

Introduction to ARIMA model and how to do point/interval estimates

“ARIMA” stands for Auto-Regressive Integrated Moving Average. An ARIMA model is classified as an “ARIMA(p, d, q)” model, where p is the number of autoregressive terms, d is the number of nonseasonal differences needed for stationarity, and q is the number of lagged forecast error in the prediction equation.

The parameters (p, d, q) are determined usually through Box-Jenkins model.

Suppose y_1, \dots, y_n are the observed values. Let Y denote the d -th difference of y ; i.e.,

$$Y_t = \sum_{i=0}^d (-1)^i \binom{d}{i} y_{t-i}.$$

The general forecasting equation of ARIMA(p, q, d) is given by

$$\hat{Y}_t = \mu + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q},$$

where \hat{Y}_t is the prediction value (of the d -th difference series) at time t and e_t is the error term at time t given by $\hat{Y}_t - Y_t$. For $t > n$, the “observed value” Y_t is replaced by the prediction value \hat{Y}_t and the error term is replaced by 0 (since there is no way to find out the true value y_t). The coefficients $\mu, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$ are determined through fitting historical data. Point estimates (of y_t) are made by using the above equation to determine Y_t and by reversing the difference process.

For the interval estimate, rewrite the ARIMA model into an infinite MA series:

$$\hat{Y}_t = \mu + \psi_1 e_{t-1} + \psi_2 e_{t-2} + \dots$$

This is always achievable when the ARIMA model is stable (or in mathematical terms, all roots of the polynomial $\mu + \phi_1 t + \dots + \phi_p t^p$ are in the open unit disk). The α prediction interval for \hat{Y}_t when $t > n$ is determined by

$$\hat{Y}_t \pm z_\alpha \sigma \sqrt{\left(1 + \sum_{i=1}^{t-n-1} \psi_i^2\right)},$$

where z_α is the $(1 - \alpha)$ critical value for standard normal distribution.

In particular, the 90% prediction interval for \hat{Y}_{n+1} is determined by

$$\hat{Y}_{n+1} \pm 1.28\sigma.$$

The prediction interval for y_t for $t > n$ is again determined by reversing the difference process. In particular, the 90% prediction interval for y_{n+1} is determined by

$$\hat{y}_{n+1} \pm 1.28\sigma.$$