Basic Time Series Concepts

January 24, 2019

The material in this set of notes is based on S&S 1.1-1.6.

Suppose we observe an $n \times 1$ vector $\boldsymbol{x} = (x_1, \dots, x_n) = \boldsymbol{\mu}_x + \boldsymbol{w}$, where $\boldsymbol{\mu}_x$ is a fixed but unknown mean, $\boldsymbol{w}$ are random errors and elements of $\boldsymbol{x}$ are ordered in time. We will refer to $\boldsymbol{x}$ as a **time series**, although the sequence of elements can also be called a **stochastic process**.

The joint distribution function of $\boldsymbol{x}$ is

$$F(c_1, \dots, c_n) = P(x_1 \leq c_1, \dots, x_n \leq c_n).$$

Often, this will be difficult to write out and work with, so it does not provide a useful means of characterizing a time series $\boldsymbol{x}$. Instead, we often characterize a time series $\boldsymbol{x}$ via its:

- **Mean Function:** $\mu_{x,t} = \mathbb{E}[x_t] = \int_{-\infty}^{\infty} x f_t(x) dx$, where $f_t(x)$ is the marginal density of x_t having integrated out all other elements of $\boldsymbol{x}$.
- **Autocovariance Function:** $\gamma_x(s,t) = \mathbb{E}[(x_s - \mu_{x,s})(x_t - \mu_{x,t})]$ for all s and t.
  - When $s = t$, gives the variance $\gamma_x(s,s) = \mathbb{V}[x_s]$.
- **Autocorrelation Function:** $\rho_x(s,t) = \gamma(s,t)/\sqrt{\gamma(s,s)\gamma(t,t)}$ for all s and t.

Without further assumptions, this is still an unwieldy way to characterize a time series because the mean function depends on t and the autocovariance and autocorrelation functions depend on both s and t. To simplify things further, we often assume that the time series is either:

- **Strongly Stationary:** The distribution of any subset of k elements of $(x_{t_1}, \ldots, x_{t_k})$ is exactly the same as the distribution of the shifted set of k elements $(x_{t_1+h}, \ldots, x_{t_k+h})$.
  - The mean function $\mu_{x,t}$ does not depend on t: $\mu_{x,t} = \mathbb{E}[x_t] = \mathbb{E}[x_{t+h}] = \mu_{x,t+h}$.
  - The autocovariance function $\gamma_x(s,t)$ depends on s and t only through their absolute difference $h = |s-t|$:

$$\begin{aligned}\gamma(s+h, s) =& \mathbb{E}[(x_{s+h} - \mu_x)(x_s - \mu_x)] \\ =& \mathbb{E}[(x_h - \mu_x)(x_0 - \mu_x)] \\ =& \gamma(h, 0).\end{aligned}$$

- **Weakly Stationary:** The mean function is constant and does not depend on time, $\mu_{x,t} = \mu_x$ and the autocovariance function $\gamma_x(s,t)$ depends on s and t only through their absolute difference $h = |s-t|$.

Note that although strong stationarity implies weak stationarity, the reverse does not hold. Strong stationarity is usually too strict to be a reasonable assumption, so from here on out we will call a time series **stationary** if it is **weakly stationary**.

When a time series is stationary, its autocovariance and autocorrelation functions can be written as functions of a single variable h. For this reason, we will drop the second arguments of the autocovariance and autocorrelation functions when a time series is stationary, writing $\gamma_x(h) = \gamma_x(h, 0)$ and $\rho_x(h) = \rho_x(h, 0)$.

When we observe a time series $\boldsymbol{x}$, we do not know the mean, autocovariance, or autocorrelation functions a priori - we need to estimate them. When $\boldsymbol{x}$ is stationary we can compute:

- The **sample mean** function:

$$\hat{\mu}_x = \bar{x} = \sum_{t=1}^{n} x_t / n. \tag{1}$$

- The **sample autocovariance function**:

$$\hat{\gamma}_x(h) = \frac{1}{n} \sum_{t=1}^{n-h} (x_{t+h} - \hat{\mu}_x)(x_t - \hat{\mu}_x), \tag{2}$$

  with $\hat{\gamma}_x(-h) = \hat{\gamma}_x(h)$ for $h = 0, 1, \ldots, n-1$.

  - We divide by n and not $n - h$ to ensure that the sample variance of a sum of elements of $\boldsymbol{x}$ computed from the $n \times n$ sample autocovariance matrix with entries $\hat{\gamma}(i - j)$ will always be nonnegative.
  - This is a biased estimate of $\gamma_x(h)$.

- The **sample autocorrelation function**:

$$\hat{\rho}_x(h) = \frac{\hat{\gamma}_x(h)}{\hat{\gamma}_x(0)}. \tag{3}$$

When we examine a sample autocorrelation function, it is natural to ask how different our estimates of the sample autocorrelation are from what we would might expect if $\boldsymbol{x}$ were a **white noise** time series with no autocorrelation at all, i.e. if $\rho_x(h) = 0$ for all $h \neq 0$. We can get a handle on this using the following result:

> If $\boldsymbol{x} = \boldsymbol{\mu}_x + \boldsymbol{w}$ where $\boldsymbol{\mu}_x = \mathbf{0}$ and $w_i \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_w^2)$ for $i = 1, \ldots n$, then $\hat{\rho}_x(h) \approx v/\sqrt{n}$, for $h = 1, \ldots H$, where $v \sim \mathcal{N}(0, 1)$ and H is fixed but arbitrary.

This result allows us to perform an approximate test of the null hypothesis that $\rho_x(h) = 0$ for any $h > 1$.