TS Module 2 Time series concepts practice problems

(The attached PDF file has better formatting.)

Time series practice problems variances and covariances

** Exercise 2.1: Random walk

The time series $Y_t = Y_{t-1} + e_t$ is a random walk with $\sigma^2_e = 0.25$ and $Y_t = 0$ for $t < 1$.

A. What is the variance of Y_N?
B. What is the standard deviation of Y_N?

Part A: A random walk is the cumulative sum of a white noise process: $Y_N = \epsilon_1 + \epsilon_2 + \dots + \epsilon_N$

The error terms ϵ_t are independent, with a variance of 0.25 each.

The variance of Y_N is $N \times 0.25$.

Part B: The standard deviation of Y_N is $\sqrt{(N \times 0.25)} = 0.5 \times \sqrt{N}$.

See equation 2.2.11 on page 13

** Question 2.2: Stationary time series

If a time series is stationary, which of the following is true?

A. $\rho_1 = 1$
B. $\rho_k = \rho_{-k}$
C. $\rho_k = \rho_{k+1}$
D. $\rho_k > \rho_{k+1}$
E. $\rho_k < \rho_{k+1}$

Answer 2.2: B

See equation 2.3.2 on page 16.

Statement A: $\rho_o = 1$, not ρ_1

Statements C, D, E: The absolute value of the autocorrelation for lag k > 0 is never greater than 1, but it can increase or decrease between any two lags. Consider an autoregressive process with $\phi_1 = 0$ and $\phi_2 = 0.5$. The autocorrelations for odd lags (k = 1, 3, 5, …) are 0; the autocorrelations for even lags (k = 2, 4, 6, …) are $0.5^{k/2}$. The autocorrelation for lag k decreases from k = 0 to 1, increases from k = 1 to 2, decreases from k = 2 to 3, and so forth.

** Exercise 2.3: Equally weighted moving average

Let $Y_t = ½ \times (\epsilon_t + \epsilon_{t-1})$. The variance of ϵ_t is σ^2_ε

A. What is $\gamma_{t,t}$, the variance of Y_t?
B. What is $\gamma_{t,t-1}$, the covariance of Y_t and Y_{t-1}?
C. What is $\rho_{t,t-1}$, the correlation of Y_t and Y_{t-1}?

Part A: The variance of Y_t is $(½)^2 \times \sigma^2_\varepsilon + (½)^2 \times \sigma^2_\varepsilon = ½\ \sigma^2_\varepsilon$.

Part B: The covariance of Y_t with Y_{t-1} = covariance $(½ \times (\epsilon_t + \epsilon_{t-1}), ½ \times (\epsilon_{t-1} + \epsilon_{t-2}))$.

Expanding the product gives four terms. Three of the terms have different ϵ's; one has ϵ^2_{t-1}.

The ϵ's are independent, so the three terms with different ϵ's have a covariance of zero.

The one term with a non-zero covariance is $(½)^2 \times \sigma^2_\varepsilon = ¼\ \sigma^2_\varepsilon$

Part C: The correlation of Y_t with Y_{t-1} is the covariance divided by the product of the standard deviations of the two terms. A moving average time series is stationary if the underlying time series is stationary. The ϵ's are a stationary white noise process, so the moving average time series is also stationary, and the variances of all terms are the same. The product of the standard deviations of two terms is the variance. The correlation $= (¼\ \sigma^2_\varepsilon)/(½\ \sigma^2_\varepsilon) = 0.500$.

See Cryer and Chan page 15, equation 2.2.16: $\rho_{t,s} = 0.5$ for $|t-s| = 1$

** Exercise 2.4: Equally weighted moving average

Let $Y_t = (\epsilon_t + \epsilon_{t-1} + \ldots + \epsilon_{t-(N-1)}) / N$. ϵ_t is a white noise process, and the variance of ϵ_t is σ^2_ε

A. What is $\gamma_{t,t}$, the variance of Y_t?
B. What is $\gamma_{t,t-1}$, the covariance of Y_t and Y_{t-1}?
C. What is $\rho_{t,t-1}$, the correlation of Y_t and Y_{t-1}?
D. What is $\gamma_{t,t-j}$, the covariance of Y_t and Y_{t-j}?
E. What is $\rho_{t,t-j}$, the correlation of Y_t and Y_{t-j}?

Part A: The Cryer and Chan text shows the analysis for a two period moving average (N = 2). The homework assignment for this module extends the analysis to longer periods. This practice problem gives the reasoning for any N. Final exam problems gives values for N and σ^2_ε and test variances, covariances, and correlations.

Y_t is a stationary time series, since it is a linear combination of white noise processes. The variance of Y_t is the expected value of $[(\epsilon_t + \epsilon_{t-1} + \ldots + \epsilon_{t-(N-1)}) / N] \times [(\epsilon_t + \epsilon_{t-1} + \ldots + \epsilon_{t-(N-1)}) / N]$.

The ϵ's are independent, normally distributed, random variables with means of zero. The expected value of $\epsilon_t \times \epsilon_s = 0$ for $t \neq s$. In the expression above, N terms have t = s and N × (N-1) terms $t \neq s$, so the expected value of the expression is $1/N^2 \times \sigma^2_\varepsilon + 1/N^2 \times \sigma^2_\varepsilon + \ldots + 1/N^2 \times \sigma^2_\varepsilon = 1/N\ \sigma^2_\varepsilon$.

We express this result as "the variance of the mean of N independent, identically distributed random variables is inversely proportional to N."

Part B: The covariance of Y_t with Y_{t-1} = covariance $[(\epsilon_t + \epsilon_{t-1} + \ldots + \epsilon_{t-(N-1)}) / N, (\epsilon_{t-1} + \epsilon_{t-2} + \ldots + \epsilon_{t-N}) / N]$.

Expanding the product gives N^2 terms. (N^2 – (N-1))terms have different ϵ's; (N - 1) have ϵ^2_{t-j}.

The ϵ's are independent, so the terms with different ϵ's have a covariance of zero.

The (N-1) terms with a non-zero covariance each has an expected value of $1/N^2 \times \sigma^2_\varepsilon$.

The covariance for a lag of one period is $(N-1)/N^2 \times \sigma^2_\varepsilon$

Part C: The correlation of Y_t with Y_{t-1} is the covariance divided by the product of the standard deviations of the two terms. A moving average time series is stationary if the underlying time series is stationary. The ϵ's are a stationary white noise process, so the moving average time series is also stationary, and the variances of all terms are the same. The product of the standard deviations of two terms is the variance. The correlation $= [(N-1)/N^2 \times \sigma^2_\varepsilon] / [N/N^2 \times \sigma^2_\varepsilon] = (N-1)/N$.

Moving averages are often used because they are more stable than the observed values. Any two observed values are uncorrelated, but two adjacent moving averages are highly correlated.

Part D: The covariance of Y_t with Y_{t-j} has (N-j) terms with a non-zero covariance for $j < N$. If $j \geq N$, all terms have a zero covariance.

The (N-j) terms with a non-zero covariance each has an expected value of $1/N^2 \times \sigma^2_\varepsilon$.

The covariance for a lag of *j* period is $(N-j)/N^2 \times \sigma^2_\varepsilon$

Part E: The correlation $= [(N-j)/N^2 \times \sigma^2_\varepsilon] / [N/N^2 \times \sigma^2_\varepsilon] = (N-j)/N$.

Jacob: Are these the moving average processes in ARIMA analysis?

Rachel: The moving average processes discussed in this course have two differences.

- They are weighted averages, where the weights are the θ_j parameters.
- The weighted average becomes the next element of the time series.

** Exercise 2.5: Random walk

- Let Y_t be a random walk with zero drift: $Y_t = Y_{t-1} + \epsilon_t$.
- The variance of the error term is σ^2_e.
- The time series starts at $t = 1$: $Y_t = 0$ for $t < 1$.
- This time series is not stationary.

A. What is $\gamma_{t,t}$, the variance of Y_t for $t > 0$?
B. What is $\gamma_{t,s}$, the covariance of Y_t and Y_s for $t < s$?
C. What is $\rho_{t,s}$, the autocorrelation of Y_t and Y_s for $t < s$?

Part A: Y_t is the sum of *t* independent identically distributed random variables, each with variance σ^2_e. The variance of Y_t is $t \times \sigma^2_e$.

Part B: The covariance is the expected value of $(\epsilon_1 + \epsilon_2 + \ldots + \epsilon_t) \times (\epsilon_1 + \epsilon_2 + \ldots + \epsilon_s)$. These random variables are independent, so only *t* terms have a non-zero covariance (since *t* < *s*). The covariance $\gamma_{t,s}$ is $t \times \sigma^2_e$.

Part C: The correlation is the covariance divided by the standard deviations of the terms, or the square roots of the variances of the terms: $(t \times \sigma^2_e) / [\, t \times \sigma^2_e \times s \times \sigma^2_e\,]^{0.5} = \sqrt{(t/s)}$.

(See Cryer and Chan equations 2.2.12 and 2.2.13 on page 13)

Page 13, Equation 2.2.12:

$$\gamma_{t,s} = t\sigma_e^2$$

Page 13, Equation 2.2.13: The autocorrelation function for a random walk is

$$\rho_{t,s} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}} = \sqrt{\frac{t}{s}} \;\; for\; 1 \le t \le s$$

Stationary processes differ from non-stationary processes. Cryer and Chan write on page 16, equation 2.3.1:

"For a stationary process, the coverage between Y_t and Y_s depends on time only through the time difference $|t - s|$ and not otherwise on the actual times *t* and *s*. For a stationary process, we can simplify our notation and write $\rho_k = \text{Corr}(Y_t, Y_{t-k})$."

ARMA processes apply to stationary processes. A random walk is not stationary, so the autocorrelation depends on the actual times *t* and *s*.

Jacob: Y_t is a set of observations, such as {100, 102, 101, 98, 101, 103, 100, …}

I understand what is meant by $\rho_{t,t-1}$: we examine pairs of consecutive observations to see if they are related.

For the time series above, $\rho_{t,t-1}$ is the correlation of two series:

{100, 102, 101, 98, 101, 103, 100, …} and {102, 101, 98, 101, 103, 100, …}.

Here *t* ranges over all integers from 2 to infinity, and *t-1* ranges from 1 to infinity.

But $\rho_{t,s}$ is the autocorrelation of two specific observations. If $t = 4$ and $s = 7$, these are observations 4 and 7. Two scalars do not have a correlation. In the illustration above, the correlation of 98 and 100 has no meaning.

Rachel: You are treating the observed values as the random variables. In fact, the observed values are one simulation of the random variables. The meaning of $\rho_{t,s}$ is as follows. We simulate the random walk Y_t 100,000 times. From each simulation, we take observations *t* and *s*. These observations are random variables, not scalars. We have two series, each with 100,000 values. The N^{th} elements of the two series are observations 4 and 7 from a random walk. All the random walks have the same parameters (drift and volatility), but they each have different values. The correlation of these two series is $\rho_{t,s}$.

Jacob: Was I correct about $\rho_{t,t-1}$?

Rachel: You explained the *sample* autocorrelation of an empirical time series. The sample autocorrelation is an estimate of the true autocorrelation of the time series process. In this chapter, Cryer and Chan deal with the theoretical autocorrelation of the process, not with estimating the autocorrelation from empirical data.

** Question 2.6: Weakly stationary time series

All but which of the following are true for a weakly stationary time series Y_t?

A. The mean function is constant over time.
B. $\gamma_{t,t} = \gamma_{t-1,t-1}$
C. Covariance ($Y_{t-2,} Y_{t-4}$) = Covariance ($Y_{t-3,} Y_{t-1}$)
D. Correlation (Y_{t+1}, Y_{t-1}) = Correlation (Y_t, Y_{t+2})
E. Correlation (Y_{t+1}, Y_{t-1}) = Correlation (Y_{t-2}, Y_{t+2})

Answer 2.6: E

See page 17, top of page, where Cryer and Chan has two attributes of weakly stationary time series:

1. The mean function is constant over time.
2. $\gamma_{t,t-k} = \gamma_{0,k}$ for all time *t* and lag *k*.

Choice E says the correlation of lag 2 = the correlation of lag 4. This is not correct.

Jacob: How might one summarize a stationary time series?

Rachel: The covariance depends only on the lag between the two observations.

Jacob: A white noise process is stationary. But the process is stochastic. Perhaps observations 1 and 2 are similar, but observations 4 and 5 are not similar. Aren't the covariances different?

Rachel: You are thinking of the time series as a set of known observations. The proper perspective is that the time series is a random draw from all sets of observations with certain attributes. If we examine all sets of observations with the given attributes, the covariance of the first and second elements is the same as the covariance of the fourth and fifth elements.

** Exercise 2.7: Random walk

Y_t is a random walk: $Y_j = \epsilon_1 + \epsilon_2 + \ldots + \epsilon_j$, with $\sigma^2_\varepsilon = 1$.

The autocorrelation of observations *t* and *s*, $\rho_{t,s}$, with $t < s$, is ½ (*t* and *s* are the t^{th} and s^{th} observations).

A. What is $\rho_{2t,s}$, the autocorrelation of observations 2t and s?
B. What is $\rho_{t,2s}$, the autocorrelation of observations t and 2s?
C. What is $\rho_{t,t+s}$, the autocorrelation of observations t and t+s?
D. What is $\rho_{s,t+s}$, the autocorrelation of observations s and t+s?

Part A: $\rho_{t,s} = \sqrt{(t/s)} = ½ \Rightarrow t/s = ¼ \Rightarrow (2t)/s = ½ \Rightarrow \rho_{2t,s} = \sqrt{(2t/s)} = \sqrt{(½)} =$ 0.707.

Jacob: Perhaps 2t is larger than s, so we should use $\sqrt{(s/2t)}$.

Rachel: t/s = ¼, so s = 4t; 2t is ½ × s.

Jacob: Do final exam problems use other parameters besides 2?

Rachel: Yes. For a parameter of k (instead of 2), use the smaller of (kt)/s and s/(kt).

- $\rho_{3t,s} = \sqrt{(3t/s)} = \sqrt{(¾)} =$ 0.866.
- $\rho_{4t,s} = \sqrt{(4t/s)} = \sqrt{1} =$ 1.000.
- $\rho_{5t,s} = \sqrt{(s/5t)} = \sqrt{(4/5)} =$ 0.894.
- $\rho_{6t,s} = \sqrt{(s/6t)} = \sqrt{(⅔)} =$ 0.817.

Part B: $\rho_{t,s} = \sqrt{(t/s)} = ½ \Rightarrow t/s = ¼ \Rightarrow t/(2s) = ⅛ \Rightarrow \rho_{2t,s} = \sqrt{(t/2s)} = \sqrt{(⅛)} =$ 0.354.

Part C: $\rho_{t,t+s} = \sqrt{(t/(t+s))} = \sqrt{(1/5)} = 0.447$.

Jacob: Can one write this as a formula?

Rachel: t/(t+s) = 1 / [(t+s) / t] = 1 / (1 + 1 / (t/s))

Part D: $\rho_{s,t+s} = \sqrt{(s/(t+s))} = \sqrt{(4/5)} = 0.894$.

Jacob: Can one write this as a formula?

Rachel: s/(t+s) = 1 / [(t+s) / s] = 1 / (1 + t/s)