_y) = d(x,y)$. It is only true when $d(x,y) \leq 1$. But there is no mention of the condition. Huber assumed the well-known fact, for any metric space $\langle X, d \rangle \; \exists$ a metric d_1 s.t. $d_1 \leq 1$ and $\langle X, d \rangle$ is homeomorphicto $\langle X, d_1 \rangle$.

Comment 5. His proof clearly indicates that this theorem has two parts:

a) T is continuous at $F \Rightarrow T_n$ is qualitative robustness (in Huber's sense) at F. We should note the difference between this result and Hampel's theorem 1. Condition ii in Hampel's theorem is required to prove, $\{h_n\}$ is equicontinuous at F, where $n \leq n_0$. We have demonstrated in the first proposition of the next section the same holds for general c.s.m.s.
b) T_n is q.r at F. and consistent at G in a nbd of $F \Rightarrow T$ is continuous at F. Here he also assumed $T_\infty(G) = T(G)$ without mentioning. All these are more clarified with an example in the next section.
The former result does not hold in Hampel's sense, while the latter does in both sense.

Comment 6. By Polya's theorem T is Kolmogorov (Kuiper) continuous at $F \Rightarrow T$ is weakly continuous F (F is continuous). So we can infer that T is Kolmogorov (Kuiper) continuous at $F \Rightarrow T_n$ is qualitative robustness (in Huber's sense) when F is continuous(Staudte 1980). It is well known that Kolmogorov metric is equivalent to Kuiper metric. We extend the result in the following subsection in the case $X = R^m$ assuming F is absolute continuous. So we can rewrite Huber's theorem 6 as follows: Assume that $T_n \to T(F)$ is consistent in an nbd of F, F is continuous. Then T is Kolmogorov-continuous at $F \Leftrightarrow T_n$ is robust at F (Parr, 1985). This reformulation might be easier to handle. These results seem very important from the viewpoint of attempt to study continuity and differentiability of T in the same topology. Staudte and Sheather (1990) gave due credit for the inception of the idea to Hampel (1968 and 1971) and briefly discussed qualitative robustness. For more detailed discussion they pointed to Staudte (1980) where he followed Hampel's definition avoiding Huber's one and described the

same relation between continuity and qualitative robustness as given by Hampel (1968 and 1971). Jureckova and Sen (1996) briefly quoted the definition of qualitative robustness of Huber and gave the comment:

"The weak continuity of $T_n(=T(G_n)$ at G (here G is the true d.f. of X_1) and its consistency at G in the sense that $T_n \to T(G)$ almost surely (a.e) as n $\to\infty$ characterize the robustness of T_n in a nbd of G". Our comment 5 and the new result in proposition 3 indicate that the sentence may lead one to some inappropriateness due to its briefness.

2.2. New results

2.2.1. Three New Propositions.
This subsection upholds three new propositions some of which are indicated above.

Proposition 1. If T is continuous at F, $T_n(=T(G_n)$ is qualitative robust at F in Huber's sense for general c.s.m.s. w.r.t. Prokhorov metric.

Proof. First part of proof of theorem 2 in Cuevas (1988) begets the result.

In this regard we can quote from Cuevas and Romo (1993), "It is known (Hampel, 1971) that if that T is continuous on some $U(F_0)$ then the sequence $\{T_n\}$ is qualitatively robust at F_0." Our discussion demonstrates, the statement is not precise. Not in Hampel's but in Huber's sense, continuity of T over a nbd is equivalent to qualitative robustness of T_n over the nbd. To get results in Hampel's line we need continuity of T over whole $S_p(X)$ or T_n is continuous over $X^n \forall n$. Now using the above proposition and Ranga Rao's result the following proposition is proved.

Proposition 2. Assume that $X= R^k$ and $T_n(=T(G_n)$ is consistent in anbd of F, F is absolute continuous. Then T is Kolmogorov-continuous at $F \Leftrightarrow$ is robust at F.

Proof .

Necessary part: Ranga Rao's (1962, p-665) result implies that T is Kolmogorov-continuous at $F \Rightarrow T$ is weakly-continuous at F, and by the above proposition, $\{T_n\}$ is robust at F.

Sufficiency part: From comment 2 after Hampel's theorem 2 (in our article) we have, $\{T_n\}$ is robust at $F \Rightarrow T$ is weakly-continuous at F, i.e. T is Kolmogorov-continuous at F. Now the following results clarify the intricacies of the above theorems and propositions to some extent:

Proposition 3
Let us define two functional T, T_∞ on $Sp(X)$;
1) $T : S_p(X) \to R, T(F) = $ sum of the jumps of F.
2) $T_\infty : S_p(X) \to R, T_\infty(F) = 1$.

Results:

a) $T_n = T(F_n) \xrightarrow{\ p\ } T_\infty(F) \ \forall \ F$.

b) $T_n = T(F_n)$ is robust at $F \ \forall \ F$.

c) T is not Kolmogorov-continuous (i.e. not weakly continuous) at F, where F is not discrete.

Proof:

a) It is obvious.

$L_F(T_n) = \delta_1 = L_G(T_n) \ \forall n, \forall F, G \in Sp(X). \ \therefore T_n = T(F_n)$ is robust at all F from the definition.

b) Let F_n be empirical d.f.. Then $F_n \xrightarrow{\ \| \ \|_\infty\ } F$, but $T(F_n) \xrightarrow{\ \times\ } T(F)$

since $T(F_n) = 1 \ \forall n$ and $T(F) < 1$.

The above results demonstrate two things – i) qualitative robustness does not imply consistency of estimators and ii) qualitative robustness and consistency should be the minimum properties of an estimator from the viewpoint of robust estimation. Cuevas (1987 and 1988) generalized some of the Hampel's and Huber's results and applied them in areas of abstract inference, such as density estimation, stochastic process. He never mentioned the fact - all the results of Hampel and Huber could be easily generalized in about identical form. All the results of Hampel and Huber can be generalized to the case of "generalized statistics" (statistics which take values in the general complete separable metric spaces"). It requires only two modifications;

1) using the metric $d(x, y)$ of parametric space in place of $|x - y|$ and adjusting the definitions with the metric.

2) applying Cantor's Intersection Theorem for general complete metric space in proving lemma 2 in Hampel (1971).

3. Contributions of Mizera (2010) and Krätschmer et al.(2012)

Mizera (2010) placing Hubers' definition of qualitative robustness of a statistics t_n (estimators or test statistics) explained the complicacy of median elaborately in order to present as well as to generalize the intricate relationship between qualitative robustness and weak continuity,

He generalized Huber's theorem 6.2 in three directions:

i) He extended area of application of the theorem using Prokhorov metric in place of Levy metric deriving an Uniform Glivenko-Cantelli property (Mizera, 2010, lemma 4)

ii) He also extended the definition of weak continuity to adjust set-valued functional:

Definition of weak continuity. A functional T is called *weakly continuous* at P, if for any $\varepsilon > 0$
there is $\delta > 0$ such that
$\pi(P, Q) \le \delta$ implies $d(\theta, \tau) < \varepsilon$
for *any* value θ and τ of T at P and Q, respectively.

Pak.j.stat.oper.res. Vol. VIII No.3 2012 pp619-634

625

His main theorem: **Theorem 1.** Suppose that a procedure t_n is represented by a functional T. If T is weakly continuous at P, then any lawful version of t_n is qualitatively robust at P. He then defining coining the term, "regular functional" presented theorem 2 in order to illustrate complicacy of converse theorem 2, i.e. to show how weak consistency and qualitative robustness Implies weak continuity.

Definition of regular functional. A representation of a procedure t_n by a functional T is called *regular*, if (i) it is consistent for every P in the domain of T; and (ii) for every P and every $\tau \in T(P)$, there is a sequence Pv of empirical probabilities weakly converging to P, the functional T is univalued at every Pv, and $T(Pv)$ converges to τ .

Theorem 2. Suppose that the representation of a procedure t_n by a functional T is regular. If some lawful version of t_n is qualitatively robust at P, then T is weakly continuous (in particular, uniquely defined) at P.

Observing the fact that though all the classical moments are nonrobust by Hampel's definition, the higher moments are more affected by outliers than lower moments, Krätschmer et al.(2012**)** introduced a new concept of qualitative robustness that applies to a very large class of tail-dependent statistical functional T. The focus of the approach lies in specifying a metric d on the set of probability models for which T becomes a continuous functional at P. For R as sample space they used a weighted Kolmogorov-type distance whereas the sum of the Prokhorov metric and a moment distance was proposed for R^n or any polish space. Then they established extensions of Hampel's theorem essentially stating that when T is continuous with respect to d then it is also qualitatively robust in the sense that Hampel (Huber) condition holds if we choose the Prokhorov metric for d_2.

The proofs of these results rely on strong uniform Glivenko-Cantelli theorems in fine topologies, They also examined the sensitivity of tail-dependent statistical functionals w.r.t. infinitesimal contaminations, and proposed a new notion of infinitesimal robustness. The theoretical results were illustrated by means of several examples including general *L*- and *V*-functionals.

Readers would certainly feel interested to understand the sentence--"Nevertheless, we emphasize that the concept of qualitative robustness depends on the specific choice of the metrics d and d^{-} and not just on the topologies generated by them" as it differs from comments of other prominent researchers on robustness including proponents in this field. Bur readers' interest would not be satisfied as this point was not illustrated as they pledeged.

4. Influence function

4.1 Definition

The most central concept in Hampel's fundamental contribution to the theory of robustness (Hampel, 1968, 1971 and 1974) is the "influence function" (originally termed as "influence curve"). In his seminal article in 1974 he first gave the definition for particular case (both sample space and functional range space are R or subsets of R)

and then for general case (sample case- X , a complete separable metric space (c.s.m.s.) and functional range space- R^k). Let T be a R^k -valued mapping from a subset of the probability measures on X , $D_T(X)$, a finitely full and convex subset of $S_p(X)$. Let $F \in D_T(X)$ and δ_x denote the atomic probability measure concentrated in any given $x \in X$. Then the vector-valued influence function of T at F (here is a measure) is defined point wise by

$$IF_T(x,F) = \lim_{\varepsilon \downarrow 0} \frac{\left[T\left[(1-\varepsilon)F + \varepsilon\delta_x\right] - T(F)\right]}{\varepsilon} \qquad 1.1$$

Though for a particular T it is generally considered as a function of x and F , later, for brevity, it is denoted by $IF(x)$. " The IF is mainly a heuristic tool, with an intuitive interpretation " (Hampel et al.,1986, p-83). It can be intuitively interpreted as a suitably normed asymptotic influence of outliers on the value of an estimate or test statistic $T(F_n)$. It is a local robustness property. Various characteristics of an influence functions are used to develop various concepts such as Gross Error Sensitivity (GES), γ^* (supremum of $IF(x)$ w.r.t. x for fixed F) and maximum-bias curve over a local neighbourhood of F (graph of GES vs F), Local -Shift Sensitivity (LSS) (sup of slope of $IF(x)$), λ^* , Rejection Point, ρ^* (related to upper limit of the range outside which influence function vanishes) etc to delineate definite but different aspects of local robustness property. As important by- products of the attempt to quantify the effect of outlier on the estimators Change of Variance Function (CVF) has been developed from $IF(x)$ to plot asymptotic variance vs. F. The heuristics of influence function are heuristics, not theorems. But tendencies to use them as theorems are not rare in literature (Davies, 1993).

We can easily prove that all moments are Kolmogorov-discontinuous at continuous models, hence nonrobust. Their influence functions are continuous but unbounded. Then one may be tempted like Koenker (2005) to infer wrongly that unbounded influence function implies non-robustness.

4.2 Relation between influence function and qualitative robustness

There exists no direct relation between influence function and qualitative robustness. The following questions and their answers mainly by examples illustrates their relation:

❑　　Does a bounded influence function imply weak continuity of the functional? No. It is well –known that the efficient *L*-estimate of location parameter for the logistic is not robust, and $b_I(\varepsilon) = \infty \ \forall \varepsilon > 0,$ even γ^* is finite.

❑ Does a continuous influence function of bounded variation imply weak continuity of the functional? No. $IF(x)$ of the above-mentioned

estimator $= \dfrac{1}{1+e^{-x}}$ + a constant, which is continuous and strictly increasing.

❑ The following new counter example shows that even two-valued almost constant influence function does not guarantee the weak continuity of the functional;

Let $T(F)$ = size of the largest atom and $F_0 = \delta_y$. Then it can be easily shown that

$$IF(x) = \left\{ \begin{array}{ll} -1 & \text{if } x \neq y \\ 0 & \text{if } x = y \end{array} \right. .$$

But T is not weakly continuous at F_0 since

$$G_n = \frac{N(y, \sigma^2)}{n} \xrightarrow{\ w\ } F_0 \text{ and } T(G_n) = 0 \xrightarrow{\ \times\ } T(F_0) = 1.$$

- Does a weak continuous functional provide finite γ^*? No. The efficient R-estimate of location parameter for the normal distribution, the normal scores estimate has $\gamma^* = \infty$ with $IF(x) = x$, but T is weakly continuous at the model.

- Does a Hadamard differentiable functional beget a bounded $IF(x)$? No. If the derivative is weakly continuous, we get bounded and continuous $IF(x)$. We also get weak continuity of T if V, the associated vector space is topologized by Kolmogorov norm and F is continuous (see Nasser, 2000; proposition 4.6.1).

- Does a Frèchet differentiable M-functional at F, which is Kolmogorov-continuous at F have always a bounded influence function? Yes (Clarke, 1983).

The discussions amply substantiate our comments made at the beginning of the subsection. We should be very cautious to comment in general about relation between influence function and qualitative robustness. In a particular class of estimators we may have clear-cut relation between the two.

5. A finite version and a simulated version of qualitative robustness index

5.1. Finite version

We have already mentioned that non-availability of finite sample version of qualitative robustness is one of the main reasons behind its less popularity than the two other concepts− influence function and breakdown point. While proposing a definition of finite-version qualitative robustness, we keep in mind that an estimator with finite breakdown point equal to zero should have empirically lower QRI whereas estimators with high breakdown point should have higher QRI. We propose two versions of SQRI(SQRI 1 and SQRI 2):

$$SQRI(veresion1) = \frac{1}{1 + \max_{j} \left| \hat{\theta} - \hat{\theta}(j) \right|}$$

$$SQRI(veresion2) = \frac{1}{1 + \max_{i \neq j} \left| \hat{\theta}(i) - \hat{\theta}(j) \right|}$$

It is easy to prove i) It's maximum value is 1. It's minimum value is zero or above zero, for example, for simple correlation co-efficient, 1/3. The more SQRI is the more qualitative robust the estimator is. Alam et al. (2008) compared 15 estimators of simple correlation co-efficient investigating the bias, standard error, MSE, length of 90

percentile interval, sensitivity curve of each estimator under a variety of situations and also employed probability plot, box plot and perspective plot to judge their performances. The normal score estimator showed the best performance overall.

We have made experiments on simulated as well as real world problems to apply our proposed SQRI method using 15 estimators of correlation coefficient. Detailed information of data sets are in appendix A. The results show that the proposed method successfully chooses the best robust estimator as Alam et al. (2008), the normal score estimator. The results are given in Table 1 and Table 2 (Appendix B). The visualization of data sets are in Figure 1-3 (Appendix B).

5.2. A simulated version of qualitative robustness index

To assess the effect of ε% contamination on the sampling distribution of the measures we define Qualitative Robustness Index, QRI(ϵ)= it should be in same line in equation

$$\frac{1}{1 + \sum_{i=1}^{k} \left| q_i^c(\varepsilon) - q_i \right|} \times 100$$

Here q_i is the ith quantile of a measure at a model and q_i^c, the ith quantile of the measure at the model contaminated by ε% contamination. A slight variation of this measure was used in Alam et al. (2010) to quantify effect of contamination on different types of canonical correlation coefficient at multivariate normal models.

6. Conclusion

"Qualitative robustness is of little help in the actual selection of a robust procedure suited for a particular application. In order to make a rational choice, we must introduce quantitative aspects as well."(Huber, 1981,p-73) As example, both for location and scale parameters, there exit three class of robust estimators – M-type, L-type and R-type – under mild conditions; and each class contains different robust subclasses (Huber,1981; Chapter 3 and 5; and Hampel et al., 1986; chapter 2). None the less we should start from a consistent and qualitative robust procedure and then seek procedures with extra robust criteria as such high breakdown point, smooth and bounded influence function, uniform asymptotic normality etc.

References

1.   Alam, M.A., Nasser, M. and Fukumizu, K. (2010). Sensitivity analysis in robust and kernel canonical correlation analysis. *Journal of Multimedia*, Vol. 5, issue 1, 3-11, Academy Publisher, Finland.

2.   Alam, M.A., Nasser, M. and RahmatullahImon, A. H. M. (2008). Sensitivity and influence analysis of estimators of correlation coefficients. *Journal of Applied Probability & Statistics*Vol. 3, no. 1, 119–136.

3.   Bahadur, R.R. and Savage, L.J. (1956). The nonexistence of certain statistical procedures in nonparametric problems. *Ann. Math. Statist.* Vol. 27, 1115-1122.

4. Bednarski, T. (1993). Frèchet differentiability of statistical functionals and implications to robust statistics. In: Morgenthaler, S, Ronchetti, E. and Stahel, W.A., (Eds.), *New Directions in Statistical Data Analysis and Robustness*, Besel, BirkhauserVerlag, 25-34.

5. Clarke, B.R.(1983). Uniqueness and Frèchet differentiability of functional solutions to maximum likelihood type equations. *The Annals of Statist.*11, 1196-1206.

6. Clarke, B.R.(2000). A remark on robustness and weak continuity of M-estimators. *J. Australian math. Soc.* (Series A) 68, 411-418.

7. Cueas, A. (1987). Density estimation; robustness versus consistency. In: M.L. Puri, J.P. Vilplana and W. Wertz, Eds., *New Perspectives in Theoretical and Applied Statistics*. Wiley, NewYork, 259-264.

8. Cueas, A. (1988), Qualitative robustness in abstract in inference. *J. Statist.Plann. Inf.*, 18, 277-289.

9. Cueas, A. and Romo, J. (1993), On robustness properties of bootstrap approximations. *Jour.Statist. Plann. Inf.* 37,181-191.

10. Cuevas, A. and Sanz, P. (1989). A class of qualitatively robust estimates. *Statistics* 20, 509–520.

11. Daouia A.and Ruiz-Gazen, A.(2004). Robust nonparametric frontier estimators: qualitative robustness and influence function, Technical report.

12. Davies, P.L. (1993). Aspects of robust regression.*The Annals of Statist.,* 21, 1843-1899.

13. Devlin, J. S., Gnanadesikan, R. and Kettenring, R. J. (1975). Robust estimation and outlier detection with correlation coefficients. *Biometrika*72, 531–545.

14. Donoho, D.L. (1982). Breakdown properties of multivariate location estimators. Ph.D. qualifying paper, Department of Statistics, Harvard University, Cambridge, Mass.

15. Donoho, D.L. (1983). The notion of breakdown point, In A Festschrift for Erich L.Lehmann, P.J. Bickel, K.A. Doksum and J.L. Hodges eds., Wadswoth, Belmont(CA):157-184.

16. Efron, B. and Tibshirani, R.J. (1993). *An Introduction to the Bootstrap.* Chapman & Hall, New York.

17. Fasano, M.V., Maronna,R.A., Sued, M and Yohai, V.J.(2012). Continuity and differentiability of regression M functionals. *Bernoulli*, Bernoulli Society, To appear.

18. Fieller, E. C., Hartley, H. O. and Pearson, E. S. (1957). Test for rank correlation coefficients I, *Biometrika*44, 470–481.

19. Fieller, E. C. and Pearson, E. S. (1961). Test for rank correlation coefficients II, *Biometrika*48, 29–40.

20.	Frėchet, M. (1906). Sur quelques points du calcul fonctionnel. RendicontiCircolo Mat. *Palermo.* 22, 1-24.

21.	Gideon, R. A. and Hollister, R. A. (1987). A Rank Correlation Coefficient Resistant to Outliers, *J. Amer. Stat. Assoc.* 82, 656–666.

22.	Gideon, R. A.(1998). A Generalized Interpretation of Pearson's r,http://www.math.umt.cdu/gideon.

23.	Gnanadesikan, R. and Kettenring, J. R. (1972). Robust estimates, residuals, and outlier detection with multiresponse data, *Biometries*28, 81–124.

24.	Hable, R. and Christmann, A. (2011). On qualitative robustness of support vector machines. In Proceedings of J. Multivariate Analysis. 993-1007. J*ournal of Multivariate Analysis* http://dl.acm.org/citation.cfm?id=1972225Vol. 102 Issue 6, July, 2011.

25.	Hampel, F.R. (1968). Contributions to the Theory of Robust Estimation. Ph. D. Thesis, University of California, Berkeley.

26.	Hampel, F.R. (1971). A General Qualitative Definition of Robustness. *Ann. Math. Stat.*, 42, 1887-1896.

27.	Hampel, F.R. (1974). The Influence Curve and Its Role in Robust Estimation. J. Amer. Statist.Assoc. 62, 1179-1186.

28.	Hampel, F.R., Rousseeuw, P.J.,Ronchtti, E.M., Stahel, W.A., (1986). *Robust Statistic: The Approach Based on Influence Functons.* Wiley, New York.

29.	Hausdorff, F (1914): Gründge *der Mengenlehre*. Von Veit, Leipzig

30.	Huber,J. (1977). *Robust Statistical Procedures.* Regional Conference Series in Applied Mathematical No. 27, Soc. Industries. Appl. Math., Philadelphia, Penn.

31.	Huber, P. J. (1972). Robust statistics: a review. *Ann.Math.Statist.*43, 1041-1067.

32.	Huber, P.J. (1981). *Robust Statistics.* John Wiley & Sons, New York.

33.	Huber, J. and Ronchetti, E. M.(2009). *Robust Statistics.* 2nd edition. John Wiley & Sons, New York

34.	Maronna, R. A. and Martin, R. D. and Yohai, V. J. (2006). *Robust Statistics: Theory and Methods.* 2nd edition. John Wiley &Sons,New York.

35.	Jureckova, J. and Sen, P.K. (1996). *Robust Statistical Procedures: Asymptotics and Interrelations.* John. Wiley and Sons, Inc., N.Y.

36.	Kendall, M. G. (1973). *Rank Correlation Methods.* 4th ed. Charles Griffin, London.

37.	Koenker, R. (2005). Introduction to Robustness, Econ 574, Lecture 16, Dept. of Economics,University of Illinois, USA., available at http://www.econ.uiuc.edu/~roger/courses/476/lectures/L15.pdf

38.	Krätschmer, V., Schied, A. and Zähle, H. (2012).Qualitative and infinitesimal robustness of tail-dependent statistical functionals. *Journal of Multivariate Analysis*, Vol.103, issue 1, 35–47.

39.   Mizera, I. (2010). Qualitative robustness and weak continuity: the extreme unction. IMS Collections Festschrift, Institute of Mathematical Statistics.

40.   Nasser, M (2000). Continuity and Differentiability of Statistical Functionals: Its Relation to Robustness in Bootstrapping. Unpublished Ph.D thesis, RCMPS,C.U., Bangladesh.

41.   Parr, W.C. (1985). The bootstrap: some large sample theory and connections with robustness. *Statist*. Probab. Lett. 3, 97-100.

42.   Rao, R.R., 1962. Relations between weak and uniform convergence of measures with applications. *Ann. Math. Statist.* 33, 659-680.

43.   Reeds, J. A.(1976). On the Definition of von Mises Functionals, Ph. D. thesis, Dept of Statistics Harvard University, Cambridge, Mass.

44.   Rieder, H. (1994). *Robust Asymptotic Statistics.* Springer-Verlag, New York.

45.   Salibian-Barrera, M.(2000). Contirbutions to the Theory of Robust Inference, Unpublished Ph. D. thesis. University of British Columbia, Department of Statistics, Vancouver.

46.   Shao, J.(1999). *Mathematical Statistics*. Springer-Verlag, New York.

47.   Staudte, R.G.(1980). Robust Estimation. Queen's Papers in Pure and Applied Mathematics, no.53. A.J. Coleman, and P. Ribenboim (Eds.). Kinston, Ont., Canada: Queen's Univ.

48.   Staudte, R.G. and Sheather, S.J.(1990). *Robust Estimation and Testing.* John Wiley Sons. Inc. New York.

Appendix A

Data-1.50 sampled schools have drown from 82 law schools (Efron and Tibshirani,1993, Table 3.2,p.21).

Data-2. 45 sampled schools have drown form 82 law schools and 5 samples have drown from bivariate normal distribution with mean vector (38.31, 599.66),unit variance for both variable and covariance 0.1.

Data-3. Biochamical data (Maronna et.al, 2006,Table 6.1, p. 177).

Data-4. 50 sampled from bivariate normal distribution with mean vector (0,0),unit variance for both variables and covariance 0.5.

Appendix B

Table1: Results of SQRI (version-1) of 15 correlation estimators

Estimators	Data-1	Data-2	Data-3	Data-4
r_p	0.98039	0.93458	0.76336	0.95238
r_{av}	0.98039	0.96154	0.82645	0.96154
r_{aum}	0.98039	0.95238	0.82645	0.97087
r_{mad}	0.96154	0.95238	0.69444	0.93458
r_s	0.94340	0.95238	0.84746	0.93458
r_{mf}	0.95238	0.95238	0.81967	0.95238
r_k	0.96154	0.97087	0.90090	0.96154
r_{gd}	0.95238	0.96154	0.81301	0.95238
$r_Q^{\cdot}$	0.98039	0.89286	0.69930	0.87719
$r_k^{\cdot}$	0.92593	0.93458	0.81301	0.91743
$r_{ns}^{\cdot}$	**0.98039**	**0.98039**	**0.95238**	**0.98039**
$r_{ssd}^{\cdot}$	0.92593	0.97087	0.81967	0.94340
$r_{bvt}^{\cdot}$	0.95238	0.89286	0.86207	0.90909
$r_{buw}^{\cdot}$	0.97087	0.91743	0.78125	0.90909
$r_{gct}^{\cdot}$	0.96154	0.68966	0.86207	0.91743

Table2: Results of SQRI (version-2) of 15 correlation estimators

Estimators	Data-1	Data-2	Data-3	Data-4
r_p	0.96154	0.90090	0.69444	0.91743
r_{av}	0.96154	0.92593	0.76923	0.93458
r_{aum}	0.96154	0.92593	0.75188	0.94340
r_{mad}	0.93458	0.92593	0.60606	0.88496
r_s	0.90909	0.91743	0.74074	0.90909
r_{mf}	0.92593	0.92593	0.72993	0.93458
r_k	0.93458	0.94340	0.81967	0.93458
r_{gd}	0.92593	0.96154	0.71429	0.92593
$r_Q^{\cdot}$	0.98039	0.85470	0.58824	0.86957
$r_k^{\cdot}$	0.87719	0.88496	0.70922	0.87719
$r_{ns}^{\cdot}$	**0.96154**	**0.97087**	**0.90909**	**0.96154**
$r_{ssd}^{\cdot}$	0.86957	0.95238	0.77519	0.90090
$r_{bvt}^{\cdot}$	0.92593	0.84034	0.86207	0.88496
$r_{buw}^{\cdot}$	0.94340	0.90090	0.65359	0.86957
$r_{gct}^{\cdot}$	0.95238	0.64103	0.86207	0.90909

(i) Pearson correlation coefficient, r_p(Pearson, 1896), (ii) An absolute value CC, r_{av}(Gideon,1998), (iii) An absolute value from median CC, r_{avm}(Gideon,1998),(iv) Median-type CC, r_{mad}(Gideon,1998)
(v) Spearman's CC, r_s(Spearman, 1904),(vi) Spearman's Modified Footrule CC, r_{mf}(Gini,1914)
(vii) Kendall's CC, r_k(Kendall,1938),(viii) The Greatest Deviation CC, r_{gd}(Gideon and Hollister,1987)
(ix) The quadrant Estimate CC estimate, r_Q : (Sheppard, 1899;Blomqvist,1950),(x) Transformation of Kendall's CC estimate, r_K(Kendall, 1970),(xi) Normal Scores CC estimate, r_{ns}(Fieller, Hartley and Pearson, 1957; FiellerandPearson, 1961),(xii) The Sum and differences of the standardized observed values CC estimate, r_{ssd},(Gnanadesikan and Kettenring, 1972),(xiii) Bivariate trimmed CC estimate, r_{bvt}(Gnanadesikan and Kettenring, 1972),(xiv) Bivariate Winsorized CC estimate, r_{bvw}(Devlin et al. 1975)
(xv) Trimming with respect to the principal components estimate, r_{pct}: (Devlin et al.1975)
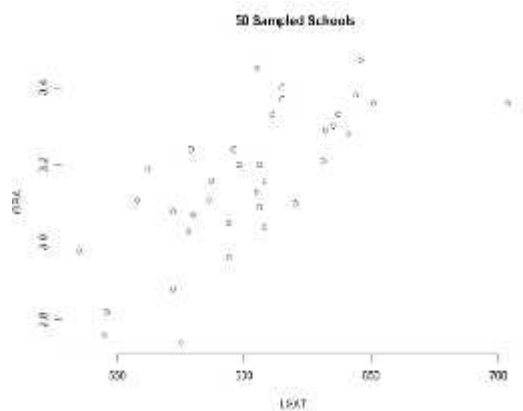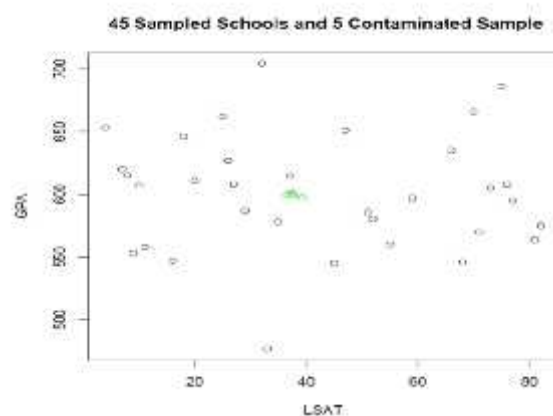

Fig1. Scatter plot of 50 sampled Schools.
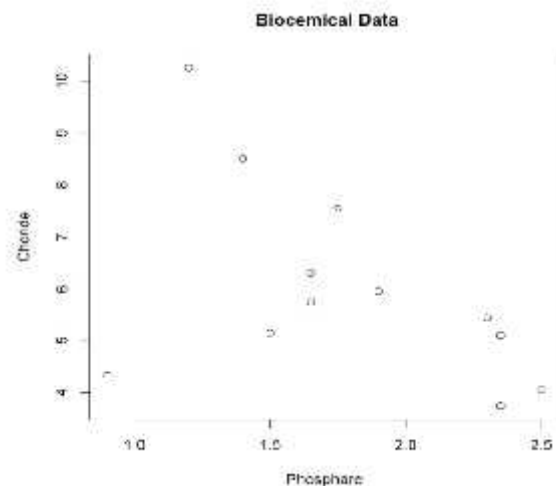

Fig2. Scatter plot of 44 sampled Schools and 5 contaminated samples.

Fig3. Scatter Biochemical data