TS Module 10 Sample autocorrelation computations practice problems

(The attached PDF file has better formatting.)

** Exercise 10.1: Sample Autocorrelation Function

The series Y_t = {8, 12, 11, 16, 10, 14, 9, 14, 13, 13, 9, 15} is modeled by an ARMA process.

A. What is the average Y value ($\bar{y}$)?
B. What are the deviations of each observation from the mean?
C. What is the sum of squared deviations?
D. What is the sum of squared deviations with a lag of one period?
E. What is r_1, the *sample* autocorrelation at lag 1?
F. What is the sum of squared deviations with a lag of two periods?
G. What is r_2, the *sample* autocorrelation at lag 2?
H. What is the sum of squared deviations with a lag of three periods?
I. What is r_2, the *sample* autocorrelation at lag 3?

Part A: Form a table with 7 columns.

T	Y_t	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
1	8	-4	16	0	4	-16
2	12	0	0	0	0	0
3	11	-1	1	-4	2	-2
4	16	4	16	-8	8	-12
5	10	-2	4	-4	6	-4
6	14	2	4	-6	4	2
7	9	-3	9	-6	-3	-3
8	14	2	4	2	2	-6
9	13	1	1	1	-3	3
10	13	1	1	-3	3	
11	9	-3	9	-9		
12	15	3	9			
Total	144	0.00	74.0	-37.0	23.0	-38.0

- Column 1 is the period (1 through 12)
- Column 2 is the observation
- Column 3 is $y_t - \bar{y}$
- Column 4 is $(y_t - \bar{y})^2$
- Column 5 is $(y_t - \bar{y}) \times (y_{t+1} - \bar{y})$
- Column 6 is $(y_t - \bar{y}) \times (y_{t+2} - \bar{y})$
- Column 7 is $(y_t - \bar{y}) \times (y_{t+3} - \bar{y})$

The sum of the 12 observations is 144, so the average $\bar{y}$ is 12: $\bar{y} = \sum y_t / 12 = 144/12 = 12.000$

Part B: The deviations $y_t - \overline{y}$ are in Column 3. The sum of the deviations is zero.

Part C: The squared deviations is in Column 4. All squares are positive, so the sum is positive (74).

Part D: The squared deviations with a one period lag is in Column 5. The first four entries are

- Period 1: –4 × 0 = 0
- Period 2: 0 × –1 = 0
- Period 3: –1 × 4 = –4
- Period 4: 4 × –2 = –8

This column has only 11 entries, not 12. There is no observation for Period 13, so we can't form a squared deviation with a one period lag for period 12.

Sums of squared deviations with lags may be positive or negative. The sum of squared deviations with a one period lag is –37.

Part E: The sample autocorrelation with a one period lag (r_1) is –37 / 74 = -0.500.

$r_1 = [\sum (y_t - \overline{y}) \times (y_{t+1} - \overline{y})] / \sum (y_t - \overline{y})^2 = -37 / 74 = -0.500$

Part F: The squared deviations with a two period lag is in Column 6. The first four entries are

- Period 1: –4 × –1 = 4
- Period 2: 0 × 4 = 0
- Period 3: –1 × –2 = 2
- Period 4: 4 × 2 = 8

This column has only 10 entries. There is no observation for Periods 13 or 14, so we can't form a squared deviation with a two periods lag for periods 11 or 12.

The sum of squared deviations with a two period lag is 20.

Part G: The sample autocorrelation with a two period lag (r_2) is 20 / 74 = 0.270.

$r_2 = [\sum (y_t - \overline{y}) \times (y_{t+2} - \overline{y})] / \sum (y_t - \overline{y})^2 = 20 / 74 = 0.270$

Part H: Continue in the same fashion for all sample autocorrelations.

Part I: The sample autocorrelation with a three period lag (r_3) is –38 / 74 = -0.514.

Jacob: Does the type of model (AR(1), AR(2), MA(1), and so forth) affect the sample autocorrelations?

Rachel: The type of model determines the actual autocorrelations, not the sample autocorrelations.

Jacob: The numerator of the ratio for r_2 has 10 entries and the denominator has 12 entries. Does this distort the sample autocorrelation?

Rachel: If the time series has many observations and the lag is not too log, the distortion is not material. At the extremes, such as the sample autocorrelations for lags 10 and 11 in a time series with 12 observations, the sample autocorrelations are too low. But these sample autocorrelations are distorted anyway by random fluctuations, so the potential distortion for the number of entries is not important.

** Exercise 10.2: Joining Two Series

A time series has 20,000 observations.

- The first 10,000 observations are a white noise process with μ = 1 and σ^2_ε = 0.0001
- The next 10,000 observations are a white noise process with μ = –1 and σ^2_ε = 0.0001

A. What is the mean of the full time series?
B. What are the deviances of the full time series?
C. What is the denominator of the sample autocorrelations?
D. What is the sample autocorrelation of lag 1?
E. What is the sample autocorrelation of lag 10,000?

Part A: The mean of the full time series is 10,000 × 1 + 10,000 × –1 = zero. The observations are stochastic, so the actual mean will be slightly different from zero. With 20,000 observations and a σ^2_ε = 0.0001, the difference is miniscule.

Part B: The deviances of the full time series are random numbers; we don't know them exactly. Since σ^2_ε = 0.0001 is so small, the deviances are close to 10,000 1's and 10,000 –1's.

Part C: The denominator of the sample autocorrelations has 20,000 terms. Each term is close to 1^2 or $(-1)^2$. The sum of the terms is about 20,000.

Part D: The numerator of the sample autocorrelation of lag 1 has 20,000 – 1 = 19,999 terms. Of these:

- 19,998 terms are close to 1 × 1 or –1 × –1.
- 1 term is close to 1 × –1.

The sum of the terms in the numerator is close to 19,998 × 1 + 1 × –1 ≈ 19,997.

The sample autocorrelation of lag 1 is approximately 19,997 / 20,000 ≈ 1.

Jacob: You seem to be ignoring the σ^2_ε.

Rachel: The variance is so small that it has little effect on a time series of 20,000 observations. The actual sample autocorrelation of lag 1 is less than 1. Final exam problems are multiple choice questions, and the choices are sufficiently different that only one choice is reasonable.

Part E: The numerator of the sample autocorrelation of lag 1 has 20,000 – 10,000 = 10,000 terms. All 10,000 terms are close to 1 × –1.

The sum of the terms in the numerator is close to 10,000 × –1 ≈ –10,000.

The sample autocorrelation of lag 10,000 is approximately –10,000 / 19,997 ≈ –0.5

Jacob: All the cross-products are –1, and all the variances are +1. Shouldn't the autocorrelation be –1?

Rachel: The numerator of this sample autocorrelation has only half as many terms as the denominator.

Jacob: What does the correlogram look like?

Rachel: Here is the correlogram. The black triangles are actually closely spaced columns of steadily decreasing height.