MONASH University

MONASH BUSINESS SCHOOL

ETF3231/5231: Business forecasting

Ch5. The forecasters’ toolbox

https://bf.numbat.space/

AACSB ACCREDITED EQUIS ACCREDITED ASSOCIATION OF MBAS AMBA ACCREDITED

Outline

Outline

1 A tidy forecasting workflow

2 Some simple forecasting methods

3 Residual diagnostics

4 Distributional forecasts and prediction intervals

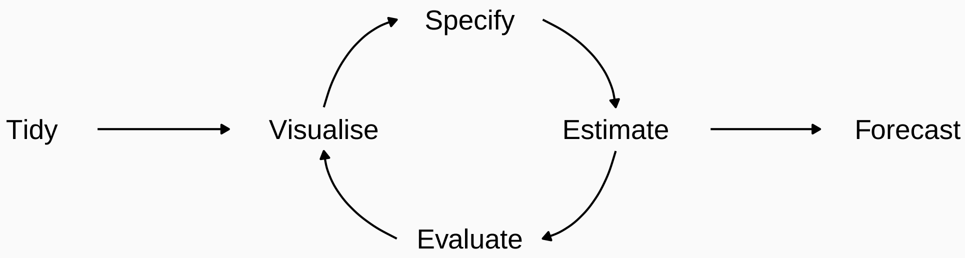
5 Forecasting with transformations

6 Forecasting and decomposition

7 Evaluating forecast accuracy

8 Time series cross-validation

A tidy forecasting workflow



Outline

Some simple forecasting methods - Benchmarks

- `MEAN(y)`: Average method
- `NAIVE(y)`: Naïve method
- `SNAIVE(y ~ lag(m))`: Seasonal naïve method
- `RW(y ~ drift())`: Drift method

Note: distinguish between a method and a model

Model fitting

- The `model()` function trains models to data.
- The `forecast()` function generates forecasts.

`SNAIVE(y ~ lag(m))`: Seasonal naïve method

- Forecasts equal to last value from same season.
- Forecasts: $\hat{y}_{T+h|T} = y_{T+h-m(k+1)}$, where m = seasonal period and k is the integer part of $(h-1)/m$.

Outline

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

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

Useful properties (for distributions & prediction intervals)

3. $\{e_t\}$ have constant variance.
4. $\{e_t\}$ 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.

Portmanteau tests

Consider a whole set of r_k values, and develop a test to see whether the set is significantly different from a zero set.

Ljung-Box test

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

where ℓ is max lag being considered and T is number of observations.

- My preferences: $\ell = 10$ for non-seasonal data, $\ell = 2m$ for seasonal data.
- Better performance, especially in small samples.

Portmanteau tests

- If data are WN, Q^* has χ^2 distribution with $(\ell - K)$ degrees of freedom where K = no. parameters in model.
- When applied to raw data, set $K = 0$.
- `lag` = ℓ, `dof` = K

```r
augment(fit) %>%
  features(.resid, ljung_box, lag=10, dof=0)
```

```
## # A tibble: 1 x 4
##   Symbol .model        lb_stat lb_pvalue
##   <chr>  <chr>           <dbl>     <dbl>
## 1 FB     NAIVE(Close)     12.1     0.276
```

Outline

Forecast distributions

Assuming residuals: have zero mean, are uncorrelated, normal, with variance = $\hat{\sigma}^2$:

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

- 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.
- Point forecasts often useless without a measure of uncertainty (such as prediction intervals).
- Prediction intervals require a stochastic model (with random errors, etc).
- Usually too narrow due to unaccounted uncertainty.

Outline

Outline

Decomposition models

`decomposition_model()` creates a decomposition model

- You must provide a method for forecasting the `season_adjust` series.
- A seasonal naive method is used by default for the `seasonal` components.
- The variances from both the seasonally adjusted and seasonal forecasts are combined.

Outline

Outline

Creating the rolling training sets

There are three main rolling types which can be used.

- Stretch: extends a growing length window with new data.
- Slide: shifts a fixed length window through the data.
- Tile: moves a fixed length window without overlap.

Three functions to roll a tsibble: `stretch_tsibble()`, `slide_tsibble()`, and `tile_tsibble()`.

For time series cross-validation, stretching windows are most commonly used.