**Dependence**

We would like to understand the relationship among all random variables in a time series. In order to do that, we would need to look at the joint distribution function.

Suppose the time series consists of the random variables . The cumulative joint distribution function for these random variables is:

F(c1, c2, …, cn) = 

This can be VERY difficult to examine over the MULTIDIMENSIONS.

Instead, it is often easier to look at the one or two dimensional distribution functions. The one-dimensional cumulative distributional function is denoted by Ft(x) = P(xt ≤ x) for a random variable xt at time t. The corresponding probability distribution function is



The mean value function is



If there is no confusion about what random variable is being used, it is o.k. to drop the subscript x from . Thus, we would have μt.

Important: The interpretation of μt is that it represents the mean taken over ALL possible events that could have produced xt. Another way to think about it is suppose that  is observed an infinite number of times. Then  represents the average value at time 1,  represents the average value at time 2, …

Example: Moving Average

Let mt = (wt + wt-1 + wt-2)/3 where wt ~ ind. N(0,1) for t = 1, …, n.

Then

μt = E(mt)

= E[(wt + wt-1 + wt-2)/3]

= (1/3) E(wt + wt-1 + wt-2)

= (1/3) [E(wt) + E(wt-1) + E(wt-2)]

= (1/3) [0 + 0 + 0]

= 0

Example: Autoregressions

Let xt = 0.7xt-1 + wt where wt ~ independent N(0,1) for t = 1, …, n.

Then

μt = E(xt)

= E(0.7xt-1 + wt)

= 0.7E(xt-1) + E(wt)

= 0.7E(0.7xt-2+wt-1) + 0

= 0

This result will be discussed more later once more notation is introduced.

Autocovariance function

To assess the dependence between two random variables, we need to examine the two-dimensional cumulative distribution function. This can be denoted as F(cs, ct) = P(xs ≤ cs, xt ≤ ct) for two different time points s and t.

In another course, you learned about the covariance function which measures the linear dependence between two random variables:

For two random variables X and Y, the covariance between them is

Cov(X,Y) = E[(X– μx)(Y – μy)] = E(XY) – μxμy

where μx = E(X) and μy = E(Y)

Because we are interested in linear dependence between two random variables in the same time series, we will examine the autocovariance function:

γx(s,t) = Cov(xs, xt)

= E[(xs – μs)(xt – μt)] for all s and t.



where  and assuming continuous random variables

Notes:

* If the autocovariance is 0, there is no linear dependence.
* If s = t, the autocovariance is the variance:
γx(t,t) = E[(xt-μt)2]
* If there is no confusion about what random variable is being used, it is o.k. to drop the subscript x from γx(s,t). Thus, we would have γ(s,t).

Example: White noise

Suppose wt ~ ind. N(0,) for t=1,…,n. What is γ(s,t) for s = t and s ≠ t?



Example: Moving Average

Let mt = (wt + wt-1 + wt-2)/3 where wt ~ ind. N(0,1) for t = 1, …, n.

γ(s,t) = E[(ms-μs)(mt-μt)] = E[msmt] because μs = μt = 0

Then

E[msmt] = E[(ws + ws-1 + ws-2)/3 × (wt + wt-1 + wt-2)/3]

 = (1/9)E[(ws + ws-1 + ws-2)(wt + wt-1 + wt-2)]

To find this, we need to examine a few different cases:

* s = t

E[mtmt] = E[] = Var(mt) + E(mt)2 because Var(mt) =

 E[] - E(mt)2

= (1/9){Var(wt + wt-1 + wt-2)} + 02

= (1/9){Var(wt) + Var(wt-1) + Var(wt-2)} because wt’s

 are independent

= (1/9)(1+1+1) = 3/9

* s = t - 1

E[mt-1mt] = (1/9)E[(wt-1 + wt-2 + wt-3)(wt + wt-1 + wt-2)]

= (1/9)E[wt-1wt + wt-1wt-1 + wt-1wt-2 + wt-2wt + wt-2wt-1 +

 wt-2wt-2 + wt-3wt + wt-3wt-1 + wt-3wt-2]

= (1/9)[E(wt-1wt) + E(wt-1wt-1) + E(wt-1wt-2) + E(wt-2wt)

 + E(wt-2wt-1) + E(wt-2wt-2) + E(wt-3wt)

+ E(wt-3wt-1) + E(wt-3wt-2)]

= (1/9)[E(wt-1)E(wt) + E() + E(wt-1)E(wt-2)

+ E(wt-2)E(wt) + E(wt-2)E(wt-1) + E()

+ E(wt-3)E(wt) + E(wt-3)E(wt-1) + E(wt-3)E(wt-2)]

=(1/9)[0×0 + 1 + 0×0 + 0×0 + 0×0 + 1 + 0×0 + 0×0

 + 0×0]; E() = 1 because Var(wt-1) = 1

=2/9

* s = t - 2

E[mt-2mt] = (1/9)E[(wt-2 + wt-3 + wt-4)(wt + wt-1 + wt-2)]

= (1/9)E[wt-2wt + wt-2wt-1 + wt-2wt-2 + wt-3wt + wt-3wt-1

+ wt-3wt-2 + wt-4wt + wt-4wt-1 + wt-4wt-2]

= (1/9)[E(wt-2wt) + E(wt-2wt-1) + E(wt-2wt-2) + E(wt-3wt)

+ E(wt-3wt-1) + E(wt-3wt-2) + E(wt-4wt)

+ E(wt-4wt-1) + E(wt-4wt-2)]

= (1/9)[0×0 + 0×0 + 0×0 + 1 + 0×0 + 0×0 + 0×0 +

+ 0×0 + 0×0]

= 1/9

* s = t – 3

E[mt-3mt] = (1/9)E[(wt-3 + wt-4 + wt-5)(wt + wt-1 + wt-2)]

= (1/9)E[wt-3wt + wt-3wt-1 + wt-3wt-2 + wt-4wt + wt-4wt-1

+ wt-4wt-2 + wt-5wt + wt-5wt-1 + wt-5wt-2]

= (1/9)E[0×0 + 0×0 + 0×0 + 0×0 + 0×0 + 0×0 + 0×0

+ 0×0 + 0×0] because E(wt) = 0 for all t=1,…,n

= 0

* Notice that s = t + 1, t + 2, … can be found in a similar manner. Also s = t - 4, t - 5,… can be found. In summary, the autocovariance function is



Notes:

* The word “lag” is used to denote the time separation between two values. For example, |s - t| = 1 denotes the lag is 1 and |s - t| = 2 denotes the lag is 2. We will use this “lag” terminology throughout this course.
* The autocovariance depends on the lag, but NOT individual times for the moving average example! This will be VERY, VERY important later!

Autocorrelation function (ACF)

In another course, the Pearson correlation coefficient was defined to be:



for two variables X and Y where Cov(X,Y) denotes the covariance between X and Y and Var(X) denotes the variance of X. The reason the correlation coefficient is examined instead of the covariance is that it is always between –1 and 1. Note the following:

* ρ close to 1 means strong, positive linear dependence
* ρ close to –1 means strong, negative linear dependence
* ρ close to 0 means weak linear dependence.

The autocorrelation is the extension of the Pearson correlation coefficient to time series analysis. The autocorrelation function (ACF) is



where s and t denote two time points. The ACF is also between –1 and 1 and has a similar interpretation as for correlation coefficient.

Notice that γ(t,t) is the variance at time t.

Example: White noise

Suppose wt ~ independent N(0,) for t = 1, …, n. What is ρ(s,t) for s = t and s ≠ t?

Example: Moving Average

Let mt = (wt + wt-1 + wt-2)/3 where wt ~ ind. N(0,1) for t = 1, …, n.

Note that 

Then 

For example, ρ(s,t) for |s - t| = 1 is



Example: Strong positive and negative linear dependence (dependence.R)

Fill in blank:

* If there is strong positive linear dependence between xs and xt, the time series will appear *smooth or choppy* in a plot of the series versus time. Choose an answer.
* If there is strong negative linear dependence between xs and xt, the time series will appear *smooth or choppy* in a plot of the series versus time. Choose an answer.

Think of plots of xt vs. xs to help answer the above questions.

Below are three plots illustrating these statements. I simulated data from different time series models. The autocorrelation for |s - t| = 1 is given for each model. The “estimated” autocorrelation, denoted by , is given for that particular data set. The calculation of this estimate will be discussed later.

1) ρ(s,t) = 0.4972,  for |s - t| = 1





2) ρ(s,t) = -0.4972,  for |s - t| = 1



3) ρ(s,t) = 0,  for |s - t| = 1



Plot 1) is the least choppy (jagged) and plot 2) is the most choppy. Remember what a correlation means. A positive correlation means that “large” values tend to occur with other “large” values and “small” values tend to occur with other “small” values. A negative correlation means that “large” values tend to occur with other “small” values and “small” values tend to occur with other “large” values.

Please see program for the code.