Test (2013) 22:442–447
DOI 10.1007/s11749-013-0334-6

DISCUSSION

Rejoinder on: An updated review of Goodness-of-Fit tests for regression models

Wenceslao González-Manteiga · Rosa M. Crujeiras

Published online: 25 July 2013

First for all, we would like to thank the discussants for reading our paper and for taking time to prepare such interesting and valuable contributions. The feeling, after revising the discussions and going back to the original version of the paper, is that there is still too much to say about Goodness-of-Fit (GoF) tests for regression models and the discussants have given a good proof of this. Although they qualify the review as *thorough* and *nearly exhaustive*, it is clear after their comments there are quite a few issues and open problems that were left untreated.

We have tried to organize this rejoinder according to the different topics that have been brought up in the discussions, instead of answering each discussion separately. Bearing this in mind, some brief comments will be made on (a) tests based on the error distribution; (b) tests with dependence between error and covariate; (c) bandwidth selection and calibration; (d) non- and semiparametric hypotheses; (e) tests in survival analysis and (f) some further topics.

(a) Tests based on the error distribution In Sect. 2.4, concerned with this topic, two main methodological approaches are presented. The first one, collected in the paper by Van Keilegom et al. (2008), is based on the empirical distribution of the residuals, whereas Huskova and Meintanis (2009) considered an alternative route

This rejoinder refers to the comments available at doi:10.1007/s11749-013-0328-4, doi:10.1007/s11749-013-0329-3, doi:10.1007/s11749-013-0330-x, doi:10.1007/s11749-013-0331-9, doi:10.1007/s11749-013-0332-8, doi:10.1007/s11749-013-0333-7.

W. González-Manteiga (✉) · R.M. Crujeiras
Dpt. of Statistics and Operations Research, Faculty of Mathematics,
University of Santiago de Compostela, Santiago de Compostela, Spain
e-mail: wenceslao.gonzalez@usc.es

R.M. Crujeiras
e-mail: rosa.crujeiras@usc.es

based on the characteristic function (CF). As M.D. Jiménez-Gamero points out, it is quite frequent that tests based on CFs require weaker conditions than their analogues using empirical distributions, but as S.G. Meintanis comments, it is always an issue to provide an intuitive interpretation in this setting. In his discussion, Meintanis remarks a simple way of rewriting the tests statistic in Eq. (15), firstly, by an interpretation of the CF in terms of a convolution with a specific weight function and secondly, as a weighted integral which turns out into a series expansion on matching moments. From these two perspectives, a bandwidth parameter can be directly attached to the weight function, resulting in an L^2 distance between two kernel density estimators for the first case, and controlling the number of moments difference entering the test statistic, for the second case. This nice interpretation sheds some light on CFs-based tests and its relation with the corresponding ones based on the error distribution.

(b) Tests with dependence between error and covariate As I. Van Keilegom mentions in her discussion, Sect. 2.4 and subsequent in the paper are stated assuming that the covariate X and the error term ε are independent, and although there are testing methods for assessing homoskedasticity, it seems necessary to design procedures for testing the full hypothesis of independence. Both Van Keilegom and Meintanis agree in their views, and refer to Einmahl and Van Keilegom (2008a, 2008b) for L^2-type tests and Hlavka et al. (2011) for tests based on CFs. How can we deal with a potential dependence between X and ε? A possible route would be to extend the testing methods accounting for a dependence model between error and covariate, being this relation modeled, for instance, by a copula. Just at the very end of Sect. 1.1 we referred to the extension of GoF tests to copulas, so not only the regression model should be tested, but also the copula function. Apart from the references included in the paper about this topic, we should mention the work by Scaillet (2007), who studied the asymptotic properties of the test statistic proposed by Fermanian (2005) with fixed smoothing parameters. It should be noted that the classical kernel estimators for copulas suffer from a corner bias problem, and Omelka et al. (2009) proposed improved estimators, studied their weak convergence and used them in GoF tests for copulas. Another challenging problem would be to test simultaneously both hypotheses, about the regression model and the independence between error and covariate.

(c) Bandwidth selection and calibration When writing the paper, we both were convinced that bandwidth selection for smooth–based tests was a really tough problem, but after reading the discussants contributions (specifically, the comments by J. de Uña and S. Sperlich), we can say that it is still far from being solved. J. de Uña provides an overview on the algorithms he developed jointly with P. Martínez-Camblor in the context of k-sample smooth tests and we agree that the Bootstrap Minimum (BM) method may deserve further investigation, although with some concerns about the required computational time. In addition, in order to extend the ideas of the BM method to GoF for regression models, the alternative should be specified. This may be possible by considering directional alternatives. In GoF calibration, the computational cost is one of the points highlighted by S. Sperlich, when presenting adaptive testing methods for a parametric null hypothesis, where the distribution of the tests depends on a random bandwidth maximizing a certain criterion. The bandwidth selection issue understood as a calibration problem is part of S. Sperlich's

discussion. We have nothing to add to his comments, rather that we basically agree, and certainly it will be worth reading his forthcoming paper where the subsampling ideas are used for calibration.

(d) Non- and semiparametric hypotheses In his discussion, S. Sperlich focuses on the extensions of GoF methods for assessing non- or semiparametric null hypotheses. This topic was reviewed in Sect. 4 in the paper, and this discussion perfectly complements the contents. From a general perspective and checking the most recent contributions in GoF methods, it is easy to see that the testing problem for non- and semiparametric hypotheses may get complicated by the consideration of complex data structures such as missing data. For instance, very recently Xu et al. (2012) considered the testing problem of a partially linear model with missing response at random and Liu et al. (2011) assessed the heteroscedasticity in a partially linear model with missing covariates. These are just some examples of testing problems for non- and semiparametric hypotheses with complex data, and although there are many other settings sharing these features, we would like to remark that a key issue here is the appearance of a nonparametric rate under the null hypothesis, with the subsequent impact on the test calibration.

(e) Tests in survival analysis At least two of the discussants have a wide experience in survival analysis. The complexity that may arise in this context was partially collected in Sects. 6.1 and 6.4, dealing with censored and/or truncated data and length-biased data. The assumption about independence between truncation and censoring times may not be realistic in some scenarios, as remarked by J. de Uña. The implications behind this problem are difficult to describe, but at least something which seems clear is that, considering the dependence between truncation and censoring as a nuisance, the calibration procedure should be adapted to account for this feature. Also in the survival setting, the extension of the existing tests to doubly truncated data or missing binary response would be also interesting. For instance, it is noted that a non-parametric estimator for the regression function in the double-truncated setting is already available. Following the ideas in Cao and González-Manteiga (2008), the design of a parametric estimator allowing an i.i.d. representation would enable the construction of a GoF test. The previous issues fit within the main scope of Sect. 6, on regression models with complex data, but an unexplored area which has been brought up by I. Van Keilegom is the development of GoF tests for cure models. The previous comments on the possible construction of a GoF test also apply here.

(f) Some further topics The testing problems collected in the review are important, but there are other different scenarios deserving attention. As an example, I. Van Keilegom discusses testing problems on the coefficient of variation, revising the existing alternatives for testing proportionality between the regression function and the scale function, with some work under development considering non-constant relations. H. Dette complements the work with deeper attention to testing problems in quantile regression and both discussants say also some words about transformation models and inverse regression.

At this point, we could not resist the temptation to contribute with some very recent ideas and extensions that have just been developed. In Sect. 5.2 we made some

comments about testing in spatial and spatio-temporal models, where some contributions have been made from the frequency domain. The use of these tests is restricted to regularly sampled data, which is not always the case in spatial analysis. Following the idea of Diblasi and Bowman (2001), which adapted the regression methods to construct a test for independence in spatial data based on a smoothed variogram, Bowman and Crujeiras (2013) proposed some tests for assessing isotropy and stationarity. The hypotheses to test are nonparametric, but they can be rewritten in terms of the variogram function. The second extension we would like to mention refers to directional data. Directional data are data collected in a sphere, and although there exist parametric density models, such as the von Mises–Fisher distribution, general GoF tests were not available. This is the contribution of Boente et al. (2013), which may state the basis for developing GoF tests in regression models involving directional covariates.

As we have already mentioned, we are aware of the incompleteness of this paper. Trying to do our best, we have proceeded with an update of the contributions in the different topics considered in the paper, so we will make a brief mention on some papers that should have been included in the main text. In Sect. 5.3 we revise some testing ideas in continuous time models. In this setting, Aït-Sahalia et al. (2009) provided some tests for the transition density of a jump-diffusion process (which also apply for pure diffusions), sampled in a discrete way. The authors noted that the direct estimation of local characteristics of the process with discrete data may lead to inconsistent estimates, but the discretization procedure cannot be avoided, and in this work and in subsequent papers (already mentioned in the review) the model specification is done using densities at the observed discrete frequency. Forman et al. (2011) took advantage from the discretization scheme and developed a GoF test for continuous time models named the downsampling test. This proposal allows to check if the parameter estimates from different sampling frequencies are consistent, comparing the estimates from the original sample with other ones from downsampled data. Recently, Lin et al. (2013) proposed a GoF test for continuous time stochastic volatility models, also using discrete samples. The test is based on CFs (see the discussion by S. Meintanis), but taking in consideration both fixed and decreasing sampling schemes. Last, but not least, we would like to remark that apart from the calibration problem, described in detail by S. Sperlich, there may be another issues to bear in mind when analyzing GoF tests in complicated settings. For instance, in the particular case of generalized partially linear models, the usual estimators are not robust, and this problem is inherited by the test statistic. Boente et al. (2013) introduced a family of robust statistics to test a parametric model vs. a semiparametric one.

Along the paper we have considered both smoothing tests and tests based on empirical processes, but most of the discussion has been focused on the first part. Although the second type of tests overcomes the problem of a bandwidth selection and these tests are asymptotically more powerful, the simulation carried out by Miles and Mora (2002) shows that this is not the case for small samples. As J. de Uña says, the simulation study in this paper is restricted to one and two covariates, and certainly this is not conclusive. However, one could check the simulation results in Stute et al. (2006) to get a confirmation with up to five covariates. To be fair, sample sizes are small (just 100 data) in both cases, so the behavior for a large sample should be investigated.

Finally, there are some other issues that cannot be left aside. The first one refers to the level of computational implementation of the proposed methods, as noted by J. de Uña. From our knowledge, there is not a generic package in R or in other software comprising the methods that have been included in the review. Smooth regression estimators (classical Nadaraya–Watson, local polynomial, splines, ...) can be found in different R packages and in other commercial programs, but there is a limited number of available tests in line with the survey. Library intRegGOF provides functions for applying the integrated regression GoF introduced by Stute (1997). We may also cite the sm library, developed by Bowman and Azzalini (see Bowman and Azzalini 1997), which includes smooth-based tests for linearity, monotonicity and significance, among others. Nevertheless, there could be other tests available that we may not know. However, if one would like to develop a single toolbox for smooth-based GoF tests (such as the ones presented in Eqs. (5), (7), and (8)) at least the following building blocks should be available: (i) a smooth regression estimator for the nonparametric component, (ii) a $\sqrt{n}$-consistent estimator for the parametric model (such as a maximum likelihood estimator) and (iii) a proper resampling mechanism which allows to mimic the null hypothesis. Apparently, the computation of the test statistics for a fixed bandwidth and a given weight function may not pose significant problems, except for some adequate reformulation for T_{2n} in (7) in order to avoid repeating operations.

Additionally, some of the discussants have suggested the possibility of writing a monograph on this topic, taking this review as a starting point. We feel flattered by their proposal and we will try to find the time and the courage to undertake such an exciting project.

References

Aït-Sahalia Y, Fan J, Peng H (2009) Nonparametric transition–based tests for jump diffusions. J Am Stat Assoc 104:1102–1116

Boente G, González-Manteiga W, Rodríguez D (2013) Goodness-of-fit test for directional data. Scand J Stat. doi:10.1111/sjos.12020

Bowman AW, Azzalini A (1997) Applied smoothing techniques for data analysis: the kernel approach with S-plus illustrations. Oxford University Press, Oxford

Bowman AW, Crujeiras RM (2013) Inference for variograms. Comput Stat Data Anal 6:19–31

Cao R, González-Manteiga W (2008) Goodness-of-fit tests for conditional models under censoring and truncation. J Econom 143:166–190

Diblasi AM, Bowman AW (2001) On the use of the variogram in checking for independence in spatial data. Biometrics 57:211–218

Einmahl J, Van Keilegom I (2008a) Tests for independence in nonparametric regression. Stat Sin 18:601–616

Einmahl J, Van Keilegom I (2008b) Specification tests in nonparametric regression. J Econom 143:88–102

Fermanian JD (2005) Goodness-of-fit tests for copulas. J Multivar Anal 95:119–152

Forman JL, Markussen B, Sorensen H (2011) Goodness-of-fit based on downsampling with applications to linear drift diffusions. Scand J Stat 38:288–310

Hlavka Z, Huskova M, Meintanis SG (2011) Tests for independence in nonparametric heteroscedastic regression models. J Multivar Anal 102:816–827

Huskova M, Meintanis S (2009) Goodness-of-fit tests for parametric regression models based on empirical characteristic functions. Kybernetika 45:960–971

Lin LC, Leeb S, Guo M (2013) Goodness-of-fit test for stochastic volatility models. J Multivar Anal 116:473–498

Liu Z, Wang Z, Hu X (2011) Testing heteroscedasticity in partially linear models with missing covariates. J Nonparametr Stat 23:321–337

Miles D, Mora J (2002) On the performance of nonparametric specification test in regression models. Comput Stat Data Anal 42:477–490

Omelka M, Gijbels I, Veraverbeke N (2009) Improved kernel estimation of copulas: weak convergence and goodness-of-fit testing. Ann Stat 37:3023–3058

Scaillet O (2007) Kernel-based goodness-of-fit tests for copulas with fixed smoothing parameters. J Multivar Anal 98:533–543

Stute W (1997) Nonparametric model checks for regression. Ann Stat 25:613–641

Stute W, Presendo-Quindimil M, González-Manteiga W, Koul HL (2006) Model checks for higher order time series. Stat Probab Lett 76:1385–1396

Van Keilegom I, González-Manteiga W, Sánchez-Sellero C (2008) Goodness-of-fit tests in parametric regression based on the estimation of the error distribution. Test 17:401–415

Xu W, Guo X, Zhu L (2012) Goodness-of-fitting for partial linear model with missing response at random. J Nonparametr Stat 24:103–118