Math 530, HW4

Due December 11, 2017

1. One of the goals of the Edinburgh Artery Study was to investigate the risk factors for peripheral arterial disease among persons 55 to 75 years of age.You wish to compare mean LDL cholesterol levels, measured in mmol/liter, among four different populations of subjects: patients with intermittent claudication or interruptions in movement, those with major asymptomatic disease, those with minor asymptomatic disease, and those with no evidence of disease at all. Samples are selected from each population; summary statistics are shown below:

   a. Use a one-way ANOVA to test the null hypothesis that the mean LDL cholesterol levels are the same for each of the four populations. Calculate this by-hand, and include your ANOVA table.

   b. If $\alpha = .05$, what do you conclude?
   c. What assumptions about the data must be true for you to use one-way ANOVA?
   d. Is it necessary to take any additional steps in this analysis? If so, what?

	n	$\bar{x}$	s
Intermittent Claudation	73	6.22	1.62
Major Asymptomatic Disease	105	5.81	1.43
Minor Asymptomatic Disease	240	5.77	1.24
No Disease	1080	5.47	1.31

2. In the early 1900s, Latter (1902) investigated the behavior of female cuckoos, that lay their eggs on the ground and then move them to the nests of other birds. In particular, Latter gathered data on the lengths of the cuckoo eggs found in these foster-nests. Data based on this work is used in (Tippett, 1952) and is located in the file cuckoos. The data contains the lengths, in millimeters, of the lengths of cuckoo eggs and the species of the nests where the eggs were placed. Get the data by installing and loading the **resampledata** R package, and use the **Cuckoos** dataset.

```
library(resampledata)
cuckoos <- Cuckoos
head(cuckoos)
```

```
##   ID  Eggs        Bird
## 1  1 19.65 MeadowPipit
## 2  2 20.05 MeadowPipit
## 3  3 20.65 MeadowPipit
## 4  4 20.85 MeadowPipit
## 5  5 21.65 MeadowPipit
## 6  6 21.65 MeadowPipit
```

   a. Create side-by-side boxplots (in R) to compare the distribution of lengths across the different foster nests.
   b. Conduct an ANOVA test (also in R) to see if the mean lengths of the cuckoo eggs are the same across the different foster nests.
   c. Perform the Tukey Honestly Significant Difference test (without p-value adjustment) to compare all pairwise means. What can you conclude from this analysis?
   d. Do the Tukey HSD test using the p-value adjustment method of your choice. Do your conclusions from "2c" change? Given the number of pairwise contrasts, without p-value adjustment, what would be your family-wise error rate if you were to conduct each pairwise contrast at $/alpha = .05$?