CHAPTER 5

EXERCISES FOR STOCHASTIC PROCESSES

Exercise 5.1. Write an R function to compute the autocorrelation function of a time series. The preamble of the function should be as follows:

```
acf.brute = function(x,lag.max=NULL) {
## COMPUTES THE AUTOCORRELATION FUNCTION OF A TIME SERIES
## INPUT:
##   X: [NTOT]-LENGTH NUMERICAL VECTOR OF THE TIME SERIES
##   LAG.MAX: MAXUMIM LAG TO COMPUTE
##        (DEFAULT = MAX(1,FLOOR(10*log10(LENGTH(X)))))
## OUTPUT:
##   X.ACF: [1:(LAG.MAX+1)] AUTOCORRELATION FUNCTION OF X
##        FOR LAGS 0 TO LAG.MAX
```

Test your function on the time series {`set.seed(1); x = rnorm(20)`} and state the correlation values for all `lag.max` lags. Show that these values equal those obtained from the built-in R function `acf(x)`, by printing them out using the commands

```
acf.R    = acf(x)
as.numeric(acf.R$acf)
```

□

Exercise 5.2. Write a function to generate a time series from the AR1 model

$$x_{t+1} = \phi_1 x_t + w_t + k \tag{5.1}$$

where $w_t \sim GWN(0,1)$. The beginning of this function should be as follows:

```
ar1.ts = function(phi1,ntot,stdv=1,constant=0,iseed=1) {
## GENERATE A REALIZATION FROM THE AR1 MODEL
## X(T+1) = PHI1 * X(T) + W(T) + CONSTANT
## WHERE W(T) IS GAUSSIAN WHITE NOISE WITH MEAN 0 AND ST. DEV. 'STDV'
## INITIAL CONDITION IS RANDOMLY DRAWN FROM STATIONARY DISTRIBUTION
## INPUT:
#    PHI1: AUTOREGRESSIVE PARAMETER
#    NTOT: LENGTH OF THE DESIRED TIME SERIES
#    STDV: STANDARD DEVIATION OF THE NOISE (DEFAULT = 1)
#    CONSTANT: CONSTANT TERM IN THE AR1 MODEL (DEFAULT = 0)
#    ISEED: SEED FOR THE RANDOM NUMBER GENERATOR (DEFAULT = 1)
#  OUTPUT:
#    X: [NTOT]- RANDOM TIME SERIES FROM AR(1) PROCESS
```

The initial condition should be drawn from the *stationary distribution* of the AR process. What is this distribution? State this distribution for general k, ϕ_1, σ_W^2. Use this equation in your function. □

Exercise 5.3. Use the above functions to generate a time series of length 50, and corresponding sample autocorrelation functions, for $\phi_1 = 0, 0.5, 0.9$. Make plots of both quantities. In addition, on the plot for the autocorrelation, superimpose the population autocorrelation function, and the approximate 95% confidence interval for zero correlation. Comment on whether the results make sense, or whether they are 'surprising.'

□

Exercise 5.4. Use the above function to generate a time series of length 50 for $\phi_1 = 0$. Split this time series into two halves, each of length 25. Use the function `mean.equal.test` to calculate the t-statistic for testing equality of means. Repeat this 1000 times, thereby generating 1000 t-values. Plot a histogram of these t-values. Superimpose on this histogram the expected distribution of the t-statistic. This can be done using the `curve` command, as follows:

```
hist(tval,col="grey",freq=FALSE)
curve(dt(x,dof),add=TRUE)
```

(In the above example, `x` does not need to be pre-defined. See the manual on R graphics or the help pages for `curve`.) You should find that the histogram is reasonably consistent with the exact t distribution. Repeat this for $\phi_1 = 0.5, 0.9$. Are the histograms still consistent with the exact t distribution? Explain the result you find. If someone concludes that there is significant difference in mean based on a standard t-test (i.e., using iid assumption), how might this conclusion be affected if autocorrelation were taken into account? Does the true significance level increase or decrease when autocorrelation is taken into account? □

Exercise 5.5 (Mean and Variance of an AR(p) Process)**.** Consider the AR(p) process

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + W_t + k, \tag{5.2}$$

where W_t is white noise with zero mean and variance σ^2, and k is a constant. Show that the expectation of the stationary AR(p) process is

$$\mathbb{E}[X_t] = \frac{k}{1 - \phi_1 - \phi_2 - \cdots - \phi_p}. \tag{5.3}$$

Show that the variance of the stationary process is

$$\mathrm{var}[X_t] = \frac{\mathrm{var}[W_t]}{1 - \rho_1\phi_1 - \rho_2\phi_2 - \cdots - \rho_p\phi_p}. \tag{5.4}$$

(Hint: multiply both sides of (5.2) by $X_t - \mathbb{E}[X_t]$ and take the expectation.) □

Exercise 5.6 (Autocorrelation of a running mean)**.** Consider the running mean

$$Y_t = \frac{1}{2K+1} \sum_{k=-K}^{K} X_{t+k}. \tag{5.5}$$

Assume $X_t \sim GWN(0, \sigma^2)$. Compute the mean and variance of Y_t. Show that the autocorrelation function of Y_t is

$$\mathrm{cor}[Y_{t+\tau}, Y_t] = \begin{cases} 1 - \frac{|\tau|}{2K+1} & \text{for } |\tau| \leq 2K+1 \\ 0 & \text{for} |\tau| > 2K+1 \end{cases}. \tag{5.6}$$

Note that even though the process X_t is uncorrelated, running averages of X_t are correlated for lags less than twice the window size K. Explain why this makes sense. □