Fundamental Concepts for Time Series: Part I

Week I: Video 3

STAT 485/685, Fall 2020, SFU

Sonja Isberg

Time Series and Stochastic Processes

Recall: A **time series** is data obtained from observations collected sequentially over time.[2]

We model time series using a **stochastic process**: a collection of random variables indexed in some way. We let this stochastic process be indexed by time, and denote it by the set of random variables $\{Y_t : t = 0, \pm 1, \pm 2, \pm 3, \dots\}$.

For example: The time series of annual LA rainfall amounts (in inches) is 21, 17, 19, 5, 11, ... This is an observation of some stochastic process $\{Y_t\}$.

Time Series and Stochastic Processes

Recall: A **time series** is data obtained from observations collected sequentially over time.[2]

We model time series using a **stochastic process**: a collection of random variables indexed in some way. We let this stochastic process be indexed by time, and denote it by the set of random variables $\{Y_t : t = 0, \pm 1, \pm 2, \pm 3, \ldots\}$.

For example: The time series of annual LA rainfall amounts (in inches) is 21, 17, 19, 5, 11, ... This is an observation of some stochastic process $\{Y_t\}$.

Just as with any set of random variables, there is a lot that can be said about $\{Y_t\}$. Each of the Y's has some (marginal) probability distribution, and any combination of two or more Y's will have a joint distribution.

However, we don't need to know everything about these distributions. Instead, we focus on some basic properties: the first and second moments (means, variances and covariances).

Intermission: Statistics Review

Before we continue any further, let's go over some important properties of means, variances and covariances. More can be found in Appendix A of the textbook (pg. 24).

These identities will be useful when deriving properties of various stochastic processes, and for assignment questions.

Means:

The mean, or **expected value**, of a random variable X is defined as

$$\mu_X = E(X) = \int_{-\infty}^{\infty} x f(x) dx,$$

where $f(\cdot)$ is the pdf of X.

For random variables X and Y, and constants a, b and c:

$$E(aX + bY + c) = aE(X) + bE(Y) + c.$$

Statistics Review (cont'd)

Variances:

The variance of a random variable X is

$$\sigma_X^2 = Var(X) = E\left[(X - E(X))^2\right] = E(X^2) - [E(X)]^2 .$$

Also:

$$Var(X) \geq 0$$

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

The positive square root of the variance is the **standard deviation** of X, and is usually denoted by σ_X.

The variance (or standard deviation) of a variable gives a measure of its degree of *spread*.

Statistics Review (cont'd)

Covariances:

The covariance between two random variables X and Y is

$$Cov(X,Y) = E\left[(X - E(X))(Y - E(Y))\right] = E(XY) - E(X)E(Y).$$

Therefore: $Cov(X,X) = Var(X)$.

If X and Y are independent, then $Cov(X,Y) = 0$.

Statistics Review (cont'd)

Covariances:

The covariance between two random variables X and Y is

$$Cov(X,Y) = E\left[(X - E(X))(Y - E(Y))\right] = E(XY) - E(X)E(Y).$$

Therefore: $Cov(X,X) = Var(X)$.

If X and Y are independent, then $Cov(X,Y) = 0$.

Properties:

$$\begin{aligned} Cov(X,Y) &= Cov(Y,X) \\ Cov(a+bX, c+dY) &= bd Cov(X,Y) \\ Cov(c_1X_1 + c_2X_2, d_1Y_1 + d_2Y_2) &= c_1d_1Cov(X_1,Y_1) + c_1d_2Cov(X_1,Y_2) \\ &\quad + c_2d_1Cov(X_2,Y_1) + c_2d_2Cov(X_2,Y_2) \end{aligned}$$

More generally:

$$Cov\left(\sum_{i=1}^{m} c_iX_i, \sum_{j=1}^{n} d_jY_j\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} c_id_jCov(X_i,Y_j).$$

Statistics Review (cont'd)

For any two random variables X and Y:

$$Var(aX + bY) = a^2 Var(X) + b^2 Var(Y) + 2abCov(X, Y).$$

Statistics Review (cont'd)

For any two random variables X and Y:

$$Var(aX + bY) = a^2 Var(X) + b^2 Var(Y) + 2abCov(X, Y).$$

Correlations:

The correlation coefficient between X and Y is

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

Similar properties for correlation can be derived. The key one is:

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

The correlation between two random variables is a measure of the *strength* and *direction* of their linear relationship.

Back To Stochastic Processes

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

$$\mu_t = E(Y_t),$$

for any $t = 0, \pm 1, \pm 2, \pm 3, \ldots$ The subscript in μ_t indicates that the mean may be different at different times t.

The **autocovariance function** is defined as

$$\gamma_{t,s} = Cov(Y_t, Y_s).$$

Some properties:

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

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

Back To Stochastic Processes (cont'd)

The **autocorrelation function** is defined as

$$\rho_{t,s} = Corr(Y_t, Y_s) = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}}.$$

Some properties:

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

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

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

Final Comments

That's all for now!

In this video, we've reviewed some important statistical concepts, and seen how they apply to stochastic processes.

Next Week in STAT 485/685: We'll see how these concepts apply to some real time series, and we'll practice deriving them. And we'll add a new concept: 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.