# Statistics 201 Exam 3 – Practice Exam
# Fall 2021
# Confidence Intervals – Chapter 19
# **With KEY**

Disclaimer:

This practice exam is provided solely for the purpose of familiarizing you with the format and style of the Stat 201 exams. There is no explicit or implicit guarantee that the upcoming exam will ask similar questions. If you use the practice exam as your only tool to help you prepare for the upcoming exam, you most likely will not do well on the exam. You should still do the things you would have done if you did not have access to this practice exam, such as re-read the text, go over your class notes, re-work the online homework problems, and look at the list of exam topics provided and make sure that you understand all the concepts listed within it.

NOTE: These questions come from actual older exams. As such, the points on this practice exam may total more than 100 points.

HASLAM
COLLEGE OF BUSINESS
THE UNIVERSITY OF TENNESSEE, KNOXVILLE

**Please note:** when asked to write out null and alternative hypotheses, if appropriate, be sure to use proper mathematical notation (i.e., use symbols such as $<, =, \leq, \geq, \neq, >, p, \hat{p}, \mu, \bar{x}$, etc.).

Exam Grade: **100**______

ON THIS EXAM, WHEN ROUNDING A FINAL NUMERICAL ANSWER , PLEASE REPORT AT LEAST 3 SIGNIFICANT DIGITS.

EXAMPLE: the value 0.000378516 should be rounded to 0.000379

1. Below is output from a hypothesis test for a single population mean based on 28 observations. The null hypothesized value, as well as the sample mean (“actual estimate”) are seen below (among other numbers). Calculate the 3 missing values below. Show your work (where necessary), but fill in your final answers in the blanklines below.

| Test Mean | |
|---|---|
| Hypothesized Value | 20 |
| Actual Estimate | 18.6286 |
| DF | |
| Std Dev | 2.46484 |
| | **t Test** |
| Test Statistic | |
| Prob > $\lvert t\rvert$ | |
| Prob > t | 0.9967 |
| Prob < t | 0.0033* |

i) (2 points) DF : $df = n - 1 = 28 - 1 = 27$

ii) (2 points) Test Statistic : $\frac{\bar{y}-\mu_0}{SD(\bar{y})} = \frac{18.6286-20}{\frac{2.46484}{\sqrt{28}}} = \frac{-1.3714}{0.46581} = -2.944$

iii) (2 points) Prob > $|t|$ : $2(0.0033) - 0.0066$

2. It is generally believed that 75% of college sophomores regularly recycle. Are UT sophomores different from this widely accepted figure? A researcher at UT got the cooperation of the registrar's office and was given the contact information for a large number of randomly selected sophomores at UT (out of the approximately 5,000 UT sophomores). The researcher got responses to a short survey from 128 of these students. Based on the answers they provided, the researcher determined that 84 of the 128 students regularly recycled.

   i) (6 points) Are the 3 conditions for creating a confidence interval for a population proportion met in this case? State each condition, indicate whether you think each condition is met, and explain your reasoning. Provide numerical justification where appropriate.

Condition 1 is: Randomization

Is condition 1 met? (Circle one): Yes No
Briefly explain your answer:

**The Problem states a random sample of UT sophomores was selected.**

Condition 2 is: 10% Condition

Is condition 2 met? (Circle one): Yes No
Briefly explain your answer:

**128 is less than 10% of the approx. 5000 sophomores.**

Condition 3 is: **Success/Failure Condition**

Is condition 3 met? (Circle one): Yes No
Briefly explain your answer:

$$\textbf{is } n\hat{p} \textbf{ and } n\hat{q} \geq 10?$$
$$n = 128, X = 84$$
$$\hat{p} = \frac{84}{128} = 0.65625$$
$$n\hat{p} = 128(0.65625) = 84$$
$$n\hat{q} = 128(1 - 0.65625) = 44$$

**Both are above 10, so the condition is met**

Question 2 (continued)

NOTE: Regardless of your answers to part [i], complete all remaining parts.

ii) (3 points) Using the information in the introduction of this question, create a 93% confidence interval for the proportion of all UT sophomores who regularly recycle. For a 93% confidence interval in this case, Z*=1.812.

$$\hat{p} = \frac{84}{128} = 0.65625 \text{ and } \hat{q} = 1 - \hat{p} = 1 - 0.65625 = 0.34375$$

$$\mathrm{SD}(\hat{p}) = \sqrt{\frac{\hat{p}\hat{q}}{n}} = \sqrt{\frac{0.65625(0.34375)}{128}} = 0.04198$$

$$\hat{p} \pm z^{*}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$0.65625 \pm 1.812(0.04198)$$

$$0.65625 \pm 0.07607$$

$$\text{Upper Limit: } 0.65625 + 0.07607 = 0.732$$

$$\text{Lower Limit: } 0.65625 - 0.07607 = 0.580$$

$$(0.580, 0.732)$$

iii) (3 points) Interpret the interval you calculated in the context of this problem. Limit your answer to one or two sentences.

**We can be 93% confident that the proportion of all UT sophomores that "regularly recycle" is between 0.58 and 0.732 (or between 58% and 73.2%)**

iv) (2 points) Suppose the researcher wishes to carry out a formal hypothesis test for their research question. Write out the null and alternative hypotheses they wish to test.

$$H_0: p = 0.75$$

$$H_A: p \neq 0.75$$

Question 2 (continued)

v) (3 points) Rather than doing all the steps necessary for a formal hypothesis test, use the confidence interval you created in part [ii] to help you decide what your decision is regarding the null hypothesis. Circle the best answer:

Reject the null hypothesis                    Fail to reject the null hypothesis

Briefly explain your reasoning for your selection above.
(NOTE: If you have NO ANSWER to part [ii], use (.234, .678) as your answer to part [ii], and use that to answer part [v]).

**Since 0.75 is not inside the confidence interval, it seems like an unlikely value for p. So I reject the idea that p=0.75 (Reject $H_0$)**

vi) (2 points) What level of alpha ($\alpha$) did you use when your used the above technique to make a decision regarding the null hypothesis?

$$\propto = \mathbf{0.07}$$

3. A researcher believes that British men are shorter, on average, than men from the USA. A random sample of 35 men from the USA (ages 20-50) was observed for height. The mean height was 68.1 inches and the sample standard deviation was 5.34 inches. Meanwhile, a random sample of 35 men from Britain (ages 20-50) was observed for height. The mean height was 64.6 and the sample standard deviation was 5.81 inches.

   i) (2 points) Which type of problem is this? Circle the best answer:

      a) One sample proportion test
      b) One sample test of a mean
      c) Two independent samples t-test
      d) Chi-Squared test of independence

   ii) (2 points) Write down the null and alternative hypotheses suggested by the first sentence of this problem (HINT: to avoid confusion later, look at the top line of the JMP output below <u>now</u>).

$$\mathbf{H_0: \mu_{USA} - \mu_{Britian} = 0}$$
$$\mathbf{H_A: \mu_{USA} - \mu_{Britian} > 0}$$

   iii) (3 points) Below is some output from JMP for a proper analysis of these data. Calculate the missing value (to the right of the word "t Ratio").

**Easy way:**

$$t = \frac{3.4971}{1.33411} = 2.621$$

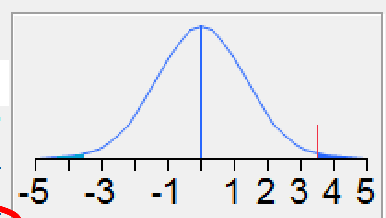
**Hard Way:**

$$t = \frac{68.1 - 64.6}{\sqrt{\frac{(5.34)^2}{35} + \frac{(5.81)^2}{35}}} = \frac{3.5}{1.3337} = 2.624$$

USA-Britain

Assuming unequal variances

| | | | |
|---|---|---|---|
| Difference | 3.49714 | t Ratio | |
| Std Err Dif | 1.33411 | DF | 67.51277 |
| Upper CL Dif | 6.15967 | Prob > \|t\| | 0.0108* |
| Lower CL Dif | 0.83461 | Prob > t | 0.0054* |
| Confidence | 0.95 | Prob < t | 0.9946 |




   iv) (1 point) There are 3 p-values displayed in the output. Circle the p-value that corresponds to your alternative hypothesis in part [ii] above.

Question 3 (continued)

v) (1 point) If we use an $\alpha$ level of 0.01, what is your conclusion regarding your null hypothesis in part (ii)? Circle the best answer:

Reject the null hypothesis Fail to reject the null hypothesis

vi) (3 points) State the remainder of your conclusion for this hypothesis test. Be sure to state your conclusion in the context of this problem.

**There is sufficient evidence to conclude, at the 0.01 level of significance, that men ages 20-50 from USA are taller, on average, than men ages 20-50 from Great Britain.**

4. (2 points) The following is some output from JMP for a one-sample test of a population mean. The 95% confidence interval for the mean is showing, but much of the other output has been erased. The researcher wanted to do a two-sided hypothesis test, and was testing to see if the population mean is 22 (as indicated in the output).

**Summary Statistics**

| | |
|---|---|
| Mean | 24.030303 |
| Std Dev | 3.4774708 |
| Std Err Mean | 0.60535 |
| Upper 95% Mean | 25.263361 |
| Lower 95% Mean | 22.797246 |
| N | 33 |

**Test Mean**

| | |
|---|---|
| Hypothesized Value | 22 |
| Actual Estimate | |
| DF | |
| Std Dev | |

**t Test**

| | |
|---|---|
| Test Statistic | |
| Prob > \|t\| | |
| Prob > t | |
| Prob < t | |

Based on the output that is still showing, what is a possible value for the p-value next to “Prob > |t|”? Circle the best answer:

a) 1.0
b) 0.9999
c) 0.5015
d) 0.0501
e) 0.0051
f) The output that is still showing offers no clue as to what that p-value might be.

**(Since 22 is <u>not</u> inside the 95% confidence interval, Prob >|t| <u>has</u> to be less than 0.05)**

5. Suppose that the Business Analytics and Statistics Department has come up with a method to determine whether or not a student cheated on their exam. We test these simple hypotheses as follows:

$H_0$: The student didn't cheat
$H_A$: The student cheated

i) (2 points) With any test there is always the risk of making the wrong decision. Describe what a type I error is in the context of this problem.

**($H_0$ is true, we reject it)**

**The student, in reality, didn't cheat, but we conclude that they did.**

ii) (2 points) Describe what a type II error is in the context of this problem.

**($H_0$ is false, we don't reject it)**

**The student, in reality, didt cheat, but we conclude that they did not.**

iii) (2 points) If we were to give a 0 to any student who we determined to have cheated using the Business Analytics and Statistics Department's detection method, which type of error would a student typically consider more serious? Briefly explain why.

**Type I. Being an honest student that is accused of cheating is muych worse from a student's perspective.**

iv)(2 points) What is the symbol used to identify the probability of making a type II error? (Circle One):

$\alpha$ $(1-\alpha)$ $\beta$ $(1-\beta)$ $\chi^2$

6. As patients arrive at a local hospital emergency room, as long as they are able to respond, and are not deemed to have an immediate life threatening condition, they are asked some standard questions before being asked to take a seat in the waiting room. The time it takes to complete this initial set of questions has come under scrutiny lately. The emergency room manager believes that the average amount of time it takes to get answers to these initial questions is 2 minutes. Most emergency room employees believe it takes longer than that, on average.

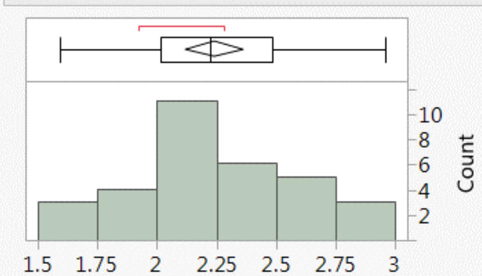
    A nursing school intern was asked to randomly select incoming patients over a 2 week period, and time how long it takes to get answers to these initial questions. The following are some of the results of the intern's data collection efforts.

    i) (2 points) Write out the null and alternative hypotheses suggested by the introduction to this problem.

$$\mathbf{H_0: \mu = 2min}$$
$$\mathbf{H_A: \mu > 2min}$$

Below is some output from the data that the intern collected:

**Time in Minutes**




**Quantiles**

| | | |
|---|---|---|
| 100.0% | maximum | 2.96 |
| 99.5% | | 2.96 |
| 97.5% | | 2.96 |
| 90.0% | | 2.782 |
| 75.0% | quartile | 2.485 |
| 50.0% | median | 2.225 |
| 25.0% | quartile | 2.0125 |
| 10.0% | | 1.717 |
| 2.5% | | 1.59 |
| 0.5% | | 1.59 |
| 0.0% | minimum | 1.59 |

**Summary Statistics**

| | |
|---|---|
| Mean | 2.2409375 |
| Std Dev | 0.3382437 |
| Std Err Mean | 0.0597936 |
| Upper 95% Mean | 2.3628873 |
| Lower 95% Mean | 2.1189877 |
| N | 32 |

Question 6 (continued)

ii) (3 points) Use the information in the output to calculate the value of the test statistic for this hypothesis test.

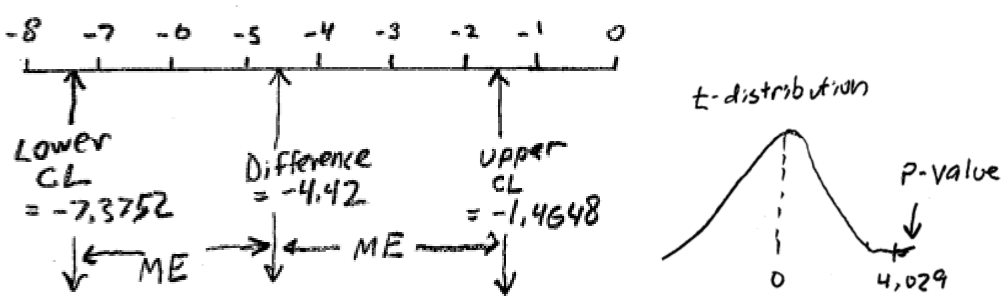
$$t = \frac{\bar{y} - \mu_0}{SD(\bar{y})} = \frac{2.2409 - 2}{\frac{0.3382437}{\sqrt{32}}} = \frac{0.2409}{0.5979} = 4.029$$

iii) (2 points) Based on your answer to part [ii] above, what would your decision be regarding the null hypothesis? Use α = 0.05. Circle the best answer:

Reject the null hypothesis    Fail to reject the null hypothesis

Briefly explain your reasoning for your selection above.
(NOTE: If you have NO ANSWER to part [ii], use (3.21) as your answer to part [ii], and use that to answer part [iii]).

**The t-distribution is a lot like the standard normal distribution. Being beyond 3 standard errors from the hypothesized values will lead to a very small p-value**




iv) (3 points) State the remainder of your conclusion for this hypothesis test. Be sure to state your conclusion in the context of this problem.

**There is sufficient evidence to conclude, at the 0.05 level of significance, that the average amount of time it takes to get answers to these initial questions is <u>greater</u> than 2 minutes.**

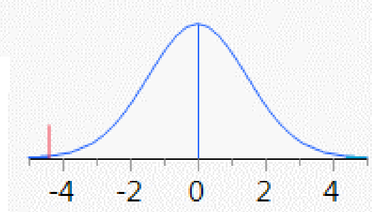
7. Below is some JMP output for a two-sample t-test. Answer the following questions based on the output provided.

**t Test**

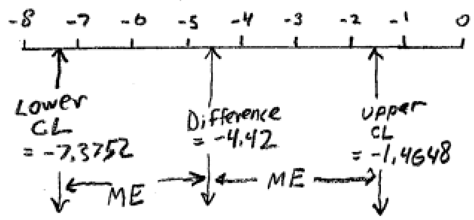
Group2-Group1

Assuming unequal variances

| | | |
|---|---|---|
| Difference | -4.4200 | t Ratio |
| Std Err Dif | 1.4713 | DF |
| Upper CL Dif | -1.4648 | Prob > \|t\| |
| Lower CL Dif | -7.3752 | Prob > t |
| Confidence | 0.95 | Prob < t |




i) (2 points) The values next to "Upper CL Dif" and "Lower CL Dif" represent the endpoints of a 95% confidence interval for the difference between the two population means. What is the "margin of error" of this confidence interval?




**Can use Upper Limit:**

$$\boldsymbol{ME = -1.4648 - (-4.42) = 2.955}$$

**OR Lower Limit**

$$\boldsymbol{ME = -4.42 - (-7.3752) = 2.955}$$

**OR use Both Upper and Lower Limits**

$$\boldsymbol{ME = \frac{-1.4648 - (-7.3752)}{2} = 2.955}$$

ii) (2 points) What was the value of $t^*_{df}$ that was used to construct this confidence interval? Please report 4 decimal places.

$$\boldsymbol{(\bar{y}_1 - \bar{y}_2) \pm t^*_{df} \times SE(\bar{y}_1 - \bar{y}_2)}$$

$$\boldsymbol{-4.42 \pm t^*_{df} \times 1.4713}$$

**Solve using Lower Limit:**

$$\boldsymbol{-7.3752 = -4.42 - t^*_{df} \times 1.4713}$$

$$\boldsymbol{-2.9552 = -t^*_{df} \times 1.4713}$$

$$\boldsymbol{t^*_{df} = 2.0086}$$

**Or Solve using Upper Limit:**

$$\boldsymbol{-1.4648 = -4.42 + t^*_{df} \times 1.4713}$$

$$\boldsymbol{2.9552 = t^*_{df} \times 1.4713}$$

$$\boldsymbol{t^*_{df} = 2.0086}$$

8. The Gallup Poll conducted a representative telephone survey during the first quarter of 2010 in the United States. Among the reported results was the following table concerning the preferred political party affiliation of respondents and their ages. We are interested in seeing if there is an association between a person's age and their political party affiliation in the United States.

**Contingency Table**

Political Party

| Age Group: Count | Democrat | Independent | Rebublican | Total |
|---|---|---|---|---|
| 18-29 | 131 | 209 | 51 | 391 |
| 30-49 | 140 | 200 | 79 | 419 |
| 50-64 | 141 | 145 | 102 | 388 |
| 65+ | 162 | 113 | 129 | 404 |
| Total | 574 | 667 | 361 | 1602 |

i) (2 points) Write the appropriate null and alternative hypothesis suggested by the statements above.

**$H_0$: Political Party affiliation is independent of which age group a person is in**

**$H_A$: Political Party affiliation is <u>not</u> independent of which age group a person is in**

ii) (4 points) Below is some output from JMP for this analysis. Two of the values in the table have been erased (A and B). Provide the necessary arithmetic for these two missing values. Write your final answers in the blanks: A <u>162.795</u>___B <u>2.0547</u>_____

**Contingency Table**

Political Party

| Age Group | Count / Expected / Cell Chi^2 | Democrat | Independent | Rebublican | Total |
|---|---|---|---|---|---|
| 18-29 | Count | 131 | 209 | 51 | 391 |
| | Expected | 140.096 | **A** | 88.1092 | |
| | Cell Chi^2 | 0.5906 | 13.1143 | 15.6294 | |
| 30-49 | Count | 140 | 200 | 79 | 419 |
| | Expected | 150.129 | 174.453 | 94.4189 | |
| | Cell Chi^2 | 0.6833 | 3.7413 | 2.5179 | |
| 50-64 | Count | 141 | 145 | 102 | 388 |
| | Expected | 139.021 | 161.546 | 87.4332 | |
| | Cell Chi^2 | 0.0282 | 1.6946 | 2.4269 | |
| 65+ | Count | 162 | 113 | 129 | 404 |
| | Expected | 144.754 | 168.207 | 91.0387 | |
| | Cell Chi^2 | **B** | 18.1195 | 15.8291 | |
| Total | | 574 | 667 | 361 | 1602 |

$$A)\ \textit{Expected Cell Count} = \frac{391\times667}{1602} = 162.795$$

$$B)\ \textit{Cell Chi-Square} = \chi^2 = \frac{(162-144.754)^2}{144.754} = 2.0547$$

Question 8 (continued)

iii) (3 points) As with any statistical technique, there are conditions that must be met for the analysis to be valid. One of the conditions of this technique (the Chi-Square test of independence) has something to do with there being enough data for the analysis to be valid. What is the name of this condition, and do you believe it is met in this case? Provide a brief explanation for your "yes" or "no" answer.

**Expected Cell Count condition. Yes, it is met. The smallest expected cell count is 87.4332. Since all expected cell counts are at least 5, the condition is met.**

iv) (2 points) From the data given, calculate the degrees of freedom for the chi-square test of independence.

$$\mathbf{df = (\#Rows - 1) \times (\#Columns - 1) = (4 - 1) \times (3 - 1) = 6}$$

v) (2 points) Below is more output from JMP for this analysis. The Pearson ChiSquare value is 76.430. Making reference to the numbers in the Contingency Table at the bottom of the previous page, briefly explain how this number was calculated (use the space to the right of the output for your answer).

**Tests**

| N | DF | -LogLike | RSquare (U) |
|---|---|---|---|
| 1602 | | 39.293192 | 0.0230 |

| Test | ChiSquare | Prob>ChiSq |
|---|---|---|
| Likelihood Ratio | 78.586 | <.0001* |
| Pearson | 76.430 | <.0001* |

vi) (1 point) Using $\alpha$=0.005, what is your conclusion regarding your null hypothesis from part [i]? Circle the best answer:

Reject the null hypothesis Fail to reject the null hypothesis

Briefly explain your reasoning for your selection above.

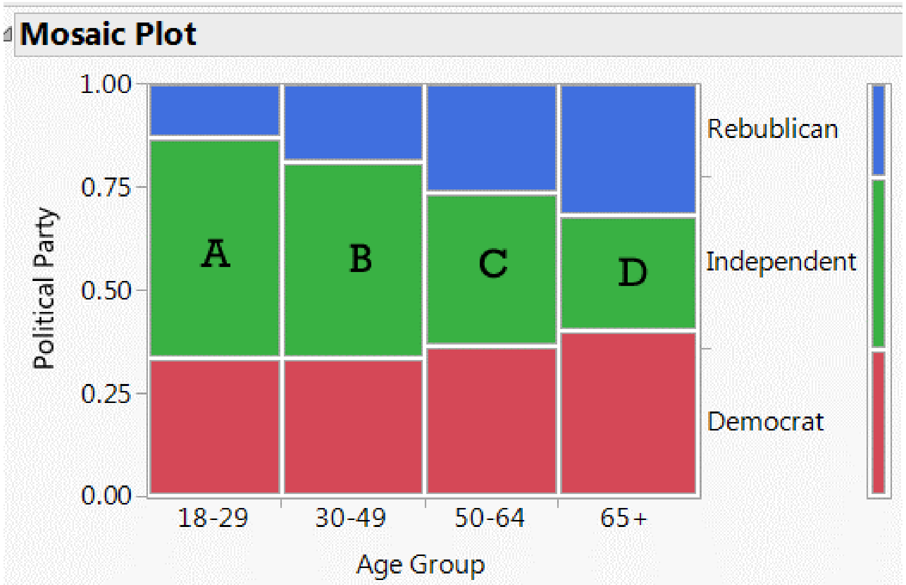
**Since the p-value of <0.0001 is less than 0.005 (less than alpha), we reject $H_0$**

vii) (2 points) State the remainder of your conclusion for this hypothesis test. Be sure to state your conclusion in the context of this problem.

**There is evidence to suggest at the 0.005 level of significance, that political party affiliation is related to a person's age in the USA (at least when this survey was taken!).**

Question 8 (continued)

viii) (4 points) The following is the Mosaic Plot for this analysis, along with the default Contingency Table from JMP. Use the output below to help you report the percentage (to 2 decimal places) of respondents in each of the four age groups that reported they were “independents”. Fill in the four blanks below with your answers. (HINT: Your answers represent the height of the bars identified as bars A, B, C and D.)




**Contingency Table**

| Age Group | Count / Total % / Col % / Row % | Democrat | Independent | Rebublican | Total |
|---|---|---|---|---|---|
| 18-29 | Count | 131 | 209 | 51 | 391 |
| | Total % | 8.18 | 13.05 | 3.18 | 24.41 |
| | Col % | 22.82 | 31.33 | 14.13 | |
| | Row % | 33.50 | 53.45 | 13.04 | |
| 30-49 | Count | 140 | 200 | 79 | 419 |
| | Total % | 8.74 | 12.48 | 4.93 | 26.15 |
| | Col % | 24.39 | 29.99 | 21.88 | |
| | Row % | 33.41 | 47.73 | 18.85 | |
| 50-64 | Count | 141 | 145 | 102 | 388 |
| | Total % | 8.80 | 9.05 | 6.37 | 24.22 |
| | Col % | 24.56 | 21.74 | 28.25 | |
| | Row % | 36.34 | 37.37 | 26.29 | |
| 65+ | Count | 162 | 113 | 129 | 404 |
| | Total % | 10.11 | 7.05 | 8.05 | 25.22 |
| | Col % | 28.22 | 16.94 | 35.73 | |
| | Row % | 40.10 | 27.97 | 31.93 | |
| Total | Count | 574 | 667 | 361 | 1602 |
| | Total % | 35.83 | 41.64 | 22.53 | |

18-29: 53.45%

30-49: 47.73%

50-64: 37.37%

65+: 27.97%

Question 8 (continued)

ix) (3 points) In one or two sentences, how would you describe the nature of the association that seems to exist between these two variables?

**As Americans get older, they seem to drift away from being "independents"/ These people seem to become both republicans and democrats, but slightly more often become republicans.**

9. (3 points) Does studying really pay off? Is there a relationship between the total number of minutes students spent preparing for all 3 exams in Stat 201 and the average score they achieved on these exams? What tool or technique would you use to address this question? Circle the best answer:

   a) Box plot
   b) Hypothesis test for a population mean
   c) Decision tree
   d) Confidence interval for the difference between two population means
   e) Simple linear regression

10. (3 points) You are the regional manager of 5 large grocery stores. You let the manager of each of these 5 stores make their own decisions regarding who to hire. Since these 5 stores are within 50 miles of each other, you would expect the ethnic background of the employees at these 5 stores to be fairly similar. Is that the case? What tool or technique would you use to address this question? Circle the best answer:

    a) Histogram
    b) Chi-Square test of independence
    c) Simple linear regression
    d) Hypothesis test for a population mean
    e) Pearson's correlation coefficient (r)

True/False Questions
Circle the best answer
(2 points each)

T F "Power" represents the probability that we correctly fail to reject a true null hypothesis.

T F If, based on the same set of data, you change your level of confidence from 90% to 95%, the resulting 95% confidence interval is considered less precise than the 90% confidence interval.

T F A p-value represents the probability of seeing the results you saw, or results more unusual than that, assuming the null hypothesis is true.

T F The only way to decrease the probability of a Type II error is to increase the probability of a Type I error.

T F In the Business Analytics presentation in class, we discussed how Target uses data to try to determine if a loyal customer is about to stop shopping there.

# [YOU MAY REMOVE THIS SHEET]

## ONE SAMPLE TESTS

**Proportion**

| | | |
|---|---|---|
| **Confidence Intervals** | $\hat{p} \pm z^* \times SE(\hat{p})$ | $SE(\hat{p}) = \sqrt{\frac{\hat{p}\hat{q}}{n}}$ |
| **Hypothesis Testing** | $z_{\hat{p}} = \frac{\hat{p} - p_0}{SD(\hat{p})}$ | $SD(\hat{p}) = \sqrt{\frac{p_0 q_0}{n}}$ |

**Mean**

| | | |
|---|---|---|
| **Confidence Intervals** | $\bar{y} \pm t^*_{df} \times SE(\bar{y})$ | $SE(\bar{y}) = \frac{s}{\sqrt{n}}$ |
| **Hypothesis Testing** | $t_{df} = \frac{\bar{y} - \mu_0}{SE(\bar{y})}$ | $SE(\bar{y}) = \frac{s}{\sqrt{n}}$ |

($df$) degrees of freedom for the t-distribution = $n - 1$

## TWO SAMPLE TESTS

**Difference Between Two Means**

| | | |
|---|---|---|
| **Confidence Intervals** | $\bar{y}_1 - \bar{y}_2 \pm t^*_{df} \times SE(\bar{y}_1 - \bar{y}_2)$ | $SE(\bar{y}_1 - \bar{y}_2) = \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}}$ |
| **Hypothesis Testing** | $t_{df} = \frac{(\bar{y}_1 - \bar{y}_2) - 0}{SE(\bar{y}_1 - \bar{y}_2)}$ | $SE(\bar{y}_1 - \bar{y}_2) = \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}}$ |

($df$) degrees of freedom for the t-distribution available from JMP output

**Test of Independence**

$$\chi^2_{df} = Sum\ of\ \frac{(obs - Exp)^2}{Exp} \quad \text{where} \quad Exp = \frac{Row\ Total \times Column\ Total}{N} \quad \text{for each cell.}$$

$$Degrees\ of\ freedom = (\#Rows - 1) \times (\#Columns - 1)$$