

Stat 565

Homework 2

Due in class Jan 21

1. Theory

1. Consider two random variables, X and Y with $E(X) = E(Y) = \mu$ and $\text{Var}(X) = \text{Var}(Y) = \sigma^2$.

a) Let $\text{Cor}(X, Y) = 0$. Show that the expectation and variance of the sample average of X and Y are:

$$E\left(\frac{X+Y}{2}\right) = \mu \quad \text{and} \quad \text{Var}\left(\frac{X+Y}{2}\right) = \sigma^2/2$$

b) Now let $\text{Cor}(X, Y) = \rho$. Find expectation and variance of the sample average of X and Y . When is the variance of the sample average of X and Y less than $\sigma^2/2$? When is the variance of the sample average of X and Y greater than $\sigma^2/2$?

2. Let ϵ_t , $t = 1, 2, \dots, T$ be a sequence of independent, identically distributed random variables with mean zero and variance one. The time series, Y_t is constructed according to:

$$Y_t = \epsilon_t + 5\epsilon_{t-1}$$

Find $E(Y_t)$, $\text{Var}(Y_t)$, $\text{Cov}(Y_t, Y_{t-1})$ and $\text{Cov}(Y_t, Y_{t-2})$.

2. Data Analysis

For the data analysis question, your answer should read like a mini-report. It will consist of some plots along with commentary that explains what the plots are and what they show. Put all your R code in an Appendix. There is generally no **right** answer. Be creative and curious! I've posted an [example](#) of how to present your analysis. Do not follow the example exactly, it is only one of many ways to approach this type of question. This example was produced with Rmarkdown, if you are interested the source is at [/homeworks/hw-example.Rmd](#)

The United States Census Bureau collects and provides data on US retail sales through the [Monthly Retail Trade Surveys](#). For this homework I have downloaded the estimates of the total monthly sales for **motor vehicle and parts dealers** in the USA from 1992 to 2014.

Use the following line of code to read the data.frame sales into R:

```
load(url("http://stat565.cwick.co.nz/data/sales.rda"))
```

There are two columns value with the estimates of monthly sales in millions of dollars, and date with the month for the estimate (this is already a Date in R, I used the 15th of the month arbitrarily).

(If you are interested, here is [the code I used to produce sales.rda](#))

Summarise and describe the trend and seasonality in the monthly sales for motor vehicle and parts dealers.

Remove the trend and seasonality to estimate residual noise.

Examine the residual noise for non-stationarity in variance and for autocorrelation.