Part 1: Stationary Time Series

Instructor: Ma, Jun

School of Economics, Renmin University of China

March 6, 2018

Time Series and Stochastic Process

Definition 1 (Time Series)

1. A collection of data (observations), collected at every $t \in T$, where T is an index set, e.g., $T = [a, b]$ with continuous observations or $T = \{t_1, t_2, ...\}$ with discrete-time observations.
2. The mathematical model that describes how the data is generated.

A stochastic process is a collection of random variables (or vectors) defined ona common probability space $(\Omega, \mathscr{F}, \mathrm{P})$, denoted by $\{X_t : t \in T\}$ with an index set T.
For a single $\omega \in \Omega$, the function t$\mapsto X_t\,(\omega) : T \to \mathbb{R}$ is called a sample path.

Time Series and Stochastic Process

Remark 2
For $T = \{1, ..., n\}$, our data $\{x_1, ..., x_n\}$, starting being observed at time 1 and stop at time n.
We use a stochastic process to describe how the data is generated, $\{X_t\}_{t \in \mathbb{Z}}$, i.e., a process that runs from infinite past to infinite future. The process run before we start observing the output (i.e., before $t = 1$) and will continue after we stop oberving at $t = n$.
Our data $\{x_1, ..., x_n\}$ is a realization of the process. I.e.,

$$\{x_1, ..., x_n\} = \{X_1(\omega), X_2(\omega), \cdots, X_n(\omega)\}.$$

Notice that $\{\cdots, X_{-1}(\omega), X_0(\omega)\}$ and $\{X_{n+1}(\omega), X_{n+2}(\omega), \cdots\}$ are unobserved.

Examples

- (I.I.D. Process) If $\{X_t\}_{t\in\mathbb{Z}}$ is a sequence of i.i.d. random variables, we often write $\{X_t\}_{t\in\mathbb{Z}} \sim \mathrm{i.i.d.}$ or $\{X_t\}_{t\in\mathbb{Z}} \sim \mathrm{i.i.d.}\left(\mu, \sigma^2\right)$ to indicate $\mathrm{E}\left[X_1\right] = \mu$ and $\mathrm{Var}\left[X_1\right] = \sigma^2$. X_s and X_t (for all $s \neq t$) are independent and also $\mathrm{P}\left[X_t \leq x\right] = \mathrm{P}\left[X_s \leq x\right]$, for all $x \in \mathbb{R}$
- (Random Walk) $\{Z_t\}_{t\in\mathbb{Z}}$ is a sequence of i.i.d. random variables. Let

$$X_t := \begin{cases} 0 & \text{if } t = 0 \\ \sum_{k=1}^{t} Z_k & \text{if } t \geq 1. \end{cases}$$

  The simplest case corresponds to

$$Z_t = \begin{cases} 1 & \text{with probability } \frac{1}{2} \\ -1 & \text{with probability } \frac{1}{2}. \end{cases}$$

Examples

- (Moving Average) $\{Z_t\}_{t\in\mathbb{Z}} \sim \mathrm{i.i.d.}\left(\mu, \sigma^2\right)$.

$$X_t := a_1 Z_t + a_2 Z_{t-1} + \cdots + a_1 Z_{t-q-1},$$

  for $(a_1, ..., a_q) \in \mathbb{R}^q$.
- (AR(1)) $\{Z_t\}_{t\in\mathbb{Z}} \sim \mathrm{i.i.d.}\left(\mu, \sigma^2\right)$.

$$X_t := \rho X_{t-1} + Z_t,\ \ \rho \in \mathbb{R}.$$

- (Signal + Noise) $\{Z_t\}_{t\in\mathbb{Z}} \sim \mathrm{i.i.d.}\left(\mu, \sigma^2\right)$.

$$X_t := \mu\left(t\right) + Z_t,$$

  where μ is a deterministic function.
- (Brownian Motion) A Brownian motion is a stochastic process with index set $T = [0, 1]$, such that (1). $\mathrm{P}\left[X\left(0\right) = 0\right] = 1$, (2). The process has independent increments, i.e., if $0 \leq t_1 \leq t_2 \leq \cdots \leq t_m$, then $X\left(t_1\right) - X\left(t_0\right), \cdots, X\left(t_m\right) - X\left(t_{m-1}\right)$ are independent. (3). For $t_2 > t_1 \geq 0$, $X\left(t_2\right) - X\left(t_1\right) \sim \mathrm{N}\left(0, t_2 - t_1\right)$.

Definitions

Definition 3 (White Noise)

A process $\{Z_t\}_{t\in\mathbb{Z}}$ is called a white noise if: (1). $\mathrm{Var}[Z_t] = \sigma^2 < \infty$ for all $t \in \mathbb{Z}$, (2). $\mathrm{E}[Z_t] = \mu$, for all $t \in \mathbb{Z}$ and (3). $\mathrm{Cov}[Z_t, Z_{t+h}] = 0$, for all $t \in \mathbb{Z}$ and $h \neq 0$. Write $\{Z_t\}_{t\in\mathbb{Z}} \sim \mathrm{WN}(\mu, \sigma^2)$.

Let $\{X_t\}_{t\in T}$ be a stochastic process, the finite-dimensional distribution (f.d.d.) for some $t_1 < t_2 < \cdots < t_n$ is the joint distribution of $(X_{t_1}, ..., X_{t_n})$.

A stochastic process $\{X_t\}_{t\in T}$ is called Gaussian if all of its f.d.d. are multivariate normal.

Definitions

Definition 4 (Autocovariance Function)

For a stochastic process $\{X_t\}_{t\in T}$, for which $\mathrm{Var}[X_1] < \infty$ (if and only if $\mathrm{E}\left[X_1^2\right] < \infty$), for all $t \in T$, the autocovariance function (ACF) $\gamma(\cdot,\cdot)$ of $\{X_t\}_{t\in T}$ is given by

$$\gamma(s,t) := \mathrm{Cov}[X_s, X_t] = \mathrm{E}\left[(X_s - \mathrm{E}[X_s])(X_t - \mathrm{E}[X_t])\right], \ (s,t) \in T^2.$$

The stochastic process $\{X_t\}_{t\in T}$ is said to be weakly stationary if (1). $\mathrm{E}\left[X_t^2\right] < \infty$ for all $t \in \mathbb{Z}$, (2). $\mathrm{E}[X_t] = \mu$, all $t \in \mathbb{Z}$ (i.e., the expectation is constant) and (3). $\gamma(s,t) = \gamma(s+r, t+r)$ for all $r \in \mathbb{Z}$ and $(s,t) \in \mathbb{Z}^2$.

Definitions

Remark 5

Apparently, $\gamma(s,t) = \gamma(t,s)$ for all $(s,t) \in \mathbb{Z}^2$. When the process is weakly stationary,

$$\gamma(0, s-t) = \gamma(t,s) = \gamma(t-s, 0) = \gamma(0, t-s).$$

And the ACF of a weakly stationary process can be expressed as a one-parameter function $\gamma(h)$ with $h = |s-t|$.

Definition 6 (Autocorrelation Function)

The autocorrelation function is

$$\rho(h) := \frac{\gamma(h)}{\gamma(0)} = \frac{\mathrm{Cov}[X_{t+h}, X_t]}{\mathrm{Var}[X_t]} = \mathrm{Corr}[X_{t+h}, X_t].$$

Remark 7

Notice that $\rho(h) = 0$ does not generally imply that X_t is independent of X_{t+h}.

Definitions

Definition 8 (Strict Stationarity)
A stochastic process $\{X_t\}_{t\in\mathbb{Z}}$ is strictly stationary if

$$F_{X_{t_1}\cdots X_{t_n}} = F_{X_{t_1+h}\cdots X_{t_n+h}}, \text{ for all } h \in \mathbb{Z}, (t_1, ..., t_n) \in \mathbb{Z}^n \text{ and } n \in \mathbb{N}$$

or

$$\mathrm{P}\left[X_{t_1} \leq x_1, \cdots, X_{t_n} \leq x_n\right] = \mathrm{P}\left[X_{t_1+h} \leq x_1, \cdots, X_{t_n+h} \leq x_n\right]$$

for all $(x_{t_1}, ..., x_{t_n}) \in \mathbb{R}^n$, $h \in \mathbb{Z}$, $(t_1, ..., t_n) \in \mathbb{Z}^n$ and $n \in \mathbb{N}$.

Remark 9
Example 1: An i.i.d. process is strictly stationary. Example 2: For a random variable X, define $X_t = X$, for all $t \in \mathbb{Z}$. $\{X_t\}_{t\in\mathbb{Z}}$ is strictly stationary.

Definitions

Remark 10 (Relation between weak and strict stationarity)

If $\{X_t\}_{t\in\mathbb{Z}}$ is strictly stationary, the distribution F_{X_t} of X_t is the same for all $t \in \mathbb{Z}$. Any pair (X_t, X_{t+h}) has a joint distribution $F_{X_t,X_{t+h}}$ that does not depend on t. Hence, $\mathrm{E}[X_t] = \int_{\mathbb{R}} x \mathrm{d}F_{X_t}(x)$ is independent of t and

$$\mathrm{Cov}[X_t, X_{t+h}] = \int_{\mathbb{R}^2} (x-\mu)(y-\mu)\,\mathrm{d}F_{X_t,X_{t+h}}(x,y),$$

where $\mu := \mathrm{E}[X_t]$ is independent of t. So strict stationarity $\Longrightarrow$ weak stationarity.

One important case when weak stationarity implies strict stationarity is that when the process is Gaussian. When $\{X_t\}_{t\in\mathbb{Z}}$ is a Gaussian process, weak stationarity of $\{X_t\}_{t\in\mathbb{Z}}$ implies its strict stationarity since $(X_{t_1}, ..., X_{t_n})$ and $(X_{t_1+h}, ..., X_{t_n+h})$ have the same mean and covariance matrix and hence the same distribution for all $h \in \mathbb{Z}$, $(t_1, ..., t_n) \in \mathbb{Z}^n$ and $n \in \mathbb{N}$.

Properties of ACF

Proposition 11

If $\gamma(\cdot)$ *is the ACF of a stationary stochastic process* $\{X_t\}_{t\in\mathbb{Z}}$ *, then*
(1). $\gamma(0) \geq 0$
(2). $|\gamma(h)| \leq \gamma(0)$ *for all* $h \in \mathbb{Z}$
(3). $\gamma(h) = \gamma(-h)$ *for all* $h \in \mathbb{Z}$*.*

Definition 12 (Positive Semi-Definite)

A real-valued function $\kappa : \mathbb{Z} \to \mathbb{R}$ positive semi-definite (p.s.d.) if

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \kappa(t_i - t_j) \geq 0$$

for all $(a_1, ..., a_n) \in \mathbb{R}^n$, $(t_1, ..., t_n) \in \mathbb{Z}^n$ and $n \in \mathbb{N}$. I.e., fixing any$(t_1, ..., t_n) \in \mathbb{Z}^n$, the $n \times n$ matrix where the (i, j) −th element is $\kappa(t_i - t_j)$ is p.s.d.

Properties of ACF

Theorem 13 (Characterization of ACF)

A function $\gamma : \mathbb{Z} \to \mathbb{R}$ *is the ACF of a stationary stochastic process* $\{X_t\}_{t \in \mathbb{Z}}$ *if and only if* γ *is an even function, i.e.,* $\gamma(h) = \gamma(-h)$ *for all* $h \in \mathbb{Z}$ *and p.s.d.*

Sample Estimates

- Assume that we have a stationary time series process $\{X_t\}_{t\in\mathbb{Z}}$ with $\mathrm{E}[X_t] = \mu$ for all $t \in \mathbb{Z}$ and ACF $\gamma(h)$, $h \in \mathbb{Z}$. In practice, μ and γ must be estimated from the data.

Definition 14 (Sample mean and Sample ACF)

Sample Mean:

$$\overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i$$

and sample ACF:

$$\widehat{\gamma}(h) := \frac{1}{n}\sum_{k=1}^{n-|h|}\left(X_{k+|h|} - \overline{X}_n\right)\left(X_k - \overline{X}_n\right).$$

Asymptotic Normality

Definition 15

Let $Y_1, Y_2, \ldots$ be a sequence of random variables. They are said to be asymptotically normal denoted as $Y_n \overset{a}{\sim} \mathrm{N}\left(\mu_n, \sigma_n^2\right)$ if $\mu_n = \mathrm{E}\left[Y_n\right]$, $\sigma_n^2 = \mathrm{Var}\left[Y_n\right]$ for each $n \in \mathbb{N}$,

$$\lim_{n\to\infty} \mathrm{P}\left[\frac{Y_n - \mu_n}{\sigma_n} \leq x\right] = \Phi\left(x\right), \text{ for all } x \in \mathbb{R},$$

where Φ is the $\mathrm{N}\left(0,1\right)$ CDF. In other words, $\frac{Y_n - \mu_n}{\sigma_n} \to_d \mathrm{N}\left(0,1\right)$.

Remark 16

For sample, let $X_1, X_2, \ldots$ be i.i.d., $\mu = \mathrm{E}\left[X_1\right]$ and $\sigma^2 = \mathrm{Var}\left[X_1\right]$. Let

$$Y_n := \frac{1}{n}\sum_{j=1}^{n} X_j,\ \mu_n := \mathrm{E}\left[Y_n\right] = \mu,\ \sigma_n := \mathrm{Var}\left[Y_n\right] = \frac{1}{n}\sigma^2.$$

Then, by standard CLT, $\frac{Y_n - \mu_n}{\sigma_n} = \frac{\sqrt{n}(Y_n - \mu)}{\sigma} \to_d \mathrm{N}\left(0,1\right)$.

Properties of Sample Mean and Sample ACF

- We can show that the sample mean and sample ACF are consistent and asymptotically normal. For now, we show only consistency of the sample mean.
- An estimator $\widehat{\theta}_n$ of some parameter θ, n is the sample size.

$$\text{consistency} \begin{cases} \text{(weak)} & \forall \epsilon > 0, \ \lim\limits_{n\to\infty} \mathrm{P}\left[\left|\widehat{\theta}_n - \theta\right| > \epsilon\right] = 0 \\ \text{(strong)} & \mathrm{P}\left(\left\{\omega \in \Omega : \lim\limits_{n\to\infty} \widehat{\theta}_n(\omega) = \theta\right\}\right) = 1 \\ \text{(MSE)} & \lim\limits_{n\to\infty} \mathrm{E}\left[\left(\widehat{\theta}_n - \theta\right)^2\right] = 0. \end{cases}$$

- MSE consistency $\Longrightarrow$ Weak consistency:
$\mathrm{P}\left[\left|\left|\widehat{\theta}_n - \theta\right| > \epsilon\right|\right] \leq \frac{\mathrm{E}\left[\left(\widehat{\theta}_n - \theta\right)^2\right]}{\epsilon^2} \xrightarrow{n\to\infty} 0$, for all $\epsilon > 0$.
- Strong consistency $\Longrightarrow$ Weak consistency: the proof is difficult.

Properties of Sample Mean and Sample ACF

Proposition 17

Suppose that $\{X_t\}_{t\in\mathbb{Z}}$ *is weakly stationary and let* γ *be its ACF and let* μ *be its mean. (1). If* $\gamma(h) \overset{h\to\infty}{\longrightarrow} 0$*, then* $\mathrm{E}\left[\left(\overline{X}_n - \mu\right)^2\right] \overset{n\to\infty}{\longrightarrow} 0$*. (2). If* $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$*,*

$$n \cdot \mathrm{Var}\left[\overline{X}_n\right] = n \cdot \mathrm{E}\left[\left(\overline{X}_n - \mu\right)^2\right] \overset{n\to\infty}{\longrightarrow} \sum_{h=-\infty}^{\infty} \gamma(h).$$

Properties of Sample Mean and Sample ACF

Remark 18
In some books, the condition $\gamma(h) \overset{h\to\infty}{\longrightarrow} 0$ is referred to as "weak dependence". We say that a stationary process is weakly dependent if it satisfies this. Note that stationarity alone does not guarantee consistency of $\overline{X}_n$. Here is a counter-example. Let $\{X_t\}_{t\in\mathbb{Z}}$ be i.i.d. mean 0. Let Z be independent of $\{X_t\}_{t\in\mathbb{Z}}$, mean 0. Define $Y_t := X_t + Z$. The process $\{Y_t\}_{t\in\mathbb{Z}}$ is strictly stationary, mean 0. But

$$\frac{1}{n}\sum_{t=1}^{n} Y_t = \frac{1}{n}\sum_{t=1}^{n} X_t + Z \to_p Z \neq \mathrm{E}[Y_1],$$

since $\{Y_t\}_{t\in\mathbb{Z}}$ has too much temporal dependence:

$$\mathrm{Cov}[Y_t, Y_{t+h}] = \mathrm{Var}[Z] \neq 0, \forall h.$$

Properties of Sample Mean and Sample ACF

Remark 19

Suppose that $\{X_t\}_{t\in\mathbb{Z}}$ is weakly stationary and let γ be its ACF and let μ be its mean. If we can show $\overline{X}_n$ is asymptotically normal, i.e.,

$$\frac{\overline{X}_n - \mu}{\sqrt{\mathrm{Var}\left[\overline{X}_n\right]}} \to_d \mathrm{N}\left(0, 1\right).$$

If $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$, since $n \cdot \mathrm{Var}\left[\overline{X}_n\right] \overset{n\to\infty}{\longrightarrow} \sum_{h=-\infty}^{\infty} \gamma(h) =: \tau^2$, we have

$$\sqrt{n}\left(\overline{X}_n - \mu\right) = \sqrt{n \cdot \mathrm{Var}\left[\overline{X}_n\right]} \cdot \frac{\overline{X}_n - \mu}{\sqrt{\mathrm{Var}\left[\overline{X}_n\right]}}$$

and thus $\sqrt{n}\left(\overline{X}_n - \mu\right) \to_d \mathrm{N}\left(0, \tau^2\right)$. We will provide a CLT for some special weakly stationary process later.