UNIVERSITY of LIMERICK

OLLSCOIL LUIMNIGH

Time Series Analysis

MS 4218

joseph.lynch@ul.ie

Outlines

Basic concepts that serve as models for time series.

- Mean, variances, covariances
- Random walk
- Moving average
- Stationarity
- White noise
- Random cosine wave
- Differencing

Stochastic Processes

A sequence of random variables: $\{Y_t : t = 0, \pm 1, \pm 2, \dots\}$ is a stochastic process.

Serves as a model for an observed time series.

It evolves over time according to probabilistic laws.

Future values have a probability distribution conditioned on past values.

Models

Many models for stochastic processes are expressed by means of an algebraic formula relating the random variable at time t to past values of the process together with an unobservable error process.

It can be difficult to specify the joint probability distribution of $Y_t; t = 1, \ldots, k$, for any set of times $t_1, \ldots, t_k$ and any value of k, i.e., $Pr(Y_{t_1} = y_{t_1}, Y_{t_2} = y_{t_2}, \ldots, Y_{t_k} = y_{t_k})$.

Easier to calculate the first and second moments.

First moment

The first moment is the expected value at a given time point.

$$\mu_t = E(Y_t).$$

If Y_t is discrete, we have

$$E(Y_t) = \sum_{y_t} y_t Pr(Y_t = y_t).$$

$$E\{g(Y_t)\} = \sum_{y_t} g(y_t) Pr(Y_t = y_t).$$

$$E\{h(Y_t, Y_s)\} = \sum_{y_t} \sum_{y_s} h(y_t, y_s) Pr(Y_t = Y_t, Y_s = y_s).$$

More properties of expectation operator

Linearity:

$$\begin{aligned} E(aY_t + bY_s + c) &= E(aY_t) + E(bY_s) + E(c) \\ &= aE(Y_t) + bE(Y_s) + c. \end{aligned}$$

Note, while Y_t and Y_s are random variables, $E(Y_t)$ and $E(Y_s)$ are constants.

We use these linearity properties when calculating the Variance and Covariance of random variables.

Second moment: Autocovariance function (acvf)

(Auto, because it is a measure of how Y co-varies with itself at different time points.)

$$
\begin{aligned}
\gamma_{t,s} &= Cov(Y_t, Y_s); \qquad t, s = 0, \pm 1, \pm 2, \ldots \\
\\
&= E[\{Y_t - E(Y_t)\}\{Y_s - E(Y_s)\}] \\
&= E[(Y_t - \mu_t)(Y_s - \mu_s)] \\
&= E(Y_t Y_s - \mu_t Y_s - \mu_s Y_t + \mu_s \mu_t) \\
&= E(Y_t Y_s) - E(\mu_t Y_s) - E(\mu_s Y_t) + E(\mu_s \mu_t) \\
&= E(Y_t Y_s) - \mu_t E(Y_s) - \mu_s E(Y_t) + \mu_s \mu_t \\
&= E(Y_t Y_s) - \mu_t \mu_s - \mu_s \mu_t + \mu_s \mu_t \\
&= E(Y_t Y_s) - \mu_t \mu_s \\
&= E(Y_t Y_s) - E(Y_t) E(Y_s) \; = \; \gamma_{s,t}.
\end{aligned}
$$

Covariance properties

$$
\begin{aligned}
Cov(a+bY_t, c+dY_s) &= E\{(a+bY_t)(c+dY_s)\} - E(a+bY_t)E(c+dY_s) \\
&= E(ac+adY_s+cbY_t+bdY_tY_s) \\
&\quad -E(a)E(c) - E(a)E(dY_s) - E(bY_t)E(c) - E(bY_t)E(dY_s) \\
&= E(ac)+adE(Y_s)+cbE(Y_t)+bdE(Y_tY_s) \\
&\quad -ac - adE(Y_s) - bcE(Y_t) - bdE(Y_t)E(Y_s) \\
&= bdE(Y_tY_s) - bdE(Y_t)E(Y_s) \\
&= bd\{E(Y_tY_s) - E(Y_t)E(Y_s)\} \\
&= bdCov(Y_t, Y_s). \qquad (1)
\end{aligned}
$$

Variance

The variance is a special case of the acvf when $s = t$.

$$
\begin{aligned}
\gamma_{t,t} &= Cov(Y_t, Y_t) \\
E[\{Y_t - E(Y_t)\}\{Y_t - E(Y_t)\}] &= E[Y_t^2 - 2Y_tE(Y_t) + \{E(Y_t)\}^2] \\
&= E(Y_t^2) - E\{2Y_tE(Y_t)\} + E[\{E(Y_t)\}^2] \\
&= E(Y_t^2) - 2E(Y_t)E(Y_t) + \{E(Y_t)\}^2 \\
&= E(Y_t^2) - \{E(Y_t)\}^2. \\
Var(aY_t + b) &= a^2 Var(Y_t). \\
Var(Y_t) &\geq 0.
\end{aligned}
$$

Variance and Covariance

$$
\begin{aligned}
Var(Y_t + Y_s) &= Var(Y_t) + Var(Y_s) + 2Cov(Y_t, Y_s) \\
Var(Y_t - Y_s) &= Var(Y_t) + Var(Y_s) - 2Cov(Y_t, Y_s) \\
Cov(Y_t + Y_s, Y_r) &= Cov(Y_t, Y_r) + Cov(Y_s, Y_r).
\end{aligned}
$$

If Y_t and Y_s are independent, $Cov(Y_t, Y_s) = 0$.

The opposite does not necessarily hold.

Covariance of sums

If $c_1, \ldots, c_m$ and $d_1, \ldots, d_n$ are constants, and $t_1, \ldots, t_m$ and $s_1, \ldots, s_n$ are time points, then

$$Cov\left(\sum_{i=1}^{m} c_i Y_{t_i}, \sum_{j=1}^{n} d_j Y_{s_j}\right) = \sum_{i=1}^{m} c_i \sum_{j=1}^{n} d_j Cov\left(Y_{t_i}, Y_{s_j}\right). \quad (2)$$

The covariance of two sums of random variables $=$ the sum of all possible covariance pairs of variables.

Variance of sums

A special case of the latter result

$$
\begin{aligned}
Var\left(\sum_{i=1}^{n} c_i Y_{t_i}\right) &= \sum_{i=1}^{n} c_i^2 Var(Y_{t_i}) \\
&\quad + 2\sum_{i=2}^{n}\sum_{j=1}^{i-1} c_i c_j Cov\left(Y_{t_i}, Y_{t_j}\right).
\end{aligned}
$$

Auto-correlation function (acf)

Measures linear dependence between 2 variables.

$$\begin{aligned}
\rho_{t,s} &= Corr(Y_t, Y_s) \\
&= \frac{Cov(Y_t, Y_s)}{\sqrt{Var(Y_t)Var(Y_s)}} \\
&= \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}} \\
\rho_{t,t} &= Corr(Y_t, Y_t) \\
&= \frac{\gamma_{t,t}}{\sqrt{\gamma_{t,t}\gamma_{t,t}}} = 1.
\end{aligned}$$

Acf properties

$$
\begin{aligned}
-1 &\leq Corr(Y_t, Y_s) = \rho_{t,s} \leq 1 \\
Corr(a + bY_t, c + dY_s) &= \text{sign } bdCorr(Y_t, Y_s). \\
\text{sign}(bd) &= \begin{cases} 1 & \text{if } bd > 0. \\ 0 & \text{if } bd = 0. \\ -1 & \text{if } bd < 0. \end{cases}
\end{aligned}
$$

Random Walk

Provides a good model as a first approximation for many time series.

Let $e_1, e_2, \ldots,$ be a sequence of independent and identically distributed (iid) random variables with expected value of 0 and variance of σ_e^2.

The observed time series Y_t is given by:

$$
\begin{aligned}
Y_1 &= e_1. \text{ (initial condition)} \\
Y_2 &= e_1 + e_2 \; = \; Y_1 + e_2. \\
Y_t &= e_1 + e_2 + \cdots + e_t \; = \; Y_{t-1} + e_t.
\end{aligned}
$$

Expectation for random walk

$$
\begin{aligned}
E(Y_t) &= E(e_1 + e_2 + \cdots + e_t) \\
&= E(e_1) + E(e_2) + \cdots + E(e_t) \\
&= 0.
\end{aligned}
$$

On average, at any given time, you expect to be where you started!

Variance for random walk

$$Var(Y_t) = Var(e_1 + e_2 + \cdots + e_t).$$

The e's are independent $\Rightarrow$ they do not co-vary, $Cov(e_t, e_s) = 0 \; \forall s \neq t$, $\therefore$

$$\begin{aligned} Var(Y_t) &= Var(e_1) + Var(e_2) + \cdots + Var(e_t) \\ &= \sigma_e^2 + \sigma_e^2 + \cdots + \sigma_e^2 \\ &= t\sigma_e^2. \end{aligned}$$

Variance is time-dependent and increases linearly as a function of time.

Auto-covariance for random walk

For $1 \leq t \leq s$, and using Equation (2) on page 11,

$$\begin{aligned}
\gamma_{t,s} &= \sum_{i=1}^{s}\sum_{j=1}^{t} Cov(e_i, e_j) \\
&= 0 \text{ if } s \neq t, \\
&= \sigma_e^2 \text{ if } s = t. \\
\therefore \gamma_{t,s} &= t\sigma_e^2 \qquad \text{for } 1 \leq t \leq s \\
&= \gamma_{s,t}.
\end{aligned}$$

Auto-correlation for random walk

$$
\begin{aligned}
\rho_{t,s} &= \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}} \\
&= \frac{t\sigma_e^2}{\sqrt{t\sigma_e^2 s\sigma_e^2}} \\
&= \frac{t\sigma_e^2}{\sigma_e^2\sqrt{ts}} \\
&= \sqrt{\frac{t}{s}} \qquad 1 \le t \le s.
\end{aligned}
$$

Values of y at neighbouring points have higher $\rho_{t,s}$ as time goes by.

Generation of a random walk

If the e_i are iid $(0, \sigma_e^2)$, the random walk can be simulated using a Normal distribution.

library(TSA) has an inbuilt dataset called ‘rwalk’ which is accessed using

```
data(rwalk)
```

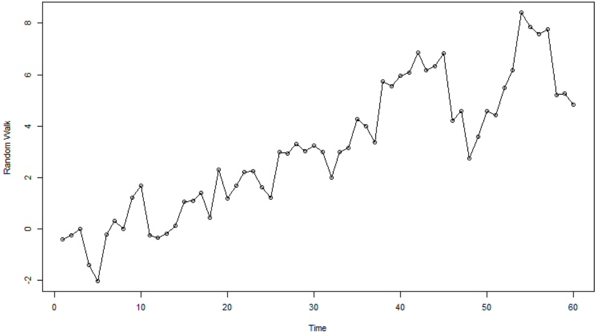
A plot of this ts object is shown overleaf.

The variance increases over time.

We expect to have $\mu = 0$ and $\rho_{s,t}$ close to 1.

Exhibit2.1

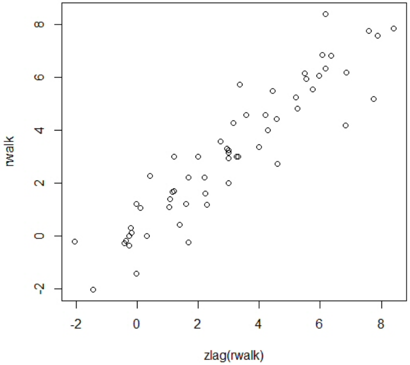
Time series plot of a simulated random walk



mean(rwalk)=3.081771 !

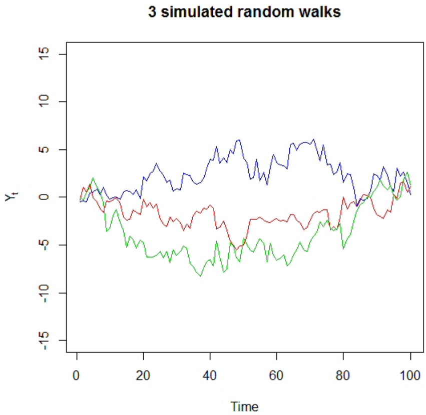
Random Walk Correlation

```
plot(y=rwalk,x=zlag(rwalk))
```




```
cor(rwalk[-1],rwalk[-length(rwalk)])= 0.9238698
```

Zero mean



Code requested in lab 2.

Moving average

Let Y_t be the average of two consecutive iid e's with zero mean and variance of σ_e^2.

Then:

$$
\begin{aligned}
Y_t &= \frac{e_t + e_{t-1}}{2}. \\
Y_{t-1} &= \frac{e_{t-1} + e_{t-2}}{2}.
\end{aligned}
$$

Expectation for moving average

$$
\begin{aligned}
E(Y_t) &= E\left(\frac{e_t + e_{t-1}}{2}\right) \\
&= \frac{1}{2}E(e_t + e_{t-1}) \\
&= \frac{1}{2}\left\{E(e_t) + E(e_{t-1})\right\} \\
&= \frac{1}{2}(0 + 0) \\
&= 0.
\end{aligned}
$$

Variance for moving average

$$
\begin{aligned}
Var(Y_t) &= Var\left(\frac{e_t + e_{t-1}}{2}\right) \\
&= \left(\frac{1}{2}\right)^2 Var(e_t + e_{t-1}) \\
&= \frac{1}{4}\{Var(e_t) + Var(e_{t-1})\} \\
&= \frac{1}{4}(\sigma_e^2 + \sigma_e^2) \\
&= 0.5\sigma_e^2.
\end{aligned}
$$

Auto-covariance for moving average

Using Equation (1) on page 8,

$$
\begin{aligned}
Cov(Y_t, Y_{t-1}) &= Cov\left(\frac{e_t + e_{t-1}}{2}, \frac{e_{t-1} + e_{t-2}}{2}\right) \\
&= \frac{1}{4} Cov(e_t + e_{t-1}, e_{t-1} + e_{t-2}) \\
&= \frac{1}{4}\{Cov(e_t, e_{t-1}) + Cov(e_t, e_{t-2}) \\
&\quad + Cov(e_{t-1}, e_{t-1}) + Cov(e_{t-1}, e_{t-2})\} \\
&= \frac{1}{4} Cov(e_{t-1}, e_{t-1}) \\
&= 0.25\sigma_e^2 = \gamma_{t,t-1}.
\end{aligned}
$$

Auto-covariance for moving average cont.

$$
\begin{aligned}
Cov(Y_t, Y_{t-2}) &= Cov\left(\frac{e_t + e_{t-1}}{2}, \frac{e_{t-2} + e_{t-3}}{2}\right) \\
&= \frac{1}{4} Cov(e_t + e_{t-1}, e_{t-2} + e_{t-3}) \\
&= \frac{1}{4}\{Cov(e_t, e_{t-2}) + Cov(e_t, e_{t-3}) \\
&\quad + Cov(e_{t-1}, e_{t-2}) + Cov(e_{t-1}, e_{t-3})\} \\
&= 0 = \gamma_{t,t-2}.
\end{aligned}
$$

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

Moving average auto-covariance summary

$$\gamma_{t,s} = \begin{cases} 0.5\sigma_e^2 & \text{if } |t-s| = 0. \\ 0.25\sigma_e^2 & \text{if } |t-s| = 1. \\ 0 & \text{if } |t-s| > 1. \end{cases}$$

Moving average auto-correlation summary

$$
\begin{aligned}
\rho_{t,s} &= \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}} \\
&= \frac{\gamma_{t,s}}{\gamma_{t,t}}
\end{aligned}
$$

$$
\rho_{t,s} = \begin{cases} \frac{0.5\sigma_e^2}{0.5\sigma_e^2} = 1 & \text{if } |t-s| = 0. \\ \frac{0.25\sigma_e^2}{0.5\sigma_e^2} = 0.5 & \text{if } |t-s| = 1. \\ \frac{0}{0.5\sigma_e^2} = 0 & \text{if } |t-s| > 1. \end{cases}
$$

This moving average process is said to be stationary.

Stationarity

Most NB assumption when making statistical inference about a stochastic process.

The statistical properties of any one section of the data are just like those of any other section of the data.

A series is **strictly stationary** if:
$Pr(Y_{t_1}, Y_{t_2}, \ldots, Y_{t_n}) = Pr(Y_{t_1-k}, Y_{t_2-k}, \ldots, Y_{t_n-k}) \forall\ t_1, \ldots, t_n$ and $\forall$ lags k.

Shifting the time origin by an amount k has no effect on the joint distribution.

Stationarity cont.

If $n = 1$ the marginal Y's are iid, and so,

$$E(Y_{t_1}) = E(Y_{t_1-k}) = \mu$$

and

$$Var(Y_{t_1}) = Var(Y_{t_1-k}) = \sigma^2$$

are time-independent.

Stationarity cont.

If $n = 2$, we have

$$Pr(Y_t, Y_s) = Pr(Y_{t-k}, Y_{s-k}) \ \forall \ t_1, \ldots t_n \text{ and } \forall \text{ lags } k.$$

$$Cov(Y_t, Y_s) = \gamma_{t,s} = Cov(Y_{t-k}, Y_{s-k}) = \gamma_{t-k,s-k}.$$

$$\text{Let } k = s, \Rightarrow Cov(Y_{t-k}, Y_{s-k}) = Cov(Y_{t-s}, Y_{s-s}) = \gamma_{t-s,0}.$$

$$\begin{aligned} \text{Let } k = t, \Rightarrow Cov(Y_{t-k}, Y_{s-k}) &= Cov(Y_{t-t}, Y_{s-t}) = \gamma_{0,s-t} \\ &= \gamma_{0,|t-s|} \end{aligned}$$

$Cov(Y_t, Y_s)$ depends only on the lag $|t-s|$ and not on t or s.

Stationarity cont.

The notation can be further simplified.

$$Cov(Y_t, Y_s) = \gamma_{t,s} = \gamma_{0,|t-s|} = \gamma_{t-s}.$$

$$Cov(Y_t, Y_{t-k}) = \gamma_{t,t-k} = \gamma_{0,t-(t-k)} = \gamma_k.$$

$$\begin{aligned}
\rho_k &= Corr(Y_t, Y_{t-k}) = \frac{Cov(Y_t, Y_{t-k})}{\sqrt{Var(Y_t)Var(Y_{t-k})}} \\
&= \frac{\gamma_k}{\sqrt{Cov(Y_t, Y_t)Cov(Y_{t-k}, Y_{t-k})}} \\
&= \frac{\gamma_k}{\sqrt{\gamma_{t-t}\gamma_{\{t-k-(t-k)\}}}} = \frac{\gamma_k}{\sqrt{\gamma_0\gamma_0}} = \frac{\gamma_k}{\gamma_0}. \\
\rho_0 &= \frac{\gamma_0}{\gamma_0} = 1 \text{ and } \rho_k = \rho_{-k} = \rho_{|k|}.
\end{aligned}$$

Weakly stationarity

This is much more commonly used than strictly stationary.

The joint probability distributions are not necessarily equivalent, but,

$$E(Y_t) = \mu.$$

$$\gamma_{t,t-k} = \gamma_k.$$

The expected value is the same no matter where we are in the series.

The auto-covariance only depends on the lag between t and $t-k$, i.e., k, and not on the actual times themselves.

White noise

Unlike the random walk where the variance increases with time, white noise is a stationary process.

White noise is a sequence of iid $\{e_t\}$ random variables with zero mean and variance of σ_e^2.

If independent, the joint distribution is the product of the marginals.

Using its distribution function, we have

$$
\begin{aligned}
Pr(e_{t_1} \leq x_1, \ldots, e_{t_n} \leq x_n) &= Pr(e_{t_1} \leq x_1), \ldots, Pr(e_{t_n} \leq x_n) \\
&= Pr(e_{t_1-k} \leq x_1), \ldots, Pr(e_{t_n-k} \leq x_n) \\
&\quad (e_t's \text{ are identically distributed}) \\
&= Pr(e_{t_1-k} \leq x_1, \ldots, e_{t_n-k} \leq x_n).
\end{aligned}
$$

Thus white noise is strictly stationary.

White noise cont.

$$
\begin{aligned}
E(e_t) &= E(e_{t-k}) \\
&= 0 = \mu. \\
\gamma_k &= Cov(e_t, e_{t-k}) \\
&= \begin{cases} Var(e_t) = \sigma_e^2 & \text{if } k = 0. \\ 0 & \text{if } k \neq 0. \end{cases}
\end{aligned}
$$

White noise cont.

$$
\begin{aligned}
\rho_k &= \frac{\gamma_k}{\gamma_0} \\
&= \begin{cases} 1 & \text{if } k = 0. \\ 0 & \text{if } k \neq 0. \end{cases}
\end{aligned}
$$

Time v frequency domains

Economic, commercial and social studies time series are often examined in the time domain.

Frequency domain methods are sometimes used in physical, engineering and meteorological areas, but the categories in either domain are in no way absolute.

In frequency domain methods, the values in the series are represented by a superposition of sine and cosine functions, as in Fourier analysis.

The properties of the series are characterized in terms of the associated frequencies and amplitudes.

Random cosine wave

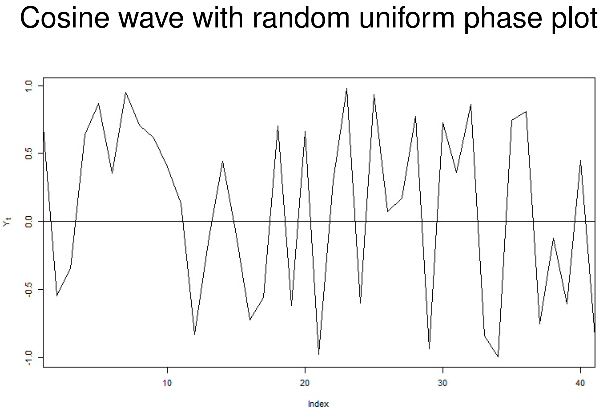
Consider the process Y_t, for $t = 0, \pm 1, \pm 2, \ldots$

$$Y_t = cos\left\{2\pi\left(\frac{t}{12} + \Phi\right)\right\}.$$

The phase ϕ, is the fraction of a complete cycle completed by time $t = 0$.

In this example, Φ is a random variable, hence we have a random cosine wave.

Plot of the cosine wave



R Code

If the random variable Φ comes from a Uniform$(a = 0, b = 1)$ distribution, then its density is given by:

$$
\begin{aligned}
f(\phi) &= \frac{1}{b-a} \\
&= \frac{1}{1-0} \\
&= 1.
\end{aligned}
$$

The expected value of the random variable Y_t is a function of the random variable Φ.

Expectation

Using the properties of expectation

$$
\begin{aligned}
E(Y_t) &= E\left[cos\left\{ 2\pi\left(\frac{t}{12} + \Phi \right) \right\} \right] = E\{g(\Phi)\} \\
&= \int_0^1 cos\left\{ 2\pi\left(\frac{t}{12} + \phi \right) \right\} f(\phi)\mathrm{d}\phi \\
&= \int_0^1 cos\left\{ 2\pi\left(\frac{t}{12} + \phi \right) \right\} 1\mathrm{d}\phi \\
&= \frac{1}{2\pi} sin\left\{ 2\pi\left(\frac{t}{12} + \phi \right) \right\}_0^1 \\
&= \frac{1}{2\pi} sin\left(2\pi\frac{t}{12} + 2\pi \right) - \frac{1}{2\pi} sin\left(2\pi\frac{t}{12} \right) \\
&= 0 \quad \forall\ t.
\end{aligned}
$$

Auto-covariance

$$
\begin{aligned}
\gamma_{t,s} &= E(Y_t Y_s) - E(Y_t)E(Y_s) = E(Y_t Y_s) - (0)(0) = E(Y_t Y_s) \\
&= E\left[cos\left\{2\pi\left(\frac{t}{12} + \Phi\right)\right\} cos\left\{2\pi\left(\frac{s}{12} + \Phi\right)\right\}\right] \\
&= \int_0^1 cos\left\{2\pi\left(\frac{t}{12} + \phi\right)\right\} cos\left\{2\pi\left(\frac{s}{12} + \phi\right)\right\} f(\phi)\mathrm{d}\phi
\end{aligned}
$$

It is easier to integrate sums than products.

Auto-covariance cont.

We now use the trigonometric identity:

$$cosAcosB = \frac{1}{2}\{cos(A-B) + cos(A+B)\}.$$

$$\begin{aligned}
cos(A-B) &= cos\left(\frac{2\pi t}{12} + 2\pi\phi - \frac{2\pi s}{12} - 2\pi\phi\right) \\
&= cos\left\{\frac{2\pi(t-s)}{12}\right\}. \\
cos(A+B) &= cos\left(\frac{2\pi t}{12} + 2\pi\phi + \frac{2\pi s}{12} + 2\pi\phi\right) \\
&= cos\left\{2\pi\left(\frac{t+s}{12} + \phi\right)\right\}.
\end{aligned}$$

Auto-covariance cont.

$$
\begin{aligned}
\gamma_{t,s} &= \frac{1}{2}\int_0^1 \left[cos\left\{\frac{2\pi(t-s)}{12}\right\} + cos\left\{2\pi\left(\frac{t+s}{12}+\phi\right)\right\}\right] f(\phi)\mathrm{d}\phi \\
&= \left[\frac{1}{2} cos\left\{\frac{2\pi(t-s)}{12}\right\}\phi\right]_0^1 + \frac{1}{2}\frac{1}{2\pi} sin\left\{2\pi\left(\frac{t+s}{12}+\phi\right)\right\}_0^1 \\
&= \frac{1}{2} cos\left\{\frac{2\pi(t-s)}{12}\right\} \\
&\quad + \frac{1}{4\pi}\left[sin\left\{\frac{2\pi(t+s)}{12} + 2\pi.1\right\} - sin\left\{\frac{2\pi(t+s)}{12} + 2\pi.0\right\}\right] \\
&= \frac{1}{2} cos\left\{\frac{2\pi(t-s)}{12}\right\} + 0.
\end{aligned}
$$

Cosine wave is stationary!

The expected value of the cosine wave is zero $\forall$ t.

The auto-covariance only depends on the lag $(t - s)$.

$$
\begin{aligned}
\rho_{t,s} &= \frac{\gamma_{t,s}}{\gamma_0} \\
&= \frac{1}{2} cos\left\{\frac{2\pi(t-s)}{12}\right\} \div \frac{1}{2} cos\left\{\frac{2\pi(t-t)}{12}\right\} \\
&= cos\left(\frac{2\pi k}{12}\right),
\end{aligned}
$$

where the lag $k = t - s$ takes values $0 \pm 1, \pm 2, \ldots$

Difference of a series

The random walk $Y_t = Y_{t-1} + e_t$ is not stationary, but the difference between successive terms is.

See Question 10 on tutorial sheet 2.

$$\begin{aligned} \nabla Y_t &= Y_t - Y_{t-1} \\ &= e_t. \end{aligned}$$

A linear sequence has a constant slope and thus a constant difference between equally spaced successive terms.

Differencing a series once removes a linear trend.

Second difference of a series

The second difference is the difference of a differenced series.

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

Second differencing can remove a quadratic trend.

Differencing

For removing trends, generally it is only necessary to difference once or at most twice.

A peak in December sales may prevent a series from being stationary.

Seasonal differencing can be used to remove seasonable patterns, e.g, it might be of interest to evaluate $Y_t - Y_{t-12}$, $Y_{t-12} - Y_{t-24}$ etc.

It is much easier to study data under the assumption of stationarity.

Next

Trends

Seasonality

Residuals and randomness