STAT 5511 (Fall 2024) Homework 0
Charles R. Doss
Assigned: Weds, Sep 4
Due: Weds, Sep 11

The usual formatting rules:

- Your homework (HW) should be formatted to be easily readable by the grader.
- You may use knitr or Sweave in general to produce the code portions of the HW. However, the output from knitr/Sweave that you include should *be only what is necessary to answer the question*, rather than just any automatic output that R produces. (You may thus need to avoid using default R functions if they output too much unnecessary material, and/or should make use of `invisible()` or `capture.output()`.)
  - For example: for output from regression, the main things we would want to see are the estimates for each coefficient (with appropriate labels of course) together with the computed OLS/linear regression standard errors and p-values. If other output is not needed to answer the question, it should be suppressed!
- Code snippets that directly answer the questions can be included in your main homework document; ideally these should be preceded by comments or text at least explaining what question they are answering. Extra code can be placed in an appendix.
- All plots produced in R should have appropriate labels on the axes as well as titles. Any plot should have explanation of what is being plotted given clearly in the accompanying text.
- Plots and figures should be *appropriately sized,* meaning they should not be too large, so that the page length is not too long. (The arguments fig.height and fig.width to knitr chunks can achieve this.)
- **Directions for "by-hand" problems:** In general, credit is given for (correct) shown work, not for final answers; so show **all** work for each problem and explain your answer fully.

1. Read, understand, and be able to execute the R material in the R appendix in Shumway and Stoffer ("Appendix R"), stopping just before the material on `lm`. (This is pages 533–539.)

   In particular,

   - Install and run R on your computer (`https://cran.r-project.org/` has the download and instructions).
   - Install the astsa package:

```
install.packages(``astsa'') ## installs on computer
library(astsa) ## loads for one interactive session
```

   - Additionally, read the help functions for the `start`, `end`, `frequency`, `cycle` (useful with boxplot), `aggregate`, `ts`, `ts.intersect`, and `window` commands. See also the `plot.ts` and `boxplot` functions.

   Then:

   (a) Do Question 1.1 (pp 38) in Shumway and Stoffer, and

   (b) There is a data file `cbe.txt` on Canvas. The column labelled "elec" gives the monthly supply of electricity (millions of kWh) in Australia from Jan 1958 to to Dec 1990. (It is available from the Australian Bureau of Statistics.)

      i. Produce a time plot (meaning plot the data against time) of the electricity production data. Then plot the aggregated annual series. Also create boxplots for the data by quarter (meaning, side-by-side box plots, one for each quarter of the year (Q1 is Jan, Feb, March, etc.). Comment on the plots.

      ii. Consider also the data available in R as the variable `AirPassengers`. This contains the number of international passenger bookings (thousands) per month of an airline (Pan Am) in the USA (obtained from the Federal Aviation Administration) for the period 1949–1960. The data is interesting for predicting future demand (so that orders of aircraft and training for air crew can be planned).
      Produce two time plots, of both the electricity and the airline data, only for the time period where both datasets are defined. Use `layout` or `split.screen` or `par(mfrow=)` to put the two plots on one screen, with one above the other, so that they could visually be compared. Provide brief comments on the plots.

You may see the R short reference card pdf posted on Canvas. In particular, note the usage of the `help` command, or the shorter form "`?`" (as in, `?plot` or `?start`).

2. Suppose that (X, Y) on $[0,1]^2$ have the bivariate cumulative distribution function

$$F(x,y) := xy\{1 + \theta(1-x)(1-y)\}$$

for $0 \le x, y \le 1$ and $|\theta| \le 1$.

   (a) Find the joint density $f(x,y)$ of (X,Y) on $[0,1]^2$.

   (b) Find the (marginal) density of X. (Which is the same as the marginal density of Y.)

   (c) Find the mean and variance of X (and thus also of Y).

   (d) Find the covariance and correlation of X and Y.

   (e) Say $X_1, \ldots, X_n$ are i.i.d. from the distribution of X above. Let $\overline{X}_n := \frac{1}{n}\sum_{i=1}^n X_i$ and $\overline{X^2}_n := \frac{1}{n}\sum_{i=1}^n X_i^2$. Find $\mathrm{Cov}(\overline{X}_n, \overline{X^2}_n)$.

   (f) What does $\overline{X}_n$ converge to in probability (as n gets large)? What does $\overline{X^2}_n$ converge to?

   (g) What is the asymptotic distribution of $\overline{X}_n$?

3. (Complex numbers / analysis) This question reviews some of the basic theory of complex numbers. Here $i = \sqrt{-1}$ and a complex number is of the form $a + bi$ for $a, b \in \mathbb{R}$.

   (a) Find the following. Write your answer in the form $a + bi$.

      i. $\overline{1 - 2i}$ (i.e., find the complex conjugate of $1 - 2i$)

      ii. $(1 - 2i) - (3 + 4i)$

      iii. $(1 - 2i)(3 + 4i)$

      iv. $(1 - 2i)/(3 - 4i)$

      v. $\sqrt{-3}\sqrt{-12}$.

   (b) If $z = 1/4 + i/4$, is $\sum_{j=0}^{\infty} z^j$ well-defined (why/why not)? If so find its value, simplified to the form $a + bi$.

   (c) Show that $|e^{i\pi t}| = 1$ for any $t \in \mathbb{R}$.

   (d) How many complex roots (or zeros) does the polynomial $a_{10}z^{10} + a_9 z^9 + \cdots + a_1 z + a_0$ have (counting non-unique roots separately) where the coefficients a_i could be any complex numbers, except $a_{10} \neq 0$?

   (e) Find all the zeros of $2z^2 + z + 3$?