STAT 135 CONCEPTS OF STATISTICS
HOMEWORK 5

Assigned July 22, 2021, due July 27, 2021

This homework pertains to materials covered in Lecture 9 and 10. The assignment can be typed or handwritten, with your name on the document, and **with properly labeled input code and computer output for those problems that require it**. To obtain full credit, please write clearly and show your reasoning. If you choose to collaborate, the write-up should be your own. Please show your work! Upload the file to the Week 5 Assignment on bCourses.

Note in this homework, we use the following abbreviations: Uniformly most powerful (UMP) test, likelihood ratio test (LRT).

Problem 1. An experiment was done to measure the effects of ozone, a component of smog. A group of 22 seventy-day-old rats were kept in an environment containing ozone for 7 days, and their weight gains were recorded. Another group of 23 rats of a similar age were kept in an ozone-free environment for a similar time, and their weight gains were recorded. The data (in grams) are given below.

$$\begin{aligned}\mathbf{X}_{\text{control}} = \{&41.0,\ 38.4,\ 24.9,\ 25.9,\ 21.9,\ 18.3,\ 13.1,\ 27.3,\ 28.5,\\ &-16.9,\ 17.4,\ 21.8,\ 15.4,\ 27.4,\ 19.2,\ 22.4,\ 17.7,\ 26.0,\\ &29.4,\ 21.4,\ 22.7,\ 26.0,\ 26.6\}\\ \mathbf{X}_{\text{ozone}} = \{&10.1,\ 6.1,\ 20.4,\ 7.3,\ 14.3,\ 15.5,\ -9.9,\ 6.8,\ 28.2,\\ &17.9,\ -12.9,\ 14.0,\ 6.6,\ 12.1,\ 15.7,\ 39.9,\ -15.9,\ 54.6,\\ &-14.7,\ 44.1,\ -9.0,\ -9.0\}\end{aligned}$$

Establish a null and an alternative hypotheses, and analyze the data using both Normal population assumption and the non-parametric Mann-Whitney test to determine the effect of ozone. Write a summary of your conclusions. [This problem is from Doksum and Sievers (1976) who provide an interesting analysis.]

Problem 2. A cross-over trial investigated whether eating oat bran lowered serum cholesterol levels. Fourteen individuals were randomly assigned a diet that included either oat bran or corn flakes. After two weeks on the initial diet, serum cholesterol were measured and the participants were then "crossed-over" to the alternate diet. After two-weeks on the second diet, cholesterol levels were once again recorded.

Data appear below. The variables CORNFLK and OATBRAN in the table represent cholesterol levels (mmol/L) of the participant on the corn flake diet and the oat bran diet respectively.

Subject	CORNFLK	OATBRAN
1	4.61	3.84
2	6.42	5.57
3	5.40	5.85
4	4.54	4.80
5	3.98	3.68
6	3.82	2.96
7	5.01	4.41
8	4.34	3.72
9	3.80	3.49
10	4.56	3.84
11	5.35	5.26
12	3.89	3.73
13	2.25	1.84
14	4.24	4.14

(*You can download this dataset from the Data_sets directory on bCourses*).

(1) Use normal theory to test the hypothesis that the cholesterol levels while on the corn flake diet is less than the oat bran diet ($\alpha = 0.05$);
(2) Perform Wilconxon signed rank test by hand to test the same hypothesis ($\alpha = 0.05$);
(3) Compare the p-value from the two tests. Are they very different?

Problem 3. Let X equal the number of alpha particles emitted from barium-133 in 0.1 second and counted by a Geiger counter. One hundred observations of X produced the following table:

Category	X	**Obs'd**
1	0,1,2*	5
2	3	13
3	4	19
4	5	16
5	6	15
6	7	9
7	8	12
8	9	7
9	10,11,12*	4
		$n = 100$

It is claimed that X follows a Poisson distribution. Use a chi-square goodness-of-fit statistic to test whether this is true.

Problem 4. During phase 3 trial of the Pfizer/BNT 162b2 vaccine, some vaccine recipients were asked to complete diaries of their symptoms during the 7 days after vaccination.

Here is a table summarizing the number of recipients who experienced headaches after the second dose:

	Pfizer-BioNTech Vaccine	Placebo
No headache	1013	1597
Mild	538	321
Moderate	480	170
Severe	67	15

TABLE 1. Systemic reactions in persons aged 18-55 years. Data from CDC.gov.

Use the test for homogeneity to examine whether two columns in Table 1 as multinomial variables have equal probabilities of having headaches of a certain severity.

Problem 5. Is age independent of the desire to ride a bicycle? A random sample of 395 people were surveyed. Each person was asked their interest in riding a bicycle (Variable A) and their age (Variable B). The data that resulted from the survey is summarized in the following table:

		Variable B (Age)				
	OBSERVED	**18-24**	**25-34**	**35-49**	**50-64**	**Total**
Variable A	**Yes**	60	54	46	41	201
	No	40	44	53	57	194
	Total	100	98	99	98	395