

## Problems Section 4B (Updated Version)

(#1-10) Use each of the following ANOVA F-test statistics and the corresponding critical values to fill out the table.

	F-test stat	Sentence to explain F-test statistic.	Critical Value	Does the F-test statistic fall in the right tail determined by the critical value? (Yes or No)	Does sample data significantly disagree with $H_0$ ?
1.	+5.573		+2.886		
2.	+1.192		+3.113		
3.	+0.664		+2.949		
4.	+4.415		+3.125		
5.	+3.718		+4.117		
6.	+0.991		+2.009		
7.	+2.652		+1.875		
8.	+1.585		+3.225		
9.	+2.447		+2.798		
10.	+8.133		+2.891		

(#11-20) Use each of the following P-values and corresponding significance levels to fill out the table.

A "Low P-value" is lower than the significance level. A "High P-value" is higher than the significance level.

Assume the P-values came from unbiased data that passed the conditions for ANOVA.

	P-value Proportion	P-value %	Significance Level Proportion	Significance Level %	Low P-value or High P-value	If $H_0$ is true, could F test stat or more extreme occur by random chance or is that unlikely?	Reject $H_0$ or Fail to reject $H_0$ ?	Significant Evidence or not?
11.	0.186		0.1					
12.	0.0042		0.01					
13.	$2.59 \times 10^{-4}$		0.05					
14.	0.006		0.01					
15.	0.353		0.05					
16.	0		0.1					
17.	0.041		0.05					
18.	0.274		0.1					
19.	$1.04 \times 10^{-8}$		0.01					
20.	0.067		0.05					

21. The F-test statistic compares the variance between the groups to the variance within the groups. Explain how the variance between the groups is calculated and what it tells us. Explain how the variance within the groups is calculated and what it tells us. How can we use the variance between and the variance within to calculate the F-test statistic?

22. If the variance between the groups were significantly larger than the variance within, would the F-test statistic be large or small? Explain why.

23. If the variance between the groups were about the same as the variance within, would the F-test statistic be large or small? Explain why.

24. The ANOVA printout involves the degrees of freedom within the groups, the degrees of freedom between the groups and the total degrees of freedom. How are the different degrees of freedom calculated?



**(#25-28) ANOVA Hypothesis Tests with Given Statcato Graphs and Statistics.**

25.

A random sample of black bears were weighed at various times of the year. Some of the bears were weighed in the spring, some in the summer and some in the fall. The bears were tagged so that the same bear was not measured more than once. Use a 1% significance level and the following Statcato statistics, graphs and ANOVA printout to test the population claim that the time of year (season) is related to the weight of the bears.

- Give the null and alternative hypothesis.
- Check the assumptions for a One-Way ANOVA test. (Assume individuals were independent.)
- Does the F-test statistic fall in the right tail starting at the critical value or not?
- What is the variance between the groups? What is the variance within the groups? Was the variance between significantly higher than the variance within? Explain how you know.
- Does the sample data significantly disagree with the null hypothesis. Explain your answer.
- Is the P-value lower or higher than the Significance Level?
- If the null hypothesis was true, could the F-test statistic or more extreme have occurred because of sampling variability or is that unlikely? Explain your answer.
- Should we reject the null hypothesis or fail to reject the null hypothesis? Explain your answer.
- Is there significant evidence or not? Explain your answer.
- Write the formal conclusion sentence for the hypothesis test addressing evidence and the claim.
- Are the categorical and quantitative variables related or not related. Explain your answer.

**One-way ANOVA: Significance level = 0.01**

Selected column variables: C1 Spring Bear Weig... C2 Summer Bear Weig... C3 Fall Bear Weight...

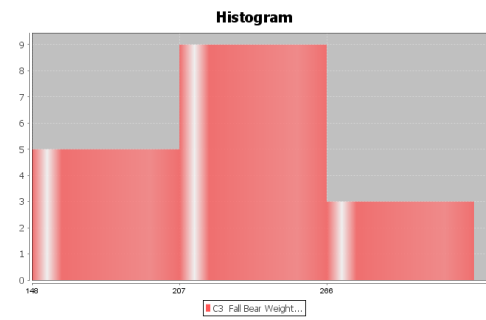
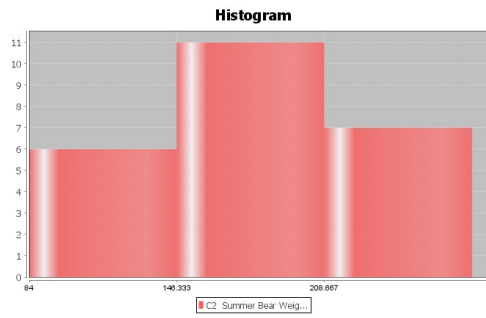
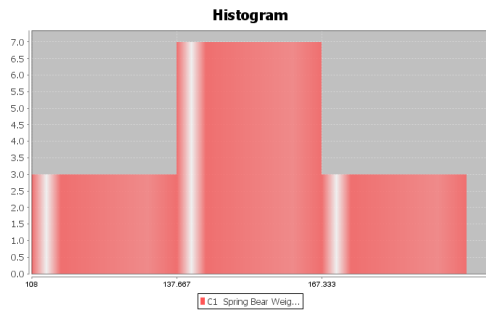
Source of Variation	DOF	SS	MS	Test statistic F	Critical value F	p-Value
Treatment (Between Groups)	2	45539.29263	22769.64632	13.55345	5.0472	0.00002
Error (Within Groups)	51	85679.46663	1679.98954			
Total	53	131218.75926				

**Descriptive Statistics**

Variable	Mean	Standard Deviation
C1 Spring Bear Weights in Pounds	151.385	22.463
C2 Summer Bear Weights in Pounds	182.125	48.017
C3 Fall Bear Weights in Pounds	228.118	40.769

Variable	N total
C1 Spring Bear Weights in Pounds	13
C2 Summer Bear Weights in Pounds	24
C3 Fall Bear Weights in Pounds	17





*This chapter is from Introduction to Statistics for Community College Students, 1<sup>st</sup> Edition, by Matt Teachout, College of the Canyons, Santa Clarita, CA, USA, and is licensed under a "CC-By" [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/) – 10/1/18*

26.

A census of Math 075 pre-stat students was taken in the fall 2015 semester. The students were separated into three sleep groups: low amount of sleep, moderate amount of sleep, high amount of sleep. They were also asked how many total units they have completed at the college. Though the data was not random, you can assume it was representative of Math 075 students at COC. Use a 10% significance level and the following Statcato statistics, graphs and ANOVA printout to test the claim that sleep is not related the total number of units completed.

- a) Give the null and alternative hypothesis.
- b) Check the assumptions for a One-Way ANOVA test. (Assume individuals were independent.)
- c) Does the F-test statistic fall in the right tail starting at the critical value or not?
- d) What is the variance between the groups? What is the variance within the groups? Was the variance between significantly higher than the variance within? Explain how you know.
- e) Does the sample data significantly disagree with the null hypothesis. Explain your answer.
- f) Is the P-value lower or higher than the Significance Level?
- g) If the null hypothesis was true, could the F-test statistic or more extreme have occurred because of sampling variability or is that unlikely? Explain your answer.
- h) Should we reject the null hypothesis or fail to reject the null hypothesis? Explain your answer.
- i) Is there significant evidence or not? Explain your answer.
- j) Write the formal conclusion sentence for the hypothesis test addressing evidence and the claim.
- k) Are the categorical and quantitative variables related or not related. Explain your answer.

**One-way ANOVA: Significance level = 0.1**

Selected column variables: C5 COC Units - Low ... C6 COC Units - Medi... C7 COC Units - High...

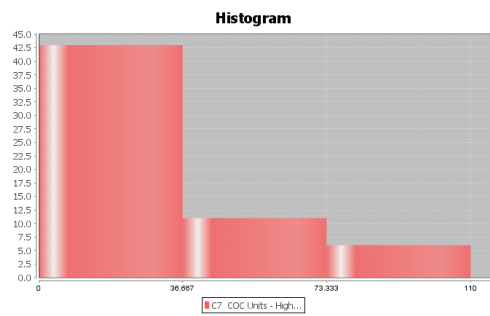
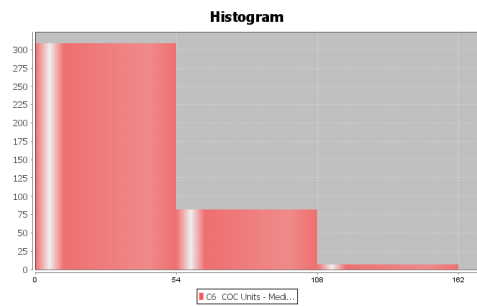
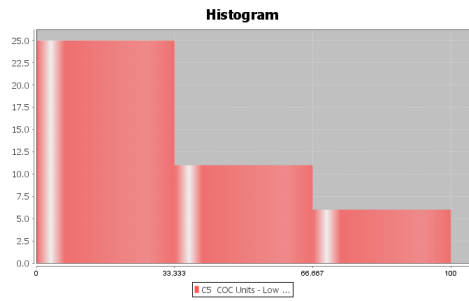
Source of Variation	DOF	SS	MS	Test statistic F	Critical value F	p-Value
Treatment (Between Groups)	2	2822.35625	1411.17813	1.83387	2.3133	0.16087
Error (Within Groups)	497	382446.38503	769.50983			
Total	499	385268.74128				

**Descriptive Statistics**

Variable	Mean	Standard Deviation
C5 COC Units - Low Sleep Group	32.952	28.586
C6 COC Units - Medium Sleep Group	32.990	27.585
C7 COC Units - High Sleep Group	25.675	28.178

Variable	N total
C5 COC Units - Low Sleep Group	42
C6 COC Units - Medium Sleep Group	398
C7 COC Units - High Sleep Group	60





*This chapter is from Introduction to Statistics for Community College Students, 1<sup>st</sup> Edition, by Matt Teachout, College of the Canyons, Santa Clarita, CA, USA, and is licensed under a "CC-By" [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/) – 10/1/18*

27.

A census of Math 075 pre-stat students was taken in the fall 2015 semester. The students were separated into four political parties: democratic, republican, independent party, and other political party. They were also asked number of alcoholic beverages they consume per week. Though the data was not random, you can assume it was representative of Math 075 students at COC. Use a 5% significance level and the following Statcato statistics, graphs and ANOVA printout to test the claim that political party is not related to the number of alcoholic beverages.

- Give the null and alternative hypothesis.
- Check the assumptions for a One-Way ANOVA test. (Assume individuals were independent.)
- Does the F-test statistic fall in the right tail starting at the critical value or not?
- What is the variance between the groups? What is the variance within the groups? Was the variance between significantly higher than the variance within? Explain how you know.
- Does the sample data significantly disagree with the null hypothesis. Explain your answer.
- Is the P-value lower or higher than the Significance Level?
- If the null hypothesis was true, could the F-test statistic or more extreme have occurred because of sampling variability or is that unlikely? Explain your answer.
- Should we reject the null hypothesis or fail to reject the null hypothesis? Explain your answer.
- Is there significant evidence or not? Explain your answer.
- Write the formal conclusion sentence for the hypothesis test addressing evidence and the claim.
- Are the categorical and quantitative variables related or not related. Explain your answer.

**One-way ANOVA: Significance level = 0.05**

Selected column variables: C9 # Drinks per Wee... C10 # Drinks per Wee... C11 # Drinks per Wee... C12 # Drinks per Wee...

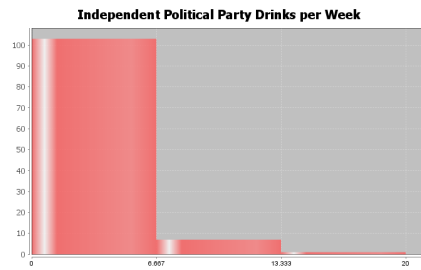
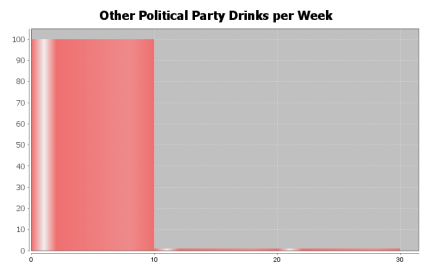
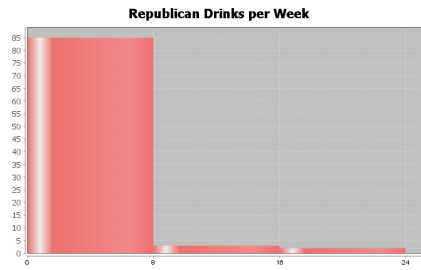
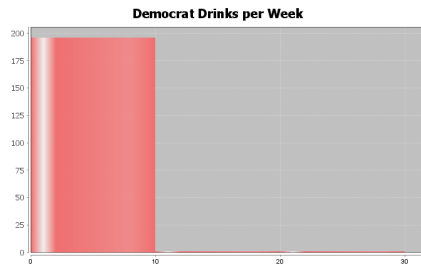
Source of Variation	DOF	SS	MS	Test statistic F	Critical value F	p-Value
Treatment (Between Groups)	3	25.44137	8.48046	0.89597	2.6228	0.44306
Error (Within Groups)	497	4704.16342	9.46512			
Total	500	4729.60479				

**Descriptive Statistics**

Variable	Mean	Standard Deviation
C9 # Drinks per Week - Democrats	0.914	2.566
C10 # Drinks per Week - Independent Political Party	1.342	2.943
C11 # Drinks per Week - Other Political Party	1.373	3.447
C12 # Drinks per Week - Republicans	1.411	3.753

Variable	N total
C9 # Drinks per Week - Democrats	198
C10 # Drinks per Week - Independent Political Party	111
C11 # Drinks per Week - Other Political Party	102
C12 # Drinks per Week - Republicans	90





This chapter is from *Introduction to Statistics for Community College Students*, 1<sup>st</sup> Edition, by Matt Teachout, College of the Canyons, Santa Clarita, CA, USA, and is licensed under a “CC-By” [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/) – 10/1/18

28.

A census of Math 075 pre-stat students was taken in the fall 2015 semester. The students were asked what their favorite social media is: Facebook, Instagram, Snapchat, or Twitter. They were also asked for the number of minutes per day spent on social media. Though the data was not random, you can assume it was representative of Math 075 students at COC. Use a 5% significance level and the following Statcato statistics, graphs and ANOVA printout to test the claim that the type of social media is related to the number of minutes per day spent on social media.

- Give the null and alternative hypothesis.
- Check the assumptions for a One-Way ANOVA test. (Assume individuals were independent.)
- Does the F-test statistic fall in the right tail starting at the critical value or not?
- What is the variance between the groups? What is the variance within the groups? Was the variance between significantly higher than the variance within? Explain how you know.
- Does the sample data significantly disagree with the null hypothesis. Explain your answer.
- Is the P-value lower or higher than the Significance Level?
- If the null hypothesis was true, could the F-test statistic or more extreme have occurred because of sampling variability or is that unlikely? Explain your answer.
- Should we reject the null hypothesis or fail to reject the null hypothesis? Explain your answer.
- Is there significant evidence or not? Explain your answer.
- Write the formal conclusion sentence for the hypothesis test addressing evidence and the claim.
- Are the categorical and quantitative variables related or not related. Explain your answer.

**One-way ANOVA: Significance level = 0.05**

Selected column variables: C14 Facebook - Social Media Minutes per day C15 Instagram - Social Media Minutes per day C16 Snapchat - Social Media Minutes per day C17 Twitter - Social Media Minutes per day

Source of Variation	DOF	SS	MS	Test statistic F	Critical value F	p-Value
Treatment (Between Groups)	3	169375.54058	56458.51353	8.20214	2.6354	0.00003
Error (Within Groups)	293	2016833.12272	6883.38950			
Total	296	2186208.66330				

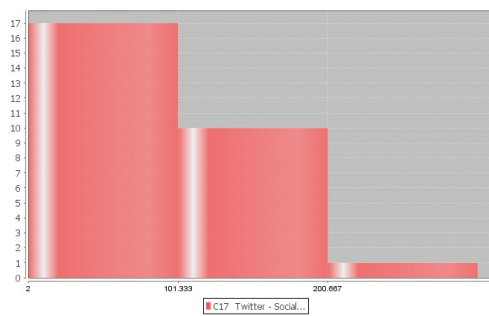
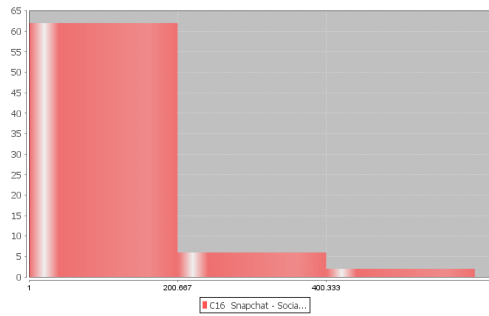
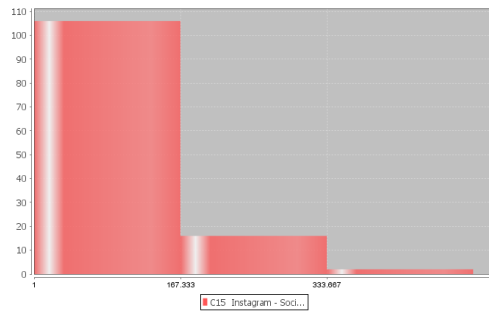
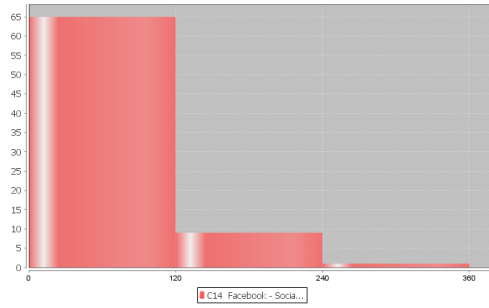
**Descriptive Statistics**

Variable	Mean	Standard Deviation
C14 Facebook - Social Media Minutes per day	43.867	58.103
C15 Instagram - Social Media Minutes per day	83.206	82.817
C16 Snapchat - Social Media Minutes per day	110.914	109.552
C17 Twitter - Social Media Minutes per day	90.964	59.408

Variable	N total
C14 Facebook - Social Media Minutes per day	75
C15 Instagram - Social Media Minutes per day	124
C16 Snapchat - Social Media Minutes per day	70
C17 Twitter - Social Media Minutes per day	28







*This chapter is from Introduction to Statistics for Community College Students, 1<sup>st</sup> Edition, by Matt Teachout, College of the Canyons, Santa Clarita, CA, USA, and is licensed under a “CC-By” [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/) – 10/1/18*

## #29-30: ANOVA Problems with Raw Excel Data and StatKey

29.

Use the car data and a 5% significance level to test the claim that the country a car is made in is related to (associated with) to its gas mileage. Assume the data met the conditions for an ANOVA hypothesis test.

Go to the data sets page at [www.matt-teachout.org](http://www.matt-teachout.org) and open the car data. Copy and paste the country and the miles per gallon columns next to each other in a new excel spreadsheet. The country should be on the left and the miles per gallon should be on the right. Go to [www.lock5stat.com](http://www.lock5stat.com) and click on the StatKey button. Under the “More advanced randomization tests” menu click on “ANOVA for Difference in Means”. Copy the two columns into the “Edit Data” menu in StatKey and push “Compute”. Next to “Original Sample” click on “ANOVA Table”. Write down the F-test statistic, the variance between (“MS groups”), the variance within (“MS error”), the total number of groups, the degrees of freedom between (“groups”), and the degrees of freedom within “Error”. Also write the null and alternative hypotheses.

- a) Total number of groups ( $k$ ) = ?
- b) Null Hypothesis ( $H_0$ )?
- c) Alternative Hypothesis ( $H_A$ )?
- d) F-test statistic = ?
- e) Variance between = ?
- f) Variance within = ?
- g) Degrees of Freedom Between = ?
- h) Degrees of Freedom Within = ?

In the “Theoretical Distributions” menu in StatKey, click on “F”, type in the degrees of freedom between in the box that says “Numerator df”, and type in the degrees of freedom within in the box that says “Denominator df”. Then click on “Ok”. Click on “Right Tail”. In the upper right proportion box, type in the given significance level. The bottom right box will now be the critical value and show where the right tail begins.

- i) Critical Value = ?
- j) Does the F-test statistic fall in the tail or not?
- k) Does the sample data significantly disagree with the null hypothesis or not?
- l) Is the variance between significantly different than the variance within or not?

Use the same Right Tail F-distribution with the same degrees of freedom. In the bottom right box of the F-distribution, type in the F-test statistic. Then the upper right box will now be P-value.

- k) P-value proportion =
- l) P-value Percentage =
- m) Was the P-value lower or higher than the significance level?
- n) Could the F-test statistic or more extreme have occurred because of sampling variability or is that unlikely?
- o) Should we reject  $H_0$  or fail to reject  $H_0$ ? (Assume the data met the conditions for an ANOVA test.)
- p) Is the P-value considered significant evidence or not? (Assume the data met the conditions for an ANOVA test.)
- q) Write the conclusion sentence for the test. (Assume the data met the conditions for an ANOVA test.)



30.

Use the car data and a 5% significance level to test the claim that the country a car is made in is not related to (not associated with) to its horsepower. Assume the data met the conditions for an ANOVA hypothesis test.

Go to the data sets page at [www.matt-teachout.org](http://www.matt-teachout.org) and open the car data. Copy and paste the country and the horsepower columns next to each other in a new excel spreadsheet. The country should be on the left and the horsepower should be on the right. Go to [www.lock5stat.com](http://www.lock5stat.com) and click on the StatKey button. Under the “More advanced randomization tests” menu click on “ANOVA for Difference in Means”. Copy the two columns into the “Edit Data” menu in StatKey and push “Compute”. Next to “Original Sample” click on “ANOVA Table”. Write down the F-test statistic, the variance between (“MS groups”), the variance within (“MS error”), the total number of groups, the degrees of freedom between (“groups”), and the degrees of freedom within “Error”. Also write the null and alternative hypotheses.

a) Total number of groups ( $k$ ) = ?

b) Null Hypothesis ( $H_0$ )?

c) Alternative Hypothesis ( $H_A$ )?

d) F-test statistic = ?

e) Variance between = ?

f) Variance within = ?

g) Degrees of Freedom Between = ?

h) Degrees of Freedom Within = ?

In the “Theoretical Distributions” menu in StatKey, click on “F”, type in the degrees of freedom between in the box that says “Numerator df”, and type in the degrees of freedom within in the box that says “Denominator df”. Then click on “Ok”. Click on “Right Tail”. In the upper right proportion box, type in the given significance level. The bottom right box will now be the critical value and show where the right tail begins.

i) Critical Value = ?

j) Does the F-test statistic fall in the tail or not?

k) Does the sample data significantly disagree with the null hypothesis or not?

l) Is the variance between significantly different than the variance within or not?

Use the same Right Tail F-distribution with the same degrees of freedom. In the bottom right box of the F-distribution, type in the F-test statistic. Then the upper right box will now be P-value.

k) P-value proportion =

l) P-value Percentage =

m) Was the P-value lower or higher than the significance level?

n) Could the F-test statistic or more extreme have occurred because of sampling variability or is that unlikely?

o) Should we reject  $H_0$  or fail to reject  $H_0$ ? (Assume the data met the conditions for an ANOVA test.)

p) Is the P-value considered significant evidence or not? (Assume the data met the conditions for an ANOVA test.)

q) Write the conclusion sentence for the test. (Assume the data met the conditions for an ANOVA test.)



### #31-32: ANOVA Problems with StatKey Data Sets

31.

Under the “ANOVA for Difference in Means” menu in StatKey, click on the button at the top left of the page and click on the “Sandwich Ants” data. In this data, we are looking at the mean average number of ants that come to three different types of sandwiches left out to spoil. Use a 5% significance level to test the claim that the number of ants is not related to (not associated with) the type of sandwich.

Next to “Original Sample” click on “ANOVA Table”. Write down the F-test statistic, the variance between (“MS groups”), the variance within (“MS error”), the total number of groups, the degrees of freedom between (“groups”), and the degrees of freedom within “Error”. Also write the null and alternative hypotheses.

a) Total number of groups ( $k$ ) = ?

b) Null Hypothesis ( $H_0$ )?

c) Alternative Hypothesis ( $H_A$ )?

d) F-test statistic = ?

e) Variance between = ?

f) Variance within = ?

g) Degrees of Freedom Between = ?

h) Degrees of Freedom Within = ?

In the “Theoretical Distributions” menu in StatKey, click on “F”, type in the degrees of freedom between in the box that says “Numerator df”, and type in the degrees of freedom within in the box that says “Denominator df”. Then click on “Ok”. Click on “Right Tail”. In the upper right proportion box, type in the given significance level. The bottom right box will now be the critical value and show where the right tail begins.

i) Critical Value = ?

j) Does the F-test statistic fall in the tail or not?

k) Does the sample data significantly disagree with the null hypothesis or not?

l) Is the variance between significantly different than the variance within or not?

Use the same Right Tail F-distribution with the same degrees of freedom. In the bottom right box of the F-distribution, type in the F-test statistic. Then the upper right box will now be P-value.

k) P-value proportion =

l) P-value Percentage =

m) Was the P-value lower or higher than the significance level?

n) Could the F-test statistic or more extreme have occurred because of sampling variability or is that unlikely?

o) Should we reject  $H_0$  or fail to reject  $H_0$ ? (Assume the data met the conditions for an ANOVA test.)

p) Is the P-value considered significant evidence or not? (Assume the data met the conditions for an ANOVA test.)

q) Write the conclusion sentence for the test. (Assume the data met the conditions for an ANOVA test.)



32.

Under the “ANOVA for Difference in Means” menu in StatKey, click on the “Pulse and Award” data. This data looks at the average pulse rates of those people that have won Olympic, Academy and Nobel awards. Use a 1% significance level to test the claim that the population mean average pulse rate is related to (associated with) the type of award the person won.

Next to “Original Sample” click on “ANOVA Table”. Write down the F-test statistic, the variance between (“MS groups”), the variance within (“MS error”), the total number of groups, the degrees of freedom between (“groups”), and the degrees of freedom within “Error”. Also write the null and alternative hypotheses.

- a) Total number of groups ( $k$ ) = ?
- b) Null Hypothesis ( $H_0$ )?
- c) Alternative Hypothesis ( $H_A$ )?
- d) F-test statistic = ?
- e) Variance between = ?
- f) Variance within = ?
- g) Degrees of Freedom Between = ?
- h) Degrees of Freedom Within = ?

In the “Theoretical Distributions” menu in StatKey, click on “F”, type in the degrees of freedom between in the box that says “Numerator df”, and type in the degrees of freedom within in the box that says “Denominator df”. Then click on “Ok”. Click on “Right Tail”. In the upper right proportion box, type in the given significance level. The bottom right box will now be the critical value and show where the right tail begins.

- i) Critical Value = ?
- j) Does the F-test statistic fall in the tail or not?
- k) Does the sample data significantly disagree with the null hypothesis or not?
- l) Is the variance between significantly different than the variance within or not?

Use the same Right Tail F-distribution with the same degrees of freedom. In the bottom right box of the F-distribution, type in the F-test statistic. Then the upper right box will now be P-value.

- k) P-value proportion =
  - l) P-value Percentage =
  - m) Was the P-value lower or higher than the significance level?
  - n) Could the F-test statistic or more extreme have occurred because of sampling variability or is that unlikely?
  - o) Should we reject  $H_0$  or fail to reject  $H_0$ ? (Assume the data met the conditions for an ANOVA test.)
  - p) Is the P-value considered significant evidence or not? (Assume the data met the conditions for an ANOVA test.)
  - q) Write the conclusion sentence for the test. (Assume the data met the conditions for an ANOVA test.)
- 

