

Homework 5  
MTH 3220, Fall 2019  
Due Thursday, Oct. 10

For **Problems 3 and 5** (but not 7), you must **show your work** to receive full credit. The other problems can be done in R.

Section in Book	Problems
10.1	3, 5, 6*, 7, 8*, 9*
10.2	Problem 1 (given below), 18**

\* For **Problems 6, 8 and 9**, you can use R to do the one-factor ANOVA. The data are in text files on the course website, and can be read into R using **read.table()**. Once the data have been read into R, use **aoov()** to carry out the ANOVA and **summary()** to look at the ANOVA table.

\*\* For Problem 18, you can use R to do the one-factor ANOVA. Then after saving the results from **aoov()** as, say, **my.anova**, you can carry out the Tukey procedure by typing:

**> TukeyHSD(my.anova)**

Additional Problem

**Problem 1.** A study investigated the productivity of secretaries with different word processing programs. An identical task was given to 18 secretaries, randomly allocated to three treatment groups. Group 1 used a primarily menu-driven program. Group 2 used a command-driven program and Group 3 used a mixture of both approaches. The secretaries all had about the same level of experience, typing speed, and computer skills. The time (in minutes) taken to complete the task was observed. The data are shown below. They're also in the file **ergonomics.txt** on the course website.

Group 1 Menu-Driven	Group 2 Command-Driven	Group 3 Mixed
12	14	10
15	11	8
12	12	10
10	11	7
11	13	9
13	14	8

We want to decide if there are any significant **differences** between the completion times for the **three treatment groups**.

Use **read.table()** (or **read.csv()**) to read to the data from **ergonomics.txt** into an R *data frame* named, say, **my.data**.

- i. Make side by side boxplots of the three sets of completion times by passing a *formula* (**Time ~ Group**) and a *data frame* (**my.data**) to **boxplot()**.
- ii. State the **hypotheses** for a one-factor ANOVA *F* test in terms of the true means  $\mu_1$ ,  $\mu_2$ , and  $\mu_3$ .
- iii. Carry out the one-factor ANOVA by passing a *formula* (**Time ~ Group**) and a *data frame* (**my.data**) to **aoov()**. Save the result as, say, **my.anova**, then report the ANOVA table obtained by:

**> summary(my.anova)**

- iv. Based on the ANOVA *F* test, is there statistically significant evidence for any differences among the mean completion times for the three word processing programs? Use level of significance  $\alpha = 0.05$ .
- v. If  $H_0$  was rejected in Part iv, carry out *Tukey's multiple comparisons procedure* to decide **which** treatment groups differ from each other by typing:

**> TukeyHSD(my.anova)**

- vi. Compute the three **sample mean completion times**  $\bar{X}_{1.}$ ,  $\bar{X}_{2.}$ , and  $\bar{X}_{3.}$  by typing:

**> aggregate(Time ~ Group, data = my.data, FUN = mean)**

Which sum of squares, **SSTr** or **SSE**, reflects variation in the observed values of  $\bar{X}_{1.}$ ,  $\bar{X}_{2.}$ , and  $\bar{X}_{3.}$ ?

- vii. If the secretaries had widely differing levels of experience, typing speed, and computer skills, how would that affect the value of the error sum of squares **SSE** and the mean squared error **MSE**?
- viii. Compute the three **sample variances**,  $S_1^2$ ,  $S_2^2$ , and  $S_3^2$ , by typing:

**> aggregate(Time ~ Group, data = my.data, FUN = var)**

Show that the **MSE** is equal to the **average** of  $S_1^2$ ,  $S_2^2$ , and  $S_3^2$ .