

STAT 3202: Practice 09

Spring 2019, OSU

Exercise 1

An experiment was designed to test the effect of nematodes (microscopic worms) on plant growth. A botanist prepares 16 identical planting pots and then introduces different numbers of nematodes into the pots. Four pots get 0 nematodes, four get 1,000, four get 5,000, and four get 10,000. A tomato seedling is transplanted into each pot. Seedlings are assigned to pots (i.e., to treatments) using a completely randomized design. Below are the data on the increase in height of the seedlings in centimeters 16 days after planting.

Number of Nematodes			
0	1,000	5,000	10,000
10.8	11.1	5.4	5.8
9.1	11.1	4.6	5.3
13.5	8.2	7.4	3.2
9.2	11.3	5.0	7.5

In R, perform the ANOVA overall F -test, at $\alpha = 0.01$. Then use the `TukeyHSD` function to determine what differences there are among the treatment groups. What levels of nematodes produce different seedling growth, at family-wise error rate (FWER) 0.01? Summarize in a sentence or two what this tells us about the treatment levels.

Exercise 2

Water samples were taken at four different locations in a river to determine whether the quantity of dissolved oxygen, a measure of water pollution, differed from one location to another. Locations 1 and 2 were selected above an industrial plant, one near the shore and the other in midstream; location 3 was adjacent to the industrial water discharge for the plant; and location 4 was slightly downriver in midstream. Five water specimens were randomly selected at each location, but one specimen, from location 4, was lost in the laboratory. The sample sizes, means, and variances are shown in the accompanying table (the greater the pollution, the lower will be the dissolved oxygen readings).

Investigate whether locations 1, 2, and 3 are different from location 4. That is, if μ_i is the mean dissolved oxygen at location i , build confidence intervals on $\mu_1 - \mu_4$, $\mu_2 - \mu_4$, and $\mu_3 - \mu_4$. Construct the three confidence intervals using a *Bonferroni correction* to ensure that the *family-wise* confidence level is 95%.

Location	Sample Size	Sample Mean	Sample Variance
1	5	6.08	0.022
2	5	6.44	0.013
3	5	4.78	0.097
4	4	6.025	0.029

Exercise 3

Do you hate Mondays? Researchers in Germany have provided another reason for you: They concluded that the risk of heart attack on a Monday for a working person may be as much as 50% greater than on any other day. The researchers kept track of heart attacks and coronary arrests over a period of 5 years among 330,000 people who lived near Augsburg, Germany. In an attempt to verify the researcher's claim, 200 working people who had recently had heart attacks were surveyed. The day on which their heart attacks occurred appear in the following table.

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
24	36	27	26	32	26	29

Do these data present sufficient evidence to indicate that there is a difference in the percentages of heart attacks that occur on different days of the week? Test using $\alpha = 0.05$.

Exercise 4

Is the chance of getting a cold influenced by the number of social contacts a person has? A study by Sheldon Cohen, a psychology professor at Carnegie Melon University, seems to show that the more social relationships a person has, the *less susceptible* the person is to colds. A group of 276 healthy men and women were grouped according to their number of relationships (such as parent, friend, church member, and neighbor). They were then exposed to a virus that causes colds. An adaptation of the results is given in the following table.

	Number of Relationships		
	3 or fewer	4 or 5	6 or more
Cold	49	43	34
No cold	31	57	62
Total	80	100	96

Do the data present sufficient evidence to indicate that susceptibility to colds is affected by the number of relationships that people have? Test at the 5% level of significance.

Exercise 5

Returning to the risk of developing a cold data in the previous exercise, calculate two confidence intervals for $p_1 - p_2$ and $p_1 - p_3$, where p_i is the probability of getting a cold in Group i , where Group 1 has 3 or fewer relationships; Group 2 has 4 or 5; and Group 3 has 6 or more. Use a Bonferroni correction to calculate these confidence intervals so as to maintain a *family-wise confidence level* of 95%.