Electronic Journal of Statistics > Vol. 2 (2008) open journal systems

Empirical likelihood based testing for regression

Ingrid Van Keilegom, *Université catholique de Louvain; Institute of Statistics*
César Sánchez Sellero, *Universidad de Santiago de Compostela*
Wenceslao González Manteiga, *Universidad de Santiago de Compostela*

Abstract

Consider a random vector (X,Y) and let $m(x)=E(Y|X=x)$. We are interested in testing $H_0 : m in {cal M}_{Theta,{cal G}} = {gamma(cdot,theta,g) : theta in Theta, g in {cal G}}$ for some known function $gamma$, some compact set $Theta subset mathbb{IR}^p$ and some function set ${cal G}$ of real valued functions. Specific examples of this general hypothesis include testing for a parametric regression model, a generalized linear model, a partial linear model, a single index model, but also the selection of explanatory variables can be considered as a special case of this hypothesis. To test this null hypothesis, we make use of the so-called marked empirical process introduced by Diebolt (1995) and studied by Stute (1997) for the particular case of parametric regression, in combination with the modern technique of empirical likelihood theory in order to obtain a powerful testing procedure. The asymptotic validity of the proposed test is established, and its finite sample performance is compared with other existing tests by means of a simulation study.

AMS 2000 subject classifications: Primary 62E20; secondary 62F03, 62F05, 62F40, 62G08, 62G10.

Keywords: Marked empirical process, Model check for regression, Nonlinear regression, Partial linear model, Residuals.

Full Text: PDF

Keilegom, Ingrid Van, Sellero, César Sánchez, Manteiga, Wenceslao González, Empirical likelihood based testing for regression, Electronic Journal of Statistics, 2, (2008), 581-604 (electronic). DOI: 10.1214/07-EJS152.

References

[1] Chen, S.X. and Cui, H.J. (2006). On Bartlett correction of empirical likelihood in the presence of nuisance parameters. Biometrika, 93, 215–220. MR2277752

[2] Chen, S.X., Härdle, W. and Li, M. (2003). An empirical likelihood goodness-of-fit test for time series. J. R. Statist. Soc. - Series B, 65, 663–678. MR1998627

[3] Delgado, M.A. and González Manteiga, W. (2001). Significance testing in nonparametric regression based on the bootstrap. Ann. Statist., 29, 1469–1507. MR1873339

[4] Diebolt, J. (1995). A nonparametric test for the regression function : asymptotic theory. J. Statist. Planning Infer., 44, 1–17. MR1323067

[5] Donald, S.G., Ihbens, G.W. and Newey, W.K. (2003). Empirical likelihood estimation and consistent tests with conditional moment restrictions. J. Econometrics, 117, 55–93. MR2002302

[6] Escanciano, J.C. (2006). Goodness-of-fit tests for linear and nonlinear time series models. J. Amer. Statist. Assoc., 101, 531–541. MR2256173

[7] Escanciano, J.C. (2007). Model checks using residual marked empirical processes. Statist. Sinica, 17, 115–138. MR2352505

[8] Escanciano, J.C. and Velasco, C. (2006). Testing the martingale difference hypothesis using integrated regression functions. Comput. Statist. Data Anal., 51, 2278–2294. MR2307501

[9] Fan, J. and Zhang, J. (2004). Sieve empirical likelihood ratio tests for nonparametric functions. Ann. Statist., 32, 1858–1907. MR2102496

[10] Hall, P. and Yatchew, A. (2005). Unified approach to testing functional hypotheses in semiparametric contexts. J. Econometrics, 127, 225–252. MR2156334

[11] Hjort, N.L., McKeague, I.W. and Van Keilegom, I. (2008). Extending the scope of empirical likelihood. Ann. Statist. (in press).

[12] Kitamura, Y., Tripathi, G. and Ahn, H. (2004). Empirical likelihood-based inference in conditional moment restriction models. Econometrica, 72, 1667–1714. MR2095529

[13] Neumeyer, N. and Dette, H. (2003). Nonparametric comparison of regression curves: an empirical process approach. Ann. Statist., 31, 880–920. MR1994734

[14] Owen, A. (2001). Empirical likelihood. Chapman and Hall, New York.

[15] Serfling, R.J. (1980). Approximation theorems of mathematical statistics. Wiley, New York. MR0595165

[16] Stute, W. (1997). Nonparametric model checks for regression. Ann. Statist., 25, 613–641. MR1439316

[17] Stute, W., González Manteiga, W. and Presedo Quindimil, M. (1998). Bootstrap approximation in model checks for regression. J. Amer. Statist. Assoc., 93, 141–149. MR1614600

[18] Stute, W., Presedo Quindimil, M., González Manteiga, W. and Koul, H.L. (2006). Model checks of higher order time series. Stat. Prob. Letters, 76, 1385–1396. MR2269649

[19] Stute, W., Thies, S. and Zhu, L.X. (1998). Model checks for regression: an innovation process approach. Ann. Statist., 26, 1916–1934. MR1673284

[20] Stute, W. and Zhu, L.X. (2002). Model checks for generalized linear models. Scand. J. Statist., 29, 535–545. MR1925573

[21] Tripathi, G. and Kitamura, Y. (2003). Testing conditional moment restrictions. Ann. Statist., 31, 2059–2095. MR2036400

[22] Van der Vaart, A.W. and Wellner, J.A. (1996). Weak convergence and empirical processes. Springer-Verlag, New York. MR1385671

[23] Van Keilegom, I., González Manteiga, W. and Sánchez Sellero, C. (2007). Goodness-of-fit tests in parametric regression based on the estimation of the error distribution. TEST (in press).

[24] Xia, Y., Li, W.K., Tong, H. and Zhang, D. (2004). A goodness-of-fit test for single-index models. Statist. Sinica, 14, 1–39. MR2036761

[25] Zhu, L. (2005). Nonparametric Monte Carlo tests and their applications. Lecture Notes in Statistics, Springer-Verlag, New York. MR2162748

[26] Zhu, L., Fujikoshi, Y. and Naito, K. (2001). Heteroscedasticity checks for regression models. Science in China (Series A), 44, 1236–1252. MR1867400

Electronic Journal of Statistics. ISSN: 1935-7524