

Assignment #1

Submit your **printed** solution at the start of class on September 12.

Short answers, please. *Explain your answer concisely.* Include a plot only if requested in the question or if it makes the point better than you can in words. Do not include a plot unless your narrative refers to the plot. Significance here refers to statistical significance. Presume necessary conditions hold for tests of significance unless conditions are specifically addressed in the question.

- Exercise 1.2 from the textbook.
- Exercise 2.9 from the textbook.
- Simulate 25 random walks with drift $\delta = 0$ of length $n = 60$, made by accumulating standard normal white noise. Plot these in a single graph, similar to what was done in Lecture 2. Modify that code as needed to compute your series.
[Be sure to start your calculations with a value for the random seed chosen by you.]
- For each of the random walks simulated in Q3, fit a linear regression of the random walk on the time index. Show a histogram of the estimated slopes and a histogram of the p-values of the usual t-test of $H_0: \beta_1 = 0$. What problem for inference occurs when testing for the slope of a random walk?
- For each of the fitted regression models, find the first autocorrelation of the residuals from the fit. Show a histogram of the estimated $\text{corr}(e_t, e_{t-1})$, where $\{e_t\}$ are the residuals.

R hint for Q4-5: Use these R commands to obtain the estimated slope and its p-value from a linear regression. You will need to put these commands inside a loop to accumulate the slopes and p-values from the different random walks.

```
time <- 1:60
regr <- lm(rw ~ time)
beta <- coefficients(summary(regr))
b <- beta[2,1]           # the estimated slope
p <- beta[2,4]           # its p-value
res <- residuals(regr)   # residuals
```

Two time series `gtemp_land` and `gtemp_ocean` which are included in `astsa` are annual temperature deviations from normal (see Example 1.2 from the textbook for more information). Use the window function in R to extract the data from 1960 onwards only:

```
land <- window(gtemp_land, start=1960)
ocean <- window(gtemp_ocean, start=1960)
```

- You could also extract the recent data by indexing, but the results won't be the same. Compare the attributes of the result from using the window function as shown above to the results obtained by indexing, as in `land <- gtemp_land[111:174]`.

7. Plot the two time series of recent land and ocean temperature deviations. Include in your plot the fitted least squared regressions of the temperature series since 1960 on time.

R hint: To show the fitted line of a simple linear regression in a plot, mimic the following code with appropriate choices for “x” and “y”. Chose an interesting color to distinguish the line from the other information.

```
plot(y ~ x)
regr <- lm(y ~ x)
abline(regr, col="you pick something")
```

8. If we accept the fitted regression models, what do the results indicate about the increases in temperature since 1960? Specifically, do the estimated slopes appear to be significantly different from zero?
9. Do the residuals from the estimated models suggest an evident flaw in the fit of either model?
10. Are the residuals from the two models related, or do they appear independent of each other? These are time series, so they may be non-contemporaneously related.