Comments on: Model-free model-fitting and predictive distributions

Carlos Velasco

Published online: 5 April 2013

© Sociedad de Estadística e Investigación Operativa 2013

I enjoyed reading this paper, which presents a new look to the problem of developing predictive distributions under very general conditions. The key idea of this new method is transforming the data to iid observations exploiting a semiparametric model (such as a mean-scale one) or the probability integral transformation (PIT) when a completely nonparametric model is preferred under a smoothness class assumption. In this last sense the term model-free prediction is coined by the author. This approach of transformation to iid (pseudo) residuals is parallel to the usual modeling strategy conducted in parametric statistics to isolate the random component in the data from the information contained in the conditioning variables. Then, resampling can be used naturally to construct valid confidence intervals using the empirical distribution of these residuals which are iid by construction. To accommodate (possibly infinite dimensional) parameter estimation effects, one-leave-out transformations are proposed by the author to replicate properly in finite samples the associated uncertainty to predicting a new design point.

This principle provides its own point predictions inverting the transformation and using an appropriate solution for a given loss function, so it is quite different from trying to build confident intervals for a given prediction mechanism, for which specific techniques or adjustments could be necessary. Though in simulations it appears that such predictions behave similarly to the usual nonparametric estimators, it could be interesting to further investigate in detail the properties of these new point estimates (of conditional expectations), both asymptotically and in finite samples.

This comment refers to the invited paper available at doi:[10.1007/s11749-013-0317-7](https://doi.org/10.1007/s11749-013-0317-7).

C. Velasco (✉)

Departamento de Economía, Universidad Carlos III de Madrid, Calle Madrid 126, 28903 Getafe, Madrid, Spain

e-mail: carlos.velasco@uc3m.es

This transformation approach is also common, as outlined in different parts of the paper, to the goodness-of-fit literature, where parametric models impose iid errors with a given marginal distribution or specify directly conditional distributions. Here the modeling target is obtaining residuals or proxies of these errors that can be confronted to such (modeling) assumptions. This checking can be done by graphical devices based on histograms of the PIT (see Dawid 1984 or Diebold et al. 1998 and Gneiting et al. 2007 for further alternatives) and also by formal tests that account for parameter estimation effects, see for instance Bai (2003) in a time series context, Delgado and Stute (2008) or Kheifets (2011), which focus on the independence assumption which is critical in time series applications and, as discussed in Sect. 4.6 of the paper, in the presence of smoothed estimates. However, the properties of such tests under nonparametric estimation are yet unknown.

As stressed in the paper, the impact of bias in the construction of predictive intervals can be important, and when g is nonlinear, special care has to be taken in the design of the predictions. Along this line, one striking point in the general procedures proposed in the setup of Sect. 4 is that of basing predictions for $g(Y)$ on the conditional distribution of Y given x , and not working directly with the distribution of $g(Y)$ given x . Then, it would be more natural to substitute

$$n^{-1} \sum_{i=1}^n g(\bar{D}_{x_f}^{-1}(u_i))$$

given in Eq. (34) of the paper by

$$n^{-1} \sum_{i=1}^n \tilde{G}_{x_f}^{-1}(u_i)$$

where $\tilde{G}_{x_f}(u)$ estimates

$$G_{x_f}(u) = \Pr\{g(Y_f) \leq u | x_f = x\}.$$

The idea of using averages of g evaluations can make sense in the case of a linear-scale model exploiting this structure, which will not hold for $g(Y)$, or in cases where the function g_m in Eq. (3) cannot be found easily for given $g(Y)$ observations, but in the general case, working directly with the PIT based on the conditional distribution of $g(Y)$ given x seems simpler and less bias-prone.

The methods proposed here seem to be able to solve also the problem of producing valid predictive distributions for time series. Thus, a simple semiparametric predictive transformation can be achieved for finite order AR and ARCH models, resembling the mean-scale model, or in a single step through the probability integral transform conditional on a finite number of previous observations, as specification tests for parametric PIT pursue. For instance, to generate h -step ahead AR(p) predictions, x_f has to be replaced by (an extrapolation of) the last p observations available. However, to account for the parameter estimation effect in a linear model, Thombs and Schucany (1990) condition on these p observations in their resampling algorithm, while Ruiz et al. (2004) recreate the estimation effect using unconditional resamples and then generate forecasts conditional on this history as in Cao et al. (1997). In the context of the bootstrap method proposed in this paper enforcing such conditioning seems complicated, though in principle it appears that conditioning is not necessary to provide

valid predictive distributions, but it would be very interesting to know the opinion of the author on this potential problem. A final issue in time series applications is that of checking the validity of a finite autoregressive structure or the need to allow for infinite dimensional conditioning sets by means of sieve methods.

Acknowledgements Research funded by the Spanish Plan Nacional de I+D+i, reference number ECO2012-31748.

References

- Bai J (2003) Testing parametric conditional distributions of dynamic models. *Rev Econ Stat* 85:531–549
- Cao R, Febrero Bande M, Gonzalez Manteiga W, Prada Sanchez JM, Garcia Jurado I (1997) Saving computer time in constructing consistent bootstrap prediction intervals for autoregressive processes. *Commun Stat B* 26:961–978
- Dawid AP (1984) Statistical theory: the prequential approach (with discussion). *J R Stat Soc A* 147:278–292
- Delgado MA, Stute W (2008) Distribution-free specification tests of conditional models. *J Econom* 143:37–55
- Diebold FX, Gunther T, Tay AS (1998) Evaluating density forecasts with applications to finance and management. *Int Econ Rev* 39:863–883
- Gneiting T, Balabdaoui F, Raftery AE (2007) Probabilistic forecasts, calibration and sharpness. *J R Stat Soc B* 69:243–268
- Kheifets I (2011) Specification tests for nonlinear time series models. Preprint, NES
- Ruiz E, Pascual L, Romo J (2004) Bootstrap predictive inference for ARIMA processes. *J Time Ser Anal* 25:449–465
- Thombs LA, Schucany WR (1990) Bootstrap prediction intervals for autoregression. *J Am Stat Assoc* 85:486–492