**Stat 191** **Assignment 4** **Winter 2018**

The assignment is due on March 5. For the first problem, you need to show your work and answers as well as plots and **R** code and output to substantiate your answers. You may find it easier to do problem 2 by hand rather than with **R**. You don't have to typeset your solution; handwritten answers together with a printout of code and output are fine.

**1.** A sampling of the acidity of rain for ten randomly selected rainfalls was recorded at three different locations in the United States, the Northeast, the middle Atlantic region, and the Southeast. The pH readings for these thirty rainfalls are shown below. (Note: pH readings range from 0 to 14; 0 is acid, 14 is alkaline. Pure water falling through the air has a pH reading of 5.7.)

| Northeast | Middle Atlantic | Southeast |
|---|---|---|
| 4.45 | 4.60 | 4.55 |
| 4.02 | 4.27 | 4.31 |
| 4.13 | 4.31 | 4.84 |
| 3.51 | 3.88 | 4.67 |
| 4.42 | 4.49 | 4.28 |
| 3.89 | 4.22 | 4.95 |
| 4.18 | 4.54 | 4.72 |
| 3.95 | 4.76 | 4.63 |
| 4.07 | 4.36 | 4.36 |
| 4.29 | 4.21 | 4.47 |

Analyze the above data to assess differences between the level of acidity in rainfalls for the three different locations. In your analysis you should try not only to discern the presence of an overall difference between the acidity of rainfall in the three regions, but also to quantify differences between the regions. Carefully state any models fitted and assumptions made and give justifications for your assumptions.

**2.** Some year ago the Los Angeles Times ran the following article under the headline **Marijuana-Hormone Study Slated – Ties to low male level sought**:

> "A report that marijuana causes a sharp drop in the male sex hormone level of heavy users is being investigated by UCLA marijuana researchers. The report, published last week in the New England Journal of Medicine, suggested that steady, intensive use of marijuana lowers the level of the male sex hormone, testosterone, and may have adverse effects on sexual functions.
>
> Drs. Robert C. Kolodny and William H. Masters from the Reproductive Biology Research Foundation, St. Louis, prepared the report that involved 20 young males who used marijuana at least four days a week for six months or longer. The St. Louis researchers periodically measured the testosterone levels in the subjects and in a similar group of young males who had never used marijuana. They found a statistically significant lower level of testosterone in the subjects who said they used 10 or more marijuana 'joints' a week as compared to those who used four to nine joints. Both groups of users had lower testosterone levels than the nonusers. Each of three subjects who stopped using for a two-week period showed a pronounced rise in testosterone levels within a week of stopping. About one-third of the users had lower-than-normal sperm counts, according to Kolodny and Masters.

Dr. Ira Frank, director of UCLA's Marijuana Research Project, said in an intervew he is now working with Kolodny on a study to be conducted at UCLA to see if the St. Louis study can be confirmed under controlled conditions. One of the criticisms of the St. Louis study is that the subjects used street marijuana, the strength and purity of which were unknown. Although the subjects said they were not using other drugs, they were not under continual surveillance and it is possible that the results attributed to marijuana may have been due to something else.

In the UCLA study, which is already underway, marijuana of known strength will be used and the subjects will be under controlled conditions 24 hours a day for a 94 day period. Theoretically, according to Koldny, it is possible that if a pregnant woman smokes marijuana, a male fetus may be adversely affected because an adequate level of testosterone appears to be necessary for certain stages of sexual differentiation. It is also possible he said, that marijuana smoking by prepuberty males may delay the onset of puberty. Kolodny also speculated that the reduced levels of testosterone may be associated with the apathy and diminished drive that have been correlated with marijuana use. This speculation is based on evidence that aggression is correlated with testosterone levels.

'Anecdotal accounts of heightened sexual drive and feelings associated with marijuana use have often been heard,' Kolodny and Masters said in the journal. 'Now the possibility of an adverse effect of frequent marijuana use on male sexual functioning must be considered.' The results of the UCLA study are expected by the end of the year, according to the researchers."

The table below gives data from Kolodny and Master's paper "Depression of Plasma Testosterone Levels after Chronic Intensive Marijuana Use", *New England Journal of Medicine*, **16**, 872–874.

**Summary Statistics on Plasma Testosterone Levels (ng/100ml)**

| | Control Group | All users Total | Mild users 5–9 joints/wk | Heavy users $\geq$ 10 joints/wk |
|---|---|---|---|---|
| $\bar{x}$ | 742 | 416 | 503 | 309 |
| $s/\sqrt{n}$ | 29 | 34 | 40 | 34 |
| $n$ | 20 | 20 | 11 | 9 |

In answering the following questions, assume normality and homoscedasticity.

(a) Test the hypothesis of equality of mean plasma testosterone levels among the control group, those who use five to nine joints per week, and those who use ten or more joints per week.

(b) Test the hypothesis that the mean plasma testosterone levels in the users and the control group are the same. Also test the hypothesis that the mean plasma testosterone levels in the "5–9 joints/wk" and the "10 or more joints/wk" groups are equal. Use $\alpha = 0.05$ in each case.

(c) Use the Tukey method of multiple comparisons to form 95% confidence intervals for all pairwise comparisons of means between the three groups (control, 5–9 joints/wk, 10 or more joints/wk). (You can get the critical values of the Studentized range distribution from `R` or from the web.)

(d) Use the Scheffé method of judging all contrasts in a one-way analysis of variance to obtain 95% simultaneous confidence intervals for the contrasts $\mu_1 - \mu_2$, $\mu_1 - \mu_3$ and $\mu_1 - \frac{1}{2}(\mu_2 + \mu_3)$, where the means $\mu_1, \mu_2$ and $\mu_3$ correspond to the three groups tested.

(e) Form 95% confidence intervals for the same contrasts listed in (d) if these were the only contrasts you would ever want to consider. (Hint: Bonferroni!)

(f) Interpret carefully the results of your multiple comparisons in (c), (d) and (e) in terms of the *LA Times* article.