**Assignment #2**

**Due:** **Wednesday January 26**, 2022 by **5 PM ET**

**Please sign** your name to the **appropriate** space below**.** Remember that you are permitted to receive (and provide) authorized assistance but must acknowledge it if you do.

I received assistance on this assignment and/or discussed it with fellow classmates or a tutor.

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

I received no assistance on this assignment and/or did not discuss it with anyone other than Professor Miller.

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**



**Brooke Wheeler**

**Homework\_2**

**Part 1: Variance**

**Instructions:** In these problems we are going to focus on variance. Provide complete justification for each problem. You may type your work or include work done by hand. Please make sure that all work is neat, organized, and shows your complete understanding of the problem.

1. Let’s start by looking at the sample variance. Generally, we think of the sample variance as

. Show that .

Schematic

Description automatically generated with low confidence

1. In class we reviewed the sampling distribution of .Show that . (Hint: Use the properties of variance we reviewed in class.)Diagram, letter, schematic

   Description automatically generated
2. What does it mean for to be an unbiased estimator of ?
   1. For to be an unbiased estimator of means that the sampling distribution of the sample variance has a mean that is equal to the population variance. This means the expected value of the sample variance is equal to the population variance.

**Part 2: Working with T Tests and F Tests**

**Instructions:** For each question below, you must show all work. Use R to find p-values and critical values. If *α* is not given and it is difficult to make conclusions, use *α* = 0*.*05. For two-sample problems, start with the F-test to decide which t-test to use. You may type your work or include work done by hand.

1. A researcher wants to examine if women lie about how much they weigh. She suspects that women will tend to give a weight less than their true weight. She looks at a large population of women whose average weight is 145 pounds. She then takes a random sample of 56 women from this population and asks them their weight (after giving them the opportunity to weigh themselves beforehand). She finds that the 56 women report an average weight of 142.2 pounds with standard deviation 9.4 pounds.
   1. Does the researcher’s data provide sufficient evidence to support her claim? Be specific.
   2. Construct a 97% confidence interval for the true mean reported weight of the women.

Graphical user interface, text, application

Description automatically generated

A piece of paper with writing on it

Description automatically generated with medium confidence

Since the p-value is .064 we fail to reject the null hypothesis. Meaning we do not have statistically sufficient evidence to say that women will give a weight less than their true weight.

1. There are two manufacturing processes, old and new, that produce the same product. The defect rate has been measured for 20 days for the old process, and for 14 days for the new process, resulting in the following sample summaries.

|  |  |  |
| --- | --- | --- |
|  | OLD | NEW |
| Average defect rate | 4.7 | 2.3 |
| Standard deviation | 6.8 | 5.0 |

The firm is interested in switching to the new process only if it can be demonstrated convincingly that the new process reduces the defect rate. Is there significant evidence of that? Use *α* = 5% and assume that the collected data represent two random samples from Normal distributions. Use the method of testing that is appropriate for this situation.Text, letter

Description automatically generated Text, letter

Description automatically generatedGraphical user interface, text, application

Description automatically generated

1. At a p-value of .135 we fail to reject the null hypothesis that the two means are equal. Meaning that we do not have statistically significant evidence to prove that there is a difference between the old and new manufacturing processes.

**Part 3: Mini-Project**

1. Data on 522 recent home sales is posted on Canvas with this assignment *(“HOME\_SALES.csv”).* The following variables are included.

|  |  |
| --- | --- |
| Column | Variable |
| 1 | Identification number 1–522 |
| 2 | Sales price of residence (×$1000 dollars) |
| 3 | Finished area of residence (square feet) |
| 4 | Total number of bedrooms in residence |
| 5 | Total number of bathrooms in residence |
| 6 | Air conditioning: present or absent |
| 7 | Number of cars that garage will hold |
| 8 | Swimming pool: present or absent |
| 9 | Year property was originally constructed |
| 10 | Quality of construction: high, medium, or low |
| 11 | Architectural style. Three styles are coded as 1, 2, and 3 |
| 12 | Lot size (square feet) |
| 13 | Location near a highway: yes or no |

Note: In order to group a quantitative variable (such as Sales price) based on a categorical variable (such as Air conditioning) we can enter something like the following into R:

t.test(SALES\_PRICE[AIR\_CONDITIONER=="YES"],SALES\_PRICE[AIR\_CONDITIONER=="NO"])

1. The sales price depends on the air conditioner in the house. Graphical user interface, text, application, email

   Description automatically generated
2. On the average, homes with an air conditioner are more expensive.Graphical user interface, text, application, email

   Description automatically generated
3. On the average, homes with an air conditioner are larger.Graphical user interface, text, application, email

   Description automatically generated
4. The sales price depends on the proximity to a highway.Graphical user interface, text, email

   Description automatically generated
5. On the average, homes are cheaper when they are close to a highway.Graphical user interface, text, application, email

   Description automatically generated
6. On the average, homes are cheaper when they are far from a highway.

**Graphical user interface, text, application, email

Description automatically generated**