5. High-frequency forecasting

MFE Elective, Forecasting & Financial Time Series
Part I: High-Frequency Forecasting

Ryoko Ito[1]

Department of Economics and Nuffield College, Oxford University
Email: ryoko.ito@economics.ox.ac.uk
Website: www.itoryoko.com

Saïd Business School, Oxford, Trinity 2017

Notations

Notations:

$\widehat{\cdot}_{t,\tau}$: estimated quantities using in-sample data.
(e.g. $\widehat{\nu}$, $\widehat{\lambda}_{t,\tau}$ etc.)

$\widetilde{\cdot}_{t,\tau}$: forecast quantities for out-of-sample period.
(e.g. $\widetilde{\varepsilon}_{t,\tau}$, $\widetilde{\lambda}_{t,\tau}$ etc.)

Out-of-sample period, set of time indices:

$$\begin{aligned} \Psi_H &= \{(t,\tau) \in \{T+1,\dots,T+H\} \times \{1,\dots,I\}\} \\ \Psi_{H,>0} &= \{(t,\tau) \in \{T+1,\dots,T+H\} \times \{1,\dots,I\} : y_{t,\tau} > 0\}. \end{aligned}$$

Producing density forecasts

Recall that we had

$$y_{t,\tau} = \exp(\lambda_{t,\tau})\varepsilon_{t,\tau}, \qquad \varepsilon_{t,\tau} \sim F(\,\cdot\,;\boldsymbol{\theta})$$

where the distribution $F(\,\cdot\,;\boldsymbol{\theta})$ is completely specified and time-invariant.

So we forecast $\lambda_{t,\tau}$ over an out-of-sample period.

Then $\exp(\widetilde{\lambda}_{t,\tau})\varepsilon_{t,\tau}$ with $\varepsilon_{t,\tau} \sim F(\,\cdot\,;\widehat{\boldsymbol{\theta}})$ is our density forecast for $(t,\tau) \in \Psi_H$.

We can produce:

- one-step ahead forecasts, or
- multi-step ahead forecasts.

Producing one-step ahead forecasts

Obtain in-sample parameter estimates using samples up to date T (i.e. $y_{t,\tau}$ for $(t,\tau) \in \Psi_{T,I}$).

Then compute one-step ahead forecast $\widetilde{\lambda}_{T+1,1}$ using in-sample results.

Using a new (out-of-sample) observation, $y_{T+1,1}$, compute $\widetilde{\lambda}_{T+1,2}$.

Repeat the above step at every new (out-of-sample) observation, $y_{T+h,\tau}$ to obtain $\widetilde{\lambda}_{T+h,\tau}$ for $h = 1,\ldots,H$, $\tau = 1,\ldots,I$.

Quality of one-step ahead forecasts

To check the quality of density forecasts, compute the predictive c.d.f. **at each positive observation** as

$$F^*(\widetilde{\varepsilon}_{t,\tau}; \widehat{\boldsymbol{\theta}}^*), \quad \widetilde{\varepsilon}_{t,\tau} = y_{t,\tau} / \exp(\widetilde{\lambda}_{t,\tau})$$

for all $(t,\tau) \in \Psi_{H,>0}$.

The predictive c.d.f. here simply gives the PIT values of forecast standardized observations.

Producing one-step ahead forecasts

Remark:

In HF, it's difficult and costly to re-estimate in-sample parameters (i.e. $\widehat{\boldsymbol{\theta}}, \widehat{\omega}, \widehat{\kappa}_\mu$ etc) at every new observation point, so a choice is not to worry about this.

Then testing the quality of one-step ahead forecasts become **a test of parameter/model stability**.

- Even if in-sample fit of the model is very good, it could be that we are over-complicating the model to get that result. (We could be "overfitting" the model.)
- But if out-of-sample forecasts are also good, the model must be stable even over the out-of-sample period.
- If one-step ahead forecasts are good, the time-invariant parameters of the model must remain at least fairly constant over the out-of-sample period.

Quality of one-step ahead forecasts

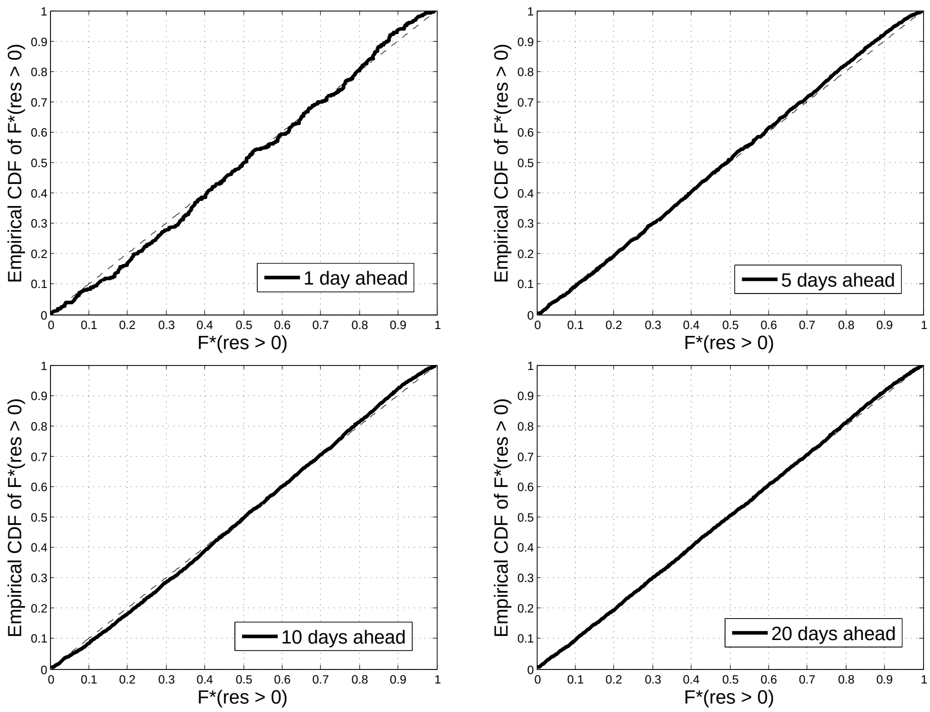

Figure: The empirical c.d.f. of the PIT values (predictive c.d.f.) of one-step ahead $\widetilde{\varepsilon}_{t,\tau}$. Forecast horizons are given in the legend. IBM30s for forecast horizons between 3 - 23 April 2000.

Quality of one-step ahead forecasts

As we have 780 observations per day for IBM30s, $H = 20$ corresponds to 156,000 steps ahead.

The distribution of the PIT values (predictive c.d.f.) is roughly $U[0,1]$ for this extended out-of-sample period.

GB2 appears to capture the empirical distribution of out-of-sample observations remarkably well for an extended out-of-sample period.

Quality of one-step ahead forecasts

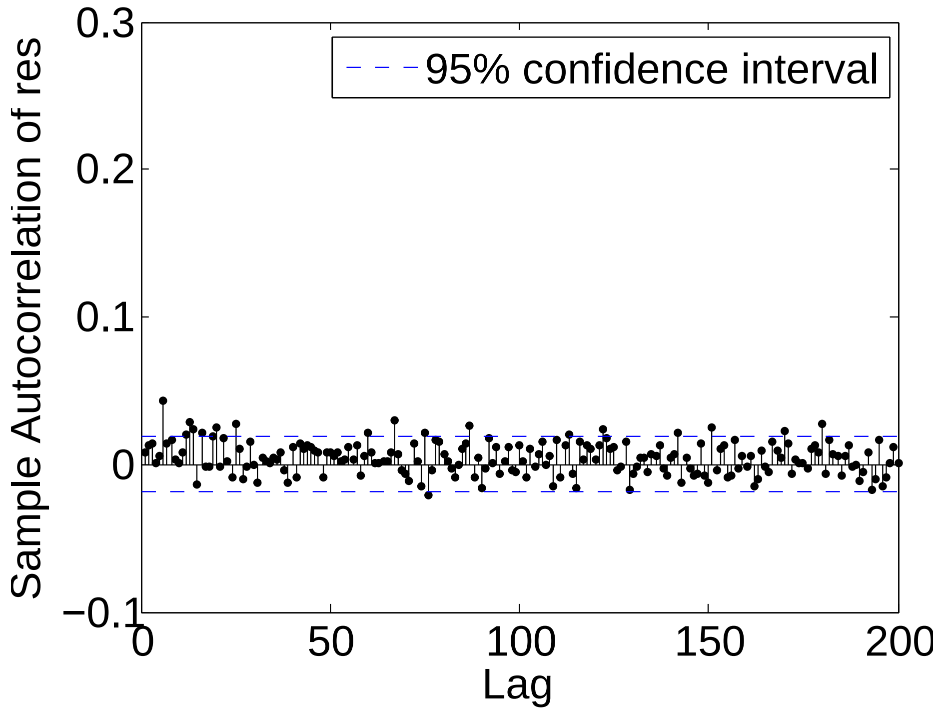

Figure: The sample autocorrelation of one-step ahead standardized observations $\tilde{\varepsilon}_{t,\tau}$. IBM30s for forecast horizons between 3 - 23 April 2000.

Quality of one-step ahead forecasts

The model captures the volume dynamics well over the extended out-of-sample period.

More formally, also check the KS statistics and the Ljung-Box statistics for selected lags.

Producing point forecasts

Note that we might be interested in point (level) forecasts, instead of density forecasts.

Level forecasts are commonly produced using:

- moments, or
- quantiles

of the error quantity, $\varepsilon_{t,\tau}$.

Since the distribution of $\varepsilon_{t,\tau}$ is fully specified, these statistical quantities are given by the theoretical moments or the quantiles of $F(\cdot;\widehat{\boldsymbol{\theta}})$.

Producing point forecasts

Note that the k-th moment of a GB2 distributed random variable X is:

$$\mathbb{E}[X^k] = \mathrm{Beta}(\xi + k/\nu, \zeta - k/\nu)/\mathrm{Beta}(\xi, \zeta),$$

which is well-defined for $-\nu\xi < k < \nu\zeta$. Beta$(\cdot,\cdot)$ here is the beta function.

So, in our case, the k-th moment of the error distribution is:

$$\begin{aligned}\mathbb{E}[\varepsilon_{t,\tau}^k] &= \mathbb{P}(\varepsilon_{t,\tau} = 0)\mathbb{E}[\varepsilon_{t,\tau}^k|\varepsilon_{t,\tau} = 0) + \mathbb{P}(\varepsilon_{t,\tau} > 0)\mathbb{E}[\varepsilon_{t,\tau}^k|\varepsilon_{t,\tau} > 0) \\ &= (1-p)\mathrm{Beta}(\xi + k/\nu, \zeta - k/\nu)/\mathrm{Beta}(\xi, \zeta)\end{aligned}$$

for $-\nu\xi < k < \nu\zeta$.

MUST check the existence of the moment being evaluated.

Producing one-step ahead point forecasts

Then one-step ahead point forecast based on the k-th moment is:

$$\mathbb{E}[y^k_{T+1,1}|\mathcal{F}_{T,I}] = \exp(\widetilde{\lambda}_{T+1,1})\mathbb{E}[\varepsilon^k_{T+1,1}].$$

This quantity can be evaluated recursively by multiplying one-step ahead forecasts $\exp(\widetilde{\lambda}_{T+h,\tau})$ by $(1-p)\mathrm{Beta}(\xi + k/\nu, \zeta - k/\nu)/\mathrm{Beta}(\xi, \zeta)$ for all $h = 1, \ldots, H$ and $\tau = 1, \ldots, I$.

A typical choice for level forecasting is $k = 1$. For volatility forecasting, $k = 2$ may be used.

Producing one-step ahead point forecasts

Denote the u-th quantile of our error distribution $F(\cdot;\boldsymbol{\theta})$ by Q_u. Then, by definition, it satisfies

$$\mathbb{P}(\varepsilon_{t,\tau} \leq Q_u) = F(Q_u;\boldsymbol{\theta}) = u$$

for $u \in [0,1]$, giving $Q_u = F^{-1}(u;\boldsymbol{\theta})$.

In our case, this becomes

$$Q_u = F^{*-1}(u - p;\boldsymbol{\theta}^*)$$

for $p \leq u$, since

- The error distribution $F(\cdot;\boldsymbol{\theta})$ has a probability mass p at zero.
- The positive support of $F(\cdot;\boldsymbol{\theta})$ is governed by $F^*(\cdot;\boldsymbol{\theta}^*)$.

The inverse c.d.f. of GB2 was discussed earlier.

Producing one-step ahead point forecasts

One-step ahead u-th quantile forecast of a given observation is Q_u multiplied by forecast scale parameter $\widetilde{\lambda}_{t,\tau}$ for $(t, \tau) \in \Psi_H$.

For risk analysis (e.g. value-at-risk), u might be a tail probability. For level forecasting, the median quantile (i.e. $u = 0.5$) might be chosen.

Producing multi-step ahead point forecasts

Multi-step ahead point forecast based on the k-th moment is:

$$\mathbb{E}[y^k_{T+h,\tau} | \mathcal{F}_{T,I}] = \mathbb{E}[\varepsilon^k_{T+h,\tau}]\mathbb{E}[e^{k\lambda_{T+h,\tau}} | \mathcal{F}_{T,I}]$$

for $h = 1, \ldots, H$ and $\tau = 1, \ldots, I$. To evaluate the second term on the RHS, the following is useful.

Consider the process $x_t = \phi x_{t-1} + \kappa v_{t-1}$ for $t \in \mathbb{N}$, where v_t is some i.i.d. random variable. The coefficients are $\phi \in (0, 1]$, and $\kappa > 0$, say.

By recursion, we have

$$\begin{aligned} x_{T+h} &= \kappa \sum_{j=1}^{T+h} \phi^{j-1} v_{T+h-j} \\ &= \kappa \sum_{j=1}^{h-1} \phi^{j-1} v_{T+h-j} + \phi^{h-1} x_{T+1}. \end{aligned}$$

Producing multi-step ahead point forecasts

Giving

$$\mathbb{E}[e^{kx_{T+h}}|\mathcal{F}_T] = \prod_{j=1}^{h-1} \mathbb{E}\left[e^{k\kappa\phi^{j-1}v_{T+h-j}}\right] e^{k\phi^{h-1}x_{T+1}}.$$

Here, the moment quantity

$$\mathbb{E}\left[e^{k\kappa\phi^{j-1}v_{T+h-j}}\right]$$

is the moment generating function (m.g.f) of the random variable v_{T+h-j}.

Producing multi-step ahead point forecasts

If a random variable X has the beta distribution with shape parameters $(\alpha, \beta) > \mathbf{0}$, the m.g.f. is

$$\begin{aligned}\mathbb{E}[e^{tX}] &= 1 + \sum_{k=1}^{\infty}\left(\prod_{r=0}^{k-1}\frac{\alpha + r}{\alpha + \beta + r}\right)\frac{t^k}{k!} \\ &= {}_1F_1(\alpha, \alpha+\beta; t)\end{aligned}$$

for any $t \in \mathbb{R}$, where ${}_1F_1(\,\cdot\,,\,\cdot\,;\,\cdot\,)$ is Kummer's confluent hypergeometric function.

Also note that, for a stationary AR(2) process of the form $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \kappa v_{t-1}$, we can obtain its infinite MA representation as follows.

Let z_1 and z_2 be the roots of the characteristic equation $z^2 - \phi_1 z - \phi_2 = 0$. If z_1 and z_2 lie inside the unit circle (i.e. x_t is stationary), then x_t permits the following ∞-MA representation:

$$x_t = \sum_{j=1}^{\infty} \alpha_j v_{t-1-j}, \qquad \alpha_j = \frac{z_1^{j+1} - z_2^{j+1}}{z_1 - z_2}.$$

These facts can be used to analytically derive multi-step ahead point forecast based on the k-th moment:

$$\mathbb{E}[y^k_{T+h,\tau}|\mathcal{F}_{T,I}] = \mathbb{E}[\varepsilon^k_{T+h,\tau}]\mathbb{E}[e^{k\lambda_{T+h,\tau}}|\mathcal{F}_{T,I}]$$
$$= e^{k(\omega+s_{T+h,\tau})}\mathbb{E}[\varepsilon^k_{T+h,\tau}]\mathbb{E}[e^{k(\mu_{T+h,\tau}+\eta^{(1)}_{T+h,\tau}+\eta^{(2)}_{T+h,\tau})}|\mathcal{F}_{T,I}].$$

Steps:

1. Rewrite $\mu_{T+h,\tau}$, $\eta^{(1)}_{T+h,\tau}$, and $\eta^{(2)}_{T+h,\tau}$ in ∞-MA forms; and
2. Evaluate m.g.f.s of the Beta random variable $u_{T+h,\tau}$

for $h = 1, \ldots, H$ and $\tau = 1, \ldots, I$.

Producing multi-step ahead point forecasts

Multi-step ahead forecasts based on quantiles is less straight forward.

We don't know the distribution of $\exp(X)$ when X is a beta distributed random variable.

Comparing forecasts

Use Diebold-Mariano test to compare forecasts of competing models.
Choice of loss function.
In-sample window when updating.