Specification testing

A model is said to be statistically adequate, or correctly specified, when its underlying assumptions are supported by the observed data. The concept of statistical adequacy was first put forward by Fisher and by Koopmans (1937) referred to as Fisher's axiom of correct specification. Once a model is selected for performing statistical inferences, the next natural step is to test if the specification of the putative model is correct. If the model is misspecified, the resulting statistical inferences are usually invalid.

Specification testing, or goodness-of-fit testing, is a classical research topic in statistics since the early thirties. The pioneering work of Kolmogorov (1933), Smirnov (1936), Cramér (1928) and von Mises (1931) for testing simple hypotheses on distribution functions were extended in the late fifties by Kac et al. (1955) and Gikhman (1953) to test composite hypothesis, where parameters are estimated. The formalization of these results was provided in the seventies by Durbin (1973) and Neuhaus (1973, 1977). Test statistics are suitable functionals of the standard empirical process resulting in omnibus tests, i.e. tests designed to detect general alternatives of nonparametric nature. The limiting null distribution of test statistics with estimated parameters is case dependent, but tests can be implemented with the assistance of a parametric bootstrap. Khmaladze (1981) proposed to use the martingale part of the empirical process with estimated parameters for constructing asymptotically distribution free tests. See D'Agostino and Stephens (1986) for an overview of classical specification tests.

As argued by Durbin and Knott (1972), see also Eubank and LaRicca (1992), classical omnibus specification tests are highly unlikely to detect many alternatives in practice. Recently, Janssen (2000) has shown that any of these tests has a preference for a finite dimensional space of alternatives. Apart from this set, the power function is almost flat on balls of alternatives. Furthermore, there exists no test which pays equal attention to an infinite number of orthogonal alternatives. As a compromise between omnibus and directional tests, smooth tests, introduced by Neyman (1937), are based on the Lagrange-Multiplier testing principle. They assume a flexible parametric model under the alternative hypothesis, usually belonging to the exponential family. If the number of parameters in the model under the alternative hypothesis increases with the sample size at a suitable rate, the smooth tests become omnibus. They are related to tests that compare the estimated parametric model under the null and a nonparametric fit using smoothers, as proposed by Rosenblatt (1975). See Rayner and Best (1989) and Hart (1997) for overviews of Neyman's smooth tests and tests using nonparametric smoothing.

The alternative testing methodologies were conceived to test the specification of distribution functions, but they have been extended to test the specification of other type of models. Specification testing of regression curves have been proposed by Härdle and Mammen (1993) using smoothers and by Stute (1997) using a type of CUSUM process. Many others have studied the properties of both procedures. The same alternative strategies have been applied for goodness-of-fit testing of survival curves, conditional variances, conditional distributions, spectral distributions and general conditional functionals.

There have also been important developments in the econometrics literature. Haavelmo (1944) discusses in his famous *Econometrica* monograph the crucial importance of a correct specification for performing valid inferences. Among the many recent contributions, it is worth mentioning Zheng (1996) on regression curves using smoothers, Fan and Li (1996) on semi-parametric models using smoothers, Andrews (1997) on conditional distributions using multi-parameter empirical processes, and Horowitz and Spokoiny (2001) on optimal smoothing based testing.

This Annals volume is edited on the occasion of an 'Explanatory Workshop of the European Science Fundation' on Specification Testing held at Santander (Spain) in December 2005. The workshop was interdisciplinary, aimed to bring together researchers working in specification testing and those with different applied interests, e.g. economics, finance, physics, medicine and engineering. Twenty-five invited participants from 13 countries presented articles in the conference. Among them, 17 have submitted their work and twelve articles have been accepted after the usual referee process in the *Journal of Econometrics*.

The papers in this volume provide an upto date perspective on the state of the art in specification testing. The majority of articles deals with omnibus specification testing under the two leading methodologies: empirical processes and nonparametric smoothing.

The articles by *Haywood and Khmaladze*, *Bai and Chan*, and *Delgado and Stute* apply martingale transforms to empirical process with estimated parameters, as suggested by Khmaladze (1981) in the standard case. *Haywood and Khmaladze* address the important problem of testing for 'lack of memory', i.e. testing that the underlying distribution is exponential. The martingale transformation is shown to be easy to implement and the resulting distribution free tests are in general more efficient than existing alternatives. In fact, they demonstrate that commonly used specification tests for exponentiality have negligible asymptotic power against a great majority of local alternatives. *Bai and Chan* propose a new distribution free specification test for GARCH models. The multivariate data set is transformed to independent components by using Rosenblatt (1952) transform, and the martingale transform is applied to the multivariate empirical process of the resulting transformed data. The tests statistics based on continuous functionals of the transformed empirical process converge to functionals of independent Brownian Motions, which can be tabulated. These transformations are also used by *Delgado and Stute* for distribution free specification testing of conditional distributions. The resulting test is in fact based on an asymptotically distribution free transformation of the empirical process considered by Andrews (1997). Unlike *Bai and Chan* problem, the interest lies in the specification of a conditional distribution, the marginal distribution of the explanatory variables is nonparametric. That is, the joint distribution under the null hypothesis is semiparametric and Rosenblatt's transformation is not directly applicable.

The tests by *Dette and Podolskij* and *Escancian*o also use empirical processes to test the specification of financial models. *Dette and Podolskij* present a new test for the form of the volatility function. When testing for homoskedasticity, the test is asymptotically distribution free, and in the general case can be implemented with the assistance of a parametric bootstrap. Bootstrap assisted tests are also proposed by *Escanciano* for simultaneous testing of the conditional mean and variance using serial dependent data, which is well motivated for specification testing of GARCH and other volatility models. In fact, the null hypothesis tested is that innovations both in the conditional mean and variance models are martingale differences. The test statistics are continuous functionals of the generalized spectral process of the residuals.

Einmahl and van Keilegom introduce a new test for location-scale models. It is based on the Hoeffding–Blum–Kiefer–Rosenblatt process of scaled nonparametric residuals and regressors. This is a test of the independence between scaled errors of a nonparametric regression model and the explanatory variables, which underlies many statistical inference procedures for nonparametric and semiparametric models.

The articles by *Lavergne and Patilea*, *Gao*, *Gijbels and van Bellegem*, *Hidalgo* and *Cao and González–Manteiga* consider tests based on the comparison of the fitted model under the null hypothesis and nonparametric fits using smoothers, in the spirit of Rosenblatt (1975) seminal paper. *Lavergne and Patilea* address the problem of the curse of dimensionality in the context of specification testing of regression models. They propose a new test which detect local alternatives converging to the null at the same rate, independently of the number of explanatory variables considered. *Gao*, *Gijbels and van Bellegem* develop a procedure for simultaneous testing the presence of structural breaks in the conditional mean and variance using serial dependent data. *Hidalgo* proposes a specification test for regression models with fixed and random explanatory variables when the data may exhibit long memory. He finds that the Central Limit Theorem for the kernel regression estimator is very different from the short memory case. The limiting distribution of the test statistic involves a convergence rate depending on the long-memory parameter, and bootstrap assisted tests are recommended. *Cao and González-Manteiga* propose specification tests of conditional models under censoring and truncation. These models are interesting in applications when the dependent variable, typically a duration, can be subject to truncation and/or censoring. Censoring appears when the individuals cannot enter the program before the follow-up period, i.e. a left truncation phenomenon. Truncation appears when the duration time is not completely observed since, for instance, the evolution of the individual could not be followed till the end of the duration for whatever reason.

In the context of ordered discrete choice models, the paper by *Mora and Moro-Ejido* compares the two main specification testing alternative methodologies: using empirical processes and smoothers.

Robinson develops sequences of parametric tests for specifying the cointegrating rank of possibly fractional multiple time series. The long-memory parameters of observables and cointegrating errors are unknown, and the test statistics are related to Hausman tests. The test exploits the fact that, when the memory parameter is common to several series, an estimate of this parameter under the restriction of no cointegration achieves an efficiency improvement over unrestricted estimates based on individual series, whereas if the series are cointegrated the former estimator is inconsistent.

Acknowledgements

I would like to take this opportunity to thank the 34 referees for their cooperation, valuable comments and helpful suggestions. I also thank Cheng Hsiao for his advice and support during the editorial process. The Conference was funded by the European Science Foundation (Ref: 01-171), Ministerio de Educación y Ciencia, Spain (Ref: SEJ2004-20241-E) and Instituto Cántabro de Estadística. Financial support from Ministerio de Educación y Ciencia, Spain, Grant no. SEJ2004-04583/ECON is also gratefully acknowledged.

References

Andrews, D.W.K., 1997. A conditional Kolmogorov test. Econometrica 65, 1097 1128.

Cramér, H., 1928. On the composition of elementary errors. Skandinavisk Aktuarietidskrift 11, 141 180.

D'Agostino, R.B., Stephens, M.A., 1986. Goodness of fit Techniques. Marcel Dekker, New York.

Durbin, J., 1973. Weak convergence of the sample distribution function when parameters are estimated. Annals of Statistics 1, 279 290.

Durbin, J., Knott, M., 1972. Components of Cramér von Mises statistics. I. Journal of the Royal Statistical Society Series B 34, 290 307.

Eubank, R.L., LaRicca, V., 1992. Asymptotic comparison of Cramér von Mises and nonparametric techniques for testing goodness of fit. Annals of Statistics 20, 2071 2086.

Fan, Y., Li, Q., 1996. Consistent model specification tests: omitted variables and semiparametric functional forms. Econometrica 64, 865 890.

Gikhman, J.I., 1953. Some remarks on A. Kolmogorov's goodness of fit test. Doklady Akademii Nauk 91, 715 718.

Haavelmo, T., 1944. The probability approach in econometrics. Econometrica 11, 1 115.

Härdle, W., Mammen, E., 1993. Comparing nonparametric versus parametric regression fits. Annals of Statistics 21, 1926 1947.

Hart, J.D., 1997. Nonparametric Smoothing and Lack Of Fit Tests. Springer, New York.

Horowitz, J., Spokoiny, V.G., 2001. An adaptive, rate optimal test of a parametric mean regression model against a nonparametric alternative. Econometrica 69, 599 631.

Janssen, A., 2000. Global power functions of goodness of fit tests. Annals of Statistics 28, 239 253.

Kac, M., Kiefer, J., Wolkowitz, J., 1955. On tests of normality and other tests of goodness of fit based on distance methods. Annals of Mathematical Statistics 26, 189 211.

Khmaladze, E.V., 1981. Martingale approach in the theory of Goodness of fit tests. Theory of Probability and its Applications 26, 240 257.

Kolmogorov, A.N., 1933. Sulla determinazione empirica di una legge di distribuzione. Giornale Instituto Italiano degli Attuari 4, 83 91.

Koopmans, T.C., 1937. Linear Regression Analysis of Economic Time Series. Netherlands Economic Institute, Haarlem.

Neuhaus, G., 1973. Asymptotic properties of the Cramér von Mises statistics when parameters are estimated. In: Hàjek, J. (Ed.), Proceedings of the Praga Symposium on Asymptotic Statistics, September 3 7, Universita Karlova Praha, vol. 2. pp. 257 297.

Neuhaus, G., 1977. . Weak convergence under contiguous alternatives when parameters are estimated: the D_k approach. Lecture Notes in Mathematics, vol. 566. Springer, Berlin, pp. 68 82.

Neyman, J., 1937. Smooth tests. Skandinavisk Aktuarietidskrift 20, 149 199.

Rayner, J.C.W., Best, D.J., 1989. Smooth Tests of Goodness of Fit. Oxford University Press, Oxford.

Rosenblatt, M., 1952. Remarks on a multivariate transformation. Annals of Mathematical Statistics 23, 470 472.

Rosenblatt, M., 1975. A quadratic measure of deviation of two dimensional density estimates and a test of independence. The Annals of Statistics 1, 1 14.

Smirnov, N.V., 1936. Sur la distribution de ω^2. Comptes Rendus de l'Académie des Sciences Paris 202, 449.
Stute, W., 1997. Nonparametric model checks for regression. Annals of Statistics 25, 613 641.
von Mises, R., 1931. Wahrscheinlichkeitsrechnung. Deuticke, Vienna.
Zheng, J.X., 1996. A consistent test of functional form via nonparametric estimation techniques. Journal of Econometrics 75, 263 289.

Miguel A. Delgado
Department of Economics Universidad Carlos III de Madrid Getafe, Spain
E-mail address: miguelangel.delgado@uc3m.es