Uniform strong consistency of robust estimators

José R. Berrendero[1]

[1] Universidad Autónoma de Madrid, Departamento de Matemáticas, 28049, Madrid, Spain

Keywords: Uniform consistency, M-estimator, GS-estimator, Contamination neighborhood.

1 Introduction

If an estimator is strongly consistent, the error of estimation will be small with arbitrarily high probability as long as the sample size is large enough (depending on F, the distribution of the data.) However, a key feature of the statistical robustness theory is the uncertainty about the probabilistic model underlying the data. The distribution F is not completely specified but it is rather assumed to belong to a family $\mathcal{F}$ of plausible distributions. From this point of view, it is required to study the uniform asymptotic behavior over $\mathcal{F}$ of robust estimators. Conditions for the uniform strong consistency of the mean are given in Shorack (2000), p. 252. Different approaches on uniform asymptotics for some kinds of estimators can be found in Davies (1998), Salibian-Barrera and Zamar (2001) and Zielinski (1998). In this paper, we give a general result on uniform strong consistency (Section 2) and apply it to study the consistency of some classes of robust estimators over contamination neighborhoods (Section 3).

2 A general result on uniform strong consistency

Assume that we have a sample $X_1, \ldots, X_n$ of independent, identically distributed random variables drawn from a distribution F, which may be any distribution of certain family $\mathcal{F}$. We will consider the general class of estimators that solve an estimating equation. That is, provided that we have an interesting score function $f(t; X_1, \ldots, X_n)$, decreasing in t, $\hat{\theta}_n$ satisfies $f(\hat{\theta}_n; X_1, \ldots, X_n) \approx 0$. To cope with discontinuous or non strictly monotone score functions, the estimator $\hat{\theta}_n$ is precisely defined as

$$\hat{\theta}_n \doteq \inf\{t : \; f(t; X_1, \ldots, X_n) < 0\}. \tag{1}$$

Let $f_n(t, F) \doteq \mathrm{E}_F[f(t; X_1, \ldots, X_n)]$ and define $\theta_n \doteq \inf\{t : \; f_n(t, F) < 0\}$. Under some conditions, it can be shown that the difference between $\hat{\theta}_n$ and θ_n vanishes a.s. uniformly over $\mathcal{F}$ as $n \to \infty$ [see (2) below]. Next, we list the precise assumptions and state the main result.

A1 Monotonicity assumption*: The score function $f(t; x_1, \ldots, x_n)$ is decreasing as a function of t, for all $x_1, \ldots, x_n$.*

A2 Bounded difference assumption*: For each $i = 1, \ldots, n$ there exists $c_i \in \mathbb{R}$ such that*

$$\sup_{\substack{x_1, \ldots, x_n \\ x_i' \in \mathbb{R}}} |f(t; x_1, \ldots, x_i, \ldots, x_n) - f(t; x_1, \ldots, x_i', \ldots, x_n)| \leq c_i,$$

and $\sum_{n=1}^{\infty} \exp(-\gamma / \sum_{i=1}^{n} c_i^2) < \infty$, for all $\gamma > 0$.

A3 Uniform lower bound assumption*: For all $\delta > 0$, $\alpha = \alpha(\delta) \doteq \inf_n \inf_{F \in \mathcal{F}} f_n(\theta_n - \delta, F) > 0$ and $\beta = \beta(\delta) \doteq -\inf_n \inf_{F \in \mathcal{F}} f_n(\theta_n + \delta, F) > 0$.*

Theorem 1 *Under assumptions* **A1**, **A2** *and* **A3**, *it holds*

$$\lim_{n \to \infty} \sup_{F \in \mathcal{F}} P_F\{\sup_{m \geq n} |\hat{\theta}_m - \theta_m| > \delta\} = 0, \quad \textit{for all } \delta > 0. \tag{2}$$

3 Some applications

Given a distribution function F_0, we define the ϵ-contamination neighborhood $\mathcal{F}_\epsilon$ as

$$\mathcal{F}_\epsilon = \{F = (1-\epsilon)F_0 + \epsilon H : H \text{ arbitrary distribution}\}.$$

In this section, we apply the general result of Section 2 to some classes of robust estimators when $\mathcal{F} = \mathcal{F}_\epsilon$. It is assumed that F_0 is symmetric and strictly increasing.

3.1 M-estimators

If we consider $f(t, X_1, \ldots, X_n) = n^{-1} \sum_{i=1}^n \psi(X_i - t)$ in (1), then $\hat{\theta}_n$ is a location M-estimator [see Huber, 1981]. Under some mild conditions on ψ, assumptions **A1**, **A2** and **A3** hold for $\mathcal{F}_\epsilon$ with $\epsilon < \epsilon^*$, where ϵ^* is the breakdown point of the estimator.

3.2 Location estimators based on U-statistics

Maritz et al. (1977) and Brown and Hettmansperger (1994) studied a class of robust location estimators defined as in (1) with $f(t, X_1, \ldots, X_n) = \binom{n}{2}^{-1} \sum_{i<j} \psi\big(X_i - cX_j - (1-c)t\big)$, where $\psi(x) = \mathrm{sgn}(x)$ and $c \in [-1, 1)$. The case $c = 0$ corresponds to the median whereas $c = -1$ yields the Hodges-Lehmann estimator $\hat{\theta}_n = \mathrm{med}_{i<j}\{(X_i + X_j)/2\}$. Assumptions **A1**, **A2** and **A3** can also be checked for these estimators.

3.3 Generalized S-estimators

The class of dispersion generalized S-estimators (GS-estimators) was defined by Croux, Rousseeuw and Hössjer (1994) in the context of linear regression. They are defined as in (1) with

$$f(t, X_1, \ldots, X_n) = \binom{n}{2}^{-1} \sum_{i<j} \chi\Big[(X_i - X_j)/t\Big] - b,$$

for a score function χ and a constant $b \in (0, 1)$. When $\chi(x)$ is even, bounded and increasing in $(0, \infty)$, with at most a finite number of discontinuities, $\chi(0) = 0$ and $\chi(\infty) \doteq \sup_x \chi(x) = 1$, it can be shown that assumptions **A1**, **A2** and **A3** hold for $\mathcal{F}_\epsilon$ and $\epsilon < \epsilon^* = \min\{b, 1-b\}$, the breakdown point of the estimator.

References

B.M. Brown and T.P. Hettmansperger (1994). Regular redescending rank estimates. *J. Amer. Statist. Assoc.* 89(426), 538–542.

C. Croux, P.J. Rousseeuw and O. Hössjer (1994). Generalized S-estimators. *J. Amer. Statist. Assoc.* 89(428), 1271–1281.

L. Davies (1998). On locally uniformly linearizable high breakdown location and scale functionals. *Ann. Statist.*, 26(3), 1103–1125.

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

J.S. Maritz, M. Wu and R.G. Staudte, Jr. (1977). A location estimator based on a U-statistic. *Ann. Statist.*, 5(4), 779–786.

M. Salibian-Barrera and R.H. Zamar (2001). Uniform asymptotics for robust location estimates when the scale is unknown. Unpublished manuscript available at `http://mathstat.carleton.ca/~matias/pubs.html`.

G.R. Shorack (2000). *Probability for Statisticians.* Springer, New York.

R. Zielinski (1998). Uniform strong consistency of sample quantiles. *Statist. Probab. Lett.*, 37(2), 115–119.

Please fill in this form and mail it together with your abstract.

My abstract fits best to topic number 9 (**"PLEASE INDICATE HERE THE NUMBER FROM THE LIST OF TOPICS BELOW WHICH BEST FITS TO YOUR ABSTRACT"**.).

I prefer to present my paper at the conference
( ) orally
(X) as a poster
(please mark the correct answer by a cross (X)).

List of Topics:

1. Algorithms
2. Applications
3. Biostatistics
4. Computing and graphics
5. Data analysis
6. Data mining
7. Economics, finance
8. Efficiency and robustness
9. Functionals and bias
10. Fuzzy statistics
11. Geostatistics
12. Inference for robust methods, model testing
13. Location depth and regression depth
14. Multivariate methods
15. Neural networks
16. Rank-based methods
17. Regression quantiles, trimming
18. Robust covariance
19. Robust designs
20. Robust regression
21. Time series analysis
22. Wavelets