Chapter 2: The Basics

· Sequence $\{Y_t : t = 0, \pm 1, \pm 2, \ldots\}$ of random variables (RVs)
  - known as stochastic process.

  · Works as model for observed time series
  · Complete probabilistic structure for process is determined by set of joint distributions of all finite collections of the Ys.
  · Don't need all of these.
  · Get most of information in these distributions from means, variances, and covariances.

Means, Variances, Covariances

Let $\{Y_t, t = 0, \pm 1, \pm 2, \ldots\}$ be a stochastic process.

Mean Function: $\mu_t = \mathbb{E}[Y_t]$, $t = 0, \pm 1, \pm 2, \ldots$

* Expectation of process at time t.
* Can be different at each time point.

Autocovariance Function

$$\gamma_{t,s} = Cov(Y_t, Y_s) = \mathbb{E}[(Y_t - \mu_t)(Y_s - \mu_s)], \quad t, s = 0, \pm 1, \pm 2, \ldots$$

$$\begin{aligned}
\mathbb{E}[(Y_t - \mu_t)(Y_s - \mu_s)] &= \mathbb{E}[Y_t Y_s] - \mu_t \mathbb{E}(Y_s) - \mu_s \mathbb{E}[Y_t] + \mu_s \mu_t \\
&= \mathbb{E}[Y_t Y_s] - \mu_t \mu_s - \mu_s \mu_t + \mu_s \mu_t \\
&= \mathbb{E}[Y_t Y_s] - \mu_s \mu_t .
\end{aligned}$$

Autocorrelation Function

$\rho_{t,s} = Corr(Y_t, Y_s), \quad s,t = 0, \pm 1, \pm 2, \ldots$

where

$$Corr(Y_t, Y_s) = \frac{Cov(Y_t, Y_s)}{\sqrt{Var(Y_t)\,Var(Y_s)}} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\,\gamma_{s,s}}}.$$

Properties of Autocorrelation / Autocovariance

1. $\gamma_{t,t} = Var(Y_t)$
2. $\rho_{t,t} = 1$
3. $\gamma_{s,t} = \gamma_{t,s}$
4. $\rho_{s,t} = \rho_{t,s}$
5. $|\gamma_{t,s}| \le \sqrt{\gamma_{t,t}\,\gamma_{s,s}}$ $\Rightarrow$ 6. $|\rho_{t,s}| \le 1$.

- If $|\rho_{s,t}| \approx 1$, Y_s, Y_t are strongly linearly related.
- If $|\rho_{s,t}| \approx 0$, weak linear relationship
- $\rho_{s,t} = 0$, then Y_s, Y_t are uncorrelated.

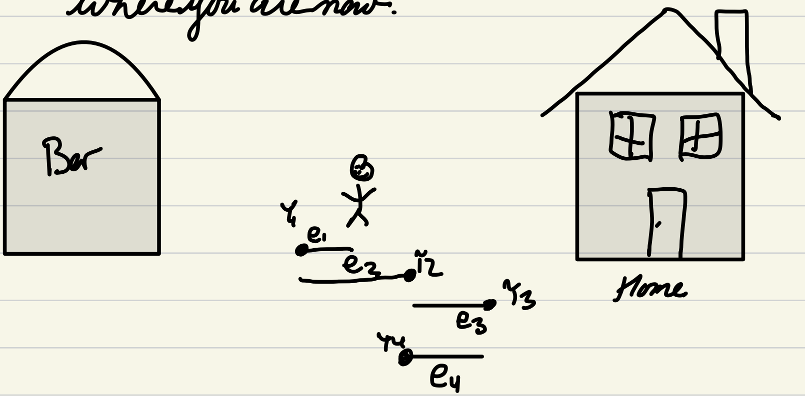
- Let $c_1, \ldots, c_m$; $d_1, \ldots, d_n$ - constants
- Let $t_1, \ldots, t_m$; $s_1, \ldots, s_n$ be time points.

- $$Cov\left[\sum_{i=1}^{m} c_i Y_{t_i}, \sum_{j=1}^{m} d_j Y_{s_j}\right] = \sum_{i=1}^{m}\sum_{j=1}^{n} c_i d_j \overbrace{Cov(Y_{t_i}, Y_{s_j})}^{\gamma_{t_i, s_j}}$$

- $$Var\left[\sum_{i=1}^{m} c_i Y_{t_i}\right] = \sum_{i=1}^{m} c_i^2 Var(Y_{t_i}) + 2\sum_{i=2}^{m}\sum_{j=1}^{i-1} c_i c_j Cov(Y_{t_i}, Y_{t_j}).$$

The Random Walk

- Kinds of like a drunk walk.
  - Where you go next is random, based only on where you are now.



Formally

$e_1, e_2, \dots$ - sequence of independent and identically distributed RVs with mean 0, variance σ_e^2.

$$Y_1 = e_1$$
$$Y_2 = e_1 + e_2$$
$$\vdots$$
$$Y_t = \sum_{i=1}^{t} e_i$$

ie: $Y_t = Y_{t-1} + e_t$ with initial condition $Y_1 = e_1$.

e_t - size of step taken at time t

Y_t - position at time t.

Mean function

$$\mu_t = \mathbb{E}[Y_t] = \mathbb{E}\left[\sum_{i=1}^{t} e_i\right] = \sum_{i=1}^{t} \mathbb{E}[e_i] = 0 \Rightarrow \mu_t = 0 \ \forall t.$$

$$Var(Y_t) = \sum_{i=1}^{t} Var(e_i) = t\sigma_e^2.$$

$\Rightarrow$ variance increases linearly in time.

Autocovariance function

$$\gamma_{t,s} = Cov(Y_t, Y_s) = Cov\left(\sum_{i=1}^{t} e_i , \sum_{j=1}^{s} e_j\right) \qquad 1 \le t \le s$$

$$= \sum_{i=1}^{t} \sum_{j=1}^{s} (1)(1)\, Cov(e_i, e_j)$$

$Cov(e_i, e_j) = 0 \ \forall i \neq j.$

$$\Rightarrow \gamma_{t,s} = \sum_{i=j}^{t} \sum_{j=1}^{s} Cov(e_i, e_j)$$

$$= Var(e_1) + Var(e_2) + \cdots + Var(e_t) = t\sigma_e^2$$

$\Rightarrow \gamma_{t,s} = t\sigma_e^2$ for $1 \le t \le s$ is the autocovariance function.

Autocorrelation Function

$$\rho_{t,s} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}} = \frac{t\sigma_e^2}{\sqrt{t\sigma_e^2 s\sigma_e^2}} = \sqrt{\frac{t}{s}} \quad \text{for } 1 \le t \le s$$

Ex. (Illustrates random Walk Behavior)

$$\rho_{1,2} = \sqrt{\tfrac{1}{2}} = 0.707 \qquad \rho_{4,9} = \sqrt{\tfrac{4}{9}} = 2/3 = 0.667$$

$$\rho_{1,25} = \sqrt{\tfrac{1}{25}} = 0.2 \qquad \rho_{1,36} = \sqrt{\tfrac{1}{36}} = 0.167.$$

Idea (Roughly): Values of Y at closer time points are more strongly and positively correlated than values at more distant time points.

Ex. RW: $e_t \overset{iid}{\sim} N(0,1)$ (Brownian Motion).
Models stock price movement, movement of particles in fluid.

Moving Average

Consider $Y_t = \frac{1}{2}(e_t + e_{t-1})$, $e_t \overset{iid}{\sim}$ with mean 0, variance σ_e^2.

$$E[Y_t] = \mu_t = \frac{1}{2}E[e_t + e_{t-1}] = \frac{1}{2}(E[e_t] + E[e_{t-1}]) = \frac{1}{2}(0+0) = 0.$$

$$Var(Y_t) = \frac{1}{4}Var(e_t + e_{t-1}) = \frac{Var(e_t) + Var(e_{t-1})}{4} = \frac{2\sigma_e^2}{4} = \boxed{\frac{\sigma_e^2}{2}}$$

$$Cov(Y_t, Y_{t-1}) = Cov\left(\frac{e_t + e_{t-1}}{2}, \frac{e_{t-1} + e_{t-2}}{2}\right)$$

$$= \frac{1}{2}\left(\frac{1}{2}\right)Cov(e_t, e_{t-1}) + \frac{1}{2}\left(\frac{1}{2}\right)Cov(e_t, e_{t-2}) + \frac{1}{2}\left(\frac{1}{2}\right)\overbrace{Cov(e_{t-1}, e_{t-1})}^{Var(e_{t-1})}$$
$$+ \frac{1}{2}\left(\frac{1}{2}\right)Cov(e_{t-1}, e_{t-2})$$

$$= \frac{1}{4}(0 + 0 + \sigma_e^2 + 0) = \frac{\sigma_e^2}{4}.$$

$$\gamma_{t,t-1} = \frac{\sigma_e^2}{4}.$$

$$\mathrm{Cov}(Y_t, Y_{t-2}) = \mathrm{Cov}\left(\frac{e_t + e_{t-1}}{2}, \frac{e_{t-2}+e_{t-3}}{2}\right)$$

$$= \frac{1}{4}\left[\overset{0}{\mathrm{Cov}(e_t, e_{t-2})} + \overset{0}{\mathrm{Cov}(e_t, e_{t-3})} + \overset{0}{\mathrm{Cov}(e_{t-1}, e_{t-2})} + \overset{0}{\mathrm{Cov}(e_{t-1}, e_{t-3})}\right]$$

$$= \frac{1}{4}(0) = 0.$$

The autocovariance function (ACVF) is given by

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

This leads to the autocorrelation function (ACF)

$$\rho_{t,s} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}} = \frac{\gamma_{t,s}}{\sqrt{\sigma_e^2\sigma_e^2(0.5)^2}} = \begin{cases} 1 & \text{if } |t-s| = 0 \\ 0.5 & \text{if } |t-s| = 1 \\ 0 & \text{if } |t-s| \geq 2 \end{cases}.$$

<u>Main Ideas</u>

- Naturally, Y_t is perfectly positively correlated with itself.
- $\mathrm{Corr}(Y_{t-h}, Y_t) = \mathrm{Corr}(Y_t, Y_{t+1}) = 0.5$ regardless of point t in time (stationarity)

- Since e_t's are independent, the distribution of $Y_t | Y_{t-1}, Y_{t-2}, \ldots, Y_1$ is the same as the distribution of $Y_t | Y_{t-1}$ - Markov property.

  - This moving average process is an example of a Markov chain.