Time Series Analysis

CHEA Makara
5th Year Engineering in Majoring Data Science
Department of Applied Mathematics and Statistics, ITC

November 2024

Acknowledgment

In this lesson, I have summarized the time series process, including additional details of formulas from the book **Time Series Analysis with Applications in R** , *by Jonathan D. Cryer and Kung-Sik Chan.* First of all, I am grateful to my lecturer, **Dr. SIM Tepmony**, for his teaching and valuable advice. Last but not least, I would like to thank my parents for their unwavering financial and emotional support, as well as all my teachers, from primary school to university. I could not have reached this point without your help.
Sincerely

Table of contents

Mean, Variance, and Covariances Overview

Suppose $(X, Y) \sim iid$ crv with distribution $f(x)$ and $f(y)$ respectively
Let X, Y have joint probability density function $f(x, y)$

- The expected value of X is defined as $E(X) = \int_{-\infty}^{\infty} xf(x)dx$
- $E[h(X)] = \int_{-\infty}^{\infty} h(x)f(x)dx$; if $\int_{-\infty}^{\infty} h(x)f(x)dx < \infty$
- Similarly, $E[h(X, Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x, y)f(x, y)dxdy$

$E(X)$ is also called mean and denoted as μ

Mean, Variance, and Covariances Overview

- The Variance of X is defined by $Var(X) = E[(X-\mu)^2]$ Consider

$$
\begin{aligned}
Var(X) &= \int_{-\infty}^{\infty} [x-\mu]^2 f(x)dx \\
Var(X) &= \int_{-\infty}^{\infty} [x^2 - 2x\mu + \mu^2] f(x)dx \\
&= \int_{-\infty}^{\infty} x^2 f(x)dx - 2\mu \int_{-\infty}^{\infty} xf(x) + \mu^2 \int_{-\infty}^{\infty} f(x)dx \\
&= E(X^2) - 2\mu E(X) + \mu^2 \\
&= E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2
\end{aligned}
$$

- Similarly, $Var(X) = E(X^2) - [E(X)]^2$

Mean, Variance, and Covariance Overview

Mean Properties

- $E(aX + b) = aE(X) + b$ for any constant a, b
  Proof:

$$
\begin{aligned}
E(aX + b) &= \int_{-\infty}^{\infty} (ax + b) f(x) dx \\
&= a \int_{-\infty}^{\infty} x f(x) dx + b \int_{-\infty}^{\infty} f(x) dx \\
&= aE(x) + b
\end{aligned}
$$

- $E(X + Y) = E(X) + E(Y)$
- $E(aX + bY) = aE(X) + bE(Y)$

Mean, Variance, and Covariance Overview

If X and Y are independent then for any function$h(x)$ and $g(y)$

- $E[h(X)g(Y)] = E[h(X)]E[(g(Y)]$
  Proof:

$$
\begin{aligned}
E(h(X)g(Y)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x)g(x)f(x,y)dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x)g(x)f_X(x)f_Y(y)dxdy \\
&= \int_{-\infty}^{\infty} h(x)f_X(x)dx \int_{-\infty}^{\infty} g(Y)f_Y(y)dy \\
&= E[h(X)]E[g(Y)]
\end{aligned}
$$

Mean, Variance, and Covariance Overview

The covariance of two random variable X and Y are denoted as $Cov(X, Y)$ and defined by

$$\begin{aligned} Cov(X, Y) &= E[(X - \mu_X)(Y - \mu_Y)] \\ &= E[XY - X\mu_y - Y\mu_x + \mu_X\mu_Y] \\ &= E[XY] - E[X]\mu_y - \mu_X E[Y] + \mu_x\mu_Y \\ &= E[XY] - \mu_X\mu_Y - \mu_X\mu_Y + \mu_X\mu_Y \\ &= E[XY] - \mu_X\mu_Y \end{aligned}$$

If X and Y are independent,

$$Cov(X, Y) = 0$$

Covariance Properties

For any random variable X, Y, Z and constant c

- $Cov(X, X) = Var(X)$
- $Cov(X, Y) = Cov(Y, X)$
- $Cov(cX, Y) = cCov(X, Y)$
- $Cov(X, Y + Z) = Cov(X, Y) + Cov(X, Z)$
- $Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(X_i, Y_j)$

Covariance Properties

Proof:

$$
\begin{aligned}
Cov(X, Y+Z) &= E[X(Y+Z)] - E[X]E[Y+Z] \\
&= E[XY+XZ] - E[X](E[Y]+E[Z]) \\
&= E[XY] + E[XZ] - E[X]E[Y] - E[X]E[Z] \\
&= E[XY] - E[X]E[Y] + E[XZ] - E[X]E[Z] \\
&= Cov(X, Y) + Cov(X, Z)
\end{aligned}
$$

Covariance Properties

We can express the generalize of variance that is obtained from covariance. Consider

$$
\begin{aligned}
Var\left(\sum_{i=1}^{n} X_i\right) &= Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} X_j\right) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} Cov(X_i, X_j) \\
&= \sum_{i=1}^{n} Cov(X_i, X_i) + \sum_{i=1}^{n} \sum_{j \neq i} Cov(X_i, X_j) \\
&= \sum_{i=1}^{n} Var(X_i) + 2\sum_{i=2}^{n} \sum_{j=1}^{i-1} Cov(X_i, X_j) \\
&= \sum_{i=1}^{n} Var(X_i) \ \text{(Because } X_i \text{ is random variable)}
\end{aligned}
$$

Time Series and Stochastic Process

- The sequence of random variables $\{Y_t : t = 0, \pm 1, \pm 2, \pm 3, \cdots\}$ is called a stochastic process and serves as a model for an observed time series.
- For a stochastic process, the mean function us defined by

$$\mu_t = E(Y_t)$$

for $t = 0, \pm 1, \pm 2, \cdots$

Mean, Variance, and Covariances

- The autocovariance function $\gamma_{t,s}$ is defined by

$$\begin{aligned}
\gamma_{t,s} &= Cov(Y_t, Y_s) \text{ for lag } t, s = 0, \pm 1, \pm 2, \cdots \\
&= E[(Y_t - \mu_t)(Y_s - \mu_s)] \\
&= E(Y_t Y_s) - \mu_t \mu_s \text{ (Assume the process is iid)}
\end{aligned}$$

- The autocorrelation function $\rho_{t,s}$ is given by

$$\begin{aligned}
\rho_{t,s} &= Corr(Y_t, Y_s) \\
&= \frac{Cov(Y_t, Y_s)}{\sqrt{Var(Y_t)Var(Y_s)}} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}}
\end{aligned}$$

Random Walk Process

Let $e_1, e_2, \cdots \sim iid$ with zero mean ($\mu = 0$) and variance $= \sigma^2$.The observed time series, $\{Y_t : t = 1, 2, \ldots\}$ is constructed as

$$Y_t = Y_{t-1} + e_t$$

with initial condition $Y_1 = e_1$. e_t is interpreted as the size of steps taken forward or backward along a number line, then Y_t is the position of the **"Random Walker"** at time t.
The alternative random walk process can write

$$Y_t = e_1 + e_2 + \cdots + e_t = \sum_{i=1}^{t} e_i$$

Random Walk Process

Mean, Variance, Covariance, and Correlation can define by

- $\mu_t = E[Y_t] = E[e_1 + e_2, \dots] = E[e_1] + E[e_2] + \cdots + E[e_t] = 0$
- $Var(Y_t) = Var(e_1, e_2, , \dots, e_t) = Var(e_1) + Var(e_2) + \cdots + Var(e_t) = \sigma_e^2 + \cdots + \sigma_e^2 = t\sigma_e^2$
- $\gamma_{t,s} = t\sigma_e^2$ for $1 \leq t \leq s$
- $\rho_{t,s} = \dfrac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}} = \dfrac{t\sigma_e^2}{\sqrt{ts(\sigma_e^2)^2}} = \sqrt{\dfrac{t}{s}}$ for $1 \leq t \leq s$

Stationary

The Stationary concept is follow that $E[Y_t] = E[Y_{t-k}]$ and $Var(Y_t) = Var(Y_{t-k})$ is constant for all t and k over time.
Autocovariance function

$$\begin{aligned}\gamma_{t,s} &= Cov(Y_t, Y_s) = Cov(Y_{t-k}, Y_{s-k}) \\ &= Cov(Y_{t-s}, Y_0) = Cov(Y_0, Y_{s-t}) \\ &= Cov(Y_0, Y_{|t-s|}) \\ \gamma_{t,s} &= \gamma_{0,|t-s|}\end{aligned}$$

For Stationary process, we can simply the notation by letting $k = |t-s|$, then

$$\gamma_k = Cov(Y_t, Y_t - k) \; \& \; \rho_k = Corr(Y_t, Y_{t-k})$$

and also $\rho_k = \dfrac{\gamma_k}{\gamma_0}$

Stationary Process

The General Properties for Stationary Process

- $\gamma_0 = Var(Y_t)$
- $\gamma_k = \gamma_{-k}$
- $|\gamma_k| \leq \gamma_0$
- $\rho_0 = 1$
- $\rho_k = \rho_{-k}$
- $|\rho_k| \leq 1$
- if the process is strictly stationary and has finite variance, the the covariance function must depend only on the time lag.
- A Stochastic process $\{Y_t\}$ is said to be weakly or second-order stationary if
  - The mean function is constant over time
  - $\gamma_{t,t-k} = \gamma_{0,k}$ for all time t and lag k

White Noise

White noise process is defined as a sequence of independent, identically distributed random variable $\{e_t\}$.
White Noise Process consists:

- $E(e_t) = \mu_t$
- $Var(e_t) = \sigma^2$
- $\gamma_k = \begin{cases} Var(e_t), & \text{for } k = 0 \\ 0, & \text{for } k \neq 0 \end{cases}$
- $\rho_k = \begin{cases} 1, & \text{for } k = 0 \\ 0, & \text{for } k \neq 0 \end{cases}$

Table of contents

Estimation of Constant Mean

Suppose the model can be performed as

$$Y_t = \mu + X_t$$

where $E(X_t) = 0$ for all t and we wish to estimate μ with time series $Y_1, Y_2, \ldots, Y_t$. So the average define as

$$\bar{Y} = \frac{1}{n} \sum_{t=1}^{n} Y_t$$

, then

$$E(\bar{Y}) = E\left[\frac{1}{n} \sum_{t=1}^{n} Y_t\right] = \mu$$

Estimation of Constant Mean

We know that $Var(Y) = \gamma_0; \gamma_k = Cov(Y_{j_k}, Y_j)\ ; \gamma_k = \rho_k \gamma_0$ where $k = i - j$

$$
\begin{aligned}
Var(\bar{Y}) &= Var\left(\frac{1}{n}\sum_{t=1}^{n} Y_t\right) \\
&= \frac{1}{n^2} Var(\sum_{t=1}^{n} Y_t) = \frac{1}{n^2} Cov\left(\sum_{i=1}^{n} Y_i, \sum_{j=1}^{n} Y_j\right) \\
&= \frac{1}{n^2}\left[\sum_{i=1}^{n} Var(Y_i) + 2\sum_{i=2}^{n}\sum_{j=1}^{i-1} Cov(Y_i, Y_j)\right] \\
&= \frac{1}{n^2}[n\gamma_0 + 2\sum_{k=1}^{n-1}(n-k)\gamma_k] = \frac{\gamma_0}{n}\left[1 + 2\sum_{k=1}^{n-1}\left(1 - \frac{k}{n}\right)\rho_k\right]
\end{aligned}
$$

Estimation of Constant Mean

For many stationary process, the autocorrelation function decays quickly enough with increasing lags such that,

$$\sum_{k=0}^{\infty} |\rho_k| < \infty$$

then

$$Var(\bar{Y}) \approx \frac{\gamma_0}{n} \left[\sum_{k=-\infty}^{\infty} \rho_k \right] \quad \text{for large } n$$

For **Random Walk Process**, we have

$$Var(\bar{Y}) = \frac{1}{n^2} Var(\sum_{i=1}^{n} \sum_{j=1}^{i} e_j) = \sigma_e^2 (2n+1) \frac{(n+1)}{6n}$$

- **Linear and Quadratic Trends in Time**

For nonconstant mean models, regression analysis is used to estimate the parameters for the trends models.

Consider the deterministic time trend expressed as

$$\mu_t = \beta_0 + \beta_1 t$$

Using Least Square Error Method to estimate the slop and intercept and we get

$$\hat{\beta}_1 = \frac{\sum_{t=1}^{n}(Y_t - \bar{Y})(t - \bar{t})}{\sum_{t=1}^{n}(t - \bar{t})^2}$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{t}$$

• **Cyclical or Seasonal Trends** Some models are modeling and estimating seasonal trends, such as for average monthly temperature data ,etc.
We assume that the observed series can be represented as

$$Y_t = \mu_t + X_t$$

where $E(X_t) = 0$ for all t.
If μ_t is monthly seasonal data, so there are 12 constant parameters, $\beta_1, \beta_2, \ldots, \beta_1 2$ such that,

$$\mu_t = \begin{cases} \beta_1 & \text{for } t = 1, 12, 25, \ldots \\ \beta_2 & \text{for } t = 2, 14, 26 \ldots \\ \vdots & \\ \beta_{12} & \text{for } t = 12, 24, 36, \ldots \end{cases}$$

Regression Methods

- **Cosine Trends**

In some cases, seasonal trends can be modeled economically with cosine curve that incorporate the smooth change expected from one time period to the next while still preserving the seasonality.

Consider the cosine curve

$$\mu_t = \beta \cos(2\pi f t + \Phi)$$

- $\beta(> 0)$ is the amplitude
- f is the frequency
- Φ is the phase of the curve

Regression Methods

Applied trigonometric identity for getting more conveniently to estimate the parameters β and Φ, such that

$$\beta\cos(2\pi ft+\Phi)=\beta_1\cos(2\pi ft)+\beta_2\sin(2\pi ft)$$

where $\beta=\sqrt{\beta_1^2+\beta_2^2}$; $\quad \Phi=\arctan(\beta_2/\beta_1)$,and conversely
$\beta_1=\beta\cos(\Phi)$; $\quad \beta_2=\beta\sin(\Phi)$
The model can be expressed as

$$\mu_t=\beta_0+\beta_1\cos(2\pi ft)+\beta_2\sin(2\pi ft)$$

$beta_0$ is constant term and meaningfully thought of as a cosine with frequency zero.

Interpreting Regression Output

- For the linear trend, suppose we have $\mu_t = \beta_0 + \beta_1 t$. If the $\{X_t\}$ process has a constant variance, then we can estimate the standard deviation of X_t, namely $\sqrt{\gamma_0}$, by the residual standard deviation.

$$s = \sqrt{\frac{1}{n-p}\sum_{t=1}^{n}(Y_t - \hat{\mu}_t)^2}$$

  where p is the number of parameters estimated in μ_t and $n-p$ is the so-call degree of freedom for s.
- R^2 interpretation, it is the square of the sample correlation coefficient between the observed series and the estimated trend.
- t-values or t-ratio are estimated regression coefficient divided by respective standard errors.

Residual Analysis

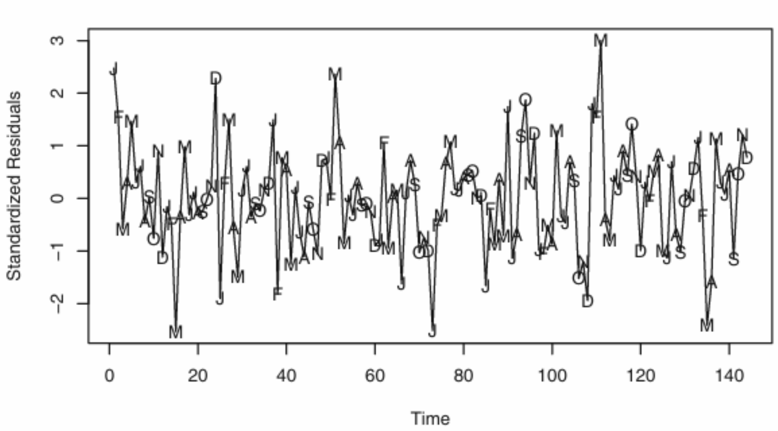
The unobserved stochastic component $\{X_t\}$ can be estimated or predicted by the residual

$$\hat{X}_t = Y_t - \hat{\mu}_t$$

- We call $\hat{X}_t$ is the residual corresponding to the tth observation.
- If the stochastic component is white noise, then the residuals should behave roughly like independent (normal) random variable with zero mean and standard deviation s.

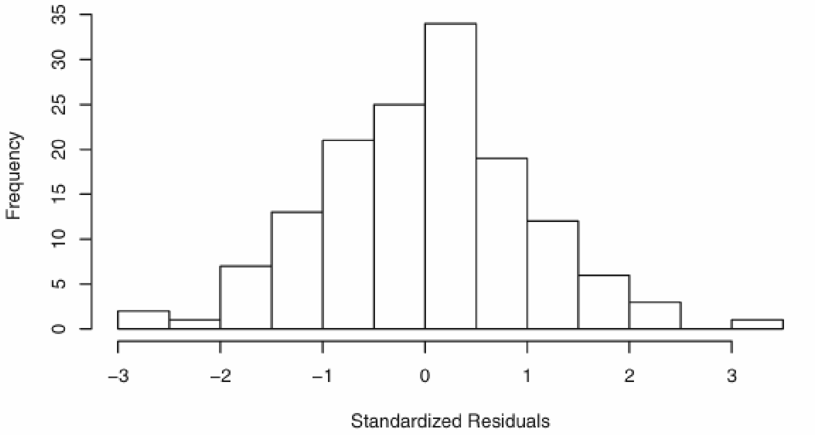
Residual Analysis

Exhibit 3.9 Residuals versus Time with Seasonal Plotting Symbols



Residual Analysis

Exhibit 3.11 Histogram of Standardized Residuals from Seasonal Means Model



Sample Autocorrelation Function

- Consider any sequences of data $Y_1, Y_2, Y_3, \ldots, Y_n$ whether residuals, standardize residuals, original data, or some transformation of data.
- We would like to estimate the autocorrelation function ρ_k for variety lags $k = 1, 2 \ldots$ in order to compute the autocorrelation between pairs k units such that $(Y_1, Y_{1+k}), (Y_2, Y_{k+2}), \ldots (Y_{n-k}, Y_n)$
- we define the sample autocorrelation function r_k at lags k as

$$r_k = \frac{\sum_{t=k+1}^{n}(Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2} \text{ for } k = 1, 2, \ldots$$

Sample Autocorrelation Function

Exhibit 3.13 Sample Autocorrelation of Residuals of Seasonal Means Model

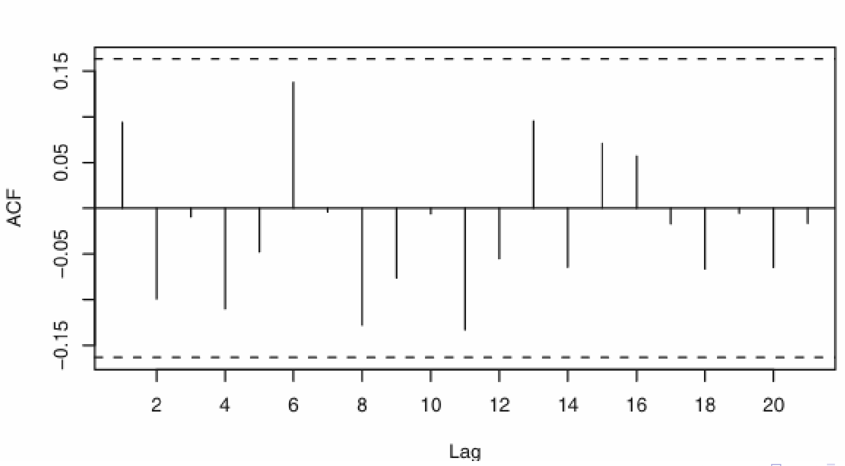


Table of contents

General Linear Processes

- A general linear process, $\{Y_t\}$, is one that can be represented as weighted linear combination of present and past white noise terms as

$$Y_t = e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots$$

- We assume the coefficient $\theta_0 = 1$, it suffices

$$\sum_{i=1}^{\infty} \theta_i^2 < \infty$$

- An important nontrivial which is returned θ's to exponentially decaying sequences such that

$$\theta_i = \phi^j$$

, where $|\phi| < 1$

General Geometric Processes

- Replace $\theta_i = \phi^j$, then the process is updated to

$$Y_t = e_t + \phi e_{t-1} + \phi^2 e_{t-2} + \dots$$

- we have,
  $E(Y_t) = E(e_t + \phi e_{t-1} + \phi^2 e_{t-2} + \dots) = E(e_t) + \phi E(e_{t-1}) + \phi^2 E(e_{t-2}) + \dots = 0$
- Variance,

$$\begin{aligned} Var(Y_t) &= Var(e_t + \phi e_{t-1} + \phi^2 e_{t-2} + \dots) \\ &= Var(e_t) + \phi^2 Var(e_{t-1}) + \phi^4 Var(e_{t-2}) + \dots \\ &= \sigma_e^2 (1 + \phi^2 + \phi^4 + \dots) \\ &= \frac{\sigma_e^2}{1 - \phi^2} \text{ (summation geometric series)} \end{aligned}$$

General Geometric Processes

- Covariance,

$$\begin{aligned}
Cov(Y_t, Y_{t-1}) &= Cov(e_t + \phi e_{t-1} + \phi^2 e_{t-2} + \dots, e_{t-1} + \phi e_{t-2} + \phi^2 e_{t-3} + \dots) \\
&= Cov(\phi e_{t-1}, e_{t-1}) + Cov(\phi^2 e_{t-2}, \phi e_{t-2}) + Cov(\phi^3 e_{t-3}, \phi^2 e_{t-3}) + \dots \\
&= \phi Cov(e_{t-1}, e_{t-1}) + \phi^3 Cov(e_{t-2}, e_{t-2}) + \phi^5 Cov(e_{t-3}, e_{t-3}) + \dots \\
&= \phi\sigma_e^2 + \phi^3\sigma_e^3 + \phi^5\sigma_e^2 + \dots \\
&= \sigma_e^2(\phi + \phi^3 + \phi^5 + \dots) \\
&= \frac{\sigma_e^2\phi}{1-\phi} \quad \text{(summation of infinite geometric sequence)}
\end{aligned}$$

- Correlation, $\rho_{t,t-1} = \dfrac{\gamma_{t,t-1}}{\gamma_0}$ or

$$Corr(Y_t, Y_{t-1}) = \frac{Cov(Y_t, Y_{t-1})}{Var(Y_t)} = (\phi\sigma_e^2\phi)/(1-\phi^2)) \; / \; (\sigma_e^2/(1-\phi^2)) = \phi$$

General Geometric Processes

- Similarly,

$$\begin{aligned} Cov(Y_t, Y_{t-k}) &= Cov(e_t + \phi e_{t-1} + \phi^2 e_{t-2} + \ldots, e_{t-k} + \phi e_{t-k-1} + \phi^2 e_{t-k-2} + \ldots) \\ &= Cov\left(\sum_{j=0}^{\infty} \phi^j e_{t-j}, \sum_{i=0}^{\infty} \phi^i e_{t-k-i}\right) \\ &= \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} \phi^j \phi^i Cov(e_{t-j}, e_{t-k-i}) \quad ; \left[Cov(e_{t-j}, e_{t-k-i}) = \sigma_e^2 \text{ if } j = k+i\right] \\ &= \sum_{i=0}^{\infty} \phi^{2i}\phi^k\sigma_e^2 = \phi^k\sigma_e^2 \sum_{i=0}^{\infty}\phi^{2i} = \frac{\phi^k\sigma_e^2}{1-\phi^2} \quad (|\phi| < 1) \end{aligned}$$

- In addition, $Corr(Y_t, Y_{t-1}) = Cov(Y_t, Y_{t-1})/Var(Y_t) = \phi^k$

Moving Average Processes

- We call a series a moving average of order q and abbreviate the name to $MA(q)$ defined as

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

- The first order moving average model $MA(1)$ is $Y_t = e_t - \theta e_{t-1}$, and clearly $E(Y_t) = 0$ and,

$$\begin{aligned} Var(Y_t) &= Var(e_t - \theta e_{t-1}) \\ &= Var(e_t) + \theta^2 Var(e_{t-1}) = \sigma_e^2(1+\theta^2) \end{aligned}$$

Consider

$$\begin{aligned} Cov(Y_t, Y_{t-1}) &= E[Y_t Y_{t-1}] - E(Y_t)E(Y_{t-1}) = E[Y_t Y_{t-1}] \\ &= E[(e_t - \theta e_{t-1})(e_{t-1} - \theta e_{t-2})] = E[-\theta e_{t-1}^2] \\ &= -\theta E[e_{t-1}^2] = -\theta \sigma_e^2 \end{aligned}$$

Moving Average Processes Cont.

- Observe that,

$$Cov(Y_t, Y_{t-2}) = Cov(e_t - \theta e_{t-1}, e_{t-2} - \theta e_{t-3}) = 0$$

- Similarly,

$$Cov(Y_t, Y_{t-k}) = 0 \text{ for all } k \geq 2$$

- Moreover,

$$\rho_1 = Corr(Y_t, Y_{t-1}) = \frac{Cov(Y_t.Y_{t-1})}{Var(Y_t)} = \frac{-\theta\sigma_e^2}{\sigma_e^2(1+\theta^2)} = -\frac{\theta}{1+\theta^2}$$

and

$$\rho_k = \gamma_k = 0 \text{ for all } k \geq 2$$

Moving Average Processes Cont.

- Consider the moving average process of order 2 , $MA(2)$:

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}$$

- Mean $E(Y_t) = 0$
- Variance,

$$\begin{aligned} Var(Y_t) &= Var(e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}) \\ &= Var(e_t) + \theta_1^2 Var(e_{t-1}) + \theta_2^2 Var(e_{t-2}) = \sigma_e^2(1 + \theta_1^2 + \theta_2^2) \end{aligned}$$

- Covariance,

$$\begin{aligned} \gamma_1 = Cov(Y_t, Y_{t-1}) &= Cov(e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}, e_{t-1} - \theta_1 e_{t-2} - \theta_2 e_{t-3}) \\ &= Cov(-\theta_1 e_{t-1}, e_{t-1}) + Cov(-\theta_2 e_{t-2}, -\theta_1 e_{t-2}) \\ &= (\theta_2 - 1)\theta_1 \sigma_e^2 \end{aligned}$$

Moving Average Processes Cont.

- Consider,

$$\begin{aligned}\gamma_2 = Cov(Y_t, Y_{t-2}) &= Cov(-\theta_2 e_{t-2}, e_{t-2}) \\ &= -\theta_2 \sigma_e^2\end{aligned}$$

- and clearly, $\gamma_k = 0$ for all $k \geq 3$
- Specifically, $\rho_k = \dfrac{\gamma_k}{\gamma_0}$, hence

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{(\theta_1\theta_2 - \theta_1)\sigma_e^2}{(1+\theta_1^2+\theta_2^2)\sigma_e^2} = \frac{-\theta_1 + \theta_1\theta_2}{1+\theta_1^2+\theta_2^2}$$

$$\rho_2 = \frac{\gamma_2}{\gamma_0} = \frac{-\theta_2}{1+\theta_1^2+\theta_2^2}$$

$$\rho_k = \gamma_k = 0, \quad \text{for all } k \geq 3$$

General Moving Average Processes

- For the $MA(q)$ process is defined by

$$Y_t = e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q}$$

- then

$$\gamma_0 = (1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)\sigma_e^2$$
$$\gamma_k = (-\theta_k + \theta_1\theta_{k+1} + \cdots + \theta_{q-k}\theta_q)\sigma_e^2$$

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \begin{cases} \dfrac{-\theta_k + \theta_1\theta_{k+1} + \cdots + \theta_{q-k}\theta_q}{1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2} & \text{for } k = 1, 2, \ldots, q \\ 0 & \text{for } k > q \end{cases}$$

Autoregressive Process

- A pth order Autoregressive denoted as $AR(p)$ and defined by

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t$$

- We assume that e_t is independent of $Y_{t-1}, Y_{t-2}, \ldots$
- The first order of autoregressive process is ,$AR(1)$:

$$Y_t = \phi Y_{t-1} + e_t$$

- Autoregressive process has zero mean, that is $E(Y_t) = 0$
- Variance,

$$\begin{aligned}
\gamma_0 = Var(Y_t) &= Var(\phi Y_{t-1} + e_t) \\
&= \phi^2 Var(Y_{t-1}) + Var(e_t) \\
\gamma_0 &= \phi^2 \gamma_0 + \sigma_e^2 \\
\implies \gamma_0 &= \frac{\sigma_e^2}{1 - \phi^2} \text{ with } \phi^2 < 1 \text{ or } |\phi| < 1
\end{aligned}$$

Autoregressive Processes Cont.

- Autocovariance,

$$\begin{aligned}\gamma_k = Cov(Y_t, Y_{t-k}) &= E[Y_t Y_{t-k}] - E[Y_t]E[Y_{t-k}] \\ &= E[(\phi Y_{t-1} + e_t)(Y_{t-k})] \\ &= E[\phi Y_{t-1} Y_{t-k}] + E[e_t(Y_{t-k})] \\ &= \phi\gamma_{k-1} = \phi^k \gamma_0 \\ \Longrightarrow \gamma_k &= \phi^2 \frac{\sigma_e^2}{1-\phi^2}\end{aligned}$$

- Autocorrelation,

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi^k \ \text{ for } k = 1, 2, 3, \ldots$$

- The process $AR(1) : Y_t = \phi Y_{t-1} + e_t$ is stationary iff $|\phi| < 1$

Autoregressice Processes Cont.

- The second order of autoregressive is denoted and defined by

$$AR(2): Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t$$

- To discuss for stationarity, we introduce the **AR characteristics polynomial** *[Hint: write in terms of e_t then observed]*:

$$\phi(x) = 1 - \phi_1 x - \phi_2 x^2$$

- The corresponding AR characteristic equation, set $\phi(x) = 0$ that is

$$1 - \phi_1 x - \phi_2 x^2 = 0$$

- The stationarity of AR process is hold iff the roots of AR characteristic equation exceed 1 in absolute value or we can say that the roots should lie outside the complex plan.

Autoregressive Processes Cont.

- In the AR(2) process the roots of quadratic characteristic equation are easily found that

$$\frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2}$$

- For AR(2) process is consider to be stationary iff three conditions are satisfied:

$$\phi_1 + \phi_2 < 1, \quad \phi_2 - \phi_1 < 1, \quad |\phi_2| < 1$$

Autoregressive Processes Cont.

- Recall that AR(2) is $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t$
- Mean $E(Y_t) = 0$ and variance

$$\begin{aligned}
\gamma_0 &= Var(\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t) \\
&= \phi_1^2 V(Y_{t-1}) + \phi_2^2 V(Y_{t-2}) + 2\phi_1\phi_2 Cov(Y_{t-1}, Y_{t-2}) + \sigma_e^2 \\
&= \phi_1^2\gamma_0 + \phi_2^2\gamma_0 + 2\phi_1\phi_2\gamma_1 + \sigma_e^2 \\
\implies \gamma_0 &= \left(\frac{1-\phi_2}{1+\phi_2}\right) \frac{\sigma_e^2}{(1-\phi_2)^2 - \phi_1^2}
\end{aligned}$$

- The autocovariance of AR(2) is defined by,

$$\begin{aligned}
\gamma_1 = Cov(Y_t, Y_{t-1}) &= E[(Y_t)(Y_{t-1})] - E[Y_t][Y_{t-1}] \\
&= E[(\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t)(Y_{t-1})] = \phi_1 E[Y_{t-1}^2] + \phi_2 E[Y_{t-1} Y_{t-2}] \\
&= \phi_1\gamma_0 + \phi_2\gamma_1 = \phi_1\gamma_0 + \phi_2\gamma_1 \text{ (Symmetric Law)}
\end{aligned}$$

Autoregressive Processes Cont.

- Similarly, $\gamma_k = \phi_1\gamma_{k-1} + \phi_2\gamma_{k-2}$ for all $k = 1, 2, 3, \ldots$
- The autocorrelation of AR(2) is consider from

$$\begin{aligned}\rho_k &= Corr(Y_t, Y_{t-k}) = \frac{\gamma_k}{\gamma_0} \\ &= \phi_1\rho_{k-1} + \phi_2\rho_{k-2} \text{ for } k = 1, 2, 3, \ldots\end{aligned}$$

  The equation of γ_k and ρ_k above are known as **Yule-Walker equations**
- If $k = 1$ the

$$\begin{aligned}\rho_1 &= \phi_1 + \phi_2\rho_{-1} \\ &= \phi_1 + \phi_2\rho_1 \\ \implies \rho_1 &= \frac{\phi_1}{1-\phi_2}\end{aligned}$$

- If $k = 2$ then $\rho_2 = \dfrac{\phi_2(1-\phi_2) + \phi_1^2}{1-\phi_2}$

Autoregessice Processes Cont.

- For $AR(2)$, there is a characteristic polynomial equation denoted as

$$1-\phi_1 x-\phi_2 x^2=0$$

Suppose G_1 and G_2 are the **Reciprocals roots** of that equation, that is

$$G_1=\frac{\phi_1-\sqrt{\phi_1^2+4\phi_2}}{2};\quad G_2=\frac{\phi_1+\sqrt{\phi_1^2+4\phi_2}}{2}$$

then $\rho_k=\dfrac{(1-G_2^2)G_1^{k+1}-(1-G_1^2)G_2^{k+1}}{(G_1-G_2)(1+G_1G_2)}$

- In case, the roots are satisfied $\phi_1^2+4\phi_2=0$, then

$$\rho_k=\left(1+\frac{1+\phi_2}{1-\phi_2}k\right)\left(\frac{\phi_1}{2}\right)^2 \quad \text{for } k=0,1,2,\ldots$$

The General Autoregressive Process

- Consider the pth order of autoregressive process

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t$$

- with corresponded AR characteristic polynomial

$$1 - \phi_1 x - \phi_2 x^2 - \cdots - \phi_p x^p = 0$$

  This process is stationary if all the roots of the AR characteristics equation each exceed 1 in absolute value (modulus) or all the roots are lie outside the unite circle of complex plan.

- This conditions below is necessary, but not sufficient, those are

$$\begin{cases} \phi_1 + \phi_2 + \cdots + \phi_p < 1 \\ |\phi_p| < 1 \end{cases}$$

The General Autoregressive Process

- Consider the model

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t \qquad (*)$$

is stationary and consists of zero mean. If we multiply this equation by Y_{t-k} and divided by γ_0, consequently we can write the eq. as

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \cdots + \phi_p \rho_{k-p} \text{ for } k \geq 1$$

This equation is known as **Yule-Walker equations**

- If we multiply Y_t to $(*)$, then we obtain

$$\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \cdots + \phi_p \gamma_p + \sigma_e^2$$

We can simplify that $\gamma_0 = \dfrac{\sigma_e^2}{1 - \phi_1 \rho_1 - \phi_2 \rho_2 - \cdots - \phi_p \rho_p}$ (by using $\rho_k = \gamma_k / \gamma_0$)

The Mixed Autoregressive Moving Average

- If the model contains partly of autoregressive and partly of moving average, we obtain a quite general time series model, $\{Y_t\}$, is called $ARMA(p,q)$ process of order p and q such that
$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q}$$
- The **ARMA(1,1)** Model can be perform as
$$Y_t = \phi Y_{t-1} + e_t - \theta e_{t-1} \quad (**)$$
- **Remark:**
$$\begin{aligned} E(e_t Y_t) &= E[e_t(\phi Y_{t-1} + e_t - \theta e_{t-1})] \\ &= E(e_t^2) = \sigma_e^2 \end{aligned}$$

ARMA(1,1) Model Cont.

- and,

$$\begin{aligned} E(e_{t-1}Y_t) &= E[e_{t-1}(\phi Y_{t-1} + e_t - \theta e_{t-1})] \\ &= \phi E[e_{t-1}Y_{t-1}] - \theta E[e^2_{t-1}]] \\ &= \phi\sigma_e^2 - \theta\sigma_e^2 = (\phi - \theta)\sigma_e^2 \end{aligned}$$

- Moreover,

$$E[e_{t-1}Y_{t-1}] = E[e_{t-1}(\theta Y_{t-2} + e_{t-1} - \theta e_{t-2})] = \sigma_e^2$$

- In addition,

$$E[e_{t-1}Y_{t-2}] = E[e_{t-1}(\phi Y_{t-3} + e_{t-2} - \theta Y_{t-3})] = 0$$

- Similarly,

$$E[e_{t-1}Y_{t-k}] = 0, \text{ for all } k \geq 2$$

ARMA(1,1) Model Cont.

- From equation $(**)$, if we multiply Y_{t-k} and take expectation, so we get:

$$E[Y_t Y_{t-k}] = \phi E[Y_{t-1} Y_{t-k}] + E[e_t Y_{t-k}] - \theta E[e_{t-1} Y_{t-k}]$$

  Since $E(Y_t) = 0$, so $E[Y_t Y_{t-k}] = \gamma_k$

  Consequently, the previous equation can be written as

$$\begin{cases} \gamma_0 = \phi\gamma_1 + [1 - \theta(\phi - \theta)]\sigma_e^2 & \text{for } k = 0 \\ \gamma_1 = \phi\gamma_0 - \theta\sigma_e^2 & \text{for } k = 1 \\ \gamma_k = \phi_{\gamma_{k-1}} & \text{for } k \geq 2 \end{cases}$$

ARMA(1,1) Model Cont.

- In cases $k = 0$ and $k = 1$ above, we can deduce the value of variance, namely, γ_0

$$\begin{aligned}
\gamma_0 &= \phi\gamma_1 + [1 - \theta(\phi - \theta)]\sigma_e^2 \\
&= \phi(\phi\gamma_0 - \theta\sigma_e^2) + [1 - \theta(\phi - \theta)]\sigma_e^2 \\
\Longrightarrow \gamma_0 &= \frac{(1 - 2\theta\phi + \theta^2)\sigma_e^2}{1 - \phi^2}
\end{aligned}$$

- autocorrelation function

$$\begin{aligned}
\rho_k = \frac{\gamma_k}{\gamma_0} &= \frac{\phi\gamma_{k-1}(1 - \phi^2)}{(1 - 2\theta\phi + \theta^2)\sigma_e^2} = \frac{\phi^{k-1}\gamma_1(1 - \phi^2)}{(1 - 2\theta\phi + \theta^2)\sigma_e^2} \\
&= \frac{(\phi\gamma_0 - \theta\sigma_e^2)(1 - \phi^2)}{(1 - 2\theta\phi + \theta^2)\sigma_e^2}\phi^{k-1} = \frac{[\phi(1 - 2\phi\theta + \theta^2) - \theta(1 - \phi^2)](1 - \phi^2)}{(1 - \phi^2)(1 - 2\phi\theta + \theta^2)}\phi^{k-1} \\
&= \frac{(1 - \theta\phi)(\phi - \theta)}{1 - 2\theta\phi + \theta^2}\phi^{k-1} \quad \text{for } k \geq 1
\end{aligned}$$

Invertibility

- Consider an $MA(1)$ model:

$$\begin{aligned} Y_t &= e_t - \theta e_{t-1} \\ \Longrightarrow e_t &= Y_t + \theta e_{t-1} \\ e_t &= Y_t + \theta(Y_{t-1} + \theta e_{t-2}) \\ e_t &= Y_t + \theta Y_{t-1} + \theta^2 Y_{t-2} \\ e_t &= Y_t + \theta Y_{t-1} + \theta^2 Y_{t-2} + \dots \\ \text{or} \quad Y_t &= (-\theta Y_{t-1} - \theta^2 Y_{t-1} - \theta^3 Y_{t-3} - \dots) + e_t \end{aligned}$$

- If $|\theta| < 1$, we can see that MA(1) model can be inverted into a infinite-order autoregressive model.
- We say that the MA(1) model is invertible if and only if $|\theta| < 1$
- For a general $ARMA(p, q)$ model, we require both stationary and invertibility.

Table of contents

Stationary Through Differencing

- Consider for $AR(1)$ model
$$Y_t = \phi Y_{t-1} + e_t$$
, which assumed e_t is independent with $Y_{t-1}, Y_{t-2}, \ldots$
- This $AR(1)$ is stationary if $|\phi| < 1$
- If $|\phi| \geq 1$ this process will be an exponential growth or explosive behavior.
- In case $\phi = 1$ then,
$$Y_t = Y_{t-1} + e_t$$
alternative way, we can rewrite this as
$$\nabla Y_t = e_t$$
where $\nabla Y_t = Y_t - Y_{t-1}$ is the first difference of Y_t.

ARIMA Models

- A time series $\{Y_t\}$ is said to follow an integrated autoregressive moving average model if the dth difference $W_t = \nabla^d Y_t$ is a stationary ARMA process.
- If $\{W_t\}$ follows an ARMA(p, q) model, we say that $\{Y_t\}$ is an ARIMA(p, d, q) process.
- If the process contains no autoregressive terms, we call it an integrated moving average, $IMA(d, q)$
- If the process consists no moving average terms, we denote the model, $ARI(p, d)$
- For the ARIMA$(p, 1, q)$, we assume that the process start at some time point $t = -m$, where $-m$ is the earlier than time $t = 1$, at which point we first observed the series. For convenience, we take $Y_t = 0$ for $t < -m$. The difference equation $Y_t - Y_{t-1} = W_t$ can be solved by summing both side from $t = -m$ to $t = t$ to get the representation

$$Y_t = \sum_{j=-m}^{t} W_j$$

The IMA(1,1) Model

- The IMA(1,1) model satisfactorily represents numerous time series can perform model as

$$Y_t = Y_{t-1} + e_t - \theta e_{t-1}$$

  and in this case $W_t = e_t - \theta e_{t-1}$, where $\nabla Y_t = W_t$
  So, after rearrangement, we can rewrite the equation into

$$Y_t = e_t + (1-\theta)e_{t-1} + (1-\theta)e_{t-2} + \cdots + (1-\theta)e_{-m} - \theta e_{-m-1}$$

- Since we are assuming that $-m < 1$ and $t > 0$ and we may usefully think of Y_t as mostly an equally weighted accumulation of a large number of white noise values.
- we can see that $Var(Y_t) = [1 + \theta^2 + (1-\theta)^2(t+m)]\sigma_e^2$ and

$$Corr(Y_t, Y_{t-k}) = \frac{1 - \theta + \theta^2 + (1+\theta^2)(t+m-k)}{\sqrt{Var(Y_t)Var(Y_{t-k})}}$$

ARI(1,1) Model

- The ARI(1,1) model can perform as

$$Y_t - Y_{t-1} = \phi(Y_{t-1} - Y_{t-2}) + e_t$$

or

$$Y_t = (1 + \phi)Y_{t-1} - \phi Y_{t-2} + e_t$$

Constant Terms in ARIMA Model

- For an $ARIMA(p, d, q)$ model, $\nabla^d Y_t = W_t$ is a stationary ARAM(p, q) process.
- Nonzero constant mean, μ, in a stationary ARMA model $\{W_t\}$ can be accommodated in either of two ways,

$$W_t - \mu = \phi_1(W_{t-1} - \mu) + \phi_2(W_{t-2} - \mu) + \cdots + \phi_p(W_{t-p} - \mu) + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

Alternative way, we can introduce a constant term θ_0 such that,

$$W_t = \theta_0 + \phi_1 W_{t-1} + \phi_2 W_{t-2} + \cdots + \phi_p W_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

Similarly,

$$E[W_t] = E\left[\theta_0 + \phi_1 W_{t-1} + \cdots + \phi_p W_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}\right]$$
$$\implies \mu = \theta_0 + (\phi_1 + + \cdots + \phi_p)\mu$$

Constant Terms in ARIMA Model Cont.

- From the previous equation, we can see that

$$\mu = \frac{\theta_0}{1 - \phi_1 - \phi_2 - \cdots - \phi_p}$$

Alternative,

$$\theta_0 = (1 - \phi_1 - \phi_2 - \cdots - \phi_p)\mu$$

- What will be the effect of a nonzero mean for W_t on the undifferenced series Y_t ? Consider case IMA$(1,1)$, the eqaution is statisfy

$$W_t = \theta_0 + e_t - \theta e_{t-1}$$

We can find that also:

$$Y_t = e_t + (1-\theta)e_{t-1} + \cdots + (1-\theta)e_{-m} - \theta e_{-m-1} + (t+m+1)\theta_0$$

We see that this equations is added linear deterministic time trend $(t+m+1)\theta_0$ with slop θ_0

Constant Terms in ARIMA Models Cont.

- An equivalent representation of the process would then be

$$Y_t = Y_t' + \beta_0 + \beta_1 t$$

  where Y_1' is an IMA$(1,1)$ series with $E(\nabla Y_t') = 0$ and $E(\nabla Y_t) = \beta_1$.
- For a general ARIMA(p,d,q) model where $E(\nabla^d Y_t) \neq 0$, it can be argued that $Y_t = Y_t' + \mu$, where μ_t is a deterministic polynomial of degree d and Y_t' is ARIMA(p,d,q) with $E[Y_t'] = 0$.
- With $d = 2$ and $\theta_0 \neq 0$, a quadratic trend would be implied.

Other Transformations Log-Transformed

- Logarithm transformation is also a useful method for series where increased dispersion seems to be associated with higher levels of the series.
- Logarithms transformation is applied if the standard deviation of the series is proportional to the level of series. This process will produce a series with approximately constant variance over time.
- Moreover, if the level of the series is changing roughly exponentially, then log-transformed series will exhibit a linear time trend.
- Suppose $Y_t > 0$ for all t and that

$$E[Y_t] = \mu_t \quad \text{and} \quad \sqrt{Var(Y_t)} = \mu_t \sigma$$

  Then

$$E[\log(Y_t)] \approx \log(\mu_t) \quad \text{and} \quad Var(\log(Y_t)) \approx \sigma^2$$

Other Transformation Log-Transformed Cont.

Proof: Using Taylor's series of logarithms $\log(1-x) \approx -(x + \frac{1}{2}x^2 + \cdots + \frac{1}{n}x^n) + o(x^n)$
Consider,

$$\log(1-(1-\frac{Y_t}{\mu_t})) \approx -(1-\frac{Y_t}{\mu_t})$$
$$\log(\frac{Y_t}{\mu_t}) = \frac{Y_t}{\mu_t} - 1 = \frac{Y_t - \mu_t}{\mu_t}$$
$$\log(Y_t) = \log(\mu_t) + \frac{Y_t - \mu_t}{\mu_t}$$
$$\Longrightarrow E[\log(Y_t))] = E[\log(\mu_t)] + \frac{1}{\mu_t}E[Y_t] - 1$$
$$\Longrightarrow E[\log(Y_t)] = \log(\mu_t)$$

for variance, $$Var(\log(Y_t)) = \frac{1}{\mu_t^2}Var(Y_t) = \sigma^2 \quad (\sqrt{Var(Y_t)} = \mu_t\sigma)$$ $\square$

Other Transformation - Power Transformations

- The power transformations was introduced by Box and Cox (1964)
- For given value of parameter λ, transformation is defined by

$$g(x) = \begin{cases} \dfrac{x^\lambda - 1}{\lambda} & \text{for } \lambda \neq 0 \\ \log(x) & \text{for } \lambda = 0 \end{cases}$$

- when $\lambda \to 0, (x^\lambda - 1)/\lambda \to \log(x)$, and when $\lambda = 1/2$, it produces a square root transformations useful with **Poisson-like** data, and $\lambda = -1$ corresponds to a reciprocal transformation.
- The power transformation applies only to positive data values. If some of the values are negative or zero, a positive constant may be added to all of the values to make them all positive before doing the power transformation.

The Backshift Operator

- The backshift operator, denoted B, operates on time index of a series and shifts time back one time unit to form a new series.
- In particular,

$$BY_t = Y_{t-1}$$

Consider,

$$\begin{aligned} Y_{t-1} &= BY_t \\ Y_{t-2} &= BY_{t-1} = B(BY_t) = B^2 Y_t \\ &\cdots \\ Y_{t-k} &= B^k Y_t, \quad \text{for any positive integer } k \end{aligned}$$

- The backshift operator is linear since for any constants a, b and c ans series Y_t and X_t

$$B(aY_t + bX_t + c) = aBY_t + bX_t + c$$

The Backshift Operator Cont.

- Consider for MA(1) model, in terms of B for noise

$$Y_t = e_t - \theta e_{t-1} = e_t - \theta B e_t = (1 - \theta B) e_t = \Theta(B) e_t$$

where $\Theta(B)$ is the MA characteristic polynomial "evaluated" at B

- In general, for $MA(q)$ model, we can write

$$\begin{aligned} Y_t &= e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \\ &= e_t - \theta_1 B e_t - \theta_2 B^2 e_t - \cdots - \theta_q B^q e_t \\ &= \left(1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q\right) e_t \\ \iff Y_t &= \Theta(B) e_t \end{aligned}$$

The Backshift Operator Cont.

- For autoregressive models AR(p), we consider

$$
\begin{aligned}
Y_t &= \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t \\
\Longrightarrow e_t &= Y_t - \phi_1 Y_{t-1} - \cdots - \phi_p Y_{t-p} \\
&= Y_t - \phi_1 B Y_t - \phi_2 B^2 Y_t - \cdots - \phi_p B^p Y_t \\
&= (Y_t - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) Y_t \\
\Longleftrightarrow e_t &= \Phi(B) Y_t
\end{aligned}
$$

- For ARMA(p, q) model can be written compactly as

$$\Phi(B) y_t = \Theta(B) e_t$$

Remark: *ARMA model is stationary, if $\Phi(X) = 0$ is the stationary condition. It mean that all the roots of characteristics polynomial are lie outside the unite cycle and $\Theta(X) = 0$ is invertibility condition if all the roots lie outside the unite circle (all roots are exceed 1)*

The Backshift Operator Cont.

- We can also express the differencing in terms of B,

$$
\begin{aligned}
\nabla Y_t &= Y_t - Y_{t-1} = Y_t - BY_t = (1-B)Y_t \\
\nabla^2 Y_t &= \nabla(\nabla Y_t) = \nabla Y_t - \nabla Y_{t-1} \\
&= (1-B)Y_t - (Y_{t-1} - Y_{t-2}) = (1 - 2B + B^2)Y_t \\
&= (1-B)^2 Y_t \\
&\cdots \\
\nabla^d Y_t &= (1-B)^d Y_t
\end{aligned}
$$

- The general ARIMA(p, d, q) model is expressed concisely as

$$\Phi(B)(1-B)^d Y_t = \Theta(B)e_t$$

Stationary case are similar to ARMA model,however, it is implied the dth differencing.

Table of contents

Properties of the Sample Autocorrelation Function

- Recall that ,the sample autocorrelation function of observed series $\{Y_t : t = 1, 2, \ldots, n\}$ is

$$r_k = \frac{\sum t = k + 1^n (Y - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^{n} (Y_t - \bar{Y})^2} \quad \text{for } k = 1, 2, ..$$

Properties of the Sample Autocorrelation Function

- We suppose that $Y_t = \mu + \sum_{j=0}^{\infty} \phi_j e_{t-j}$ (ARMA Model)
- where $e_t \sim iid\ f(Y)$ with $\mu = 0$ and $Var(Y_t) = \sigma_e^2$, it is satisfied

$$\sum_{j=0}^{\infty} |\phi_j| < \infty \ \ \& \ \ \sum_{j=0}^{\infty} j\phi_j^2 < \infty \longrightarrow \text{ Stationary for ARMA Model}$$

- As n tends to ∞ the variance c_{jj} and covariance c_{ij} is defined as

$$c_{ij} = \sum_{k=-\infty}^{\infty} (\rho_{k+i}\rho_{k+j} + \rho_{k-i}\rho_{k+j} - 2\rho_i\rho_k\rho_{k+j} - 2\rho_j\rho_k\rho_{k+i} + 2\rho_i\rho_j\rho_k^2)$$

Properties of the Sample Autocorrelation Function Cont.

- For AR(1) process with $\rho_k = \phi^k$ for $k > 0$. then

$$Var(Y_t) \approx \frac{1}{n}\left[\frac{(1+\phi^2)(1-\phi^{2k})(1-\phi^{2k})}{1-\phi^2} - 2k\phi^{2k}\right]$$

- In particular,

$$Var(r_1) \approx \frac{1-\phi^2}{n}$$

- Generally, for large lags,

$$Var(r_k) \approx \frac{1}{n}\left[\frac{1+\phi^2}{1-\phi^2}\right] \qquad \text{for large } k$$

- For $0 < i < j$

$$c_{ij} = \frac{(\phi^{j-i} - \phi^{j+i})}{1-\phi^2} + (j-i)\phi^{j-i} - (j+i)\phi^{j+i} \qquad \text{and} \qquad Corr(r_1, r_2) \approx 2\phi\sqrt{\frac{1-\phi^2}{1+2\phi^2-3\phi^4}}$$

Properties of the Sample Autocorrelation Function Cont.

- For MA(1), We get

$$c_{11} = 1 - 3\rho_1^2 + 4\rho_1^4 \quad \text{and} \quad c_{kk} = 1 + 2\rho_1^2 \quad for\, k > 1$$

- Furthermore,

$$c_{12} = 2\rho_1(1 - \rho_1^2)$$

- For more general $MA(q)$ process.

$$c_{kk} = 1 + 2\sum_{j=1}^{q} \rho_j^2 \quad \text{for } k > q$$

- Moreover,

$$Var(r_k) = \frac{1}{n}\left[1 + 2\sum_{j=1}^{q} \rho_j^2\right] \quad \text{for } k > q$$

The Partial and Extended Autocorrelation Functions

- If $\{Y_t\}$ is a normally distributed time series, we denote ϕ_{kk} (the coefficient of partial autocorrelation at lag k) and defined as

$$\phi_{kk} = Corr(Y_t, Y_{t-k} | Y_{t-1}, Y_{t-2}, \ldots, Y_{t-2}, Y_{t-k+1}$$

- Consider Y_t is a linear function of the intervening variables $Y_{t-1}, Y_{t-2}, .., Y_{t-k+1}$, says, $\beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_{k-1} Y_{t-k+1}$. The partial autocorrelation function at lag k is then defined to be the correlation between the prediction errors; that is,

$$\begin{aligned}\phi_{kk} = Corr(&Y_t - \beta_1 Y_{t-1} - \beta_2 Y_{t-2} - \cdots - \beta_{k-1} Y_{t-2},\\ &Y_{t-k} - \beta_1 Y_{t-k+1} - \beta_2 Y_{t-k} - \cdots - \beta_{k-1} Y_{t-1})\end{aligned}$$

- For $\phi_{11} = 1$
- and $\phi_{22} = \dfrac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$
- In case AR(1), model $\phi_{22} = \dfrac{\phi^2 - \phi^2}{1 - \phi^2} = 0$, (Recall that $\rho_k = \phi^k$)

PACF Cont.

- For MA(1) case, the equation can be shown as

$$\phi_{kk} = -\frac{\theta^k(1-\theta^2)}{1-\theta^{2(k+1)}} \quad \text{for } k \geq 1$$

- The behavior of ACF and PACF can be shown as

Exhibit 6.3 General Behavior of the ACF and PACF for ARMA Models

	AR(p)	MA(q)	ARMA(p,q), p>0, and q>0
ACF	Tails off	Cuts off after lag q	Tails off
PACF	Cuts off after lag p	Tails off	Tails off

Figure: ACF and PACF for Model Specification *Chapter 6, Model Specification, Page 116*

The Extended Autocorrelation Function

- The EACF method uses the fact that if the AR part of a mixed ARMA model is known, "filtering out" the autoregression from the observed time series results in a pure MA process that enjoys the cutoff property in its ACF.
- In the cases of ARMA(1,1) model, then $Y_t = \phi Y_{t-1} + e_t - \theta e_{t-1}$
- The ECAF of ARMA(1,1) is illustrated as

Exhibit 6.4 Theoretical Extended ACF (EACF) for an ARMA(1,1) Model

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	x	x	x	x	x	x	x	x	x	x	x	x	x
1	x	0*	0	0	0	0	0	0	0	0	0	0	0	0
2	x	x	0	0	0	0	0	0	0	0	0	0	0	0
3	x	x	x	0	0	0	0	0	0	0	0	0	0	0
4	x	x	x	x	0	0	0	0	0	0	0	0	0	0
5	x	x	x	x	x	0	0	0	0	0	0	0	0	0
6	x	x	x	x	x	x	0	0	0	0	0	0	0	0
7	x	x	x	x	x	x	x	0	0	0	0	0	0	0

Specification of Some Simulated Time Series

Exhibit 6.5 **Sample Autocorrelation of an MA(1) Process with θ = 0.9**

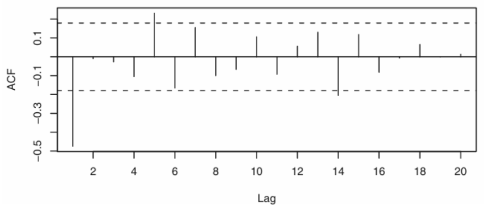
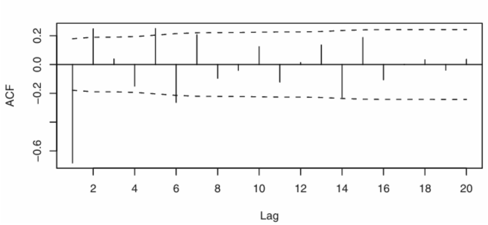


Exhibit 6.9 **Alternative Bounds for the Sample ACF for the MA(2) Process**



Specification of Some Simulated Time Series Cont.

Exhibit 6.10 Sample ACF for an AR(1) Process with $\phi = 0.9$

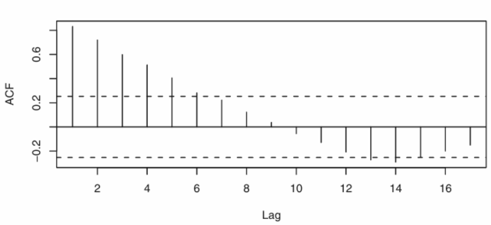
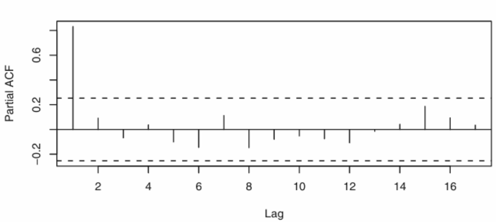


Exhibit 6.11 Sample Partial ACF for an AR(1) Process with $\phi = 0.9$



Specification of Some Simulated Time Series Cont.

Exhibit 6.12 Sample ACF for an AR(2) Process with $\phi_1 = 1.5$ and $\phi_2 = -0.75$

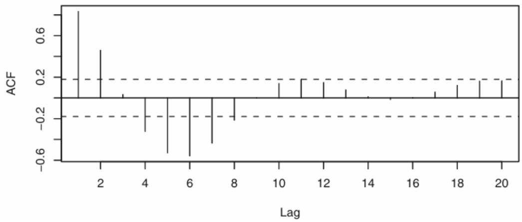
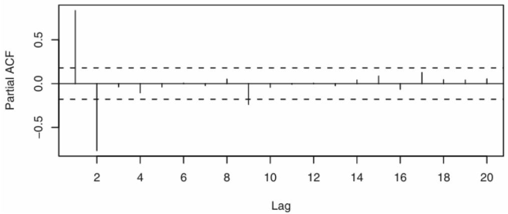


Exhibit 6.13 Sample PACF for an AR(2) Process with $\phi_1 = 1.5$ and $\phi_2 = -0.75$



Specification of Some Simulated Time Series Cont.

Exhibit 6.14 Simulated ARMA(1,1) Series with $\phi = 0.6$ and $\theta = -0.3$.

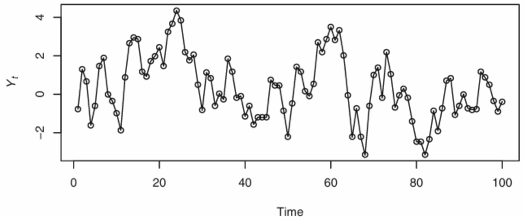
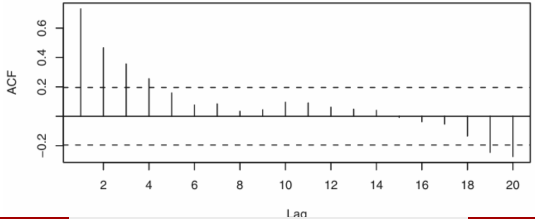


Exhibit 6.15 Sample ACF for Simulated ARMA(1,1) Series



Specification of Some Simulated Time Series Cont.

Exhibit 6.16 Sample PACF for Simulated ARMA(1,1) Series

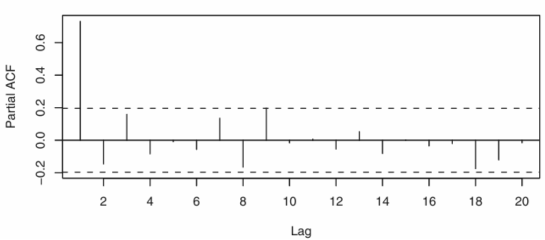


Exhibit 6.17 Sample EACF for Simulated ARMA(1,1) Series

AR / MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	x	x	x	o	o	o	o	o	o	o	o	o	o
1	x	o	o	o	o	o	o	o	o	o	o	o	o	o
2	x	o	o	o	o	o	o	o	o	o	o	o	o	o
3	x	x	o	o	o	o	o	o	o	o	o	o	o	o
4	x	o	x	o	o	o	o	o	o	o	o	o	o	o
5	x	o	o	o	o	o	o	o	o	o	o	o	o	o
6	x	o	o	o	x	o	o	o	o	o	o	o	o	o
7	x	o	o	o	x	o	o	o	o	o	o	o	o	o