Forecasting Models with R

Section 4: Auto Regressive Integrated Moving Average Models

EXFINSIS
www.exfinsis.com

Auto Regressive Integrated Moving Average (ARIMA) Models

- **Auto Regressive Integrated Moving Average (ARIMA) models** specify the conditional mean of a process. They are stochastic or random models specified as sum of deterministic function of lagged dependent or explained variable, stochastic or random lagged model error and stochastic or random error term. (George E.P. Box and Gwilym M. Jenkins. "Time Series Analysis: Forecasting and Control" *San Francisco: Holden-Day*. 1970).

First Order Stationary Time Series

- **First order stationary time series** are stochastic or random processes that have constant mean which doesn´t exhibit trend or seasonality.

First Order Stationary Time Series

- **Augmented Dickey-Fuller Unit Root Test (ADF)** evaluates whether the mean is constant throughout the time series. It can be done for both trend and seasonality. (David A. Dickey and Wayne A. Fuller. "Distribution of the Estimators for Autoregressive Time Series with a Unit Root". *Journal of the American Statistical Association*. 1979)

$$\boldsymbol{trend}: \Delta\hat{\boldsymbol{y}}_{\boldsymbol{t}} = \boldsymbol{c} + \boldsymbol{\beta} * \boldsymbol{t} + \boldsymbol{\gamma}_{\boldsymbol{1}} * \boldsymbol{y}_{\boldsymbol{t-1}} + \boldsymbol{\theta}_{\boldsymbol{1}} * \Delta\boldsymbol{y}_{\boldsymbol{t-1}} + \cdots + \boldsymbol{\theta}_{\boldsymbol{k}} * \Delta\boldsymbol{y}_{\boldsymbol{t-k}}$$

  - *If $\boldsymbol{\gamma_1}$ coefficient p-value < 0.05: first order trend or seasonal stationary time series with 95% confidence.*

First Order Stationary Time Series

- **Time series differentiation** is a de-trending or de-seasonalization procedure which consists of subtracting previous period or season data from current period. Logarithmic differentiation is commonly used as by definition it is approximately equal to the rate of return.

$$\boldsymbol{trend}: \Delta \boldsymbol{y_t} = \boldsymbol{y_t} - \boldsymbol{y_{t-1}} \; ; \; \Delta^2 \boldsymbol{y_t} = \Delta \boldsymbol{y_t} - \Delta \boldsymbol{y_{t-1}}$$

ARIMA Models Specification

- **Auto Regressive Integrated Moving Average (ARIMA) models specification** consists of determining autoregressive (p), integration (d) and moving average (q) order parameters. Integration or differentiation order (d) parameter obtained when time series became first order stationary.

$$\Delta^{d}\hat{y}_t = c + \sum_{i=1}^{p} \alpha_i * \Delta^{d} y_{t-i} + \sum_{i=1}^{q} \beta_i * \Delta^{d} \varepsilon_{t-i}$$

ARIMA Models Specification

- **Normal and partial autocorrelation functions (ACF and PACF)** are used to determine autoregressive (p) and moving average (q) order parameters.

$$\rho(y_t, y_{t-l}) = \frac{\sum_{t=1}^{n-l}(y_t - \mu y_t) * (y_{t-l} - \mu y_{t-l})}{\sqrt{\sum_{t=1}^{n-l}(y_t - \mu y_t)^2 * \sum_{t=1}^{n-l}(y_{t-l} - \mu y_{t-l})^2}}$$

$$\rho_p(y_t, y_{t-l}) = \frac{\rho_i(y_t, y_{t-l})}{-\sqrt{|\rho_i(y_t, y_t) * \rho_i(y_{t-1}, y_{t-l})|}}$$

- ***PARAMETERS: ACF & PCF Functions:***
  - Upper and lower limits $= \pm\frac{1.96}{\sqrt{n}}$
  - *AR(p): ACF (exponential decay), PACF (p significant lags before dropping to zero).*
  - *MA(q): ACF (q significant lags before dropping to zero), PACF (exponential decay).*
  - *ARMA(p,q): ACF (decay after q lag), PACF (decay after p lag).*

Random Walk Model

- **Random walk model** consists of forecasted values equal to previous period's data.
- ***ARIMA (0,1,0) without constant***

$$\hat{y}_t = y_{t-1}$$

Geometric Random Walk Model

- **Geometric random walk model** consists of forecasted values equal to previous period's logarithmic data.
- ***ARIMA (0,1,0) without constant***

$$\hat{y}_t = ln(y_{t-1})$$

Random Walk with Drift Model

- **Random walk with drift model** consists of forecasted values equal to previous period's data plus the arithmetic mean from differentiated time series.
- ***ARIMA (0,1,0) with constant***

$$\hat{y}_t = c + y_{t-1}$$
$$\Delta y_t = y_t - y_{t-1}$$

Geometric Random Walk with Drift Model

- **Geometric random walk with drift model** consists of forecasted values equal to previous period's data multiplied by the exponential of arithmetic mean from differentiated logarithmic time series.
- ***ARIMA (0,1,0) with constant***

$$\hat{y}_t = e^c * y_{t-1}$$

$$\Delta y_t = ln(y_t) - ln(y_{t-1}) = ln\left(\frac{y_t}{y_{t-1}}\right)$$

First Order Autoregressive Model

- **First order autoregressive model** consists of forecasted values equal to deterministic positive simple linear relationship with previous period's data plus a constant intercept.
- ***ARIMA (1,0,0) with constant***

$$\hat{y}_t = c + \alpha_1 * y_{t-1}$$

Differentiated First Order Autoregressive Model

- **Differentiated first order autoregressive model** consists of forecasted difference values equal to deterministic positive simple linear relationship with previous period's differentiated data plus a constant intercept.
- ***ARIMA (1,1,0) with constant***

$$\hat{y}_t = c + y_{t-1} + \alpha_1 * (y_{t-1} - y_{t-2})$$

Brown's Simple Exponential Smoothing Model

- **Brown's simple exponential smoothing model** is appropriate for forecasting data with no trend or seasonal patterns. It consists of forecasted difference values equal to stochastic or random positive simple non-linear relationship with previous period's residuals or forecasting errors.
- ***ARIMA (0,1,1) without constant***

$$\hat{y}_t = y_{t-1} + \beta_1 * \varepsilon_{t-1}$$

Simple Exponential Smoothing with Growth Model

- **Simple exponential smoothing with growth model** consists of forecasted difference values equal to stochastic or random positive simple non-linear relationship with previous period's residuals or forecasting errors plus a constant intercept.
- ***ARIMA (0,1,1) with constant***

$$\hat{y}_t = y_{t-1} + c + \beta_1 * \varepsilon_{t-1}$$

Holt's Linear Trend Model

- **Holt's linear trend model** is appropriate for forecasting data with linear trend pattern. It consists of forecasted difference of difference values equal to stochastic or random positive simple non-linear relationship with previous period's residuals or forecasting errors plus a constant intercept.
- ***Option 1: ARIMA (0,2,1) with constant***

$$\hat{y}_t = c + 2 * y_{t-1} - y_{t-2} + \beta_1 * \varepsilon_{t-1}$$

Holt's Linear Trend Model

- **Holt's linear trend model** is appropriate for forecasting data with linear trend pattern. It consists of forecasted difference of difference values equal to stochastic or random positive multiple non-linear relationships with previous and previous of previous periods' residuals or forecasting errors.
- ***Option 2: ARIMA (0,2,2) without constant***

$$\hat{y}_t = 2 * y_{t-1} - y_{t-2} + \beta_1 * \varepsilon_{t-1} + \beta_2 * \varepsilon_{t-2}$$

Gardner's Additive Damped Trend Model

- **Gardner's additive damped trend model** is appropriate for forecasting data with damped linear trend pattern. It consists of forecasted difference values equal to deterministic positive simple linear relationship with previous difference data, and stochastic or random positive multiple non-linear relationships with previous and previous of previous periods' residuals or forecasting errors.
- ***ARIMA (1,1,2) without constant***

$$\hat{y}_t = y_{t-1} + \alpha_1 * (y_{t-1} - y_{t-2}) + \beta_1 * \varepsilon_{t-1} + \beta_2 * \varepsilon_{t-2}$$

Seasonal Random Walk with Drift Model

- **Seasonal random walk with drift model** consists of forecasted values equal to previous season's data plus the arithmetic mean of seasonally differentiated time series.
- ***ARIMA (0,0,0)x(0,1,0)$_m$ with constant***

$$\hat{y}_t = c + y_{t-m}$$

Seasonal Random Trend Model

- **Seasonal random trend model** consists of forecasted values equal to previous period's data plus previous seasonal difference.
- ***ARIMA (0,1,0)x(0,1,0)$_m$ without constant***

$$\hat{y}_t = y_{t-m} + y_{t-1} - y_{t-m-1}$$

General Seasonal Model

- **General seasonal model** consists of forecasted difference of seasonal difference values equal to stochastic or random positive multiple non-linear relationships with previous period and seasons' residuals or forecasting errors.
- ***ARIMA (0,1,1)x(0,1,1)m without constant***

$$\hat{y}_t = y_{t-m} + y_{t-1} - y_{t-m-1} + \beta_1 * \varepsilon_{t-1} + \gamma_1 * \varepsilon_{t-m} - \beta_1 * \gamma_1 * \varepsilon_{t-m-1}$$

General First Order Auto Regressive Seasonal Model

- **General first order autoregressive seasonal model** consists of forecasted seasonal difference values equal to deterministic positive linear relationship with previous seasonal differences and stochastic or random positive multiple non-linear relationships with previous period and seasons' residuals or forecasting errors.
- ***ARIMA (1,0,1)x(0,1,1)$_m$ with constant***

$$\hat{y}_t = c + y_{t-m} + \alpha_1 * (y_{t-1} - y_{t-m-1}) + \beta_1 * \varepsilon_{t-1} + \gamma_1 * \varepsilon_{t-m} - \beta_1 * \gamma_1 * \varepsilon_{t-m-1}$$

Seasonally Differentiated First Order Autoregressive Model

- **Seasonally differentiated first order autoregressive model** consists of forecasted seasonal difference values equal to positive simple linear relationship with previous period's difference data plus a constant intercept.
- ***ARIMA (1,0,0)x(0,1,0)$_m$ with constant***

$$\hat{y}_t = c + y_{t-m} + \alpha_1 * (y_{t-1} - y_{t-m-1})$$

Holt-Winters Additive Seasonality Model

- **Holt-Winters additive seasonality model** consists of forecasted difference of seasonal difference values equal to stochastic or random positive multiple non-linear relationships with previous periods' residuals or forecasting errors.
- ***ARIMA (0,1,m+1)x(0,1,0)$_m$ without constant***

Model Selection

- **Model selection** is done by evaluating which one minimizes the amount of data or information lost when representing the underlying time series. Therefore, it is a tradeoff between goodness of fit and complexity.
- **Akaike information criterion (AIC)** (Hirotugu Akaike. "A New Look at the Statistical Model Identification". *IEEE Transactions and Automatic Control*. 1974.)

$$AIC = n * ln\left(\frac{SSE}{n}\right) + 2 * (p + q + c + 2)$$

Model Selection

- **Corrected Akaike information criterion (AICc)** (N. Sigura. "Further Analysis of the Data by Akaike's Information Criterion and the Finite Corrections". *Communications in Statistics*. 1978).

$$AIC_c = AIC + \frac{2 * (p + q + c + 2) * (p + q + c + 3)}{n - (p + q + c) - 3}$$

- **Schwarz bayesian information criterion (BIC)** (Gideon E. Schwarz. "Estimating the Dimension of a Model". *Annals of Statistics*. 1978).

$$BIC = n * ln\left(\frac{SSE}{n}\right) + (p + q + c + 2) * ln(n)$$

Residuals White Noise

- **Residuals white noise** or random signal is required for model selection so that these residuals or forecasting errors don't include any information relevant for the prediction.
- **Ljung-Box Q Test** evaluates whether there is any lagged correlation or several order autocorrelation within previously fitted current and previous residuals or forecasting errors. (Greta M. Ljung and George P. Box. "On a Measure of a Lack of Fit in Time Series Models". *Biometrika*. 1978).

$$Q = n * (n + 2) * \sum_{k=1}^{l} \frac{\rho(\varepsilon_t, \varepsilon_{t-l})^2}{n - k}$$

  - *If Q coefficient p-value > 0.05: residuals white noise with 95% confidence*