ROBUST INFERENCE

Elvezio Ronchetti [1]
Professor, Department of Econometrics
University of Geneva, CH-1211 Geneva, Switzerland

Robust statistics deals with deviations from ideal parametric models and their dangers for the statistical procedures derived under the assumed model. Its primary goal is the development of procedures which are still reliable and reasonably efficient under small deviations from the model, i.e. when the underlying distribution lies in a neighborhood of the assumed model. Robust statistics is then an extension of parametric statistics, taking into account that parametric models are at best only approximations to reality. The field is now some 50 years old. Indeed one can consider Tukey (1960), Huber (1964), and Hampel (1968) the fundamental papers which laid the foundations of modern robust statistics. Book-length expositions can be found in Huber (1981, 2nd edition by Huber and Ronchetti 2009), Hampel, Ronchetti, Rousseeuw, Stahel (1986), Maronna, Martin, Yohai (2006).

More specifically, in robust testing one would like the level of a test to be stable under small, arbitrary departures from the distribution at the null hypothesis (*robustness of validity*). Moreover, the test should still have good power under small arbitrary departures from specified alternatives *(robustness of efficiency)*. For confidence intervals, these criteria correspond to stable coverage probability and length of the confidence interval.

Many classical tests do not satisfy these criteria. An extreme case of non-robustness is the F-test for comparing two variances. Box (1953) showed that the level of this test becomes large in the presence of tiny deviations from the normality assumption; see Hampel *et al.* (1986), p. 188-189. Well known classical tests exhibit robustness problems too. The classical t-test and F-test for linear models are relatively robust with respect to the level, but they lack robustness of efficiency with respect to small departures from the normality assumption on the errors; cf. Hampel (1973), Schrader and Hettmansperger (1980), Ronchetti (1982), Heritier *et al.* (2009), p. 35. Nonparametric tests are attractive since they have an exact level under symmetric distributions and good robustness of efficiency. However, the distribution free property of their level is affected by asymmetric contamination, cf. Hampel *et al.* (1986), p. 201. Even randomization tests which keep an exact level, are not robust with respect to the power if they are based on a non-robust test statistic like the mean.

[1]Professor Elvezio Ronchetti is Past Vice-President of the Swiss Statistical Association (1988—91). He was Chair, Department of Econometrics, University of Geneva (2001—2007). He is an Elected Fellow of the American Statistical Association (2001) and of the International Statistical Institute (2008). Currently he is an Associate Editor, *Journal of the American Statistical Association* (2005—present) and Director of the Master of Science and PhD Program in Statistics, University of Geneva (2009—present). He is the co-author (with F.R. Hampel, P.J. Rousseeuw and W.A. Stahel) of the well known text *Robust Statistics : The Approach Based on Influence Functions* (Wiley, New York, 1986, translated also into Russian), and of the 2nd edition of Huber's classic *Robust Statistics* (with P. J. Huber, Wiley, 2009).

The first approach to formalize the robustness problem was Huber's (1964, 1981) minimax theory, where the statistical problem is viewed as a game between the Nature (which chooses a distribution in the neighborhood of the model) and the statistician (who chooses a statistical procedure in a given class). The statistician achieves robustness by constructing a minimax procedure which minimizes a loss criterion at the worst possible distribution in the neighborhood. More specifically, in the problem of testing a simple hypothesis against a simple alternative, Huber (1965, 1981) found the test which maximizes the minimum power over a neighborhood of the alternative, under the side condition that the maximum level over a neighborhood of the hypothesis is bounded. The solution to this problem which is an extension of Neyman-Pearson's Lemma, is the censored likelihood ratio test. It can be interpreted in the framework of capacities (Huber and Strassen, 1973) and it leads to exact finite sample minimax confidence intervals for a location parameter (Huber, 1968). While Huber's minimax theory is one of the key ideas in robust statistics and leads to elegant and exact finite sample results, it seems difficult to extend it to general parametric models, when no invariance structure is available.

The infinitesimal approach introduced in Hampel (1968) in the framework of estimation, offers an alternative for more complex models. The idea is to view the quantities of interest (for instance the bias or the variance of an estimator) as functionals of the underlying distribution and to use their linear approximations to study their behavior in a neighborhood of the ideal model. A key tool is a derivative of such a functional, the influence function (Hampel, 1974) which describes the local stability of the functional.

To illustrate the idea in the framework of testing, consider a parametric model $\{F_\theta\}$, where θ is a real parameter and a test statistic T_n which can be written (at least asymptotically) as a functional $T(F_n)$ of the empirical distribution function F_n. Let $H_0 : \theta = \theta_0$ be the null hypothesis and $\theta_n = \theta_0 + \Delta/\sqrt{n}$ a sequence of alternatives. We consider a neighborhood of distributions $F_{\epsilon,\theta,n} = (1-\epsilon/\sqrt{n})F_\theta + (\epsilon/\sqrt{n})G$, where G is an arbitrary distribution and we can view the asymptotic level α of the test as a functional of a distribution in the neighborhood. Then by a von Mises expansion of α around F_{θ_0}, where $\alpha(F_{\theta_0}) = \alpha_0$, the nominal level of the test, the asymptotic level and (similarly) the asymptotic power under contamination can be expressed as

$$\lim_{n\to\infty} \alpha(F_{\epsilon,\theta_0,n}) = \alpha_0 + \epsilon \int IF(x;\alpha,F_{\theta_0})dG(x) + o(\epsilon), \tag{1}$$

$$\lim_{n\to\infty} \beta(F_{\epsilon,\theta_n,n}) = \beta_0 + \epsilon \int IF(x;\beta,F_{\theta_0})dG(x) + o(\epsilon), \tag{2}$$

where

$$IF(x;\alpha,F_{\theta_0}) = \phi(\Phi^{-1}(1-\alpha_0))IF(x;T,F_{\theta_0})\Big/[V(F_{\theta_0},T)]^{1/2},$$

$$IF(x;\beta,F_{\theta_0}) = \phi(\Phi^{-1}(1-\alpha_0)-\Delta\sqrt{E})IF(x;T,F_{\theta_0})\Big/[V(F_{\theta_0},T)]^{1/2},$$

$\alpha_0 = \alpha(F_{\theta_0})$ is the nominal asymptotic level, $\beta_0 = 1-\Phi(\Phi^{-1}(1-\alpha_0)-\Delta\sqrt{E})$

is the nominal asymptotic power, $E = [\xi'(\theta_0)]^2 \Big/ V(F_{\theta_0}, T)$ is Pitman's efficacy of the test, $\xi(\theta) = T(F_\theta)$, $V(F_{\theta_0}, T) = \int IF(x; T, F_{\theta_0})^2 dF_{\theta_0}(x)$ is the asymptotic variance of T, and $\Phi^{-1}(1-\alpha_0)$ is the $1-\alpha_0$ quantile of the standard normal distribution Φ and ϕ is its density; see Ronchetti (1979), Rousseeuw and Ronchetti (1979). More details can be found in Markatou and Ronchetti (1997) and Huber and Ronchetti (2009), Ch. 13.

Therefore, bounding the influence function of the the test statistic T from *above* will ensure *robustness of validity* and bounding it from *below* will ensure *robustness of efficiency* This is in agreement with the exact finite sample result about the structure of the censored likelihood ratio test obtained using the minimax approach.

In the multivariate case and for general parametric models, the classical theory provides three asymptotically equivalent tests, Wald, score, and likelihood ratio test, which are asymptotically uniformly most powerful with respect to a sequence of contiguous alternatives. If the parameter of the model is estimated by a robust estimator such as an $M-$estimator T_n defined by the estimating equation $\sum_{i=1}^n \psi(x_i; T_n) = 0$, natural extensions of the three classical tests can be constructed by replacing the score function of the model by the function ψ. This leads to formulas similar to (1) and (2) and to optimal bounded influence tests; see Heritier and Ronchetti (1994).

References

[1] Box G.E.P. (1953). Non-normality and Tests on Variances'. *Biometrika.* **40**, 318–335.

[2] Hampel F. R. (1968). Contribution to the Theory of Robust Estimation. Ph.D Thesis, University of California, Berkeley.

[3] Hampel F.R. (1973). Robust Estimation: a Condensed Partial Survey. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **27**, 87–104.

[4] Hampel F.R. (1974). The Influence Curve and its Role in Robust Estimation. *Journal of the American Statistical Association* **69**, 383–393.

[5] Hampel F.R., Ronchetti E.M., Rousseeuw P.J. and Stahel W.A. (1986). *Robust Statistics: The Approach Based on Influence Functions.* Wiley, New York.

[6] Heritier S. and Ronchetti E. (1994). Robust Bounded-influence Tests in General Parametric Models. *Journal of the American Statistical Association*, **89**, 897–904.

[7] Heritier S., Cantoni E., Copt S. and Victoria-Feser M.-P. (2009). *Robust Methods in Biostatistics.* Wiley, Chichester.

[8] Huber P.J. (1964). Robust Estimation of a Location Parameter. *Annals of Mathematical Statistics* **35**, 73–101.

[9] Huber P.J. (1965). A Robust Version of the Probability Ratio Test. *Annals of Mathematical Statistics* **36**, 1753–1758.

[10] Huber P.J. (1968). Robust Confidence Limits. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **10**, 269–278.

[11] Huber P. J. and Strassen V. (1973). Minimax Tests and the Neyman-Pearson Lemma for Capacities. *The Annals of Statistics* **1**, 251–263; **2**, 223–224.

[12] Huber P.J. (1981). *Robust Statistics.* Wiley, New York.

[13] Huber P.J. and Ronchetti E. M. (2009). *Robust Statistics.* 2nd edition, Wiley, New York.

[14] Markatou M. and Ronchetti E. (1997). Robust Inference: The Approach Based on Influence Functions. In Maddala G. S. and Rao C. R. eds., *Handbook of Statistics.* **15**, North Holland, 49–75.

[15] Maronna R. A., Martin R. D. and Yohai V. J. (2006). *Robust Statistics: Theory and Methods.* Wiley, New York.

[16] Ronchetti E. (1979). Robustheitseigenschaften von Tests. Diploma Thesis, ETH Zürich, Switzerland.

[17] Ronchetti E. (1982). Robust Testing in Linear Models: The Infinitesimal Approach. PhD Thesis, ETH Zürich, Switzerland.

[18] Rousseeuw, P.J. and Ronchetti E. (1979). The Influence Curve for Tests. Research Report 21, Fachgruppe für Statistik, ETH Zürich, Switzerland.

[19] Schrader R. M. and Hettmansperger T. P. (1980). Robust Analysis of Variance Based Upon a Likelihood Ratio Criterion. *Biometrika*, **67**, 93–101.

[20] Tukey J. W. (1960). A Survey of Sampling from Contaminated Distributions. In Olkin I. ed., *Contributions to Probability and Statistics*, Stanford University Press, 448–485.