MONASH University

MONASH BUSINESS SCHOOL

What makes a good prediction interval or probabilistic forecast?

Beinan Xu

(supervisor:Professor Rob Hyndman)

28 May 2018

Introduction

Statistical forecasting models usually provide an estimate of the forecast distribution, or at least a prediction interval, for each forecast horizon.

Scoring Rules

- Interval scoring rules
- Distribution scoring rules

Case study

- Financial data: ASX 200
- M3 datasets

Scoring Rules

- Proper scoring rules provide summary measures of the predictive performance that allow for the joint assessment of calibration and sharpness.
- The scores to be negatively oriented penalties that forecasts wish to minimize.

Interval forecast

- Occasionally, full predictive distributions are difficult to specify, and the forecaster might quote predictive quantiles, such as value at risk in financial applications or prediction intervals only.(Gneiting and Raftery, 2007)
- Interval forecasts is a special case of quantile prediction.

Interval scoring rule

Winkler loss scoring rule is selected to evaluate interval forecasts.

- It is the most commonly used interval forecast loss function.The forecaster is rewarded for narrow prediction intervals, and he or she incurs a penalty, the size of which depends on α, if the observation misses the interval.

- $(1-\alpha) \times 100\%$ is represent the central prediction interval.

Winkler loss scoring rules

$$S_{\alpha}^{int}(l, u; x) = (u - l) + \frac{2}{\alpha}(l - x)1\{x < l\} + \frac{2}{\alpha}(x - u)1\{x > u\}$$

Probabilistic forecast

- A probabilistic forecast takes the form of a predictive probability distribution over future quantities or events of interest. Probabilistic forecasting aims to maximize the sharpness of the predictive distributions, subject to calibration, on the basis of the available information set. (Gneiting and Katzfuss, 2014)

Distribution scoring rules

Three scoring are chosen to evaluate probabilistic forecasts under Gaussian predictive distribution.

Logarithmic score

$$LogS(F, y) = logF(y)$$

Continuous Ranked Probability Score

$$CRPS(F, y) = \int_{-\infty}^{\infty} (F(x) - 1\{y \leq x\})^2 dx$$

Dawid-Sebastianti score

$$DSS(F, y) = \frac{(y - \mu_F)^2}{\sigma_F^2} + 2log\sigma_F$$

Model selection

ARIMA by auto.arima

- This function in R uses a variation of the Hyndman Khandakar algorithm (Hyndman and Khandakar 2008), which combines unit root tests, minimization of the AICc and MLE to obtain an ARIMA model.
- By setting the model and finding the smallest AICc, to obtain the most suitable model.

GARCH by fGarch

- This model has become important in the analysis of time series data, particularly in financial applications when the goal is to analyze and forecast volatility.

Model selection

ETS by ets()

- Information criteria can be used for model selection, AIC and AICc.
- All ETS models are non-stationary, so they cannot be used in financial data.

Random Walk model by rwf()

- The forecasts from a random walk model are equal to the last observation, as future movements are unpredictable, and are equally likely to be up or down.
- Random walk models are widely used for non-stationary data.

Case study one: ASX200

- Data information
  - The raw data comes from YahooFinance (2018), it is the daily data over 10 years period until the beginning of 2018.
  - Features
    - ★ The unconditional distribution is leptokurtic
    - ★ The return series appears to have a constant unconditional mean
    - ★ The volatility of return changes over time and volatility tends to arrive in clusters
- Evaluating by scoring rules
  - Interval forecasts
  - Probabilistic forecasts

Evaluating for interval score

- Models selection
- Evaluation results
  - Interval forecasts by two models
  - Setting different prediction intervals
  - Winkler loss scoring rule

ARIMA model select

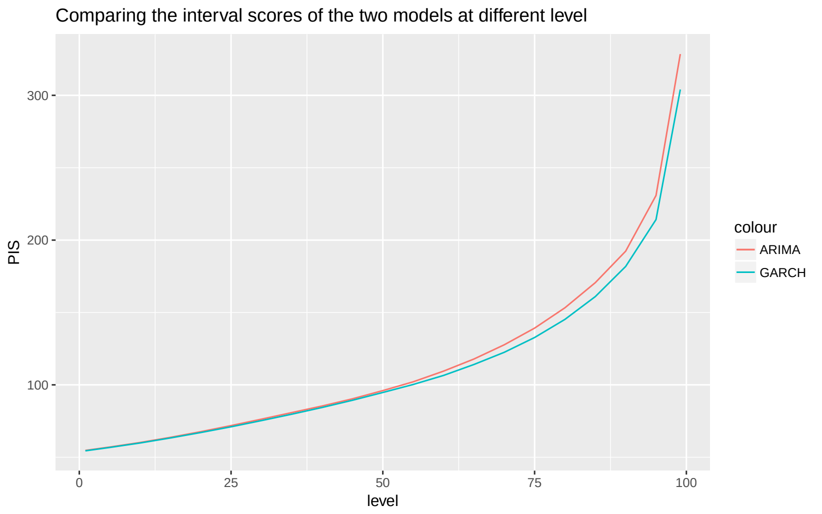
```
## Series: dftrain
## ARIMA(0,0,3) with zero mean
##
## Coefficients:
##           ma1     ma2      ma3
##       -0.0399  0.0064  -0.0505
## s.e.   0.0189  0.0190   0.0194
##
## sigma^2 estimated as 3021:  log likelihood=-15184.98
## AIC=30377.95   AICc=30377.97   BIC=30401.7
```

GARCH model select

Table 1: Garch model select

	AIC	BIC	SIC	HQIC
garch11	10.608	10.623	10.608	10.614
garch12	10.609	10.626	10.609	10.615
garch21	10.609	10.626	10.609	10.616
garch22	10.610	10.629	10.610	10.617
arch1	10.779	10.791	10.779	10.783
arch2	10.729	10.744	10.729	10.734

Evaluating interval forecasts for two models



Evaluation of probabilisitic forecasts

- Same models
- Evaluation results
  - Probabilistic forecasts by two models
  - Three scoring rules

Result of evaluation

Table 2: Scoring Rules for MA model and GARCH model

	CRPS	LogS	DSS
GARCH	20.70	5.10	8.36
ARIMA	21.13	5.14	8.45

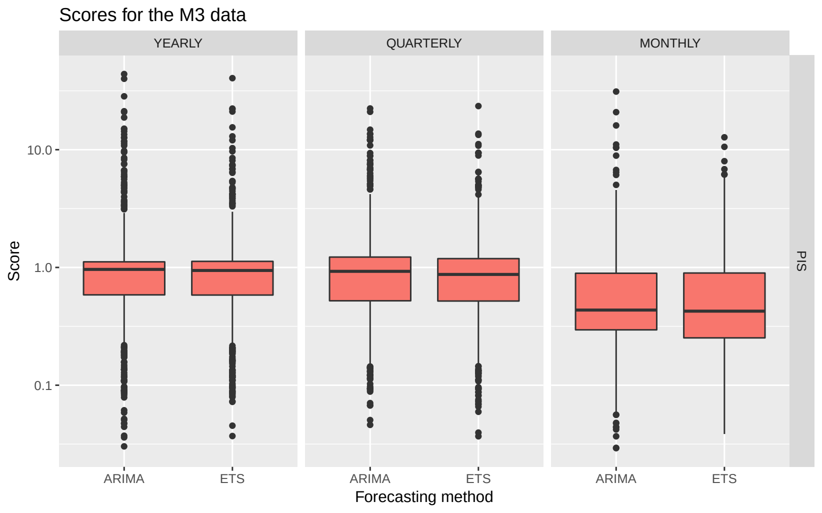
Case study two: M3 data set

- The M3 dataset includes 3003 different type time series, it is from R packages “Mcomp” (Hyndman, 2018).

Evaluating for interval forecast

- Models selection
  - ARIMA and ETS
- Standardization
  - Random walk model
  - Winkler loss scoring rule

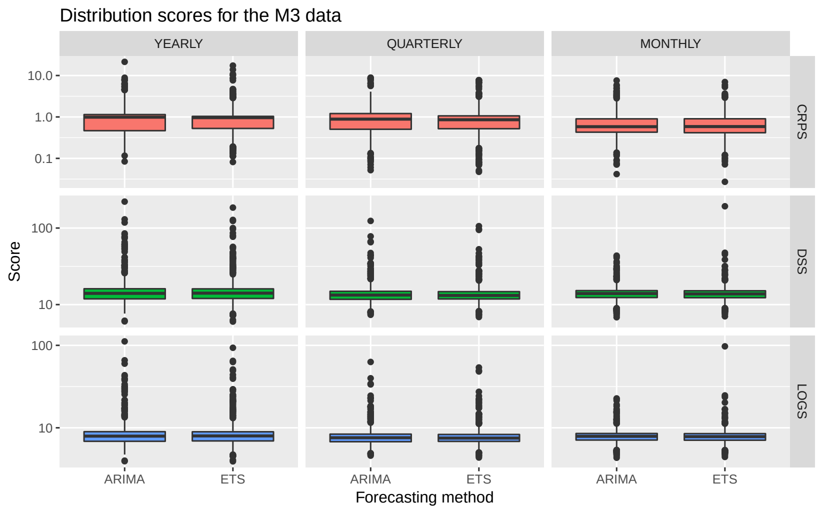
Evaluating interval scores



Evaluating for probabilistic forecasts

- Models selection
  - ARIMA and ETS
- Standardization
  - Random walk model
  - CRPS scoring rules

Evaluating distribution scores



conclusion

- Conclusion

This project has focus on the measures for evaluating prediction intervals and probabilistic forecasts using scoring rules. After using two case study, we compared the evaluation results obtained from a range of statistical models.

- Further discussion
  - The evaluation of probabilistic forecasts of multivariate variables.
  - The evaluation in both parametric and nonparametric settings.

Q & A

Question and Answer

Reference

- Almutaz, Ibrahim ; Ajbar, Abdelhamid ; Khalid, Yasir ; Ali, Emad Desalination, A probabilistic forecast of water demand for a tourist and desalination dependent city: Case of Mecca, Saudi Arabia, 2012*Peer Reviewed Journal* Vol.294, pp.53-59
- Edgar C. Merkle, Mark Steyvers (2013) Choosing a Strictly Proper Scoring Rule. *Decision Analysis* 10(4):292-304.
- Gneiting T, Balabdaoui F, Raftery AE. 2007. Probabilistic forecasts, calibration and sharpness. *J. R. Stat. Soc. B* 67:243-68
- Gneiting, T., & Katzfuss, M. (2014). Probabilistic forecasting. *Annual Review of Statistics and Its Application*, 1(1), 125-151.
- Gneiting T, Raftery AE. 2007. Strictly proper scoring rules, prediction, and estimation. *J. Am. Stat. Assoc.* 102:359-78
- Hersbach, H. (2000), “Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems, *Weather and Forecasting*, 15,559-570.
- Hyndman, R. J. (2018). forecast: Forecasting functions for time series and linear models (R package version 8.3). https://CRAN.R-project.org/package=forecast
- Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: principles and practice. 2nd ed., Melbourne, Australia: OTexts. https://OTexts.org/fpp2/

Reference

- Lowe, Rachel ; Coelho, Caio As ; Barcellos, Christovam ; Carvalho, Marilia Sá ; Catão, Rafael De Castro ; Coelho, Giovanini E ; Ramalho, Walter Massa ; Bailey, Trevor C ; Stephenson, David B ; Rodó, Xavier eLife, 2016, *Peer Reviewed Journal* Vol.5
- Matheson, J. E., and Winkler, R. L. (1976), "Scoring Rules for Continuous Probability Distributions, *Management Science*, 22, 1087-1096.
- Raftery, A. E. (2016). Use and communication of probabilistic forecasts. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 9(6), 397-410.
- Roulston, M. S., and Smith, L. A. (2002), "Evaluating Probabilistic Forecasts Using Information Theory, *Monthly Weather Review*, 130, 1653-1660.