Fundamental Concepts for Time Series: Stationarity

Week II: Video 5

STAT 485/685, Fall 2020, SFU

Sonja Isberg

In This Video

1. Some motivating ideas behind stationarity
2. Two types of stationarity
   - Strict stationarity
   - Weak (or second-order) stationarity
3. Testing for stationarity, in a White Noise process and a Random Walk process

Review: Stochastic Process Properties

Recall: A stochastic process $\{Y_t\}$ is a set of random variables indexed in some way (in our case, by time). 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 values of the mean μ_t and variance $\gamma_{t,t}$ may be different at different times t.

   The different Y_t's may even come from entirely different distributions (e.g., Normal vs. Chi-squared).

Review: Stochastic Process Properties

Recall: A stochastic process $\{Y_t\}$ is a set of random variables indexed in some way (in our case, by time). 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 values of the mean μ_t and variance $\gamma_{t,t}$ may be different at different times t.

   The different Y_t's may even come from entirely different distributions (e.g., Normal vs. Chi-squared).

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

   Example: The joint distribution of Y_t and Y_s may depend on the two times t and s.

   e.g., $\gamma_{t,s} = Cov(Y_t, Y_s)$ may be a function of t and s.

We saw this with the example of monthly precipitation in Vancouver (see Video 4 for examples).

Ideas Behind Stationarity

In order to be able to conduct statistical analyses, we have to make some assumptions about the process $\{Y_t\}$.

One of the most important assumptions, which we will often be making, is that the process is **stationary**.

The idea behind stationarity: The probability laws that govern the behaviour of the process do not change over time.[2]

Two types of stationarity: **strict stationarity** and **weak (or second-order) stationarity**.

We will learn both, but throughout the rest of the course "stationarity" will usually be used to refer to weak stationarity.

Strict Stationarity

Definition: A stochastic process $\{Y_t\}$ is said to be **strictly stationary** if the joint distribution of $Y_{t_1}, Y_{t_2}, \ldots, Y_{t_m}$ is the same as the joint distribution of $Y_{t_1 - k}, Y_{t_2 - k}, \ldots, Y_{t_m - k}$, for any set of time points $t_1, t_2, \ldots, t_m$ and any lag k.

In other words: No matter how much you move forwards or backwards in time (k), the process still "behaves the same way".

Strict Stationarity

Definition: A stochastic process $\{Y_t\}$ is said to be **strictly stationary** if the joint distribution of $Y_{t_1}, Y_{t_2}, \ldots, Y_{t_m}$ is the same as the joint distribution of $Y_{t_1-k}, Y_{t_2-k}, \ldots, Y_{t_m-k}$, for any set of time points $t_1, t_2, \ldots, t_m$ and any lag k.

In other words: No matter how much you move forwards or backwards in time (k), the process still "behaves the same way".

For m=1: For any time t: The distribution of Y_t is the same as the distribution of Y_{t-k}, for any lag k.
i.e., Each of the Y_t's are identically distributed.
Where have we seen such a process before?

Strict Stationarity

Definition: A stochastic process $\{Y_t\}$ is said to be **strictly stationary** if the joint distribution of $Y_{t_1}, Y_{t_2}, \ldots, Y_{t_m}$ is the same as the joint distribution of $Y_{t_1-k}, Y_{t_2-k}, \ldots, Y_{t_m-k}$, for any set of time points $t_1, t_2, \ldots, t_m$ and any lag k.

In other words: No matter how much you move forwards or backwards in time (k), the process still "behaves the same way".

For m=1: For any time t: The distribution of Y_t is the same as the distribution of Y_{t-k}, for any lag k.
i.e., Each of the Y_t's are identically distributed.
Where have we seen such a process before?

For m=2: For any times t and s: The joint distribution of Y_t and Y_s is the same as the joint distribution of Y_{t-k} and Y_{s-k}, for any lag k.
i.e., The way that Y_t and Y_s are related depends only on how close the times are, not what the times actually are.

Strict Stationarity

Definition: A stochastic process $\{Y_t\}$ is said to be **strictly stationary** if the joint distribution of $Y_{t_1}, Y_{t_2}, \ldots, Y_{t_m}$ is the same as the joint distribution of $Y_{t_1-k}, Y_{t_2-k}, \ldots, Y_{t_m-k}$, for any set of time points $t_1, t_2, \ldots, t_m$ and any lag k.

In other words: No matter how much you move forwards or backwards in time (k), the process still "behaves the same way".

For m=1: For any time t: The distribution of Y_t is the same as the distribution of Y_{t-k}, for any lag k.
i.e., Each of the Y_t's are identically distributed.
Where have we seen such a process before?

For m=2: For any times t and s: The joint distribution of Y_t and Y_s is the same as the joint distribution of Y_{t-k} and Y_{s-k}, for any lag k.
i.e., The way that Y_t and Y_s are related depends only on how close the times are, not what the times actually are.

This is a very strict assumption!

Weak (or Second-Order) Stationarity

For most results in this course to hold, we only need the following weaker form of stationarity:

Definition: A stochastic process $\{Y_t\}$ is said to be **weakly** (or **second-order**) **stationary** if *both* of the following are true:

1. The mean of each Y_t is the same.
2. The covariance between any Y_t and Y_s depends only on how far apart the times are.

Weak (or Second-Order) Stationarity

For most results in this course to hold, we only need the following weaker form of stationarity:

Definition: A stochastic process $\{Y_t\}$ is said to be **weakly** (or **second-order**) **stationary** if *both* of the following are true:

1. The mean of each Y_t is the same.
2. The covariance between any Y_t and Y_s depends only on how far apart the times are.

This is called *weak* stationarity because its assumptions are weaker than for strict stationarity. Why?

- It only focuses on marginal distributions ($m = 1$), and joint distributions between 2 variables ($m = 2$).
- It only pertains to the means and covariances, not other aspects of the distributions.

If a process is strictly stationary, this implies that it must also be weakly stationary!

Understanding Weak (or Second-Order) Stationarity

Let's go over the definition a bit more carefully. The two conditions are:

Understanding Weak (or Second-Order) Stationarity

Let's go over the definition a bit more carefully. The two conditions are:

1. The mean of each Y_t is the same.

Understanding Weak (or Second-Order) Stationarity

Let's go over the definition a bit more carefully. The two conditions are:

1. The mean of each Y_t is the same.

   i.e.., $E(Y_1) = E(Y_2) = E(Y_3) = \dots$

Understanding Weak (or Second-Order) Stationarity

Let's go over the definition a bit more carefully. The two conditions are:

1. The mean of each Y_t is the same.

   i.e.., $E(Y_1) = E(Y_2) = E(Y_3) = \dots$

   i.e., The mean function μ_t is always the same value, for all t.

Understanding Weak (or Second-Order) Stationarity

Let's go over the definition a bit more carefully. The two conditions are:

1. The mean of each Y_t is the same.

   i.e.., $E(Y_1) = E(Y_2) = E(Y_3) = \dots$

   i.e., The mean function μ_t is always the same value, for all t.

   i.e., Once you derive μ_t, it should be "constant with respect to t". In other words, there should be no t in the expression.

Understanding Weak (or Second-Order) Stationarity (cont'd)

2. The covariance between any Y_t and Y_s depends only on how far apart the times are.

Understanding Weak (or Second-Order) Stationarity (cont'd)

2. The covariance between any Y_t and Y_s depends only on how far apart the times are.

   i.e., $Cov(Y_1, Y_2) = Cov(Y_2, Y_3) = Cov(Y_3, Y_4) = \ldots$ (all 1 unit apart)
   $Cov(Y_1, Y_3) = Cov(Y_2, Y_4) = Cov(Y_3, Y_5) = \ldots$ (all 2 units apart)
   ... etc.

Understanding Weak (or Second-Order) Stationarity (cont'd)

2. The covariance between any Y_t and Y_s depends only on how far apart the times are.

   i.e., $Cov(Y_1, Y_2) = Cov(Y_2, Y_3) = Cov(Y_3, Y_4) = \ldots$ (all 1 unit apart)
   $Cov(Y_1, Y_3) = Cov(Y_2, Y_4) = Cov(Y_3, Y_5) = \ldots$ (all 2 units apart)
   $\ldots$ etc.

   i.e., The covariance function $\gamma_{t,s}$ depends only on the distance between t and s, not on their actual values.

Understanding Weak (or Second-Order) Stationarity (cont'd)

2. The covariance between any Y_t and Y_s depends only on how far apart the times are.

   i.e., $Cov(Y_1, Y_2) = Cov(Y_2, Y_3) = Cov(Y_3, Y_4) = \ldots$ (all 1 unit apart)
   $Cov(Y_1, Y_3) = Cov(Y_2, Y_4) = Cov(Y_3, Y_5) = \ldots$ (all 2 units apart)
   $\ldots$ etc.

   i.e., The covariance function $\gamma_{t,s}$ depends only on the distance between t and s, not on their actual values.

   i.e., Once you derive $\gamma_{t,s}$, it should not be a function of t or s. They may show up as $k = |t - s|$, or not at all.

Understanding Weak (or Second-Order) Stationarity (cont'd)

2. The covariance between any Y_t and Y_s depends only on how far apart the times are.

   i.e., $Cov(Y_1, Y_2) = Cov(Y_2, Y_3) = Cov(Y_3, Y_4) = \ldots$ (all 1 unit apart)
   $Cov(Y_1, Y_3) = Cov(Y_2, Y_4) = Cov(Y_3, Y_5) = \ldots$ (all 2 units apart)
   $\ldots$ etc.

   i.e., The covariance function $\gamma_{t,s}$ depends only on the distance between t and s, not on their actual values.

   i.e., Once you derive $\gamma_{t,s}$, it should not be a function of t or s. They may show up as $k = |t - s|$, or not at all.

   Note: For any lag k:

   $$Cov(Y_1, Y_{1-k}) = Cov(Y_2, Y_{2-k}) = Cov(Y_3, Y_{3-k}) = \ldots$$

   So, we introduce the new notation: $\gamma_k = Cov(Y_t, Y_{t-k})$.

Some New Notation

We just saw that, if a process is stationary:

For any lag k:

$$Cov(Y_1, Y_{1-k}) = Cov(Y_2, Y_{2-k}) = Cov(Y_3, Y_{3-k}) = \dots$$

This allows us to introduce some short-hand notation:

Some New Notation

We just saw that, if a process is stationary:

For any lag k:

$$Cov(Y_1, Y_{1-k}) = Cov(Y_2, Y_{2-k}) = Cov(Y_3, Y_{3-k}) = \dots$$

This allows us to introduce some short-hand notation:

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

$$\gamma_0 = Var(Y_t)$$

Some New Notation

We just saw that, if a process is stationary:

For any lag k:

$$Cov(Y_1, Y_{1-k}) = Cov(Y_2, Y_{2-k}) = Cov(Y_3, Y_{3-k}) = \dots$$

This allows us to introduce some short-hand notation:

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

$$\gamma_0 = Var(Y_t)$$

Also:

$$\rho_k = Corr(Y_t, Y_{t-k}) = \frac{\gamma_k}{\gamma_0}$$

$$\rho_0 = Corr(Y_t, Y_t) = 1$$

Some New Notation

We just saw that, if a process is stationary:

For any lag k:

$$Cov(Y_1, Y_{1-k}) = Cov(Y_2, Y_{2-k}) = Cov(Y_3, Y_{3-k}) = \dots$$

This allows us to introduce some short-hand notation:

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

$$\gamma_0 = Var(Y_t)$$

Also:

$$\rho_k = Corr(Y_t, Y_{t-k}) = \frac{\gamma_k}{\gamma_0}$$

$$\rho_0 = Corr(Y_t, Y_t) = 1$$

Some properties from before:

$$\gamma_k = \gamma_{-k}$$

$$\rho_k = \rho_{-k}$$

$$-1 \leq \rho_k \leq 1$$

Is a White Noise Stationary?

Recall: A white noise is a sequence of independent, identically distributed random variables, usually denoted by $\{e_t\}$.

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.

Is a White Noise Stationary?

Let's check the conditions for weak stationarity:

1. **Check:** The mean function does not depend on t.

2. **Check:** The covariance between any e_t and e_{t-k} does not depend on t.

Is a Random Walk Stationary?

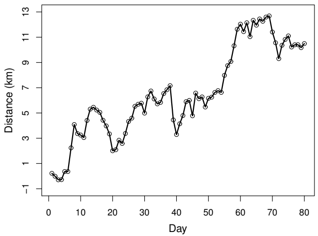
Recall: Let $e_1, e_2, e_3, \ldots$ be a white noise process with mean 0 and variance σ_e^2.

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

Recall the elephant example:



Is a Random Walk Stationary?

Let's check the conditions for weak stationarity:

1. **Check:** The mean function does not depend on t.

2. **Check:** The covariance between any Y_t and Y_{t-k} does not depend on t.

Final Comments

That's all for now!

In this video, we've introduced the important concept of stationarity, and tested whether it applies to White Noise processes and Random Walk processes.

In Chapter 5, we'll learn about a technique that can sometimes be used to turn a non-stationary process into a stationary one: *differencing*.

Coming Up Next: A few more examples, and the end of Chapter 2.

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.