Stat 445/545: Assignment 1 Due Feb 5 2020 Wed in class

Problem 1 Install R, Rstudio, Rmarkdown, and also Texworks `https://softfamous.com/texworks/` in your local computer. Download hw0.Rmd or hw0sweave.Rnw from my website, modify the files a little bit to get familiar with the code, then turn in the generated pdf file or html file for credits.

Problem 16.3 In a study of intentions to get flu-vaccine shots in an area threatened by an epidemic, 90 persons were classified into three groups of 30 according to the degree of risk of getting flue. Each group was together when the persons were asked about the likelihood of getting the shots, on a probability scale ranging from 0 to 1.0. Unavoidably, most persons overheard the answers of nearby respondents. An analyst wishes to test whether the mean intent scores are the same for the three risk groups. Consider each assumption for ANOVA model (16.2) and explain whether this assumption is likely to hold in the present situation.

$$Y_{ij} = \mu_i + \epsilon_{ij}, \text{ for } i = 1, \cdots, r; j = 1, \cdots, n_i \text{ Model (16.2)}$$

where Y_{ij} is the value of the response variable in the jth trial for the ith factor level or treatment, μ_i are parameters, and ϵ_{ij} are independent $N(0, \sigma^2)$

Problem 16.5 In a study of length of hospital stay (in number of days) of persons in four income groups, the parameters are as follows: $\mu_1 = 5.1, \mu_2 = 6.3, \mu_3 = 7.9, \mu_4 = 9.5, \sigma = 2.8$. Assume that ANOVA model (16.2) is appropriate.

(a) Draw a representation of this model in the format of Figure 16.2 (a figure as the one in slide 10 from chap 16).

(b) Suppose 100 persons from each income group are randomly selected fro the study. Find $E(MSTR)$ and $E(MSE)$. Is $E(MSTR)$ substantially larger than $E(MSE)$ here? What is the implication of this?

(c) If $\mu_2 = 5.6$ and $\mu_3 = 9.0$, everything else remaining the same, what would $E(MSTR)$ be? Why is $E(MSTR)$ substantially larger here than in part (b) even though the range of the factor level means is the same?

Problem 16.6 A student asks:"Why is the F test for equality of factor level means not a two-tail test since any differences among the factor level means can occur in either direction?"

Explain, utilizing the expressions for the expected mean square in the following:

$$E(MSE) = \sigma^2$$

$$E(MSTR) = \sigma^2 + \frac{\sum n_i(\mu_i - \mu_\cdot)^2}{r-1}$$

where $\mu_\cdot = \dfrac{\sum n_i \mu_i}{n_T}$

Problem 16.9 (Do by hand, you can use calculator though, write down steps to find the values) Rehabilitation therapy. A rehabilitation center researcher was interested in examining the relationship between physical fitness prior to surgery of persons undergoing corrective knee surgery and time required in physical therapy until successful rehabilitation. Patient records in the rehabilitation center were examined, and 24 male subjects ranging in age from 18 to 30 years who had undergone similar corrective knee surgery during the past year were selected for the study. The number of days required for successful completion of physical therapy and the prior physical fitness status (below average, average, above average) for each patient follow:

Table 1: Data for problem 16.9

		j									
	i	1	2	3	4	5	6	7	8	9	10
1	Below average	29	42	38	40	43	40	30	42		
2	Average	30	35	39	28	31	31	29	35	29	33
3	Above average	26	32	21	20	23	22				

Assume that ANOVA cell mean model (16.2) is appropriate:

(a) Do the factor level means appear to differ? Does the variability of the observations within each factor level appear to be approximately the same for all factor levels?

(b) Obtain the fitted values.

(c) Obtain the residuals. Do they sum to zero?

(d) Obtain the ANOVA table.

(e) Test whether or not the mean number of days required for successful rehabilitations is the same for the three fitness groups. Control the α risk at 0.01. State the alternatives, decision rule, and conclusion.

(f) Obtain the P-value for the test in part (e). Explain how the same conclusion reached in part (e) can be obtained by knowing the P-value.
(g) What appears to the nature of the relationship between physical fitness status and duration of required physical therapy?

Stat 445/545: **Assignment 2, due Feb 19, Wednesday**

5th edition:
Chapter 17, 17.1, 17.4, 17.10
Chapter 18, 18.15, 18.16

Stat 445/545: **Assignment 3 Due Mar 4th, Wednesday**

5th edition:
Chapter 19, 19.4, 19.7, 19.14, 19.15, 19.26, 19.32(a-d)

Stat 445/545: **Assignment 4 Due Mar 25, Wednesday**

5th edition:
Chapter 20, 20.5, 20.7
Chapter 21, 21.5, 21.6 (a, c, d)
Chapter 23: 23.1, 23.4 (a, d, e, f)