The Fundamental Concepts in the theory of time series

YENNI ANGRAINI

Outline

- Time Series and Stochastic Processes
- Means, Variances, and Covariances
- Stationarity

Time Series and Stochastic Processes

- **A stochastic process** is a collection or ensemble of random variables indexed by a variable t, usually representing time.
- A *stochastic process* means that one has a system for which there are observations at certain times, and that the outcome, that is, the observed value at each time is a random variable.
- The sequence of random variables $\{Y_t : t = 0, \pm 1, \pm 2, \pm 3, \ldots\}$ is called a **stochastic process**
- Another simple **stochastic process** is the random walk, $Y_t = Y_{t-1} + e_t$

Means

For a stochastic process $\{Y_t: t = 0, \pm 1, \pm 2, \pm 3, \ldots\}$, the **mean function** is defined by

$$\mu_t = E(Y_t) \qquad \text{for } t = 0, \pm 1, \pm 2, \ldots \qquad (2.2.1)$$

That is, μ_t is just the expected value of the process at time t. In general, μ_t can be different at each time point t.

Properties of Variance

$$Var(X) \geq 0$$

$$Var(a + bX) = b^2 Var(X)$$

If X and Y are independent, then

$$Var(X + Y) = Var(X) + Var(Y)$$

In general, it may be shown that

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

Properties of Covariance

$$Cov(a + bX, c + dY) = bdCov(X, Y)$$

$$Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$$

$$Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z)$$

$$Cov(X, X) = Var(X)$$

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

If X and Y are independent,

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

Properties of Correlation

The **correlation coefficient** of X and Y, denoted by $Corr(X, Y)$ or ρ, is defined as

$$\rho = Corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}}$$

$$-1 \leq Corr(X, Y) \leq 1$$

$$Corr(a + bX, c + dY) = sign(bd)Corr(X, Y)$$

$$\text{where } sign(bd) = \begin{cases} 1 \text{ if } bd > 0 \\ 0 \text{ if } bd = 0 \\ -1 \text{ if } bd < 0 \end{cases}$$

Variances, and Covariances

The **autocovariance function**, $\gamma_{t,s}$, is defined as

$$\gamma_{t,s} = Cov(Y_t, Y_s) \qquad \text{for } t, s = 0, \pm 1, \pm 2, \ldots \tag{2.2.2}$$

where $Cov(Y_t, Y_s) = E[(Y_t - \mu_t)(Y_s - \mu_s)] = E(Y_t Y_s) - \mu_t \mu_s$.

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

$$\rho_{t,s} = Corr(Y_t, Y_s) \qquad \text{for } t, s = 0, \pm 1, \pm 2, \ldots \tag{2.2.3}$$

where

$$Corr(Y_t, Y_s) = \frac{Cov(Y_t, Y_s)}{\sqrt{Var(Y_t)Var(Y_s)}} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}} \tag{2.2.4}$$

The following important properties of Variances and Covariances

$$\left.\begin{array}{ll} \gamma_{t,t} = Var(Y_t) & \rho_{t,t} = 1 \\ \gamma_{t,s} = \gamma_{s,t} & \rho_{t,s} = \rho_{s,t} \\ |\gamma_{t,s}| \le \sqrt{\gamma_{t,t}\gamma_{s,s}} & |\rho_{t,s}| \le 1 \end{array}\right\}$$

- Values of $\rho_{t,s}$ near ±1 indicate strong (linear) dependence,
- Values of $\rho_{t,s}$ near zero indicate weak (linear) dependence.
- If $\rho_{t,s} = 0$, we say that Y_t and Y_s are *uncorrelated*.

The Random Walk

Let $e_1, e_2, \ldots$ be a sequence of independent, identically distributed random variables each with zero mean and variance σ_e^2.

The observed time series, $\{Yt : t = 1, 2, \ldots\}$, is constructed as follows:

$$\left.\begin{aligned} Y_1 &= e_1 \\ Y_2 &= e_1 + e_2 \\ &\vdots \\ Y_t &= e_1 + e_2 + \cdots + e_t \end{aligned}\right\} \Rightarrow \quad Y_t = Y_{t-1} + e_t$$

Mean of Random Walk

$$\mu_t = E(Y_t) = E(e_1 + e_2 + \cdots + e_t) = E(e_1) + E(e_2) + \cdots + E(e_t)$$
$$= 0 + 0 + \cdots + 0$$

$$\mu_t = 0 \quad \text{for all } t$$

Variance of Random Walk

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

$$Var(Y_t) = t\sigma_e^2$$

Notice that the process variance increases linearly with time.

Covariance of Random Walk

To investigate the covariance function, suppose that $1 \le t \le s$. Then we have

$$\gamma_{t,s} = Cov(Y_t, Y_s) = Cov(e_1 + e_2 + \cdots + e_t, e_1 + e_2 + \cdots + e_t + e_{t+1} + \cdots + e_s)$$

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

However, these covariances are zero unless $i = j$, in which case they equal $Var(e_i) = \sigma_e^2$. There are exactly t of these so that $\gamma_{t,s} = t\sigma_e^2$.

Autocovariance and Autocorrelation function for Random Walk

Since $\gamma_{t,s} = \gamma_{s,t}$, this specifies the autocovariance function for all time points t and s and we can write

$$\boxed{\gamma_{t,s} = t\sigma_e^2 \qquad \text{for } 1 \le t \le s}$$

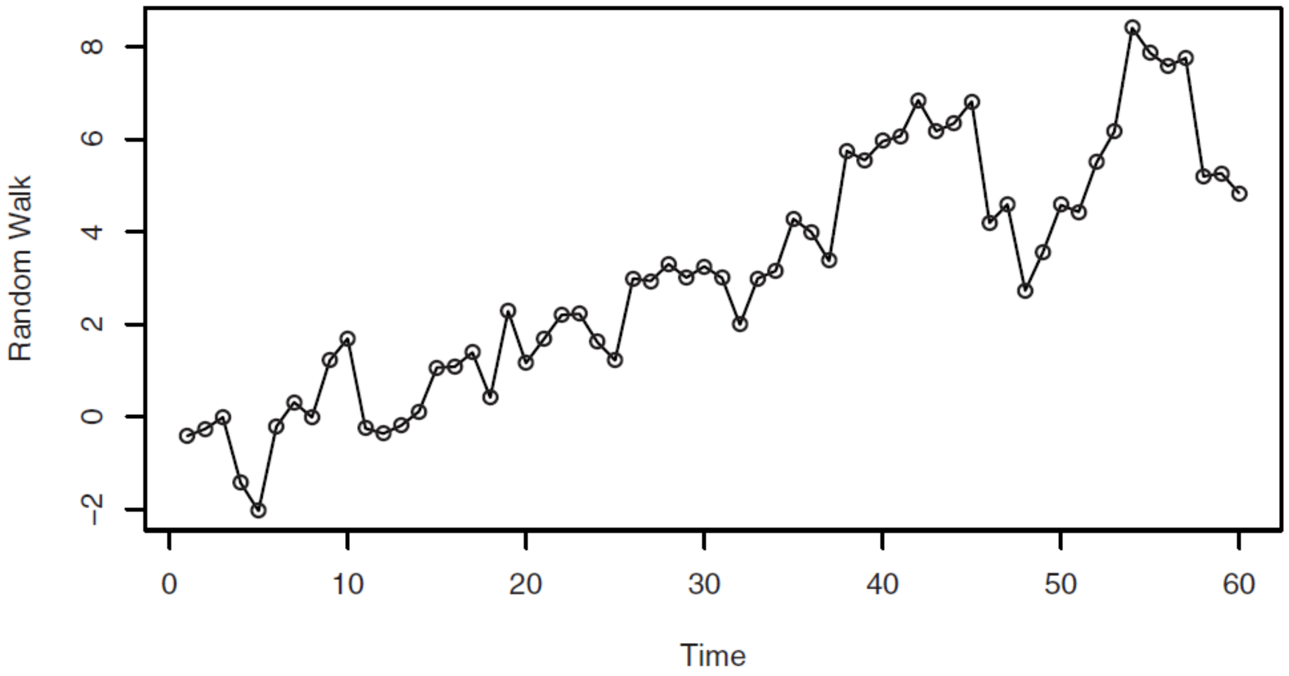
The autocorrelation function for the random walk is now easily obtained as

$$\rho_{t,s} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}} = \sqrt{\frac{t}{s}} \qquad \text{for } 1 \le t \le s$$

The Random Walk

The simple random walk process provides a good model (at least to a first approximation) for phenomena as diverse as the movement of common stock price, and the position of small particles suspended in a fluid—so-called Brownian motion.

Time Series Plot of a Random Walk



A Moving Average

Suppose that $\{Y_t\}$ is constructed as

$$Y_t = \frac{e_t + e_{t-1}}{2}$$

the e's are assumed to be independent and identically distributed with zero mean and variance σ_e^2

Mean of A Moving Average

$$\boxed{Y_t = \frac{e_t + e_{t-1}}{2}}$$

$$\mu_t = E(Y_t) = E\left\{\frac{e_t + e_{t-1}}{2}\right\} = \frac{E(e_t) + E(e_{t-1})}{2}$$

$$= 0$$

Variance of A Moving Average

$$Y_t = \frac{e_t + e_{t-1}}{2}$$

$$Var(Y_t) = Var\left\{\frac{e_t + e_{t-1}}{2}\right\} = \frac{Var(e_t) + Var(e_{t-1})}{4}$$

$$= 0.5\sigma_e^2$$

Covariance of A Moving Average

$$\boxed{Y_t = \frac{e_t + e_{t-1}}{2}}$$

$$\begin{aligned}
Cov(Y_t, Y_{t-1}) &= Cov\left\{\frac{e_t + e_{t-1}}{2}, \frac{e_{t-1} + e_{t-2}}{2}\right\} \\
&= \frac{Cov(e_t, e_{t-1}) + Cov(e_t, e_{t-2}) + Cov(e_{t-1}, e_{t-1})}{4} \\
&\qquad + \frac{Cov(e_{t-1}, e_{t-2})}{4} \\
&= \frac{Cov(e_{t-1}, e_{t-1})}{4} \qquad \text{(as all the other covariances are zero)} \\
&= 0.25\sigma_e^2
\end{aligned}$$

$$\gamma_{t,\,t-1} = 0.25\sigma_e^2 \qquad \text{for all } t$$

Covariance of A Moving Average

$$Y_t = \frac{e_t + e_{t-1}}{2}$$

Furthermore,

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

$$= 0 \qquad \text{since the } e\text{'s are independent.}$$

Similarly, $Cov(Y_t, Y_{t-k}) = 0$ for $k > 1$, so we may write

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

Autocorrelation function for A Moving Average

$$\rho_{t,s} = \begin{cases} 1 & \text{for } |t-s| = 0 \\ 0.5 & \text{for } |t-s| = 1 \\ 0 & \text{for } |t-s| > 1 \end{cases}$$

Stationarity

- To make statistical inferences about the structure of a stochastic process on the basis of an observed record of that process, we must usually make some simplifying (and presumably reasonable) assumptions about that structure.
- The most important such assumption is that of **stationarity**.
- A process $\{Y_t\}$ is said to be **strictly stationary** if the joint distribution of $Y_{t_1}, Y_{t_2}, \ldots, Y_{t_n}$ is the same as the joint distribution of $Y_{t_1-k}, Y_{t_2-k}, \ldots, Y_{t_n-k}$ for all choices of time points $t_1, t_2, \ldots, t_n$ and all choices of time lag k.
- When $n = 1$, the (univariate) distribution of Y_t is the same as that of Y_{t-k} for all t and k
- It then follows that $E(Y_t) = E(Y_{t-k})$ for all t and k so that the mean function is constant for all time
- $Var(Y_t) = Var(Y_{t-k})$ or all t and k so that the variance is also constant over time.

Stationarity

Setting $n = 2$ in the stationarity definition we see that the bivariate distribution of Y_t and Y_s must be the same as that of Y_{t-k} and Y_{s-k} from which it follows that $Cov(Y_t, Y_s) = Cov(Y_{t-k}, Y_{s-k})$ for all t, s, and k. Putting $k = s$ and then $k = t$, we obtain

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

Stationarity

That is, the covariance between Y_t and Y_s depends on time only through the time difference $|t - s|$ and not otherwise on the actual times t and s. Thus, for a stationary process, we can simplify our notation and write

$$\gamma_k = Cov(Y_t, Y_{t-k}) \qquad \text{and} \qquad \rho_k = Corr(Y_t, Y_{t-k})$$

Note also that

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

$$\left.\begin{array}{ll} \gamma_0 = Var(Y_t) & \rho_0 = 1 \\ \gamma_k = \gamma_{-k} & \rho_k = \rho_{-k} \\ |\gamma_k| \le \gamma_0 & |\rho_k| \le 1 \end{array}\right\}$$

Important example of as stationarity process – White Noise

- **White noise** process is defined as a sequence of independent, identically distributed random variables $\{e_t\}$
- Many useful processes can be constructed from **white noise**
- The fact that $\{e_t\}$ is strictly stationary

$\mu_t = E(e_t)$ is constant and

$$\gamma_k = \begin{cases} Var(e_t) & \text{for } k = 0 \\ 0 & \text{for } k \neq 0 \end{cases} \qquad \rho_k = \begin{cases} 1 & \text{for } k = 0 \\ 0 & \text{for } k \neq 0 \end{cases}$$

$$\boxed{Y_t = \frac{e_t + e_{t-1}}{2}} \Rightarrow$$

is another example of a stationary process constructed from white noise.

Many Thanks