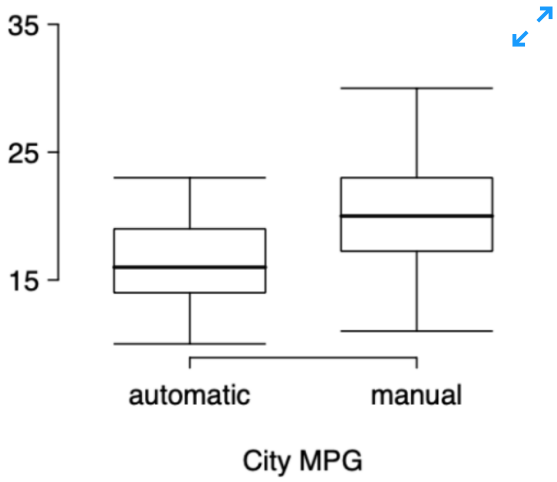
Problem Set #7 - Comparing Means of Multiple Groups

Part 1

Each year the US Environmental Protection Agency (EPA) releases fuel economy data on cars manufactured in that year. Below are summary statistics on fuel efficiency (in miles/gallon) from random samples of cars with manual and automatic transmissions.

	City MPG	
	Automatic	Manual
Mean	16.12	19.85
SD	3.58	4.51
n	26	26



Question 1

Review

The sample sizes of each of the two samples collected are a bit smaller than the general cutoff we use to ensure normality of the sampling distributions; however, after visualizing the data, we determine it's okay to continue with the hypothesis testing procedure. Why?

Select an answer and submit. For keyboard navigation, use the up/down arrow keys to select an answer.

a	The sample size conditions not met because the samples are both smaller than size 30 and the sampled distributions appear to be strongly skewed.
b	The boxplots of the sampled data indicate that both samples came from approximately symmetric distributions and therefore it is okay to move forward with hypothesis testing procedures.
c	We are certain that the population distributions of MPG for both automatic and manual cars are normal.

Correct Answer:

✓ **b** The boxplots of the sampled data indicate that both samples came from approximately - symmetric distributions and therefore it is okay to move forward with hypothesis testing procedures.

Show Submitted Answer | Hide Correct Answer | Check My Answer

Since we are interested in comparing the two populations, MPG of manual and automatic cars, the parameter of interest is the difference in the means of these two populations: $\mu_{\text{man}} - \mu_{\text{auto}}$.

Our point estimate for this parameter is $\mu_{\text{man}} - \mu_{\text{auto}}$.

Question 2

Review

Calculate the point estimate.

Type your numeric answer and submit

Correct Answer:

✓ 3.73

Show Submitted Answer | Hide Correct Answer | Check My Answer

The standard error for this point estimate is calculated using the sample standard deviations and sample sizes of the two samples.

Question 3
Review

Calculate the standard error. Round your answer to three decimal places.

Type your numeric answer and submit

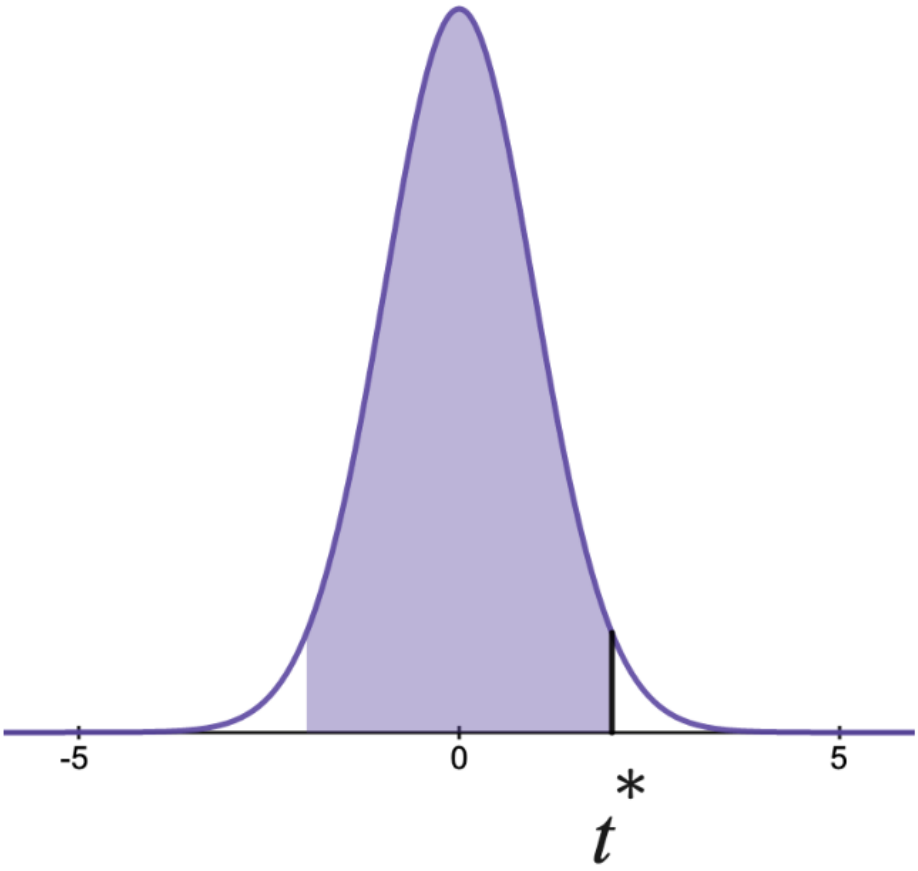
Correct Answer:

✓ 1.129

Show Submitted Answer | Hide Correct Answer | Check My Answer

If the conditions for inference are met, the calculation $t = \frac{\overline{x}_{\text{man}-\overline{x}_{\text{auto}}}}{SE}$ follows a *t* distribution with v degrees of freedom. The calculation for v is rather complicated. Use one of the tools available in Canvas to calculate the Satterthwaite degrees of freedom to find this value.

The *t* distribution below has the corresponding Satterthwaite degrees of freedom. The purple-shaded region represents the middle 95% of the distribution. To construct the 95% confidence interval for the difference in average MPG between manual and automatic cars, we first need to determine the critical value t^*.



Question 4
Review

Using R and the qt() function, find the critical value needed to construct the 95% confidence interval. Round to three decimal places.

Type your numeric answer and submit

Correct Answer:

✓ 2.011

Show Submitted Answer | Hide Correct Answer | Check My Answer

Question 5
Review

Fill in the Blanks

Type your answers in all of the blanks and submit

X_2 X^2 Ω

Using your answers from questions 2, 3, and 4, calculate the lower and upper bounds of the 95% confidence interval for the difference in average MPG between manual and automatic transmissions. Round your answers to two decimal places. Report the lower bound in the first space and upper bound in the second.

95% Confidence Interval: (Type your answer here , Type your answer here)

Correct Answers:

Using your answers from questions 2, 3, and 4, calculate the lower and upper bounds of the 95% confidence interval for the difference in average MPG between manual and automatic transmissions. Round your answers to two decimal places. Report the lower bound in the first space and upper bound in the second. 95% Confidence Interval: (

✓ 1.46

...,

✓ 6

Show Submitted Answer | Hide Correct Answer | Check My Answer

Part 2

People who eat lots of fruits and vegetables may have lower rates of colon cancer than those who eat little of these foods. Fruits and vegetables are rich in "antioxidants" such as vitamins A, C, and E. Will taking antioxidants help prevent colon cancer? A medical experiment studied this question with 864 people who were at risk of colon cancer. The subjects were divided into four groups: daily beta-carotene, daily vitamins C and E, all three vitamins every day, or daily placebo. After four years, the researchers were surprised to find no significant difference in colon cancer among the groups.

Question 6
Review

What is the treatment in this experiment?

Select an answer and submit. For keyboard navigation, use the up/down arrow keys to select an answer.

a	type of food (fruit and vegetables or neither)
b	whether or not colon cancer develops
c	at risk or not of colon cancer
d	whether or not subjects consumed beta-carotene, vitamins C and E or both each day

Correct Answer:

✓ **d** - whether or not subjects consumed beta-carotene, vitamins C and E or both each day

Show Submitted Answer | Hide Correct Answer | Check My Answer

Question 7
Review

What is the response in this experiment?

Select an answer and submit. For keyboard navigation, use the up/down arrow keys to select an answer.

a	type of food (fruit and vegetables or neither)

b	whether or not subjects consumed beta-carotene, vitamins C and E or both each day
c	at risk or not of colon cancer
d	whether or not colon cancer develops

Correct Answer:

✓ **d** - whether or not colon cancer develops

Show Submitted Answer | Hide Correct Answer | Check My Answer

Question 8
Review

What does "no significant difference" mean in describing the outcome of the study?

Select an answer and submit. For keyboard navigation, use the up/down arrow keys to select an answer.

a	The outcome of the study was not conclusive enough to draw significant conclusions regarding the population considered.
b	The observed differences were no more than what might reasonably occur by chance even if there is not effect due to the treatments.
c	The observed differences were not considered significantly large or significantly small.
d	There is no difference between the treatments for the population of everyone that might take the same treatments.

Correct Answer:

✓ **b** The observed differences were no more than what might reasonably occur by chance even if - there is not effect due to the treatments.

Show Submitted Answer | Hide Correct Answer | Check My Answer

Part 3

The lumen output was determined for each of $k = 3$ different brands of lightbulbs having the same wattage, with $n_i = 50$ bulbs of each brand tested (this is the number of observations in each treatment group). The mean squares were computed as **MSG = 299.1** and **MSE = 101.5**.

Question 9
Review

State the hypotheses of interest. Let μ_i = the true average lumen output for brand *i* bulbs

Select an answer and submit. For keyboard navigation, use the up/down arrow keys to select an answer.

a	$H_0 : \mu_1 = \mu_2 = \mu_3$ vs. H_A : at least one of the μ_i's differs from the other population means
b	$H_0 : \mu_1 \neq \mu_2 \neq \mu_3$ vs. H_A : all three μ_i's are equal
c	$H_0 : \mu_1 \neq \mu_2 \neq \mu_3$ vs. H_A : at least two of the μ_i's are equal

Correct Answer:

✓ a - $H_0 : \mu_1 = \mu_2 = \mu_3$ vs. H_A : at least one of the μ_i's differs from the other population means

Show Submitted Answer | Hide Correct Answer | Check My Answer

Question 10
Review

Use the Single Factor ANOVA *F* test with $\alpha = 0.05$ to decide whether there are any differences in true average lumen outputs among the three brands for this type of bulb. State the *F* statistic below. Round your answer to two decimal places.

Type your numeric answer and submit

Correct Answer:

✓ 2.95

Show Submitted Answer | Hide Correct Answer | Check My Answer

Question 11
Review

State the p-value. Round your answer to three decimal places.

Calculate the F test statistic then use R to find your p-value using the CDF of the f distribution pf(). Recall the p-value from an F test is always the area to the right of the F test statistic. Example: 1-pf(f test statistic, df1, df2)

Type your numeric answer and submit

Correct Answer:

✓ 0.056

Show Submitted Answer | Hide Correct Answer | Check My Answer

Question 12
Review

State the conclusion in the problem context.

There was previously an issue with the options for this question. It is now graded based on participation (if you answer you will get credit). Please try your best to find the correct answer as you are expected to know how to make the correct hypothesis testing conclusion.

Select an answer and submit. For keyboard navigation, use the up/down arrow keys to select an answer.

a	Fail to reject H_0. There is only slightly suggestive evidence that the average lumen output differs between the three brands.
b	Reject H_0. There is only slightly suggestive evidence that the average lumen output differs between the three brands.
c	Fail to reject H_0. There is convincing evidence that the average lumen output differs between the three brands.
d	Reject H_0. There is convincing evidence that the average lumen output differs between the three brands.

Correct Answer:

✓ **a** Fail to reject H_0. There is only slightly suggestive evidence that the average lumen output

Show Submitted Answer | Hide Correct Answer | Check My Answer

Part 4

Six samples of each of four types of cereal grain grown in a certain region (wheat, barley, maize, and oats) were analyzed to determine thiamin content (μg/g). A single factor ANOVA F test was performed. The following R output contains the ANOVA table.

	df	Sum Sq	Mean Sq	F	p-value
Grain	3	8.525	2.8415	***Answer to question 9***	***Answer to question 10***
Residuals	20	15.445	0.7722		

Question 13
Review

Do these data suggest that at least two of the grains differ with respect to true average thiamin content? State the appropriate hypotheses.

Select an answer and submit. For keyboard navigation, use the up/down arrow keys to select an answer.

a	H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$ vs. H_A: all four μ_i's are unequal
b	H0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$ vs. HA: at least one μ_i differs from the other population means
c	H0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$ vs. HA: at least two μ_i's are equal
d	H0: $\mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4$ vs. HA: at least two μ_i's are equal

Correct Answer:

✓ b - H0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$ vs. HA: at least one μ_i differs from the other population means

Show Submitted Answer | Hide Correct Answer | Check My Answer

Question 14
Review

Compute the test statistic value.(Round your answer to three decimal places.)

$$F = \frac{MSG}{MSE}$$

Type your numeric answer and submit

Correct Answer:

✓ 3.68

Show Submitted Answer | Hide Correct Answer | Check My Answer

Question 15
Review

Using the test statistic and the degrees of freedom for this ANOVA F test, calculate the p-value. Which of the following statements is true?

Use R to find your p-value using the CDF of the f distribution pf(). Recall the p-value from an F test is always the area to the right of the F test statistic. Example: 1-pf(f test statistic, df1, df2)

Select an answer and submit. For keyboard navigation, use the up/down arrow keys to select an answer.

a	p-value > 0.1
b	0.05 < p-value < 0.1
c	0.01 < p-value <0.05
d	p-value < 0.01

Correct Answer:

✓ **c** - 0.01 < p-value <0.05

Show Submitted Answer | Hide Correct Answer | Check My Answer

Question 16
Review

State the conclusion in the problem context. Use a significance level of 0.05.

Select an answer and submit. For keyboard navigation, use the up/down arrow keys to select an answer.

a	Reject H_0. There is moderately suggestive evidence that at least two of the grains differ in average thiamin content.
b	Fail to reject H_0. There is moderately suggestive evidence that at least two of the grains differ in average thiamin content.
c	Reject H_0. There is not significant evidence that at least two of the grains differ in average thiamin content.

d	Fail to reject H_0. There is not significant evidence that at least two of the grains differ in average thiamin content.

Correct Answer:

✓ **a** Reject H_0. There is moderately suggestive evidence that at least two of the grains differ in average thiamin content.

Show Submitted Answer | Hide Correct Answer

Open in Reading View