1. Which of the following statements in NOT an assumption of inference for a regression model?
   A) The dependent variable is linearly related to the explanatory variable.
   B) The errors around the idealized regression line follow a Normal model.
   C) The errors around the idealized regression line have equal variability.
   D) The errors around the idealized regression line are independent of each other.
   E) The errors around the idealized regression line are linearly related.

2. A researcher found that a 98% confidence interval for the mean hours per week spent studying by college students was (13, 17). Which is true?
   I. There is a 98% chance that the mean hours per week spent studying by college students is between 13 and 17 hours.
   II. 98% of college students study between 13 and 17 hours a week.
   III. Students average between 13 and 17 hours per week studying on 98% of the weeks.
   A) none   B) I only   C) II only   D) III only   E) I and III

3. A professor was curious about her students’ grade point averages (GPAs). She took a random sample of 15 students and found a mean GPA of 3.01 with a standard deviation of 0.534. Which of the following formulas gives a 99% confidence interval for the mean GPA of the professor’s students?
   A) $3.01 \pm 2.947(0.534/\sqrt{15})$
   B) $3.01 \pm 2.977(0.534/\sqrt{15})$
   C) $3.01 \pm 2.576(0.534/\sqrt{15})$
   D) $3.01 \pm 2.947(0.534/14)$
   E) $3.01 \pm 2.977(0.534/14)$

4. A philosophy professor wants to find out whether the mean age of the men in his large lecture class is equal to the mean age of the women in his classes. After collecting data from a random sample of his students, the professor tested the hypothesis $H_0: \mu_M = \mu_W$ against the alternative $H_A: \mu_M - \mu_W \neq 0$. The $P$-value for the test was 0.003. Which is true?
   A) There is a 0.3% chance that the mean ages for the men and women are equal.
   B) There is a 0.3% chance that the mean ages for the men and women are different.
   C) It is very unlikely that the professor would see results like these if the mean age of men was equal to the mean age of women.
   D) There is a 0.3% chance that another sample will give these same results.
   E) There is a 99.7% chance that another sample will give these same results.

5. Absorption rates into the body are important considerations when manufacturing a generic version of a brand-name drug. A pharmacist read that the absorption rate into the body of a new generic drug (G) is the same as its brand-name counterpart (B). She has a researcher friend of hers run a small experiment to test $H_0: \mu_G = \mu_B = 0$ against the alternative $H_A: \mu_G - \mu_B \neq 0$. Which of the following would be a Type I error?
   A) Deciding that the absorption rates are different, when in fact they are not.
   B) Deciding that the absorption rates are different, when in fact they are.
   C) Deciding that the absorption rates are the same, when in fact they are not.
   D) Deciding that the absorption rates are the same, when in fact they are.
   E) The researcher cannot make a Type I error, since he has run an experiment.

6. The two samples whose statistics are given in the table are thought to come from populations with equal variances. What is the pooled estimate of the population standard deviation?

<table>
<thead>
<tr>
<th>$n$</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>32</td>
<td>5</td>
</tr>
<tr>
<td>20</td>
<td>30</td>
<td>6</td>
</tr>
</tbody>
</table>


__7. At one vehicle inspection station, 13 of 52 trucks and 11 of 88 cars failed the emissions test. Assuming these vehicles were representative of the cars and trucks in that area, what is the standard error of the difference in the percentages of all cars and trucks that are not in compliance with air quality regulations?

A) 0.025  B) 0.032  C) 0.049  D) 0.070  E) 0.095

__8. At one SAT test site students taking the test for a second time volunteered to inhale supplemental oxygen for 10 minutes before the test. In fact, some received oxygen, but others (randomly assigned) were given just normal air. Test results showed that 42 of 66 students who breathed oxygen improved their SAT scores, compared to only 35 of 63 students who did not get the oxygen. Which procedure should we use to see if there is evidence that breathing extra oxygen can help test-takers think more clearly?

A) 1-proportion z-test  B) 2-proportion z-test  C) 1-sample t-test
D) 2-sample t-test  E) matched pairs t-test

__9. A survey asked people “On what percent of days do you get more than 30 minutes of vigorous exercise?” Using their responses we want to estimate the difference in exercise frequency between men and women. We should use a

A) 1-proportion z-interval  B) 2-proportion z-interval  C) 1-sample t-interval
D) 2-sample t-interval  E) matched pairs t-interval

__10. Two agronomists analyzed the same data, testing the same null hypothesis about the proportion of tomato plants suffering from blight. One rejected the hypothesis but the other did not. Assuming neither made a mistake in calculations, which of these possible explanations could account for this apparent discrepancy?

I. One agronomist wrote a one-tailed alternative hypothesis, but the other used 2 tails.
II. They wrote identical hypotheses, but the one who rejected the null used a higher $\alpha$ – level.
III. They wrote identical hypotheses, but the one who rejected the null used a lower $\alpha$ – level.

A) I only  B) II only  C) III only  D) I or II  E) I or III

11. Cloning A random sample of 800 adults was asked the following question: “Do you think current laws concerning the use of cloning for medical research are too strict, too lenient, or about right?” The pollsters also classified the respondents with respect to highest education level attained: high school, 2-year college degree, 4-year degree, or advanced degree. We wish to know if attitudes on cloning are related to education level. (All the conditions are satisfied – don’t worry about checking them.) a. Write appropriate hypotheses.

<table>
<thead>
<tr>
<th></th>
<th>Strict</th>
<th>Lenient</th>
<th>Right</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>High school</td>
<td>93</td>
<td>107</td>
<td>182</td>
<td>382</td>
</tr>
<tr>
<td></td>
<td>106.01</td>
<td>87.38</td>
<td>188.61</td>
<td></td>
</tr>
<tr>
<td>2-year</td>
<td>27</td>
<td>19</td>
<td>56</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>28.31</td>
<td>23.33</td>
<td>50.36</td>
<td></td>
</tr>
<tr>
<td>4-year</td>
<td>82</td>
<td>50</td>
<td>140</td>
<td>272</td>
</tr>
<tr>
<td></td>
<td>75.48</td>
<td>62.22</td>
<td>134.30</td>
<td></td>
</tr>
<tr>
<td>Adv. degree</td>
<td>20</td>
<td>7</td>
<td>17</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>12.21</td>
<td>10.07</td>
<td>21.73</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>222</td>
<td>183</td>
<td>395</td>
<td>800</td>
</tr>
</tbody>
</table>

$\chi^2 = 1.60 + 4.40 + 0.23 + 0.06 + 0.80 + 0.63 + 0.56 + 2.40 + 0.24 + 4.97 + 0.93 + 1.03 = 17.86$

$P = 0.0066$
b. Suppose the expected counts had not been given. Show how to calculate the expected count in the first cell (106.01).


d. State your complete conclusion in context.

12. **Exercise** A random sample of 150 men found that 88 of the men exercise regularly, while a random sample of 200 women found that 130 of the women exercise regularly.
   a. Based on the results, construct and interpret a 95% confidence interval for the difference in the proportions of women and men who exercise regularly.

b. A friend says that she believes that a higher proportion of women than men exercise regularly. Does your confidence interval support this conclusion? Explain.
13. **Bedrooms** A random sample of 76 apartments is collected near a university. All of the apartments in the sample have between 1 and 6 bedrooms. The variables recorded for each apartment are **Rent** (in dollars) and the number of **Bedrooms**. The regression output is:

The dependent variable is Rent

\[ R \text{ squared} = 62.0\% \quad R \text{ squared (adjusted)} = 61.5\% \]
\[ s = 364.4 \quad \text{with} \quad 76 - 2 = 74 \quad \text{degrees of freedom} \]

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coeff</th>
<th>SE(Coeff)</th>
<th>t-ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>357.795</td>
<td>111.6</td>
<td>3.2</td>
<td>0.0020</td>
</tr>
<tr>
<td>Bedrooms</td>
<td>400.554</td>
<td>36.42</td>
<td>11.0</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>

a. Write out the regression equation.

b. Compute a 95% confidence interval for the coefficient of Bedrooms. Explain your confidence interval in the context of the problem.

c. Based on your interval is the number of bedrooms a significant predictor of rent? Explain how you reached your answer.

14. **Haircuts** You need to find a new hair stylist and know that there are two terrific salons in your area, Hair by Charles and Curl Up & Dye. You want a really good haircut, but you do not want to pay too much for the cut. A random sample of costs for 10 different stylists was taken at each salon (each salon employs over 100 stylists).

a. Indicate what inference procedure you would use to see if there is a significant difference in the costs for haircuts at each salon. Check the appropriate assumptions and conditions and indicate whether you could or could not proceed. (Do not do the actual test.)
14. **Haircuts** (continued)
   
   b. A friend tells you that he has heard that Curl Up & Dye is the more expensive salon.
      
      i. Write hypotheses for your friend’s claim.

      ii. The following are computer outputs. Which output is the correct one to use for this test? Explain.

      iii. Use the appropriate computer output to make a conclusion about the hypothesis test based on the data. Make sure to state your conclusion in context.
Statistics Test A – Part VI – Key


   a. \( H_0: \) People’s opinions on cloning are independent of education level.
   \( H_A: \) There is an association
   b. \( (382/800)(222) = 106.005 \)
   c. \( (4 - 1)(3 - 1) = 6 \)
   d. Reject null \( (P < 0.05); \) There is strong evidence that opinion varies with education level.
   It appears that high school grads are more likely to think regulations are too lenient,
   people with advanced degrees too strict.

12. Exercise: A random sample of 150 men found that 88 of the men exercise regularly, while a
random sample of 200 women found that 130 of the women exercise regularly.
   a. Conditions:
      * Randomization Condition: We are told that we have random samples.
      * 10% Condition: We have less than 10% of all men and less than 10% of all women.
      * Independent samples condition: The two groups are independent of each other.
      * Success/Failure Condition: Of the men, 88 exercise regularly and 62 do not; of the
women, 130 exercise regularly and 70 do not. The observed number of both successes
and failures in both groups is at least 10.
   With the conditions satisfied, the sampling distribution of the difference in proportions is
approximately Normal with a mean of \( \hat{p}_M - \hat{p}_W \), the true difference between the
population proportions. We can find a two-proportion z-interval.

   We know: \( n_M = 150, \ \hat{p}_M = \frac{88}{150} = 0.587, \ n_W = 200, \ \hat{p}_W = \frac{130}{200} = 0.650. \)

   We estimate \( SD(\hat{p}_M - \hat{p}_W) \) as

   \[
   SE(\hat{p}_M - \hat{p}_W) = \sqrt{\frac{\hat{p}_M\hat{q}_M}{n_M} + \frac{\hat{p}_W\hat{q}_W}{n_W}} = \sqrt{\frac{(0.587)(0.413)}{150} + \frac{(0.65)(0.35)}{200}} = 0.0525
   \]

   \[
   ME = z^\ast \times SE(\hat{p}_M - \hat{p}_W) = 1.96(0.0525) = 0.1029
   \]

   The observed difference in sample proportions = \( \hat{p}_M - \hat{p}_W = 0.587 - 0.650 = -0.063 \), so
the 95% confidence interval is \(-0.063 \pm 0.1029\), or -16.6% to 4.0%.
We are 95% confident that the proportion of women who exercise regularly is between
4.0% lower and 16.6% higher than the proportion of men who exercise regularly.

   b. Since zero is contained in my confidence interval, I cannot say that a higher proportion of
women than men exercise regularly. My confidence interval does not support my friend’s
claim.
13. Bedrooms
   a. \( \hat{\text{Rent}} = 357.80 + 400.55 \) Bedrooms
   b. \( b_1 \pm t^*_{n-2} \times SE(b_1) \)  The degrees of freedom are \( n - 2 = 74 \). For a 95% C.I., \( t_{73}^* \approx 2 \).
      
      \[ 400.55 \pm 2(36.42) = (327.71, 473.39) \] dollars per bedroom
      We are 95% confident that, on average, each additional bedroom is associated with an increase of between $327.71 and $473.39 in the rent of an apartment.
   c. To test \( H_0 : \beta_1 = 0 \) vs. \( H_A : \beta_1 \neq 0 \), we can compare our confidence interval from part b to the hypothesized value of zero. Since zero is below our interval, we conclude that there is strong evidence that the number of bedrooms is positively associated with the amount of rent charged.

14. Haircuts
   a. I would use a two-sample \( t \)-test for the difference of means.
      Conditions:
      * Independent group assumption: Stylists from two different salons are definitely independent groups.
      * Randomization condition: We are told that these are random samples of stylists from each salon.
      * 10% condition: The sample represents less than 10% of all possible stylists from each salon.
      * Nearly Normal condition: We do not have the data, so we do not know about this condition. We would proceed with caution.
   b. A friend tells you that he has heard that Curl Up & Dye is the more expensive salon.
      i. Write hypotheses for your friend’s claim.
         Let \( H = \) Hair by Charles and \( C = \) Curl Up & Dye.
         \( H_0 : \mu_H - \mu_C = 0 \) (There is no difference in the mean cost of haircuts at the two salons.)
         \( H_A : \mu_H - \mu_C < 0 \) (The mean cost of haircuts is higher at Curl Up & Dye.)
      ii. We would use Output A, since we are doing a two-sample \( t \)-test. (Output B is for a paired-\( t \) test.)
      iii. The \( P \)-value of 0.070 is high, so I fail to reject the null hypothesis. There is no evidence that Curl Up & Dye is any more expensive on average than Hair by Charles.
1. Which statement correctly compares \( t\)-distributions to the normal distribution?
   I. \( t\) distributions are also mound shaped and symmetric.
   II. \( t\) distributions have less spread than the normal distribution.
   III. As degrees of freedom increase, the variance of \( t\) distributions becomes smaller.
   A) I only  B) II only  C) I and II only  D) I and III only  E) I, II, and III

2. A marketing company reviewing the length of television commercials monitored a random sample of commercials over several days. They found that a 95% confidence interval for the mean length (in seconds) of commercials aired daily was (23, 27). Which is true?
   A) 95% of the commercials they checked were between 23 and 27 seconds long.
   B) 95% of all the commercials aired were between 23 and 27 seconds a day.
   C) Commercials average between 23 and 27 seconds long on 95% of the days.
   D) 95% of all samples would show mean commercial length between 23 and 27 seconds.
   E) We’re 95% sure that the mean commercial length is between 23 and 27 seconds.

3. A random sample of 120 classrooms at a large university found that 70% of them had been cleaned properly. What is the standard error of the sample proportion?
   A) 0.028  B) 0.042  C) 0.046  D) 0.082  E) 0.458

4. A relief fund is set up to collect donations for the families affected by recent storms. A random sample of 400 people shows that 28% of those 200 who were contacted by telephone actually made contributions compared to only 18% of the 200 who received first class mail requests. Which formula calculates the 95% confidence interval for the difference in the proportions of people who make donations if contacted by telephone or first class mail?
   A) \((0.28 - 0.18) \pm 1.96 \sqrt{\frac{(0.23)(0.77)}{200}}\)  
   B) \((0.28 - 0.18) \pm 1.96 \sqrt{\frac{(0.23)(0.77)}{200} + \frac{(0.23)(0.77)}{200}}\)
   C) \((0.28 - 0.18) \pm 1.96 \sqrt{\frac{(0.23)(0.77)}{400}}\)
   D) \((0.28 - 0.18) \pm 1.96 \sqrt{\frac{(0.28)(0.72)}{200} + \frac{(0.18)(0.82)}{200}}\)
   E) \((0.28 - 0.18) \pm 1.96 \sqrt{\frac{(0.28)(0.72)}{400} + \frac{(0.18)(0.82)}{400}}\)

5. Doctors at a technology research facility randomly assigned equal numbers of people to use computer keyboards in two rooms. In one room a group of people typed a manuscript using standard keyboards, while in the other room people typed the same manuscript using ergonomic keyboards to see if those people could type more words per minute. After collecting data for several days the researchers tested the hypothesis \( H_0 : \mu_1 - \mu_2 = 0 \) against the one-tail alternative and found \( P = 0.22 \). Which is true?
   A) The people using ergonomic keyboards type 22% more words per minute.
   B) There’s a 22% chance that people using ergonomic keyboards type more words per minute.
   C) There’s a 22% chance that there’s really no difference in typing speed.
   D) There’s a 22% chance another experiment will give these same results.
   E) None of these.
6. It’s common for a movie’s ticket sales to open high for the first couple of weeks, then gradually taper off as time passes. Hoping to be able to better understand how quickly sales decline, an industry analyst keeps track of box office revenues for a new film over its first 20 weeks. What inference method might provide useful insight?
   A) 1-proportion z-test   B) t-Interval for a mean   C) χ² goodness-of-fit test   D) t-test for linear regression   E) t-Interval for slope

7. Trainers need to estimate the level of fat in athletes to ensure good health. Initial tests were based on a small sample but now the trainers double the sample size for a follow-up test. The main purpose of the larger sample is to…
   A) reduce response bias.   B) decrease the variability in the population.   C) reduce non-response bias.   D) reduce confounding due to other variables.   E) decrease the standard deviation of the sampling model.

8. Based on data from two very large independent samples, two students tested a hypothesis about equality of population means using α = 0.02. One student used a one-tail test and rejected the null hypothesis, but the other used a two-tail test and failed to reject the null. Which of these might have been their calculated value of t?
   A) 1.22   B) 1.55   C) 1.88   D) 2.22   E) 2.66

9. The two samples whose statistics are given in the table thought to come from populations with equal variances. What is the pooled estimate of the population standard deviation?
   A) 1.87   B) 3.50   C) 3.52   D) 3.56   E) 5.00

<table>
<thead>
<tr>
<th>n</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>22</td>
<td>3</td>
</tr>
<tr>
<td>55</td>
<td>25</td>
<td>4</td>
</tr>
</tbody>
</table>

10. A contact lens wearer read that the producer of a new contact lens boasts that their lenses are cheaper than contact lenses from another popular company. She collected some data, then tested the null hypothesis \( H_0 : \mu_{\text{old}} - \mu_{\text{new}} = 0 \) against the alternative \( H_A : \mu_{\text{old}} - \mu_{\text{new}} > 0 \). Which of the following would be a Type II error?
   A) Deciding that the new lenses are cheaper, when in fact they really are.
   B) Deciding that the new lenses are cheaper, when in fact they are not.
   C) Deciding that the new lenses are not really cheaper, when in fact they are.
   D) Deciding that the new lenses are not really cheaper, when in fact they are not.
   E) Applying these results to all contact lenses, old and new.
11. **College admissions** According to information from a college admissions office, 62% of the students there attended public high schools, 26% attended private high schools, 2% were home schooled, and the remaining students attended schools in other countries. Among this college’s Honors Graduates last year there were 47 who came from public schools, 29 from private schools, 4 who had been home schooled, and 4 students from abroad. Is there any evidence that one type of high school might better equip students to attain high academic honors at this college? Test an appropriate hypothesis and state your conclusion.

12. **Gas mileage** Hoping to improve the gas mileage of their cars, a car company has made an adjustment in the manufacturing process. Random samples of automobiles coming off the assembly line have been measured each week that the plant has been in operation. The data from before and after the manufacturing adjustments were made are in the table. It is believed that measurements of gas mileage are normally distributed. Write a complete conclusion about the manufacturing adjustments based on the statistical software printout shown below.

| SET M1  | 24  | 21  | 26  | 25  | 23  | 24  | 19  | 22  | 20  | 24  | 20  | 21  | 27  | 22  |
| SET M2  | 22  | 24  | 28  | 28  | 27  | 24  | 22  | 24  | 27  | 25  | 27  | 23  | 28  |

Two Sample T for M1 vs M2

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>StDev</th>
<th>SEMean</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>14</td>
<td>22.71</td>
<td>2.40</td>
</tr>
<tr>
<td>M2</td>
<td>13</td>
<td>25.31</td>
<td>2.29</td>
</tr>
</tbody>
</table>

95% CI for $\mu_1 - \mu_2$ (0.74, 4.45)

T-Test $\mu_1 = \mu_2$ (vs. $\mu_1 < \mu_2$): T = 2.88, P = 0.0041, DF = 24.98
13. **Test identification** *(NOTE: Do not do these problems!)* For each, indicate which procedure you would use, the test statistic (\(z\), \(t\), or \(\chi^2\) “chi-squared”), and, if \(t\) or \(\chi^2\), the number of degrees of freedom. A choice may be used more than once.

<table>
<thead>
<tr>
<th>Type</th>
<th>(z/t/\chi^2)</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. proportion</td>
<td>(z)</td>
<td>1 sample</td>
</tr>
<tr>
<td>2. difference of proportions</td>
<td>(t)</td>
<td>2 samples</td>
</tr>
<tr>
<td>3. mean – 1 sample</td>
<td>(\chi^2)</td>
<td></td>
</tr>
<tr>
<td>4. difference of means – independent samples</td>
<td>(t)</td>
<td></td>
</tr>
<tr>
<td>5. mean of differences – matched pairs</td>
<td>(\chi^2)</td>
<td></td>
</tr>
<tr>
<td>6. goodness of fit</td>
<td>(t)</td>
<td></td>
</tr>
<tr>
<td>7. homogeneity</td>
<td>(\chi^2)</td>
<td></td>
</tr>
<tr>
<td>8. independence</td>
<td>(z)</td>
<td></td>
</tr>
<tr>
<td>9. regression, inference for slope</td>
<td>(t)</td>
<td></td>
</tr>
</tbody>
</table>

a. A union organization would like to represent the employees at the local market. A sample of the employees revealed 74 of 120 were in favor of the union. Does the union have the required 3 to 2 majority?

b. An oral surgeon is interested in estimating how long it takes to extract all four wisdom teeth. The doctor records the times for 24 randomly chosen surgeries. Estimate the time it takes to perform the surgery with a 95% confidence interval.

c. A microwave manufacturing company receives large shipments of thermal shields from two suppliers. A sample from each supplier’s shipment is selected and tested for the rate of defects. The microwave manufacturing company’s contract with each supplier states the shipment with the smallest rate of defect will be accepted. Do the shipments’ defect rates vary from each other?

d. The owner of a construction company would like to know if his current work teams can build room additions quicker than the time allotted for by the contract. A random sample of 15 room additions completed recently revealed an average completion time of 0.32 days faster than contracted. Is this strong evidence that the teams can complete room additions in less than the contract times?

e. A farmer would like to know if a new fertilizer increases his crop yield. In an effort to decide this, the farmer recorded the yield for 10 different fields prior to adding fertilizer and after adding the fertilizer. The farmer assumes the crop yields are approximately normal. Does the fertilizer work as advertised?

f. A manufacturer gets parts from four suppliers (call them A, B, C, and D). A random sample of 1000 parts is selected from shipments by each supplier. In the samples, Supplier A has 21 defects, Supplier B has 14 defects, Supplier C has 8 defects, and Supplier D has 17 defects. Does this suggest any difference between the quality of parts provided by these suppliers?

g. In a study to determine whether there is a difference between the average jail time convicted bank robbers and car thieves are sentenced to, the law students randomly selected 20 cases of each type that resulted in jail sentences during the previous year. A 90% confidence interval was created from the results.

h. Doctors offer small candies to sixty teenagers, recording the number of candies consumed by each. One hour later they test the blood sugar level for each person. Is there evidence that blood sugar levels in teenagers are related to the amount of candy eaten?
14. **Improving productivity** A packing company considers hiring a national training consultant in hopes of improving productivity on the packing line. The national consultant agrees to work with 18 employees for one week as part of a trial before the packing company makes a decision about the training program. The training program will be implemented if the average product packed increases by more than 10 cases per day per employee. The packing company manager will test a hypothesis using $\alpha = 0.05$.

a. Write appropriate hypotheses (in words *and* in symbols).

b. In this context, which do you consider to be more serious – a Type I or a Type II error? Explain briefly.

c. After this trial produced inconclusive results the manager decided to test the training program again with another group of employees. Describe two changes he could make in the trial to increase the power of the test, and explain the disadvantages of each.
15. **Student progress** The Comprehensive Test of Basic Skills (CTBS) is used by school district to assess student progress. Two of the areas tested are math and reading. A random sample of student results was reviewed to determine if there is an association between math and reading scores on the CTBS. Here are the scatterplot, the residuals plot, a histogram of the residuals, and the regression analysis of the data. Use this information to analyze the association between the math and reading scores on the CTBS.

Dependent variable is: **Reading CTBS**

No Selector

R squared = 79.2%  R squared (adjusted) = 79.1%

$s = 6.574$ with $20 - 2 = 18$ degrees of freedom

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>2966.3</td>
<td>1</td>
<td>2966.3</td>
<td>68.7</td>
</tr>
<tr>
<td>Residual</td>
<td>777.905</td>
<td>18</td>
<td>43.2159</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>s.e. of Coeff</th>
<th>t-ratio</th>
<th>prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>5.23273</td>
<td>5.971</td>
<td>0.875</td>
<td>0.3924</td>
</tr>
<tr>
<td>Math CTBS</td>
<td>0.8555635</td>
<td>0.1045</td>
<td>8.29</td>
<td>≤ 0.0001</td>
</tr>
</tbody>
</table>

a. Is there an association? Write appropriate hypotheses.

b. Are the assumptions for regression satisfied? Explain.

c. What do you conclude?

d. Create a 95% confidence interval for the true slope.

e. Explain in context what your interval means.
11. College admissions
   \( H_0: \) Distribution of school type among honors grads is the same as for whole college.
   \( H_A: \) Distribution of school type among honors grads is different.
   These are counts; we assume this group is representative of other years; after combining home
schoolers and students from abroad as “other,” expected counts of 52.08, 21.84, and 10.08 are all \( \geq 5. \)
   OK to do a chi-square goodness of fit test with 2 \( \text{df} \).
   \[
   \chi^2 = \sum \frac{(\text{Obs} - \text{Exp})^2}{\text{Exp}} = \frac{(47 - 52.08)^2}{52.08} + \frac{(29 - 21.84)^2}{21.84} + \frac{(8 - 10.08)^2}{10.08} = 3.27
   \]
   \( P = 0.195. \) With such a large \( P \)-value we do not reject the null hypothesis. There is no evidence that
students who graduate with honors came from different high school backgrounds than others.

12. Gas mileage
   \( P = 0.0041 \) is strong evidence that the gas mileage of automobiles coming off the assembly line after
the manufacturing adjustment has been increased. We are 95% confident that the mean gas mileage
has increased between 0.74 and 4.45 miles per gallon.

13. Test identification

<table>
<thead>
<tr>
<th>Type</th>
<th>( z )/( \chi^2 )</th>
<th>( df )</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>1 or 6 ( z ) or ( \chi^2 )</td>
<td>n/a or 1</td>
</tr>
<tr>
<td>b.</td>
<td>3 ( t )</td>
<td>23</td>
</tr>
<tr>
<td>c.</td>
<td>2 or 7 ( z ) or ( \chi^2 )</td>
<td>n/a or 1</td>
</tr>
<tr>
<td>d.</td>
<td>5 ( t )</td>
<td>14</td>
</tr>
<tr>
<td>e</td>
<td>5 ( t )</td>
<td>9</td>
</tr>
<tr>
<td>f</td>
<td>7 ( \chi^2 )</td>
<td>3</td>
</tr>
<tr>
<td>g</td>
<td>4 ( t )</td>
<td>19 (or tech)</td>
</tr>
<tr>
<td>h</td>
<td>9 ( t )</td>
<td>58</td>
</tr>
</tbody>
</table>

14. Improving productivity
   a. \( H_0: \mu_d = 10 \) The mean difference in the number of cases packed is (not more than) 10.
   \( H_A: \mu_d > 10 \) The mean difference in the number of cases packed is more than 10.
   b. A Type I error would be very expensive for the packing company. A Type I error would mean
      that the manager rejected the null hypothesis when in fact the null hypothesis is true. In this
      situation, by rejecting the null hypothesis the company thought the training improved
      productivity, so they paid for the consultant to train all employees. In reality, the training did not
      improve productivity so the company wasted money on training that did not help.
   c. To increase the power of the test, we could increase the level of significance (\( \alpha \)), or increase the
      sample size. Increasing the level of significance, could lead to adopting a training method that
      actually does not improve productivity. By increasing the sample size, the trial cost would
      increase and the trial might take more time.
15. Student progress
   a. $H_0$: There is no association between Math and Reading CTBS scores. $\beta_1 = 0$
      $H_A$: There is an association between Math and Reading CTBS scores. $\beta_1 \neq 0$
   b. *Straight Enough Condition: There is no obvious bend in the scatterplot.
      *Independence Condition: The residuals show no clear pattern.
      *Does the Plot Thicken? Condition: The residual plot shows reasonably consistent spread.
      *Nearly Normal condition: A histogram of the residuals is unimodal and roughly symmetric.
   c. The $P$-value is very small, so we reject the null hypothesis. There is strong evidence of a positive association between CTBS scores Math and Reading.
   d. A 95% confidence interval for $\beta_1$ is: $\beta_1 \pm t_{18} \times SE(b_1) = 0.866 \pm 2.101(0.1045)$
      or (0.646, 1.086)
   e. We are 95% confident that the Reading CTBS score will be higher, on average, between 0.646 and 1.086 points for each additional CTBS point scored on the Math CTBS test.
1. Which statement correctly compares \( t \)-distributions to the Normal distribution?
   I. \( t \) distributions are also mound shaped and symmetric.
   II. \( t \) distributions are more spread out than the normal distribution.
   III. As degrees of freedom increase, the variance of \( t \) distributions becomes larger.
   A) I only      B) II only      C) I and II only      D) I and III only      E) I, II, and III

2. A company checking the productivity of its assembly line monitored a random sample of workers for several days. They found that a 95% confidence interval for the mean number of items produced daily by each worker was (23,27). Which is true?
   A) 95% of the workers sampled produced between 23 and 37 items a day.
   B) 95% of all the workers average between 23 and 27 items a day.
   C) Workers produce an average of 23 to 27 items on 95% of the days.
   D) 95% of samples would show mean production between 23 and 27 items a day.
   E) We’re 95% sure that the mean daily worker output is between 23 and 27 items.

3. A random sample of 120 college seniors found that 30% of them had been offered jobs. What is the standard error of the sample proportion?
   A) 0.028     B) 0.042     C) 0.046     D) 0.082     E) 0.458

4. A college alumni fund appeals for donations by phoning or emailing recent graduates. A random sample of 300 alumni shows that 40% of the 150 who were contacted by telephone actually made contributions compared to only 30% of the 150 who received email requests. Which formula calculates the 98% confidence interval for the difference in the proportions of alumni who may make donations if contacted by phone or by email?
   A) \((0.40 - 0.30) \pm \frac{2.33(0.35)(0.65)}{150}\)  
   B) \((0.40 - 0.30) \pm \frac{2.33(0.35)(0.65)}{150} + \frac{2.33(0.35)(0.65)}{150}\)  
   C) \((0.40 - 0.30) \pm 2.33\frac{(0.35)(0.65)}{300}\)  
   D) \((0.40 - 0.30) \pm 2.33\frac{(0.40)(0.60) + (0.30)(0.70)}{150} + \frac{2.33(0.35)(0.65)}{150}\)  
   E) \((0.40 - 0.30) \pm 2.33\frac{(0.40)(0.60) + (0.30)(0.70)}{300}\)

5. Investigators at an agricultural research facility randomly assigned equal numbers of chickens to be housed in two rooms. In one room, the chickens experienced normal day/night cycles, while in the other room lights were left on 24 hours a day to see if those chickens would lay more eggs. After collecting data for several days the researchers tested the hypothesis \( H_0 : \mu_1 - \mu_2 = 0 \) against the one-tail alternative and found \( P = 0.22 \). Which is true?
   A) The chickens in the lighted room averaged 0.22 more eggs per day
   B) There’s a 22% chance that chickens housed in a lighted room produce more eggs.
   C) There’s a 22% chance that there’s really no difference in egg production.
   D) There’s a 22% chance another experiment will give these same results.
   E) None of these
6. We want to know the mean winning score at the US Open golf championship. An internet search gives us all the scores for the history of that tournament, and we create a 95% confidence interval based on a *t*-distribution. This procedure was not appropriate. Why?
   A) Since these are the best players in the world, the scores are probably skewed.
   B) The entire population of scores was gathered so there is no reason to do inference.
   C) The recent record-setting score is probably an outlier.
   D) The population standard deviation is known, so we should have used a *z*-model.
   E) In big golf tournaments the players are not randomly selected.

7. Food inspectors need to estimate the level of contaminants in food products packaged at a certain factory. Initial tests were based on a small sample but now the inspectors double the sample size for a follow-up test. The main purpose of the larger sample is to…
   A) decrease the standard deviation of the sampling model.
   B) reduce confounding due to other variables.
   C) reduce response bias.
   D) decrease the variability in the population.
   E) reduce non-response bias.

8. Based on data from two very large independent samples, two students tested a hypothesis about equality of population means using $\alpha = 0.05$. One student used a one-tail test and rejected the null hypothesis, but the other used a two-tail test and failed to reject the null. Which of these might have been their calculated value of *t*?
   A) 1.22     B) 1.55     C) 1.88     D) 2.22     E) 2.66

9. The two samples whose statistics are given in the table are thought to come from populations with equal variances. What is the pooled estimate of the population standard deviation?
   A) 2.65     B) 7       C) 7.14     D) 7.22     E) 10

10. You could win a $1000 prize by tossing a coin in one of two games. To win Game A, you must get exactly 50% heads. To win Game B, you must get between 45% and 55% heads. Although which game you must play will be chosen randomly, then you may decide whether to toss the coin 20 times or 50 times. How many tosses would you choose to make?
    A) It does not matter. B) 20 tosses for either game. C) 50 tosses for either.
    D) 20 tosses for A, 50 tosses for B. E) 50 tosses for A, 20 tosses for B.

11. **Peanut M&Ms** According to Mars, Incorporated, peanut M&M’s are 12% brown, 15% yellow, 12% red, 23% blue, 23% orange, and 15% green. On a Saturday when you have run out of statistics homework, you decide to test this claim. You purchase a medium bag of peanut M&M’s and find 39 browns, 44 yellows, 36 red, 78 blue, 73 orange, and 48 greens. Test an appropriate hypothesis and state your conclusion.
12. **Test identification** Suppose you were asked to analyze each of the situations described below. *(NOTE: Do not do these problems!)* For each, indicate which procedure you would use (pick the appropriate number from the list), the test statistic \((z, t, \text{ or } \chi^2 \text{ “chi-squared”})\), and, if \(t\) or \(\chi^2\), the number of degrees of freedom. A procedure may be used more than once.

<table>
<thead>
<tr>
<th>Type</th>
<th>(z/t/\chi^2)</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>g.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. proportion – 1 sample
2. difference of proportions – 2 samples
3. mean – 1 sample
4. difference of means – independent samples
5. mean of differences – matched pairs
6. goodness of fit
7. homogeneity
8. independence

a. Among randomly selected pets, 27% of the 188 dogs and 18% of the 167 cats had fleas. Does this indicate a significant difference in rates of flea problems for these two pets?

b. Are there more broken bones in summer or winter? We get records about the number of fractures treated in January and July at a random sample of 25 emergency rooms.

c. A random sample of 600 high school seniors reported their grade point averages and the amount of financial aid offered them by colleges. We wonder if there is an association between academic success and college aid.

d. For a random sample of 200 drivers at a gas station, we record the driver’s gender (male or female) and the type of gasoline purchased (regular, plus, or premium). We wonder if there is an association between a driver’s gender and the type of gasoline they buy.

e. The school newspaper wants a 95% confidence interval for the road test failure rate. In a random sample of 65 student drivers, 37 said they failed their driver’s test at least once.

f. A supermarket chain wants to know which of two merchandise display methods is more effective. They randomly assign 15 stores to use display type A and 15 others to use display type B, then collect data about the number of items sold at each store.

g. Tags placed on garbage cans allow the disposal of up to 30 pounds of garbage. A random sample of 22 cans averaged 33.2 pounds with a standard deviation of 3.2 pounds. Is this strong evidence that residents overload their garbage cans?

h. Researchers offer small cookies to nine nursery school children and record the number of cookies consumed by each. Forty-five minutes later they observe these children during recess, and rate each child for hyperactivity on a scale from 1 – 20. Is there any evidence that sugar contributes to hyperactivity in children?
13. Scrubbers A factory recently installed new pollution control equipment (“scrubbers”) on its smokestacks in hopes of reducing air pollution levels at a nearby national park. Randomly timed measurements of sulfate levels (in micrograms per cubic meter) were taken before (Set C1) and after (Set C2) the installation. We believe that measurements of sulfate levels are normally distributed. Write a complete conclusion about the effectiveness of these scrubbers based on the statistical software printout shown.

<table>
<thead>
<tr>
<th>Set</th>
<th>10.0</th>
<th>8.0</th>
<th>8.0</th>
<th>7.0</th>
<th>6.0</th>
<th>9.0</th>
<th>11.5</th>
<th>8.0</th>
<th>9.5</th>
<th>7.5</th>
<th>5.0</th>
<th>10.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scrubbers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Two Sample T for C1 vs C2</th>
<th>N</th>
<th>Mean</th>
<th>StDev</th>
<th>SEMean</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>12</td>
<td>8.29</td>
<td>1.83</td>
<td>0.53</td>
</tr>
<tr>
<td>C2</td>
<td>10</td>
<td>5.00</td>
<td>2.84</td>
<td>0.90</td>
</tr>
</tbody>
</table>

95% CI for mu1 – mu2: (1.20, 5.38)
T-Test \( \mu_1 = \mu_2 \) (vs. not =): \( T = 3.29 \) \( P = 0.0037 \) \( DF = 20 \)

14. Blood pressure Researchers developing new drugs must be concerned about possible side effects. They must check a new medication for arthritis to be sure that it does not cause an unsafe increase in blood pressure. They measure the blood pressures of a group of 12 subjects, then administer the drug and recheck the blood pressures one hour later. The drug will be approved for use unless there is evidence that blood pressure has increased an average of more than 20 points. They will test a hypothesis using \( \alpha = 0.05 \).

a. Write appropriate hypotheses (in words and in symbols).

b. In this context, which do you consider to be more serious – a Type I or a Type II error? Explain briefly.

c. After this experiment produced inconclusive results the researchers decided to test the drug again another group of patients. Describe two changes they could make in their experiment to increase the power of their test, and explain the disadvantages of each.
15. **Auto repairs** An insurance company hopes to save money on repairs to autos involved in accidents. Two body shops in town seem to do most of the repairs, and the company wonders whether one of them is generally cheaper than the other. From their files of payments made during the past year they select a random sample of ten bills they paid at each repair shop. The data are shown in the table.

Indicate what inference procedure you would use to see if there is a significant difference in the costs of repairs done at these two body shops, then decide if it is okay to actually perform that inference procedure. (Check the appropriate assumptions and conditions and indicate whether you could or could not proceed. You do not have to do the actual test.)

<table>
<thead>
<tr>
<th>Bodies by Bock</th>
<th>Velleman’s Automagic</th>
</tr>
</thead>
<tbody>
<tr>
<td>2130</td>
<td>2570</td>
</tr>
<tr>
<td>980</td>
<td>1120</td>
</tr>
<tr>
<td>3400</td>
<td>2950</td>
</tr>
<tr>
<td>2190</td>
<td>1880</td>
</tr>
<tr>
<td>1100</td>
<td>1660</td>
</tr>
<tr>
<td>1450</td>
<td>1700</td>
</tr>
<tr>
<td>4590</td>
<td>4030</td>
</tr>
<tr>
<td>3090</td>
<td>3970</td>
</tr>
<tr>
<td>1050</td>
<td>1130</td>
</tr>
<tr>
<td>2530</td>
<td>3660</td>
</tr>
</tbody>
</table>

16. **Height and weight** Height and weight data was collected from a group of randomly selected male students.

Dependent variable is: WT(lb)
R squared = 56.6%
s = 14.16 with 25 - 2 = 23

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>s.e. of Coeff</th>
<th>t-ratio</th>
<th>prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Const</td>
<td>-364.403</td>
<td>94.61</td>
<td>-3.85</td>
<td>0.0008</td>
</tr>
<tr>
<td>HT(in)</td>
<td>7.29993</td>
<td>1.333</td>
<td>5.48</td>
<td>≤ 0.0001</td>
</tr>
</tbody>
</table>

a. Is there an association? Write appropriate hypotheses.

b. Are the assumptions for regression satisfied? Explain.

c. What do you conclude?
Statistics Test C – Part VI – Key


7. Peanut M&Ms

We want to know if the distribution of colors in the bag matches the distribution stated by Mars, Incorporated.

H₀ : The distribution of colors in the bag matches the distribution stated by Mars, Incorporated.

Hₐ : The distribution of colors in the bag does not match the distribution stated by Mars, Incorporated.

Conditions:
* Counted data: We have the counts of the number of peanut M&Ms of each color.
* Randomization: We will assume that each bag of peanut M&Ms represents a random sample of peanut M&Ms.
* Expected cell frequency: There are a total of 318 peanut M&Ms. The smallest percentage of any particular color is 12% (brown and red), and we expect 318(0.12) = 38.16. Since the smallest expected count exceeds 5, all expected counts will exceed 5, so the condition is satisfied.

Under these conditions, the sampling distribution of the test statistic is χ² with 6 – 1 = 5 degrees of freedom, and we will perform a chi-square goodness-of-fit test.

<table>
<thead>
<tr>
<th>Color</th>
<th>brown</th>
<th>yellow</th>
<th>red</th>
<th>blue</th>
<th>orange</th>
<th>green</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed Frequency</td>
<td>39</td>
<td>44</td>
<td>36</td>
<td>78</td>
<td>73</td>
<td>48</td>
</tr>
<tr>
<td>Expected Frequency</td>
<td>38.16</td>
<td>47.7</td>
<td>38.16</td>
<td>73.14</td>
<td>73.14</td>
<td>47.7</td>
</tr>
</tbody>
</table>

\[
\chi^2 = \sum \frac{(Obs - Exp)^2}{Exp} = \frac{(39 - 38.16)^2}{38.16} + \frac{(44 - 47.7)^2}{47.7} + \ldots = 0.7528
\]

P-value = P(χ² > 0.7528) = 0.980

A P-value this large says that if the distribution of colors in the bag matches the distribution stated by Mars, Incorporated, an observed chi-square value of 0.7528 would happen about 98% of the time. Thus, we fail to reject the null hypothesis. These data do not show evidence that the distribution of colors in the bag differs from the distribution stated by Mars, Incorporated.

12. Test identification

<table>
<thead>
<tr>
<th>Type</th>
<th>z/t/χ²</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>2 or 7</td>
<td>z or χ²</td>
</tr>
<tr>
<td>b.</td>
<td>5</td>
<td>t</td>
</tr>
<tr>
<td>c.</td>
<td>6</td>
<td>n/a</td>
</tr>
<tr>
<td>d.</td>
<td>8</td>
<td>χ²</td>
</tr>
<tr>
<td>e.</td>
<td>1 or 6</td>
<td>z or χ²</td>
</tr>
<tr>
<td>f.</td>
<td>4</td>
<td>t</td>
</tr>
<tr>
<td>g.</td>
<td>3</td>
<td>t</td>
</tr>
<tr>
<td>h.</td>
<td>9</td>
<td>t</td>
</tr>
</tbody>
</table>
13. Scrubbers

$P < 0.05$ is strong evidence that the scrubbers have changed the mean pollution level. We are 95% confident that the mean sulfate level has decreased between 1.20 and 5.38 micrograms per cubic meter.

14. Blood pressure

a. $H_0: \mu_d = 20$ The mean increase in blood pressure is safe.
   $H_A: \mu_d > 20$ The mean increase in blood pressure exceeds the safe limit.

b. Type II is dangerous; the medication is approved even though blood pressure increases too much. Type I means that an acceptable medication is not approved; that’s too bad, but not dangerous.

c. Increase alpha; could lead to rejecting a medication that’s actually okay. Increase $n$; more costly and difficult, and could endanger more subjects.

15. Auto repairs

$t$-test for difference of means – the samples are independent, each is an SRS from its body shop and less than 10% of all repairs done there. The Bock data are unimodal and roughly symmetric, but a histogram of the Velleman bills is too skewed. We cannot proceed.

16. Height and weight

a. $H_0$: There is no association between height and weight.
   $H_A$: There is an association between height and weight.

b. The scatterplot looks straight enough, residuals are random and display consistent spread, the histogram of residuals looks roughly unimodal and symmetric.

c. Reject $H_0$ because of the small $P$-value; there is strong evidence of an association between height and weight.