Fundamental Concepts for Time Series: Part II

Week II: Video 4

STAT 485/685, Fall 2020, SFU

Sonja Isberg

In This Video

1. Review of last week's concepts, with extra examples:
    - Stochastic process
    - Mean function
    - Autocovariance function
    - Autocorrelation function
2. White Noise process, and deriving some properties for it
3. Random Walk process, and deriving some properties for it

Review: The Big Picture

A **stochastic process** is a collection of random variables indexed in some way:

$$\{Y_t : t = 0, \pm 1, \pm 2, \pm 3, \ldots\}$$

$$\text{i.e., } \{\ldots, Y_{-3}, Y_{-2}, Y_{-1}, Y_0, Y_1, Y_2, Y_3, \ldots\}$$

Suppose:

- The indices are time points (e.g., days, months or years).
- We have a set of n observations, at the times $t = 1, 2, 3, \ldots, n$, i.e., $\{Y_1, Y_2, Y_3, \ldots, Y_n\}$.

This set of n observations from the stochastic process is our **time series**.

Review: The Big Picture

A **stochastic process** is a collection of random variables indexed in some way:

$$\{Y_t : t = 0, \pm 1, \pm 2, \pm 3, \ldots\}$$

$$\text{i.e., } \{\ldots, Y_{-3}, Y_{-2}, Y_{-1}, Y_0, Y_1, Y_2, Y_3, \ldots\}$$

Suppose:

- The indices are time points (e.g., days, months or years).
- We have a set of n observations, at the times $t = 1, 2, 3, \ldots, n$, i.e., $\{Y_1, Y_2, Y_3, \ldots, Y_n\}$.

This set of n observations from the stochastic process is our **time series**.

Example: Monthly precipitation in Vancouver, Canada for each month of 2020 (in mm)[3]: 257, 91, 41, 23, 56, 53, 27, 47.

This set of numbers is the observed time series, with $n = 8$ observations. It comes from some stochastic process for monthly precipitation $\{Y_t\}$, where t is an index for the month.

Review: Properties of the Stochastic Process

The stochastic process $\{Y_t\}$ is a set of random variables. Just as with any set of random variables, there is a lot that can be said about $\{Y_t\}$:

1. Each of the Y_t's has some (marginal) probability distribution.

   Example: The expected monthly precipitation is higher for January than for July.

   The variance of monthly precipitation for March is higher than for November.

   They may even follow entirely different distributions: The March monthly precipitation might come from a Normal distribution, while the November monthly precipitation might come from a Chi-square distribution.

Review: Properties of the Stochastic Process

The stochastic process $\{Y_t\}$ is a set of random variables. Just as with any set of random variables, there is a lot that can be said about $\{Y_t\}$:

1. Each of the Y_t's has some (marginal) probability distribution.

   Example: The expected monthly precipitation is higher for January than for July.

   The variance of monthly precipitation for March is higher than for November.

   They may even follow entirely different distributions: The March monthly precipitation might come from a Normal distribution, while the November monthly precipitation might come from a Chi-square distribution.

2. Any combination of two or more Y_t's will have a joint distribution.

   Example: The correlation between monthly precipitation for January & February is higher than for January & August.

Review: Properties of the Stochastic Process

The stochastic process $\{Y_t\}$ is a set of random variables. Just as with any set of random variables, there is a lot that can be said about $\{Y_t\}$:

1. Each of the Y_t's has some (marginal) probability distribution.

   Example: The expected monthly precipitation is higher for January than for July.

   The variance of monthly precipitation for March is higher than for November.

   They may even follow entirely different distributions: The March monthly precipitation might come from a Normal distribution, while the November monthly precipitation might come from a Chi-square distribution.

2. Any combination of two or more Y_t's will have a joint distribution.

   Example: The correlation between monthly precipitation for January & February is higher than for January & August.

We don't need to know everything about each of these distributions. We focus on some basic properties: the means, variances and covariances.

Review: Mean Function of a Stochastic Process

The **mean function** of a stochastic process $\{Y_t : t = 0, \pm 1, \pm 2, \pm 3, \dots\}$ is

$$\mu_t = E(Y_t), \quad \text{for any } t = 0, \pm 1, \pm 2, \pm 3, \dots$$

The subscript t indicates that the mean may be different at different times t.

Example: Estimate of the mean monthly precipitation for January (based on 30 years of data): $\mu_1 = E(Y_1) = 179$ mm.[4]

For July: $\mu_7 = E(Y_7) = 53$ mm.

Review: Autocovariance Function of a Stochastic Process

The **autocovariance function** of a stochastic process is

$$\begin{aligned} \gamma_{t,s} &= Cov(Y_t, Y_s), \quad \text{for any } t \;\&\; s \\ &= E\left[(Y_t - \mu_t)(Y_s - \mu_s)\right] \\ &= E(Y_t Y_s) - \mu_t \mu_s. \end{aligned}$$

This is just the covariance between any two Y's in the same stochastic process.

Review: Autocovariance Function of a Stochastic Process

The **autocovariance function** of a stochastic process is

$$\begin{aligned}\gamma_{t,s} &= Cov(Y_t, Y_s), \quad \text{for any } t \,\&\, s \\ &= E\left[(Y_t - \mu_t)(Y_s - \mu_s)\right] \\ &= E(Y_t Y_s) - \mu_t \mu_s.\end{aligned}$$

This is just the covariance between any two Y's in the same stochastic process.

Note:

$$\gamma_{t,t} = Var(Y_t)$$

$$\gamma_{t,s} = \gamma_{s,t}$$

If Y_t and Y_s are independent: $\gamma_{t,s} = 0$.

Review: Autocorrelation Function of a Stochastic Process

The **autocorrelation function** of a stochastic process is

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

This is just the correlation between any two Y's in the same stochastic process.

It is a measure of the *strength* and *direction* of their linear relationship.

Review: Autocorrelation Function of a Stochastic Process

The **autocorrelation function** of a stochastic process is

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

This is just the correlation between any two Y's in the same stochastic process.

It is a measure of the *strength* and *direction* of their linear relationship.

Note:

$$-1 \leq \rho_{t,s} \leq 1$$

$$\rho_{t,t} = 1$$

$$\rho_{t,s} = \rho_{s,t}$$

White Noise

The first type of stochastic process we will learn about is the **white noise process**. This will be our most important building block for other, more complicated processes.

Definition: A sequence of independent, identically distributed random variables.

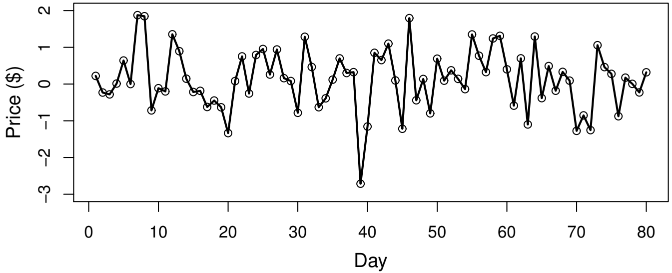
We usually use the notation $\{e_t\}$ when referring to a white noise process (instead of $\{Y_t\}$).

Properties:

- The random variables are independent, so: For any $t \neq s$, $\gamma_{t,s} = Cov(e_t, e_s) = 0$.
- Each e_t has the same distribution.
- For any t: $\mu_t = E(e_t) =$ (some constant value). Usually we assume: $\mu_t = E(e_t) = 0$.
- For any t: $\gamma_{t,t} = Var(E_t) =$ (some constant value). We denote this value by σ_e^2.

White Noise Example

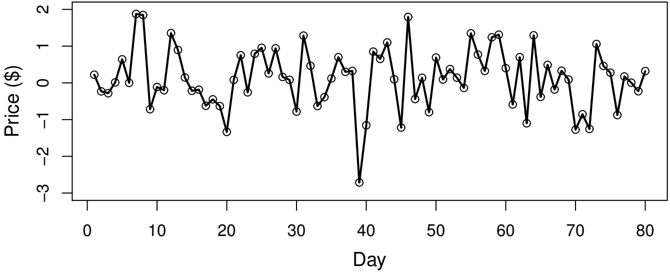
Example: White noise process, where the e_t's are independent and $N(0,1)$ (i.e, $\sigma_e^2 = 1$):



1. What is the mean function for this process?

2. What is the autocovariance function for this process?

White Noise Example (cont'd)



3. What is the autocorrelation function for this process?

Random Walk

The **random walk** is another very common type of stochastic process.

Definition: Let $e_1, e_2, e_3, \ldots$ be a white noise process with mean 0 and variance σ_e^2. The random walk process is constructed as:

$$\left.\begin{aligned} Y_1 &= e_1 \\ Y_2 &= Y_1 + e_2 \\ Y_3 &= Y_2 + e_3 \\ &\vdots \end{aligned}\right\}$$

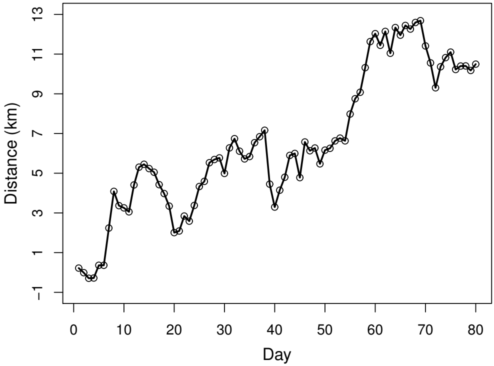
In general, for any t: $Y_t = Y_{t-1} + e_t$.

In other words: $Y_t = e_1 + e_2 + \ldots + e_t$.

Random Walk Example

Suppose a herd of elephants usually tends to stay near a specific lake, but sometimes wanders around. Researchers place a GPS tracker on the elephants to track their movement. Each day, the elephants wander a random distance from the lake (sometimes forward, sometimes backward). This random distance is independent and identically distributed each day, so it is a white noise process.

At the end of day t, the elephants' total distance from the lake will be governed by a **random walk**:



Random Walk Properties

1. What is the mean function for a random walk process?

2. What is the variance for a random walk process?
   Hint: If X, Y and Z are independent:
   $$Var(aX + bY + cZ) = a^2 Var(X) + b^2 Var(Y) + c^2 Var(Z).$$

Random Walk Properties

1. What is the mean function for a random walk process?

2. What is the variance for a random walk process?
   Hint: If X, Y and Z are independent:
   $$Var(aX + bY + cZ) = a^2 Var(X) + b^2 Var(Y) + c^2 Var(Z).$$

Note that the variance increases with t!

Random Walk Properties (cont'd)

3. What is the autocovariance function for a random walk process?

4. What is the autocorrelation function for a random walk process?

Random Walk Properties (cont'd)

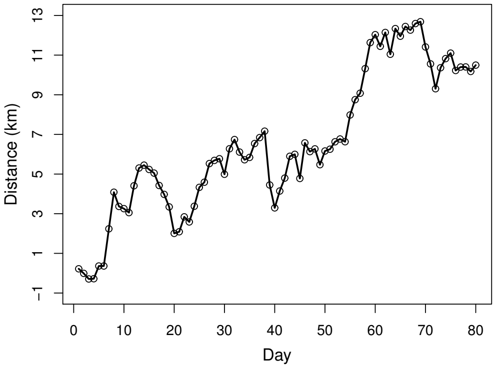
3. What is the autocovariance function for a random walk process?

4. What is the autocorrelation function for a random walk process?

Note: Neighbouring time points become more positively correlated as time goes by.

Back to the Elephants

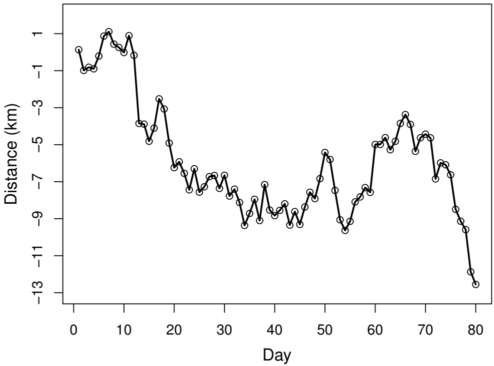
Recall the elephant random walk example:



For a random walk, $\mu_t = 0$ for all t. However, these observations are clearly deviating from 0 as time goes on. Why is this?

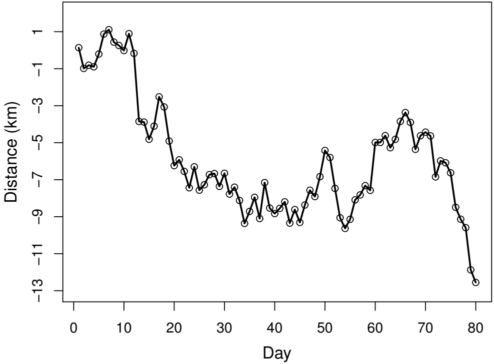
Back to the Elephants (cont'd)

Note: A different observation of the same random walk may have looked something like this:



Back to the Elephants (cont'd)

Note: A different observation of the same random walk may have looked something like this:



For a random walk, the variance is increasing as time goes on, and points are becoming more and more strongly positively correlated.

Final Comments

That's all for now!

In this video, we've reviewed some important concepts of stochastic processes, and we've practiced applying them to two types of processes: the White Noise and the Random Walk.

Coming Up Next: The concept of **stationarity**.

References

[1] Chan, K. S., & Ripley, B. (2020). TSA: Time Series Analysis. R package version 1.2.1.

[2] Cryer, J. D., & Chan, K. S. (2008). *Time series analysis: with applications in R.* Springer Science and Business Media.

[3] *Total Precipitation - Monthly data for Vancouver.* Canada Weather Stats. Retrieved from `https://vancouver.weatherstats.ca/charts/precipitation-monthly.html`.

[4] *Climate of Vancouver.* Wikipedia. Retrieved from `https://en.wikipedia.org/wiki/Climate_of_Vancouver`.