Suggested Problems 1
DSC 534

Problems

Problem 1 Suppose that $T = 1, 2$ and a process is defined by tossing a dice at $t = 1$ and by tossing another dice at time $t = 2$. The sample space is given by the set $\Omega = \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}$. We define the stochastic process

$$X(t, \omega) = t + [\text{dice outcome at time } t]^2.$$

So for a given ω, say $\omega = (2, 4)$ the realization of the process at time $t = 1$ is equal to $1 + 2^2 = 5$ and at time $t = 2$ is equal to $2 + 4^2 = 18$. So we can easily calculate all realizations since the sample space and T are both finite. Calculate the mean value and the autocovariance function of this process.

Problem 2 Consider a signal-plus-noise model of the general form $X_t = S_t + W_t$, where W_t is Gaussian white noise with $\sigma_W^2 = 1$. Simulate and plot $n = 200$ observations from each of the following two models.

1. $X_t = S_t + W_t$, where

$$S_t = \begin{cases} 0 & \text{t=1,2, \ldots,100,} \\ 10\exp(-\frac{(t-100)}{20})\cos(\frac{2\pi t}{4}) & \text{t=101,102,\ldots,200.} \end{cases}$$

2. $X_t = S_t + W_t$, where

$$S_t = \begin{cases} 0 & \text{t=1,2, \ldots,100,} \\ 10\exp(-\frac{(t-100)}{200})\cos(\frac{2\pi t}{4}) & \text{t=101,102,\ldots,200.} \end{cases}$$

3. Compare the general appearance of the series (1) and (2) with the earthquake series and the explosion series that can be found in `astsa` R package.

Problem 3 Let X be a random variable with mean μ and variance σ^2. Let $Y_t = X$ for all t.

1. Show that $\{Y_t\}$ is strictly and weakly stationary.
2. Calculate the acf of Y_t.
3. Sketch a time series plot of Y for different choices of X random variables (for instance you can use normal, chi-square with ν degrees of freedom and so on).

Problem 4 Suppose that $\{W_t\}$ is a white noise sequence with mean 0 and variance σ^2. Define $X_t = W_t + \theta_1 W_{t-1}$. We have seen in the class that this is a MA(1) process. Suppose that you observe the series and you can produce an excellent estimator of the $\rho(h)$. Can you estimate θ_1 based on the estimator of your acf?

Problem 5 Suppose that X_t is stationary and consider $Y_t = \nabla X_t \equiv X_t - X_{t-1}$. In addition, consider $Z_t = \nabla^2 X_t$. If the mean of X_t is zero and its acf is given by $\rho_X(h)$ show that both Y_t and Z_t are stationary and calculate their mean and acf.

Problem 6 Suppose that X_t is a stationary time series with mean μ and auto-covariance function $\gamma_X(h)$. Denote by

$$S^2 = \frac{1}{n-1}\sum_{t=1}^{n}(Y_t - \bar{Y})^2,$$

the sample version. Show that

$$E(S^2) = \frac{n}{n-1}\gamma_X(0) - \frac{n}{n-1}Var(\bar{X}) = \gamma_X(0) - \frac{2}{n-1}\sum_{h=1}^{n-1}\Big(1-\frac{h}{n}\Big)\gamma_X(h).$$

What does the above identity show for the case of X_t being a white noise process.

Problem 7 Suppose that R and Φ are independent random variables with Φ being uniformly distributed in (0,1). For simplicity, you can assume that R is Normal random variable with mean 0 and variance 1. But the following works for and random variable R, as long as R is independent of U. Simulate the following process

$$X_t = R\cos\Big(2\pi(\omega t + \Phi)\Big)$$

with $0 < \omega < 1/2$ and $t = 1, 2, \ldots, n$, $n = 100, 500, 1000$. You can choose the frequence ω. Plot the series and the acf for each sample size. What do you observe? Can you give a theoretical justification of your observations based on the definition of the process.

Problem 8 The data file `wages` in the R package `TSA` contains monthly values of the average hourly wages (in dollars) for workers in the U.S. apparel and textile products industry for July 1981 through June 1987. In what follows the standardized residuals refer to ordinary residuals divided by the estimator of the standard deviation of the noise process in a regression model. Use `rstudent` in `R` to obtain them.

1. Display and interpret the time series plot for these data.
2. Use least squares to fit a linear time trend to this time series. Interpret the regression output. Save the standardized residuals from the fit for further analysis.
3. Construct and interpret the time series plot of the standardized residuals from part (2).
4. Use least squares to fit a quadratic time trend to the wages time series. Interpret the regression output. Save the standardized residuals from the fit for further analysis.
5. Construct and interpret the time series plot of the standardized residuals from part (4).
6. Calculate and interpret the sample autocorrelation function for the standardized residuals obtained from part (4).
7. Investigate the normality of the standardized residuals obtained from part (4). Consider histograms and normal probability plots. Interpret the plots.

Problem 9 The data file `prescrip` in `TSA` package gives monthly U.S. prescription costs for the months August 1986 to March 1992. These data are from the State of New Jersey's Prescription Drug Program and are the cost per prescription claim.

1. Display and interpret the time series plot for these data. Use plotting symbols that permit you to look for seasonality.
2. Calculate and plot the sequence of month-to-month percentage changes in the prescription costs. Again, use plotting symbols that permit you to look for seasonality.
3. Use least squares to fit a cosine trend with fundamental frequency 1/12 to the percentage change series. Interpret the regression output. Save the standardized residuals.
4. Plot the sequence of standardized residuals to investigate the adequacy of the cosine trend model. Interpret the plot.
5. Calculate and interpret the sample autocorrelations for the standardized residuals. Investigate the normality of the standardized residuals.

Problem 10 The file named `gold` in package `TSA` contains the daily price of gold (in dollars per troy ounce) for the 252 trading days of year 2005.

1. Display the time series plot of these data. Interpret the plot.
2. Display the time series plot of the differences of the logarithms of these data. Interpret this plot.
3. Calculate and display the sample ACF for the differences of the logarithms of these data and argue that the logarithms appear to follow a random walk model.
4. Display the differences of logs in a histogram and interpret.
5. Display the differences of logs in a quantile-quantile normal plot and interpret.