

MONASH BUSINESS SCHOOL

# ETC3550/ETC5550 Applied forecasting

Week 5: Accuracy evaluation



# **Outline**

- 1 Workshop tomorrow
- 2 Residuals
- 3 Forecast distributions
- 4 Forecast evaluation
- 5 Time series cross-validation

# **Outline**

- 1 Workshop tomorrow
- 2 Residuals
- 3 Forecast distributions
- 4 Forecast evaluation
- 5 Time series cross-validation

# **Workshop tomorrow**

The workshop is online tomorrow.

Zoom link on Ed Discussion Forum

# **Outline**

- 1 Workshop tomorrow
- 2 Residuals
- 3 Forecast distributions
- 4 Forecast evaluation
- 5 Time series cross-validation

#### **Fitted values**

- $\hat{y}_{t|t-1}$  is the forecast of  $y_t$  based on observations  $y_1, \dots, y_{t-1}$ .
- We call these "fitted values".
- Sometimes drop the subscript:  $\hat{y}_t \equiv \hat{y}_{t|t-1}$ .
- Often not true forecasts since parameters are estimated on all data.

#### For example:

- $\hat{y}_t = \bar{y}$  for average method.
- $\hat{y}_t = y_{t-1} + (y_T y_1)/(T-1)$  for drift method.

6

# **Forecasting residuals**

**Residuals in forecasting:** difference between observed value and its fitted value:  $e_t = y_t - \hat{y}_{t|t-1}$ .

# Forecasting residuals

**Residuals in forecasting:** difference between observed value and its fitted value:  $e_t = y_t - \hat{y}_{t|t-1}$ .

#### **Assumptions**

- $\{e_t\}$  uncorrelated. If they aren't, then information left in residuals that should be used in computing forecasts.
- $\{e_t\}$  have mean zero. If they don't, then forecasts are biased.

# Forecasting residuals

**Residuals in forecasting:** difference between observed value and its fitted value:  $e_t = y_t - \hat{y}_{t|t-1}$ .

#### **Assumptions**

- $\{e_t\}$  uncorrelated. If they aren't, then information left in residuals that should be used in computing forecasts.
- $\{e_t\}$  have mean zero. If they don't, then forecasts are biased.

#### **Useful properties** (for distributions & prediction intervals)

- $\{e_t\}$  have constant variance.
- $\{e_t\}$  are normally distributed.

## **ACF of residuals**

- We assume that the residuals are white noise (uncorrelated, mean zero, constant variance). If they aren't, then there is information left in the residuals that should be used in computing forecasts.
- So a standard residual diagnostic is to check the ACF of the residuals of a forecasting method.
- We *expect* these to look like white noise.

# **Ljung-Box tests**

 $r_k$  = autocorrelation of residual at lag k

Test whole set of  $r_k$  values simultaneously.

$$Q^* = T(T + 2) \sum_{k=1}^{\ell} (T - k)^{-1} r_k^2$$

 $Q^* = T(T+2) \sum_{k=1}^{\ell} (T-k)^{-1} r_k^2$  where  $\ell$  is max lag being considered and T is number of observations.

- My preferences:  $\ell$  = 10 for non-seasonal data,  $\ell$  = 2m for seasonal data (where *m* is seasonal period).
- If data are WN,  $Q^*$  has  $\chi^2$  distribution with  $\ell$  degrees of freedom.

# **Outline**

- 1 Workshop tomorrow
- 2 Residuals
- 3 Forecast distributions
- 4 Forecast evaluation
- 5 Time series cross-validation

## **Forecast distributions**

- A forecast  $\hat{y}_{T+h|T}$  is (usually) the mean of the conditional distribution  $y_{T+h} \mid y_1, \dots, y_T$ .
- Most time series models produce normally distributed forecasts.
- The forecast distribution describes the probability of observing any future value.

# **Forecast distributions**

Assuming residuals are normal, uncorrelated, sd =  $\hat{\sigma}$ :

**Mean:** 
$$y_{T+h|T} \sim N(\bar{y}, (1+1/T)\hat{\sigma}^2)$$

Naïve: 
$$y_{T+h|T} \sim N(y_T, h\hat{\sigma}^2)$$

Seasonal naïve: 
$$y_{T+h|T} \sim N(y_{T+h-m(k+1)}, (k+1)\hat{\sigma}^2)$$

**Drift:** 
$$y_{T+h|T} \sim N(y_T + \frac{h}{T-1}(y_T - y_1), h^{\frac{T+h}{T}}\hat{\sigma}^2)$$

where k is the integer part of (h-1)/m.

Note that when h=1 and T is large, these all give the same approximate forecast variance:  $\hat{\sigma}^2$ .

## **Prediction intervals**

- A prediction interval gives a region within which we expect  $y_{T+h}$  to lie with a specified probability.
- Assuming forecast errors are normally distributed, then a 95% PI is

$$\hat{y}_{T+h|T} \pm 1.96\hat{\sigma}_h$$

where  $\hat{\sigma}_h$  is the st dev of the *h*-step distribution.

■ When h = 1,  $\hat{\sigma}_h$  can be estimated from the residuals.

# **Prediction intervals**

- Point forecasts are often useless without a measure of uncertainty (such as prediction intervals).
- Prediction intervals require a stochastic model (with random errors, etc).
- For most models, prediction intervals get wider as the forecast horizon increases.
- Use level argument to control coverage.
- Check residual assumptions before believing them.
- Prediction intervals are usually too narrow due to unaccounted uncertainty.

# **Outline**

- 1 Workshop tomorrow
- 2 Residuals
- 3 Forecast distributions
- 4 Forecast evaluation
- 5 Time series cross-validation

# **Training and test sets**



- A model which fits the training data well will not necessarily forecast well.
- A perfect fit can always be obtained by using a model with enough parameters.
- Over-fitting a model to data is just as bad as failing to identify a systematic pattern in the data.
- The test set must not be used for *any* aspect of model development or calculation of forecasts.
- Forecast accuracy is based only on the test set.

#### **Forecast errors**

Forecast "error": the difference between an observed value and its forecast.

$$e_{T+h} = y_{T+h} - \hat{y}_{T+h|T},$$

where the training data is given by  $\{y_1, \ldots, y_T\}$ 

- Unlike residuals, forecast errors on the test set involve multi-step forecasts.
- These are *true* forecast errors as the test data is not used in computing  $\hat{y}_{T+h|T}$ .

# **Measures of forecast accuracy**

```
y_{T+h} = (T+h)th observation, h = 1, ..., H
\hat{y}_{T+h|T} = its forecast based on data up to time T.
 e_{T+h} = y_{T+h} - \hat{y}_{T+h|T}
    MAE = mean(|e_{T+h}|)
                                             RMSE = \sqrt{\text{mean}(e_{T+h}^2)}
    MSE = mean(e_{T+h}^2)
   MAPE = 100mean(|e_{T+h}|/|v_{T+h}|)
```

# **Measures of forecast accuracy**

```
y_{T+h} = (T+h)th observation, h = 1, ..., H
\hat{y}_{T+h|T} = its forecast based on data up to time T.
 e_{T+h} = y_{T+h} - \hat{y}_{T+h|T}
    MAE = mean(|e_{T+h}|)
                                              RMSE = \sqrt{\text{mean}(e_{T+h}^2)}
    MSE = mean(e_{T+h}^2)
   MAPE = 100 \text{mean}(|e_{T+h}|/|y_{T+h}|)
```

- MAE, MSE, RMSE are all scale dependent.
- MAPE is scale independent but is only sensible if  $y_t \gg 0$  for all t, and y has a natural zero.

#### **Scaled Errors**

Proposed by Hyndman and Koehler (IJF, 2006).

For non-seasonal time series, scale errors using naïve forecasts:

$$q_{T+h} = \frac{e_{T+h}}{\frac{1}{T-1} \sum_{t=2}^{T} |y_t - y_{t-1}|}.$$

#### **Scaled Errors**

Proposed by Hyndman and Koehler (IJF, 2006).

■ For non-seasonal time series, scale errors using naïve forecasts:

$$q_{T+h} = \frac{e_{T+h}}{\frac{1}{T-1}\sum_{t=2}^{T}|y_t-y_{t-1}|}.$$

■ For seasonal time series, scale forecast errors using seasonal naïve forecasts:

$$q_{T+h} = \frac{e_{T+h}}{\frac{1}{T-m} \sum_{t=m+1}^{T} |y_t - y_{t-m}|}.$$

## **Scaled errors**

#### **Mean Absolute Scaled Error**

 $\mathsf{MASE} = \mathsf{mean}(|q_{T+h}|)$ 

## **Scaled errors**

#### **Mean Absolute Scaled Error**

$$\mathsf{MASE} = \mathsf{mean}(|q_{T+h}|)$$

#### **Root Mean Squared Scaled Error**

RMSSE = 
$$\sqrt{\text{mean}(q_{T+h}^2)}$$

where

$$q_{T+h}^2 = \frac{e_{T+h}^2}{\frac{1}{T-m} \sum_{t=m+1}^T (y_t - y_{t-m})^2},$$

and we set m = 1 for non-seasonal data.

# **Outline**

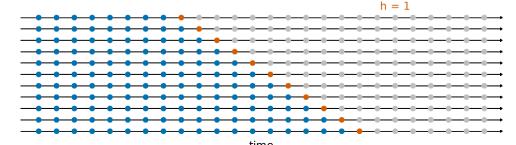
- 1 Workshop tomorrow
- 2 Residuals
- 3 Forecast distributions
- 4 Forecast evaluation
- 5 Time series cross-validation

#### **Traditional evaluation**



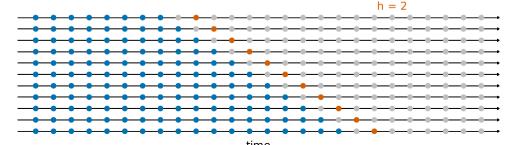
#### **Traditional evaluation**





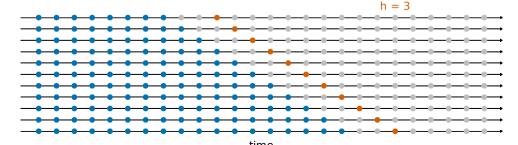
#### **Traditional evaluation**





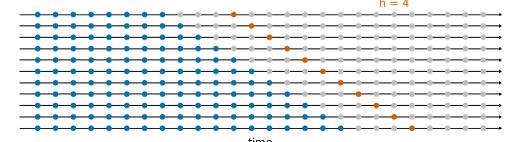
#### **Traditional evaluation**





#### **Traditional evaluation**





#### **Traditional evaluation**



# Time series cross-validation



Also known as "evaluation on a rolling forecasting origin"

# **Workshop tomorrow**

The workshop is online tomorrow.

Zoom link on Ed Discussion Forum