AMS 315 Data Analysis: Homework 3
Due: February 25, 2022

1. (10 pts) Conduct a test of $H_0 : \mu_1 \leq \mu_2$ versus $H_a : \mu_1 > \mu_2$ for the sample data summarized here. Use $\alpha = 0.05$. Assume that we have normal populations with $\sigma_1 = \sigma_2$.

Group	1	2
Sample size	15	12
Sample mean	70.1	75.5
Sample Standard Deviation	7.23	4.68

2. (10 pts) Let $\Delta = \mu_1 - \mu_2$. Conduct a test of $H_0 : \Delta = 0$ versus $H_a : \Delta > 0$ for the sample data given here. Use $\alpha = 0.05$.

Treatment 1	5.5	4.9	5.3	5.1	5.5	4.6	5.8
Treatment 2	4.3	3.8	4.7	3.9	5.3	4.4	5.2

3. (10 pts) Consider the paired data given here.

Pair	1	2	3	4	5	6
y_1	48.2	41.3	47.4	50.6	43.5	49.1
y_2	38.5	39.1	43.0	47.8	40.2	41.7

   (a) Conduct a paired t test of $H_0 : \mu_d \leq 0$ versus $H_a : \mu_d > 0$ with $d = y_1 - y_2$. Use $\alpha = 0.05$.
   (b) Using a testing procedure related to the binomial distribution, test the hypotheses in (a).

   Does your conclusion agree with the conclusion reached in part (a)?

   Hint: Try to compute p-value using binomial distribution under assumption that $P(y_1 > y_2) = 0.5$. You will need to build new T.S. following that binomial distribution.

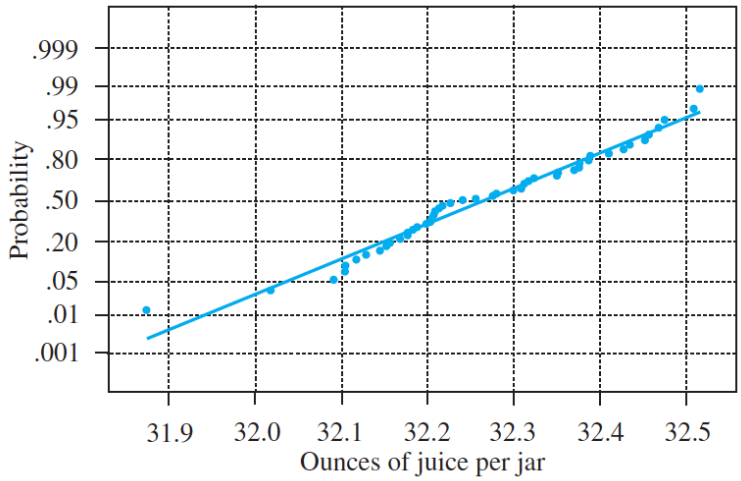
4. (10 pts) A study is being planned to evaluate the possible side effects of an anti-inflammatory drug. It is suspected that the drug may lead to an elevation in the blood pressure of users of the drug. A preliminary study of two groups of patients, one receiving the drug and the other receiving a placebo, provides the following information on the systolic blood pressure (in mm Hg) of the two groups:

Group	Mean	Standard Deviation
Placebo	125.0	19.5
Users of drug	129.5	19.7

Assume that both groups have systolic blood pressures that have a normal distribution with standard deviations relatively close to the values obtained in the pilot study. Suppose the study plan provides for the same number of patients in the placebo as in the treatment group. Determine the sample size necessary for an $\alpha = 0.05$ t test to have power of 0.90 to detect an increase of 3 mm Hg in the blood pressure of the treatment group relative to the placebo group.

5. (15 pts) Random samples of $n_1 = 21$ and $n_2 = 21$ were selected from populations 1 and 2, respectively. The corresponding sample standard deviations were $s_1 = 17.5$ and $s_2 = 10.5$.
(a) Do the data provide sufficient evidence ($\alpha = 0.05$) to indicate a difference in σ_1 and σ_2?
(b) Place a 95% confidence interval on the ratio of the variances σ_1/σ_2.

6. (15 pts) A packaging line fills nominal 32-ounce tomato juice jars with a quantity of juice having a normal distribution with a mean of 32.20 ounces. The process should have a standard deviation smaller than 0.15 ounces per jar. (A larger standard deviation leads to too many underfilled and overfilled jars.) A random sample of 51 jars is taken every hour to evaluate the process. The data from one such sample are summarized here.



Variable	N	Mean	Median	TrMean	StDev	SE Mean	Min	Max	Q1	Q3
Juice Jars	51	32.2	32.2	32.3	0.14	0.02	31.8	32.5	32.1	32.3

(a) If the process yields jars having a normal distribution with a mean of 32.20 ounces and a standard deviation of 0.15 ounces, what proportion of the jars filled on the packaging line will be underfilled?

(b) Does the plot suggest any violation of the conditions necessary to use the chi-square procedures for generating a confidence interval and a test of hypotheses about σ?

(c) Construct a 95% confidence interval on the process standard deviation.

(d) Do the data indicate that the process standard deviation is greater than 0.15?

Use $\alpha = 0.05$.

(e) Place bounds on the p-value of the test.

Note: Following two Problems are R programing problem.

7. (15 pts) The Texas A&M extension service wanted to investigate if the mean yield per acre of soybeans (in bushels) was greater than 500 bushels. In a random sample of 36 1-acre soybean plots, the sample mean and standard deviation were computed to be $\bar{y} = 553$ and $s = 124$, respectively. Set up all the parts of a statistical test for the soybean example, and use the sample data to reach a decision on whether to accept or reject the null hypothesis. Set $\alpha = 0.025$. Assume that σ can be estimated by s.

8. (15 pts) A massive multistate outbreak of foodborne illness was attributed to *Salmonella enteritidis*. Epidemiologists determined that the source of the illness was ice cream. They sampled ten production runs from the company that had produced the ice cream to determine the level of *Salmonella enteritidis* in the ice cream. These levels (MPN/g) are as follows, which are assumed as normal population:

   0.814, 0.208, 0.446, 0.591, 0.335, 0.476, 0.424, 0.229, 0.588, 0.514

   (a) What is your point estimation on mean level of *Salmonella enteritidis*?
   (b) Place a 95% confidence interval on mean level.
   (c) Use these data to determine whether the average level of *Salmonella enteritidis* in the ice cream is greater than 0.3 MPN/g, a level that is considered to be very dangerous. Set $\alpha = 0.01$.