MTHSTAT 564/564G/764–Time Series Analysis
Spring 2024
Problem Solving Set 8

Please think about the problems below from the textbook in advance of our problem solving sessions on them. These problems cover Chapter 6, and pre-recorded lectures 16-21.

1. For the white noise process, verify that $\mathrm{Var}(r_k) \approx \frac{1}{n}$ and that $\mathrm{Corr}(r_k, r_j) \approx 0$ for $k \neq j$.

2. For Exhibit 6.1 on Page 111, add the following new entries to the table and compute all four quantities: $\phi = \pm 0.99$, $\phi = \pm 0.5$, $\phi = \pm 0.1$.

3. Verify that for the MA(1) process, $c_{11} = 1 - 3\rho_1^2 + 4\rho_1^4$, and $c_{kk} = 1 + 2\rho_1^2$ for $k > 1$. Verify also that $c_{12} = 2\rho_1(1 - \rho_1^2)$.

4. For a time series of 100 observations, we calculate $r_1 = -0.49$, $r_2 = 0.31$, $r_3 = -0.21$, $r_4 = 0.11$, and $|r_k| < 0.09$ for $k > 4$. On this basis alone, what ARIMA model would we tentatively specify for the series?

5. For a series of length 169, we find that $r_1 = 0.41$, $r_2 = 0.32$, $r_3 = 0.26$, $r_4 = 0.21$, and $r_5 = 0.16$. What ARIMA model fits this pattern of autocorrelations?

6. Suppose $\{X_t\}$ is a stationary AR(1) process with parameter θ, but that we can only observe $Y_t = X_t + N_t$, where $\{N_t\}$ is a white noise measurement error independent of $\{X_t\}$.

   (a) Find the autocorrelation function for the observed process in terms of ϕ, σ_X^2, and σ_N^2.

   (b) Which ARIMA model might we specify for $\{Y_t\}$?

7. Simulate an AR(2) time series of length $n = 72$ with $\phi_1 = 0.7$ and $\phi_2 = -0.4$.

   (a) Calculate and plot the theoretical autocorrelation for this model. Plot sufficient lags until the correlations are negligible.

   (b) Calculate and plot the sample ACF for your simulated series. How well do the values and patterns match the theoretical ACF from part (a)?

   (c) What are the theoretical partial autocorrelations for this model?

   (d) Calculate and plot the sample PACF for your simulated series. How well do the values and patterns match the theoretical PACF from part (c)?

8. Simulate an MA(2) time series of length $n = 36$ with $\theta_1 = 0.7$ and $\theta_2 = -0.4$.

   (a) What are the theoretical autocorrelations for this model.

(b) Calculate and plot the sample ACF for your simulated series. How well do the values and patterns match the theoretical ACF from part (a)?

(c) Plot the theoretical partial autocorrelation for this model. Plot sufficient lags until the correlations are negligible. (We do not have a formula for this PACF. Instead, perform a very large sample simulation, say $n = 1,000$, for this model and calculate and plot the sample PACF for this simulation).

(d) Calculate and plot the sample PACF for your simulated series from part (a). How well do the values and patterns match the "theoretical" PACF from part (c)?

9. Simulate an AR(1) time series with $n = 48$ and $\phi = 0.7$.

   (a) Calculate the theoretical autocorrelations at lag 1 and lag 5 for this model.

   (b) Calculate the sample autocorrelations at lag 1 and lag 5 and compare the values with the theoretical autocorrelations. Use Equations (6.1.5) and (6.1.6) on page 111 to quantify the comparisons.

   (c) Repeat part (b) with a new simulation. Describe how the precision of the estimate varies with different samples selected under identical conditions.

   (d) If software permits, repeat the simulation of the series and calculation of r_1 and r_5 many times and form the sampling distributions of r_1 and r_5. Describe how the precision of the estimate varies with different samples selected under identical conditions. How well does the large-sample variance given in Equation (6.1.5) approximate the variance in your sampling distribution?

10. Simulate an MA(1) time series with $n = 60$ and $\theta = 0.5$.

    (a) Calculate the theoretical autocorrelation at lag 1 for this model.

    (b) Calculate the sample autocorrelation at lag 1, and compare the value with its theoretical value. Use Exhibit 6.2 on page 112 to quantify the comparisons.

    (c) Repeat part (b) with a new simulation. Describe how the precision of the estimate varies with different samples selected under identical conditions.

    (d) If software permits, repeat the simulation of the series and calculation of r_1 and r_5 many times and form the sampling distributions of r_1 and r_5. Describe how the precision of the estimate varies with different samples selected under identical conditions. How well does the large-sample variance given in Exhibit 6.2 on page 112 approximate the variance in your sampling distribution?

11. Simulate a nonstationary time series with $n = 60$ according to the ARIMA(0,1,1) model with $\theta = 0.8$.

    (a) Perform the (augmented) Dickey-Fuller test on the series with $k = 0$ in Equation (6.4.1) on page 128. (With $k = 0$, this is the Dickey-Fuller test and is not augmented). Comment on the results.

    (b) Perform the augmented Dickey-Fuller test on the series with k chosen by the software–that is, the "best" value for k. Comment on the results.

    (c) Repeat parts (a) and (b), but use the differences of the simulated series. Comment on the results. (Here, of course, you should reject the unit-root hypothesis).

12. The data file named "deere1" contains 82 consecutive values for the amount of deviation (in 0.000025 inch units) from a specified target value that an industrial machining process at Deere & Co. produced under certain specified operating conditions.

(a) Display the time series plot of this series and comment on any unusual points.

(b) Calculate the sample ACF for this series and comment on the results.

(c) Now replace the unusual value by a much more typical value and recalculate the sample ACF. Comment on the change from what you saw in part (b).

(d) Calculate the sample PACF based on the revised series that you used in part (c). What model would you specify for the revised series? (Later, we will investigate other ways to handle outliers in time series modeling).

13. The data file named "robot" contains a time series obtained from an industrial robot. The robot was put through a sequence of maneuvers, and the distance from a desired ending point was recorded in inches. This was repeated 324 times to form a time series.

(a) Display the time series plot of the data. Based on this information, do these data appear to come from a stationary process, or a nonstationary process?

(b) Calculate and plot the sample ACF and PACF for these data. Based on this additional information, do these data appear to come from a stationary process, or a nonstationary process?

(c) Calculate and interpret the sample EACF.

(d) Use the best subsets ARMA approach to specify a model for these data. Compare these results with what you discovered in parts (a), (b), and (c).