3. Independent random samples of 500 households were taken from a large metropolitan area in the United States for the years 1950 and 2000. Histograms of household size (number of people in a household) for the years are shown below.
(a) Compare the distributions of household size in the metropolitan area for the years 1950 and 2000.

(b) A researcher wants to use these data to construct a confidence interval to estimate the change in mean household size in the metropolitan area from the year 1950 to the year 2000. State the conditions for using a two-sample *t-procedure*, and explain whether the conditions for inference are met.
1. The scatterplot below displays the price in dollars and quality rating for 14 different sewing machines.

(a) Describe the nature of the association between price and quality rating for the sewing machines.

(b) One of the 14 sewing machines substantially affects the appropriateness of using a linear regression model to predict quality rating based on price. Report the approximate price and quality rating of that machine and explain your choice.
(c) Chris is interested in buying one of the 14 sewing machines. He will consider buying only those machines for which there is no other machine that has both higher quality and lower price. On the scatterplot reproduced below, circle all data points corresponding to machines that Chris will consider buying.
5. A biologist is interested in studying the effect of growth-enhancing nutrients and different salinity (salt) levels in water on the growth of shrimps. The biologist has ordered a large shipment of young tiger shrimps from a supply house for use in the study. The experiment is to be conducted in a laboratory where 10 tiger shrimps are placed randomly into each of 12 similar tanks in a controlled environment. The biologist is planning to use 3 different growth-enhancing nutrients (A, B, and C) and two different salinity levels (low and high).

(a) List the treatments that the biologist plans to use in this experiment.

(b) Using the treatments listed in part (a), describe a completely randomized design that will allow the biologist to compare the shrimps’ growth after 3 weeks.

(c) Give one statistical advantage to having only tiger shrimps in the experiment. Explain why this is an advantage.

(d) Give one statistical disadvantage to having only tiger shrimps in the experiment. Explain why this is a disadvantage.
2. An administrator at a large university wants to conduct a survey to estimate the proportion of students who are satisfied with the appearance of the university buildings and grounds. The administrator is considering three methods of obtaining a sample of 500 students from the 70,000 students at the university.

(a) Because of financial constraints, the first method the administrator is considering consists of taking a convenience sample to keep the expenses low. A very large number of students will attend the first football game of the season, and the first 500 students who enter the football stadium could be used as a sample. Why might such a sampling method be biased in producing an estimate of the proportion of students who are satisfied with the appearance of the buildings and grounds?

(b) Because of the large number of students at the university, the second method the administrator is considering consists of using a computer with a random number generator to select a simple random sample of 500 students from a list of 70,000 student names. Describe how to implement such a method.

(c) Because stratification can often provide a more precise estimate than a simple random sample, the third method the administrator is considering consists of selecting a stratified random sample of 500 students. The university has two campuses with male and female students at each campus. Under what circumstance(s) would stratification by campus provide a more precise estimate of the proportion of students who are satisfied with the appearance of the university buildings and grounds than stratification by gender?
6. In order to monitor the populations of birds of a particular species on two islands, the following procedure was implemented.

Researchers captured an initial sample of 200 birds of the species on Island A; they attached leg bands to each of the birds, and then released the birds. Similarly, a sample of 250 birds of the same species on Island B was captured, banded, and released. Sufficient time was allowed for the birds to return to their normal routine and location.

Subsequent samples of birds of the species of interest were then taken from each island. The number of birds captured and the number of birds with leg bands were recorded. The results are summarized in the following table.

<table>
<thead>
<tr>
<th></th>
<th>Island A</th>
<th>Island B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number Captured in Subsequent Sample</td>
<td>180</td>
<td>220</td>
</tr>
<tr>
<td>Number with Leg Bands in Subsequent Sample</td>
<td>12</td>
<td>35</td>
</tr>
</tbody>
</table>

Assume that both the initial sample and the subsequent samples that were taken on each island can be regarded as random samples from the population of birds of this species.

(a) Do the data from the subsequent samples indicate that there is a difference in proportions of the banded birds on these two islands? Give statistical evidence to support your answer.

(b) Researchers can estimate the total number of birds of this species on an island by using information on the number of birds in the initial sample and the proportion of banded birds in the subsequent sample. Use this information to estimate the total number of birds of this species on Island A. Show your work.

(c) The analyses in parts (a) and (b) assume that the samples of birds captured in both the initial and subsequent samples can be regarded as random samples of the population of birds of this species that live on the respective islands. This is a common assumption made by wildlife researchers. Describe two concerns that should be addressed before making this assumption.
6. Every year, each student in a nationally representative sample is given tests in various subjects. Recently, a random sample of 9,600 twelfth-grade students from the United States were administered a multiple-choice United States history exam. One of the multiple-choice questions is below. (The correct answer is C.)

In 1935 and 1936 the Supreme Court declared that important parts of the New Deal were unconstitutional. President Roosevelt responded by threatening to

(A) impeach several Supreme Court justices
(B) eliminate the Supreme Court
(C) appoint additional Supreme Court justices who shared his views
(D) override the Supreme Court's decisions by gaining three-fourths majorities in both houses of Congress

Of the 9,600 students, 28 percent answered the multiple-choice question correctly.

(a) Let $p$ be the proportion of all United States twelfth-grade students who would answer the question correctly. Construct and interpret a 99 percent confidence interval for $p$.

Assume that students who actually know the correct answer have a 100 percent chance of answering the question correctly, and students who do not know the correct answer to the question guess completely at random from among the four options.

Let $k$ represent the proportion of all United States twelfth-grade students who actually know the correct answer to the question.
(c) Based on the completed tree diagram, express the probability, in terms of \( k \), that a randomly selected twelfth-grade student would correctly answer the history question.

(d) Using your interval from part (a) and your answer to part (c), calculate and interpret a 99 percent confidence interval for \( k \), the proportion of all United States twelfth-grade students who actually know the answer to the history question. You may assume that the conditions for inference for the confidence interval have been checked and verified.
3. Each full carton of Grade A eggs consists of 1 randomly selected empty cardboard container and 12 randomly selected eggs. The weights of such full cartons are approximately normally distributed with a mean of 840 grams and a standard deviation of 7.9 grams.

(a) What is the probability that a randomly selected full carton of Grade A eggs will weigh more than 850 grams?

(b) The weights of the empty cardboard containers have a mean of 20 grams and a standard deviation of 1.7 grams. It is reasonable to assume independence between the weights of the empty cardboard containers and the weights of the eggs. It is also reasonable to assume independence among the weights of the 12 eggs that are randomly selected for a full carton.

Let the random variable $X$ be the weight of a single randomly selected Grade A egg.

i) What is the mean of $X$?

ii) What is the standard deviation of $X$?
2012

2. A charity fundraiser has a Spin the Pointer game that uses a spinner like the one illustrated in the figure below.

A donation of $2 is required to play the game. For each $2 donation, a player spins the pointer once and receives the amount of money indicated in the sector where the pointer lands on the wheel. The spinner has an equal probability of landing in each of the 10 sectors.

(a) Let $X$ represent the net contribution to the charity when one person plays the game once. Complete the table for the probability distribution of $X$.

<table>
<thead>
<tr>
<th>$x$</th>
<th>$2$</th>
<th>$1$</th>
<th>$-8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(x)$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(b) What is the expected value of the net contribution to the charity for one play of the game?
(c) The charity would like to receive a net contribution of $500 from this game. What is the fewest number of times the game must be played for the expected value of the net contribution to be at least $500?

(d) Based on last year’s event, the charity anticipates that the Spin the Pointer game will be played 1,000 times. The charity would like to know the probability of obtaining a net contribution of at least $500 in 1,000 plays of the game. The mean and standard deviation of the net contribution to the charity in 1,000 plays of the game are $700 and $92.79, respectively. Use the normal distribution to approximate the probability that the charity would obtain a net contribution of at least $500 in 1,000 plays of the game.
3. A humane society wanted to estimate with 95 percent confidence the proportion of households in its county that own at least one dog.

   (a) Interpret the 95 percent confidence level in this context.

   The humane society selected a random sample of households in its county and used the sample to estimate the proportion of all households that own at least one dog. The conditions for calculating a 95 percent confidence interval for the proportion of households in this county that own at least one dog were checked and verified, and the resulting confidence interval was 0.417 ± 0.119.

   (b) A national pet products association claimed that 39 percent of all American households owned at least one dog. Does the humane society’s interval estimate provide evidence that the proportion of dog owners in its county is different from the claimed national proportion? Explain.

   (c) How many households were selected in the humane society’s sample? Show how you obtained your answer.
6. Hurricane damage amounts, in millions of dollars per acre, were estimated from insurance records for major hurricanes for the past three decades. A stratified random sample of five locations (based on categories of distance from the coast) was selected from each of three coastal regions in the southeastern United States. The three regions were Gulf Coast (Alabama, Louisiana, Mississippi), Florida, and Lower Atlantic (Georgia, South Carolina, North Carolina). Damage amounts in millions of dollars per acre, adjusted for inflation, are shown in the table below.

<table>
<thead>
<tr>
<th>HURRICANE DAMAGE AMOUNTS IN MILLIONS OF DOLLARS PER ACRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance from Coast</td>
</tr>
<tr>
<td>&lt; 1 mile</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Gulf Coast</td>
</tr>
<tr>
<td>Florida</td>
</tr>
<tr>
<td>Lower Atlantic</td>
</tr>
</tbody>
</table>

(a) Sketch a graphical display that compares the hurricane damage amounts per acre for the three different coastal regions (Gulf Coast, Florida, and Lower Atlantic) and that also shows how the damage amounts vary with distance from the coast.
(b) Describe differences and similarities in the hurricane damage amounts among the three regions.

Because the distributions of hurricane damage amounts are often skewed, statisticians frequently use rank values to analyze such data.

(c) In the table below, the hurricane damage amounts have been replaced by the ranks 1, 2, or 3. For each of the distance categories, the highest damage amount is assigned a rank of 1 and the lowest damage amount is assigned a rank of 3. Determine the missing ranks for the 10-to-20-miles distance category and calculate the average rank for each of the three regions. Place the values in the table below.

<table>
<thead>
<tr>
<th>Distance from Coast</th>
<th>&lt; 1 mile</th>
<th>1 to 2 miles</th>
<th>2 to 5 miles</th>
<th>5 to 10 miles</th>
<th>10 to 20 miles</th>
<th>Average Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gulf Coast</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Florida</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower Atlantic</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(d) Consider testing the following hypotheses.

\[ H_0: \text{There is no difference in the distributions of hurricane damage amounts among the three regions.} \]

\[ H_a: \text{There is a difference in the distributions of hurricane damage amounts among the three regions.} \]

If there is no difference in the distribution of hurricane damage amounts among the three regions (Gulf Coast, Florida, and Lower Atlantic), the expected value of the average rank for each of the three regions is 2. Therefore, the following test statistic can be used to evaluate the hypotheses above:

\[ Q = 5 \left[ (\bar{R}_G - 2)^2 + (\bar{R}_F - 2)^2 + (\bar{R}_A - 2)^2 \right] \]

where \( \bar{R}_G \) is the average rank over the five distance categories for the Gulf Coast (and \( \bar{R}_F \) and \( \bar{R}_A \) are similarly defined for the Florida and Lower Atlantic coastal regions).

Calculate the value of the test statistic \( Q \) using the average ranks you obtained in part (c).

(e) One thousand simulated values of this test statistic, \( Q \), were calculated, assuming no difference in the distributions of hurricane damage amounts among the three coastal regions. The results are shown in the table below. These data are also shown in the frequency plot where the heights of the lines represent the frequency of occurrence of simulated values of \( Q \).

<table>
<thead>
<tr>
<th>Q</th>
<th>Frequency</th>
<th>Cumulative Frequency</th>
<th>Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>48</td>
<td>48</td>
<td>4.80</td>
<td>4.80</td>
</tr>
<tr>
<td>0.4</td>
<td>262</td>
<td>310</td>
<td>26.20</td>
<td>31.00</td>
</tr>
<tr>
<td>1.2</td>
<td>168</td>
<td>478</td>
<td>16.80</td>
<td>47.80</td>
</tr>
<tr>
<td>1.6</td>
<td>153</td>
<td>631</td>
<td>15.30</td>
<td>63.10</td>
</tr>
<tr>
<td>2.8</td>
<td>186</td>
<td>817</td>
<td>18.60</td>
<td>81.70</td>
</tr>
<tr>
<td>3.6</td>
<td>59</td>
<td>876</td>
<td>5.90</td>
<td>87.60</td>
</tr>
<tr>
<td>4.8</td>
<td>33</td>
<td>909</td>
<td>3.30</td>
<td>90.90</td>
</tr>
<tr>
<td>5.2</td>
<td>52</td>
<td>961</td>
<td>5.20</td>
<td>96.10</td>
</tr>
<tr>
<td>6.4</td>
<td>16</td>
<td>977</td>
<td>1.60</td>
<td>97.70</td>
</tr>
<tr>
<td>7.6</td>
<td>15</td>
<td>992</td>
<td>1.50</td>
<td>99.20</td>
</tr>
<tr>
<td>8.4</td>
<td>6</td>
<td>998</td>
<td>0.60</td>
<td>99.80</td>
</tr>
<tr>
<td>10.0</td>
<td>2</td>
<td>1000</td>
<td>0.20</td>
<td>100.00</td>
</tr>
</tbody>
</table>
Use these simulated values and the test statistic you calculated in part (d) to determine if the observed data provide evidence of a significant difference in the distributions of hurricane damage amounts among the three coastal regions. Explain.
4. The Behavioral Risk Factor Surveillance System is an ongoing health survey system that tracks health conditions and risk behaviors in the United States. In one of their studies, a random sample of 8,866 adults answered the question “Do you consume five or more servings of fruits and vegetables per day?” The data are summarized by response and by age-group in the frequency table below.

<table>
<thead>
<tr>
<th>Age-Group (years)</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>18–34</td>
<td>231</td>
<td>741</td>
<td>972</td>
</tr>
<tr>
<td>35–54</td>
<td>669</td>
<td>2,242</td>
<td>2,911</td>
</tr>
<tr>
<td>55 or older</td>
<td>1,291</td>
<td>3,692</td>
<td>4,983</td>
</tr>
<tr>
<td>Total</td>
<td>2,191</td>
<td>6,675</td>
<td>8,866</td>
</tr>
</tbody>
</table>

Do the data provide convincing statistical evidence that there is an association between age-group and whether or not a person consumes five or more servings of fruits and vegetables per day for adults in the United States?
2007

6. A study was designed to explore subjects’ ability to judge the distance between two objects placed in a dimly lit room. The researcher suspected that the subjects would generally overestimate the distance between the objects in the room and that this overestimation would increase the farther apart the objects were. The two objects were placed at random locations in the room before a subject estimated the distance (in feet) between those two objects. After each subject estimated the distance, the locations of the objects were rerandomized before the next subject viewed the room.

After data were collected for 40 subjects, two linear models were fit in an attempt to describe the relationship between the subjects’ perceived distances (\(y\)) and the actual distance, in feet, between the two objects.

Model 1: \(\hat{y} = 0.238 + 1.080 \times (\text{actual distance})\)

The standard errors of the estimated coefficients for Model 1 are 0.260 and 0.118, respectively.

Model 2: \(\hat{y} = 1.102 \times (\text{actual distance})\)

The standard error of the estimated coefficient for Model 2 is 0.393.

(a) Provide an interpretation in context for the estimated slope in Model 1.

(b) Explain why the researcher might prefer Model 2 to Model 1 in this context.

(c) Using Model 2, test the researcher’s hypothesis that in dim light participants overestimate the distance, with the overestimate increasing as the actual distance increases. (Assume appropriate conditions for inference are met.)
The researchers also wanted to explore whether the performance on this task differed between subjects who wear contact lenses and subjects who do not wear contact lenses. A new variable was created to indicate whether or not a subject wears contact lenses. The data for this variable were coded numerically (1 = contact wearer, 0 = noncontact wearer), and this new variable, named “contact,” was included in the following model.

Model 3: \[ \hat{y} = 1.05 \times (\text{actual distance}) + 0.12 \times (\text{contact}) \times (\text{actual distance}) \]

The standard errors of the estimated coefficients for Model 3 are 0.357 and 0.032, respectively.

(d) Using Model 3, sketch the estimated regression model for contact wearers and the estimated regression model for noncontact wearers on the grid below.

(e) In the context of this study, provide an interpretation of the estimated coefficients for Model 3.
(a) Compare the distributions of household size in the metropolitan area for the years 1950 and 2000.

- **Shape:** Both distributions are skewed right; distribution of Year 2000 has a stronger skew.
- **Center:** The center has moved slightly lower from 1950's to 2000's distribution.
- **Spread:** The spread of 1950's distribution is slightly greater than that of 2000's.

(a) The distribution of household size in 1950 is slightly skewed to the right, but does not show any extreme outliers. It also includes households with sizes between 1 and 14.

The distribution of household size in 2000 is extremely skewed to the right and households with 10-12 people may be considered outliers. This distribution includes households with sizes between 1 and 12, which is different than the 1950's distribution.
2012 #1: (a) Describe the nature of the association between price and quality rating for the sewing machines.

The association seems to have a weak positive correlation between price and quality rating of the sewing machines. There are several influential points, but some may increase the sum of the squared residuals, therefore creating a weaker association (lower R²). However, the weak positive correlation indicates that an increase in price will increase the quality rating.

The low value of R² also tells that very little percent of variation in the dependent variable can be explained by the price.

(b) One of the 14 sewing machines substantially affects the appropriateness of using a linear regression model to predict quality rating based on price. Report the approximate price and quality rating of that machine and explain your choice.

\[
\begin{align*}
\text{Price:} & \quad 2000 \text{ dollars} \\
\text{Quality:} & \quad 65
\end{align*}
\]

This point denotes a substantial reason why not to use a linear regression model because a rapid increasing linear trend is shown with the removal of the point. However, with the outlier, the trend is more along the lines of a curve regression rather than a linear regression.

(c) Chris is interested in buying one of the 14 sewing machines. He will consider buying only those machines for which there is no other machine that has both higher quality and lower price. On the scatterplot reproduced below, circle all data points corresponding to machines that Chris will consider buying.
(a) Describe the nature of the association between price and quality rating for the sewing machines.

The scatter plot shows a somewhat linear relationship between price and quality rating. The majority of the observations falls between $0-$500 with an average quality rating of 60. As you go beyond the $500-$1000 range, the quality rating goes up. There is an outlier in the $1000-$2500 range. This will affect the regression line since it is an outlier in the X direction. If the outlier were removed the chart would be a better predictor of quality rating given any price from $0-$2500.

(b) One of the 14 sewing machines substantially affects the appropriateness of using a linear regression model to predict quality rating based on price. Report the approximate price and quality rating of that machine and explain your choice.

The price of the influential point is approximately $21,200 with a quality rating of 65. This point affects the regression line by changing the slope of the line. If the point were to be left in the graph, than the slope of the regression line will not predict the quality ratings as accurately.

(c) Chris is interested in buying one of the 14 sewing machines. He will consider buying only those machines for which there is no other machine that has both higher quality and lower price. On the scatterplot reproduced below, circle all data points corresponding to machines that Chris will consider buying.
(c) Give one statistical advantage to having only tiger shrimps in the experiment. Explain why this is an advantage.

Because tiger shrimp are more similar to other tiger shrimp than to other types of shrimp, using only tiger will create less variability based on the type of shrimp for the biologist, therefore he may have a smaller standard error for his experiment with not that many shrimps.

(d) Give one statistical disadvantage to having only tiger shrimps in the experiment. Explain why this is a disadvantage.

Tiger shrimp may differ from other types of shrimp, so the results of an experiment only containing tiger shrimp cannot be extended to other types of shrimp.
(c) Give one statistical advantage to having only tiger shrimps in the experiment. Explain why this is an advantage.

One statistical advantage to having only tiger shrimps is that their ambient rate of growth should be relatively the same. Also, they should be affected by the nutrients and salinity the same as the other tiger shrimps in the tanks. This will decrease the effect of varying variables on the experiment allowing for more precise conclusions to be drawn.

(d) Give one statistical disadvantage to having only tiger shrimps in the experiment. Explain why this is a disadvantage.

A disadvantage is that the experiment can only help to explain the effects on tiger shrimp. Because different kinds of shrimp may be affected differently, this experiment cannot be generalized to all shrimp, only tiger shrimp.
2013 #2:

(a) Because of financial constraints, the first method the administrator is considering consists of taking a
convenience sample to keep the expenses low. A very large number of students will attend the first football
game of the season, and the first 500 students who enter the football stadium could be used as a sample.
Why might such a sampling method be biased in producing an estimate of the proportion of students who
are satisfied with the appearance of the buildings and grounds?

This sampling method would be biased because students
who enjoy attending football games may have a more
satisfied perspective of the buildings and grounds, since
they enjoy going out onto the grounds for the football game,
and this satisfied perspective may be especially abundant
in those attending the football game causing an
inaccurate portrayal of the overall student population's
opinion.

(b) Because of the large number of students at the university, the second method the administrator is
considering consists of using a computer with a random number generator to select a simple random
sample of 500 students from a list of 70,000 student names. Describe how to implement such a method.

Get an alphabetical list of all the students and assign
each individual a number, 1–70,000. Then use a random
number generator to select 500 numbers, disregarding
repeats, and use the people whose names correspond
with the 500 numbers selected to participate in the
study.
(c) Because stratification can often provide a more precise estimate than a simple random sample, the third method the administrator is considering consists of selecting a stratified random sample of 500 students. The university has two campuses with male and female students at each campus. Under what circumstance(s) would stratification by campus provide a more precise estimate of the proportion of students who are satisfied with the appearance of the university buildings and grounds than stratification by gender?

If student satisfaction is generally the same regardless of gender for the campus and the two campuses were very different from having different designs or one was newer or nicer than the other, one another, stratification by campus would provide a more precise estimate of the proportion of students who are satisfied with the appearance of the buildings and grounds than stratification by gender.

Student “C”

(a)

This could be biased because the students who are attending the football game are obviously very proud of their school and therefore support their team and school. These students most likely will not have a problem with their school’s appearance because they’re so proud of the school itself and what it stands for.
(b) First each student should be assigned a number that falls between 00000 and 69,999. Now all 70,000 students have a number assigned to them. Then use a computer to generate a random digits table and pick out the first 500 numbers that fall between 00000 and 69,999. When you have your 500 numbers, correspond the correct number to the correct student and you have a sample of 500 random students.

(c) Stratification by campus would give a more precise estimate than gender because if one campus is nicer than another, that proportion will obviously be more satisfied with the appearance of the university, compared to the less appealing campus.
Part (c):

Possible concerns are:

1. Are some birds more likely to be captured than others? If, for example, slower birds are more likely to be captured in both the initial and subsequent samples, we would tend to underestimate the population size, thinking that a larger proportion of birds has been banded than was actually the case (because the birds that were caught and banded in the initial sample were also more likely to be the ones caught in the subsequent sample).

2. It may be the case that birds that are caught and banded in the initial sample learn from the experience and are less likely to be caught as part of the second sample. This would cause us to overestimate the population size.

3. There must be enough time between the samples so that there is adequate mixing of the banded and unbanded birds.

4. By banding the birds the researchers might make them more susceptible to their predators. In order to have a reasonable estimate, we must assume that the death rate of the banded birds is the same as the death rate of the unbanded birds so the banding procedure should not harm the birds or make them more conspicuous. For example, using large fluorescent bands is not a good idea.

5. Differentiable catchability. For example, birds that spend most of their time on the nest may be much less likely to be captured than other birds, and young birds may be more likely to be captured.

6. If the time between samples is too long, births could occur in the populations. Obviously, the new arrivals will not be banded.
2011 #6: Student “A”

Of the 9,600 students, 28 percent answered the multiple-choice question correctly.

(a) Let \( p \) be the proportion of all United States twelfth-grade students who would answer the question correctly. Construct and interpret a 99 percent confidence interval for \( p \).

\[
\hat{p} = 0.28 \quad n = 9600
\]

\[
\hat{p} \pm z^{*} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.28 \pm 2.576 \sqrt{\frac{0.28(1-0.28)}{9600}} = (0.2682, 0.2918)
\]

I am 99\% confident that the true proportion \( p \), of twelfth-grade students who would answer the question correctly, is between 0.2682 and 0.2918.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knows and answers correctly</td>
<td>Probability = ( k )</td>
</tr>
<tr>
<td>Knows and answers incorrectly</td>
<td>Probability = 0</td>
</tr>
<tr>
<td>Answers correctly</td>
<td>Conditional probability = 1</td>
</tr>
<tr>
<td>Answers incorrectly</td>
<td>Conditional probability = 0</td>
</tr>
<tr>
<td>Answers correctly</td>
<td>Conditional probability = ( \frac{1}{4} )</td>
</tr>
<tr>
<td>Answers incorrectly</td>
<td>Conditional probability = ( \frac{3}{4} )</td>
</tr>
<tr>
<td>Probability = 0.25</td>
<td></td>
</tr>
<tr>
<td>Probability = ( \frac{3}{4} (1-k) )</td>
<td></td>
</tr>
<tr>
<td>Probability = ( \frac{1}{4} (1-k) )</td>
<td></td>
</tr>
</tbody>
</table>

(c) Based on the completed tree diagram, express the probability, in terms of \( k \), that a randomly selected twelfth-grade student would correctly answer the history question.

\[ k + \frac{1}{4} (1-k) \]
(d) Using your interval from part (a) and your answer to part (c), calculate and interpret a 99 percent confidence interval for \( k \), the proportion of all United States twelfth-grade students who actually know the answer to the history question. You may assume that the conditions for inference for the confidence interval have been checked and verified.

\[
.2482 = k + \frac{1}{4}(1-k) \quad .2918 = k + \frac{1}{4}(1-k)
\]

\[
k = .0243 \quad k = .0557
\]

\[
(.0243, .0557)
\]

I am 99\% confident that the true value for \( k \), the proportion of all U.S. 12th graders who actually know the answer to the question, is between .0243 and .0557.
2013 #3:

3. Each full carton of Grade A eggs consists of 1 randomly selected empty cardboard container and 12 randomly selected eggs. The weights of such full cartons are approximately normally distributed with a mean of 840 grams and a standard deviation of 7.9 grams.

(a) What is the probability that a randomly selected full carton of Grade A eggs will weigh more than 850 grams?

\[ z = \frac{x - \mu}{\sigma} = \frac{850 - 840}{7.9} = 1.2658 \]

\[ P(Z > 1.2658) = 0.102792 \]

\[ P = 0.1028 \]

(b) The weights of the empty cardboard containers have a mean of 20 grams and a standard deviation of 1.7 grams. It is reasonable to assume independence between the weights of the empty cardboard containers and the weights of the eggs. It is also reasonable to assume independence among the weights of the 12 eggs that are randomly selected for a full carton.

Let the random variable \( X \) be the weight of a single randomly selected Grade A egg.

i) What is the mean of \( X \)?

\[ \text{Mean of full carton} = 840 \text{g} \quad \text{Mean of empty carton} = 20 \text{g} \]

\[ \mu_X = \frac{840 - 20}{12} = 68.3333 \text{g} \]

\[ \text{Mean of } X = 68.3333 \text{g} \]

ii) What is the standard deviation of \( X \)?

\[ \sigma \text{ of full carton} = 7.9 \text{g} \quad \sigma^2 \text{ of full carton} = 62.41 \]

\[ \sigma \text{ of empty carton} = 1.7 \text{g} \quad \sigma^2 \text{ of empty carton} = 2.89 \]

\[ \sigma_X^2 = \frac{62.41 - 2.89}{12} = 4.96 \text{g} \Rightarrow \sqrt{4.96} = 2.2271 \]

\[ \text{Standard deviation of } X = 2.2271 \text{g} \]
3. Each full carton of Grade A eggs consists of 1 randomly selected empty cardboard container and 12 randomly selected eggs. The weights of such full cartons are approximately normally distributed with a mean of 840 grams and a standard deviation of 7.9 grams.

(a) What is the probability that a randomly selected full carton of Grade A eggs will weigh more than 850 grams?

\[ P = 0.1028 \]

(b) The weights of the empty cardboard containers have a mean of 20 grams and a standard deviation of 1.7 grams. It is reasonable to assume independence between the weights of the empty cardboard containers and the weights of the eggs. It is also reasonable to assume independence among the weights of the 12 eggs that are randomly selected for a full carton.

Let the random variable \( X \) be the weight of a single randomly selected Grade A egg.

i) What is the mean of \( X \)?

\[ \mu_{\text{container}} = 20 \text{g} \]

\[ \mu_{\text{container} + \text{eggs}} = 840 \text{g} \]

\[ \mu_X = \frac{(\mu_{\text{c+e}} - \mu_{\text{c}})}{12} = \frac{840 - 20}{12} = 68.3333 \text{g} \]

ii) What is the standard deviation of \( X \)?

\[ \sigma_{\text{container}} = 1.7 \text{g} \]

\[ \sigma_{\text{container} + \text{eggs}} = 7.9 \text{g} \]

\[ \sigma_X = \sqrt{\left(\frac{(17)^2 + (7.9)^2}{12}\right)} = \sqrt{8.0808} = 0.7349 \]
2012 #2:

(a) Let $X$ represent the net contribution to the charity when one person plays the game once. Complete the table for the probability distribution of $X$.

<table>
<thead>
<tr>
<th>$x$</th>
<th>$S2$</th>
<th>$S1$</th>
<th>$-S8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(x)$</td>
<td>.6</td>
<td>.3</td>
<td>1</td>
</tr>
</tbody>
</table>

(b) What is the expected value of the net contribution to the charity for one play of the game?

$$E(\text{net}) = (\cdot 4)(3) + (\cdot 3)(1) - (1)(8) = \$0.70$$

The expected value of the net contribution to the charity for one play of the game is $\$0.70$.

(c) The charity would like to receive a net contribution of $500 from this game. What is the fewest number of times the game must be played for the expected value of the net contribution to be at least $500$?

Get $500$

Each game 0.70

$\#$ games = $\frac{\text{total out}}{\text{each game}}$

$\#$ games = $\frac{500}{0.70}$

$\#$ games = 714.2857

The fewest number of games played would be 715.

(d) Based on last year's event, the charity anticipates that the Spin the Pointer game will be played 1,000 times. The charity would like to know the probability of obtaining a net contribution of at least $500 in 1,000 plays of the game. The mean and standard deviation of the net contribution to the charity in 1,000 plays of the game are $700$ and $92.79$, respectively. Use the normal distribution to approximate the probability that the charity would obtain a net contribution of at least $500 in 1,000 plays of the game.

$$N(\bar{X} = 700, \sigma = 92.79)$$

$$z = \frac{\bar{X} - \mu}{\sigma} = \frac{500 - 700}{92.79} = -2.1554$$

$$P(z > -2.1554) = .9844$$

The approximate probability that the charity would obtain a net contribution of at least $500 in 1,000 plays of the game is .9844.
In part (d) the student uses calculator notation to answer the question. The student indicates use of a normal distribution but does not clearly identify the mean and standard deviation. Calculator notation in itself does not identify parameters, so the response does not satisfy this component of section 3. The student does not identify the bound or the direction. Unless the student identifies the lower bound and the upper bound of the interval, the calculator notation itself is not sufficient, and another component is not satisfied. The student reports the correct normal probability of 0.984, satisfying the third component. Only one of the three components of section 3 is satisfied, so section 3 was scored as incorrect. Because two sections were scored as essentially correct and one section was scored as incorrect, the response earned a score of a 2.
2010 #3:
3. A humane society wanted to estimate with 95 percent confidence the proportion of households in its county that own at least one dog.

(a) Interpret the 95 percent confidence level in this context.

If this test were repeated many times, 95% of the resulting intervals would capture the true proportion of households in its county that own at least one dog.

The humane society selected a random sample of households in its county and used the sample to estimate the proportion of all households that own at least one dog. The conditions for calculating a 95 percent confidence interval for the proportion of households in this county that own at least one dog were checked and verified, and the resulting confidence interval was 0.417 ± 0.119.

(b) A national pet products association claimed that 39 percent of all American households owned at least one dog. Does the humane society’s interval estimate provide evidence that the proportion of dog owners in its county is different from the claimed national proportion? Explain.

No, because 0.39 is captured in the interval that the humane society achieved.

(c) How many households were selected in the humane society’s sample? Show how you obtained your answer.

\[
\text{Margin of Error} = z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}
\]

\[
1.119 = 1.96 \sqrt{\frac{0.5(1-0.5)}{n}}
\]

\[
\left(\frac{1.119}{1.96}\right)^2 = \frac{\text{1.5(0.5)}}{n}
\]

\[
100^2 \times 0.086 \approx \frac{0.75}{n}
\]

\[
n \approx 67.82
\]

69 households
3. A humane society wanted to estimate with 95 percent confidence the proportion of households in its county that
own at least one dog.

(a) Interpret the 95 percent confidence level in this context.

The real proportion of households in the county that
own at least one dog will be in the confidence interval
95% of the time.

The humane society selected a random sample of households in its county and used the sample to estimate the
proportion of all households that own at least one dog. The conditions for calculating a 95 percent confidence
interval for the proportion of households in this county that own at least one dog were checked and verified, and
the resulting confidence interval was 0.417 ± 0.119.

(b) A national pet products association claimed that 39 percent of all American households owned at least one
dog. Does the humane society’s interval estimate provide evidence that the proportion of dog owners in its
county is different from the claimed national proportion? Explain.

\[ H_0: \ p = 0.39 \]
\[ H_a: \ p \neq 0.39 \]

\[ 0.417 \pm 0.119 \approx (0.298, 0.536) \]

Because the interval contains 0.39, we fail to reject
the null hypothesis. There is not sufficient evidence to
suggest that the proportion of dog owners in its county
is different from the claimed national proportion.

(c) How many households were selected in the humane society’s sample? Show how you obtained your answer.

\[ 0.119 = \text{margin of error} \]
\[ 0.119 = \text{critical value} \cdot \text{S.D.} \]
\[ 0.119 = (1.959) \sqrt{\frac{0.39(1-0.39)}{n}} \]
\[ 0.119 = 1.959 \sqrt{0.240} \]
\[ 0.00557 = \frac{0.3929}{n} \]
\[ n = \frac{0.3929}{0.00557} = 710.77 \]

GO ON TO THE NEXT PAGE.
Sample: 3B Score: 3
In part (a) the student provides a reasonable interpretation of the confidence level, including the concept of repeated sampling. The student’s use of “test” in the response is interpreted to mean procedure. It would have been ideal if the student had added the word *approximately* before “95% of the resulting intervals,” but the omission was considered to be minor in this situation. The response includes context as well, so part (a) was scored as essentially correct. In part (b) the response answers the question and provides the minimal justification required to be scored as essentially correct. In part (c) the student gives a correct equation but incorrectly substitutes 0.5 instead of 0.417 for $p^\ast$. The value 0.5 would be correct if the question had asked for the minimum sample size required to obtain a margin of error of 0.119 in a future study. But in this case 0.119 is the obtained margin of error, and the point of the question is to work backward to find out what sample size must have been used to obtain it. Therefore, part (c) was scored as partially correct. With two parts essentially correct and one part partially correct, the response earned a score of 3.

Sample: 3C Score: 2
In part (a) the response notes that “[t]he real proportion of households ... will be in the confidence interval 95% of the time” but does not specify what is meant by “95% of the time.” There is no mention that the intervals conceptually come from repeated sampling, so part (a) was scored as partially correct. In part (b) the student illustrates what is being tested by writing hypotheses and then uses the confidence interval to make the correct conclusion. This is a nice illustration of an essentially correct response for part (b). In part (c) the student provides the correct equation but uses the claimed value of 0.39 instead of the observed sample proportion of 0.417. In addition, there is an arithmetic error made in completing the calculation. Making either one of those errors would result in the same outcome as making both of them, which is why part (c) was scored as partially correct. With one part essentially correct and two parts partially correct, the response earned a score of 2.
2010 #6: (Exemplary Solution: Score = 4)

### Hurricane Damage Amounts in Millions of Dollars Per Acre

<table>
<thead>
<tr>
<th>Distance from Coast</th>
<th>&lt; 1 mile</th>
<th>1 to 2 miles</th>
<th>2 to 5 miles</th>
<th>5 to 10 miles</th>
<th>10 to 20 miles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gulf Coast</td>
<td>24.7</td>
<td>21.0</td>
<td>12.0</td>
<td>7.3</td>
<td>1.7</td>
</tr>
<tr>
<td>Florida</td>
<td>35.1</td>
<td>31.7</td>
<td>20.7</td>
<td>6.4</td>
<td>3.0</td>
</tr>
<tr>
<td>Lower Atlantic</td>
<td>21.8</td>
<td>15.7</td>
<td>12.6</td>
<td>1.2</td>
<td>0.3</td>
</tr>
</tbody>
</table>

(a) Sketch a graphical display that compares the hurricane damage amounts per acre for the three different coastal regions (Gulf Coast, Florida, and Lower Atlantic) and that also shows how the damage amounts vary with distance from the coast.

(b) Describe differences and similarities in the hurricane damage amounts among the three regions.

Florida had the most hurricane damage in 4 of the 5 strata. (Gulf Coast had more damage 5-10 miles from the coast). A similarity for all three regions, as expected, the damage decreased as we went farther from the coast, and it was greatest for all 3 regions when we went less than one mile from the coast.
(c) In the table below, the hurricane damage amounts have been replaced by the ranks 1, 2, or 3. For each of the distance categories, the highest damage amount is assigned a rank of 1 and the lowest damage amount is assigned a rank of 3. Determine the missing ranks for the 10-to-20-miles distance category and calculate the average rank for each of the three regions. Place the values in the table below.

<table>
<thead>
<tr>
<th>Distance from Coast</th>
<th>&lt; 1 mile</th>
<th>1 to 2 miles</th>
<th>2 to 5 miles</th>
<th>5 to 10 miles</th>
<th>10 to 20 miles</th>
<th>Average Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gulf Coast</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Florida</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1, 2</td>
</tr>
<tr>
<td>Lower Atlantic</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2, 8</td>
</tr>
</tbody>
</table>

(d) Consider testing the following hypotheses.

\[ H_0: \] There is no difference in the distributions of hurricane damage amounts among the three regions.

\[ H_a: \] There is a difference in the distributions of hurricane damage amounts among the three regions.

If there is no difference in the distribution of hurricane damage amounts among the three regions (Gulf Coast, Florida, and Lower Atlantic), the expected value of the average rank for each of the three regions is 2. Therefore, the following test statistic can be used to evaluate the hypotheses above:

\[
Q = 5 \left[ (\bar{R}_G - 2)^2 + (\bar{R}_F - 2)^2 + (\bar{R}_A - 2)^2 \right]
\]

where \( \bar{R}_G \) is the average rank over the five distance categories for the Gulf Coast (and \( \bar{R}_F \) and \( \bar{R}_A \) are similarly defined for the Florida and Lower Atlantic coastal regions).

Calculate the value of the test statistic \( Q \) using the average ranks you obtained in part (c).

\[
5 \left[ (2 - 2)^2 + (1.2 - 2)^2 + (2.8 - 2)^2 \right]
\]

\[
5 \left[ 0 + .64 + .64 \right]
\]

\[ Q = 6.4 \]
Use these simulated values and the test statistic you calculated in part (d) to determine if the observed data provide evidence of a significant difference in the distributions of hurricane damage amounts among the three coastal regions. Explain.

The $Q$ statistic for the observed data was 6.4, a $Q$-value of 6.4 or greater only occurred 39 times out of 1000 simulations.

Our $P$-value would be $\frac{39}{1000} = 0.039$.

At the significance level $\alpha = 0.05$, we reject the null hypothesis with a $P$-value $\leq 0.039$. Based on this data, there is evidence to suggest that there is a difference in the distribution of hurricane damage in these 3 regions.
Student Sample C:

(a) Sketch a graphical display that compares the hurricane damage amounts per acre for the three different coastal regions (Gulf Coast, Florida, and Lower Atlantic) and that also shows how the damage amounts vary with distance from the coast.

```
GO ON TO THE NEXT PAGE.
```

(b) Describe differences and similarities in the hurricane damage amounts among the three regions.

The graphs of all three regions appear to have similar non-linear shapes.

The center of damage in Florida is the highest, and the Lower Atlantic the lowest.

Florida appears to have the greatest spread and the Lower Atlantic and Gulf Coast appear to have similar spreads.

There do not appear to be any outliers.
ASSIGNED RANKS WITHIN DISTANCE CATEGORIES

<table>
<thead>
<tr>
<th>Distance from Coast</th>
<th>&lt; 1 mile</th>
<th>1 to 2 miles</th>
<th>2 to 5 miles</th>
<th>5 to 10 miles</th>
<th>10 to 20 miles</th>
<th>Average Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gulf Coast</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Florida</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1.2</td>
</tr>
<tr>
<td>Lower Atlantic</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2.8</td>
</tr>
</tbody>
</table>

Calculate the value of the test statistic $Q$ using the average ranks you obtained in part (c).

$$Q = 5 \left[ (2-2)^2 + (1.2-2)^2 + (2.8-2)^2 \right]$$

$$Q = 6.4$$

Use these simulated values and the test statistic you calculated in part (d) to determine if the observed data provide evidence of a significant difference in the distributions of hurricane damage amounts among the three coastal regions. Explain.

This does not provide significant evidence that there is a significant difference in the distributions of hurricane damage amounts among the three regions. Although most of the $Q$ test statistic values appear clustered between zero and three, 6.4 is not that uncommon, and there are many values farther away from the center that it is. This $Q = 6.4$ does not appear that uncommon to have occurred by random sampling variability with there being no difference among the regions.
Sample: 6C Score: 2
In part (a) the graph is a well-labeled plot of the categorical distances on the vertical axis and damage amounts on the horizontal axis, and a key is provided for ease in identifying the three regions. Section 1, consisting of part (a), was scored as essentially correct. In part (b), although the phrase “similar non linear shapes” is correct for the graphs, the response does not convey the nature of relationship (decreasing damages as distance from the coast increases), and none of the statements provides a complete response for differences among the regions for all distances. Section 2, consisting of part (b), was scored as incorrect. In parts (c) and (d) the ranks, average ranks and test statistic, $Q$, were correctly calculated. Section 3, consisting of parts (c) and (d), was scored as essentially correct. In part (e) the simulated $Q$ values are described, but an approximate $p$-value is not identified, and the test statistic $Q = 6.4$ is determined to be “not that uncommon.” Hence, section 4, consisting of part (e), was scored as incorrect. With two sections essentially correct and two sections incorrect, the response earned a score of 2.

(The following commentary is for the next problem: 2013 #4)

Sample: 4C Score: 2
The hypotheses are correctly stated but lack context. Due to the lack of context, step 1 was scored as partially correct. The correct test procedure is not identified by name or formula, and the first component of step 2 is not satisfied. The response indicates that “random” is a condition and has verified this condition by placing a check mark next to it. Although this satisfies the second component, it is considered minimal weak communication. The response says “expected cell” followed by a check mark, but the student does not indicate that all expected counts must be greater than or equal to five and does not verify that this condition has been satisfied; the third component is not satisfied. In addition, the response states “linearity” as a condition. The inclusion of this word in itself prevents the response from satisfying the third component even if the expected counts had been verified. Only one of three components is satisfied, which resulted in a score of incorrect for step 2. Step 3 includes the minimum amount of communication required by stating the correct test statistic and $p$-value, resulting in a score of essentially correct. Step 4 has the correct conclusion stated in context; however, the conclusion is not justified by linkage to the $p$-value, and the response does not satisfy the second component. Step 4 was scored as partially correct. Because one step was scored as essentially correct, two steps were scored as partially correct, and one step was scored as incorrect, the response earned a score of 2.
2013 #4: (Student Sample A—Score = 4)

<table>
<thead>
<tr>
<th>Age-Group (years)</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>18–34</td>
<td>231</td>
<td>741</td>
<td>972</td>
</tr>
<tr>
<td>35–54</td>
<td>669</td>
<td>2,242</td>
<td>2,911</td>
</tr>
<tr>
<td>55 or older</td>
<td>1,291</td>
<td>3,692</td>
<td>4,983</td>
</tr>
<tr>
<td>Total</td>
<td>2,191</td>
<td>6,675</td>
<td>8,866</td>
</tr>
</tbody>
</table>

Do the data provide convincing statistical evidence that there is an association between age-group and whether or not a person consumes five or more servings of fruits and vegetables per day for adults in the United States?

**State:** I want to perform a $\chi^2$ test for association.

**For the following at $\alpha = 0.05$:**

- $H_0$: There is no association between age group and whether or not a person consumes five or more servings of fruits and vegetables per day.
- $H_a$: There is an association between age group and whether or not a person consumes 5 or more servings of fruit and vegetables per day.

**Plan:** If conditions are met, I will perform a $\chi^2$ test for association.

- Random: Random sample (stated in question)
- Expected Counts: The smallest expected count is 240.2, which is greater than 5.
- Independent: It's reasonable to assume that the responses don't influence each other.

**Do:** $\chi^2 = 8.983$  Degrees of Freedom = 2  $p = 0.0112$

**Conclude:** The $p$-value of 0.0112 is smaller than $\alpha$, so I reject $H_0$. There is sufficient evidence to conclude that an association exists between age group and whether or not a person consumes at least 5 servings of fruits and vegetables per day for adults in the United States.
Student Sample C: Score = 2

<table>
<thead>
<tr>
<th>Age-Group (years)</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>18–34</td>
<td>231</td>
<td>741</td>
<td>972</td>
</tr>
<tr>
<td>35–54</td>
<td>669</td>
<td>2,242</td>
<td>2,911</td>
</tr>
<tr>
<td>55 or older</td>
<td>1,291</td>
<td>3,692</td>
<td>4,983</td>
</tr>
<tr>
<td>Total</td>
<td>2,191</td>
<td>6,675</td>
<td>8,866</td>
</tr>
</tbody>
</table>

Do the data provide convincing statistical evidence that there is an association between age-group and whether or not a person consumes five or more servings of fruits and vegetables per day for adults in the United States?

\[
\begin{align*}
H_0 &: \text{There is no association} \\
H_A &: \text{There is an association}
\end{align*}
\]

\[
\chi^2 = 8.983
\]

\[
p = 0.011201
\]

We may reject \( H_0 \). There is an association between age group and whether or not a person consumes 5 or more servings of fruits and vegetables per day for adults in the US.
(a) Provide an interpretation in context for the estimated slope in Model 1.

For every one foot further away the two objects actually are, our best estimate is that the perceived distance will increase by 1.080 feet on average.

(b) Explain why the researcher might prefer Model 2 to Model 1 in this context.

The researcher may believe that the true relationship is directly linear, and that if the objects were in the same place they would not be perceived as 0.238 feet apart or anything near that large.

(c) Using Model 2, test the researcher’s hypothesis that in dim light participants overestimate the distance, with the overestimate increasing as the actual distance increases. (Assume appropriate conditions for inference are met.)

\[ H_0: \beta = 1 \]
\[ H_a: \beta > 1 \]

where \( \beta \) is the true slope of the linear relationship in model 2.

Assume the sample data were independent (sufficiently randomized), the true relationship is linear, has a consistent standard deviation, for any actual distance and is normally distributed in the residuals (\( Y \)).

\[ t - test \ for \ slope \ of \ regression \ line \]

\[ t = \frac{1.102 - 1}{0.393} = 0.2595 \]

\[ df = 38 \]
\[ P(t > 0.2595 | H_0) = 0.3983 \]

There is virtually no evidence that the researcher's hypothesis is correct, because if subjects were in fact unbiased perceivers of the distance, a result indicating at least this much of overestimation would occur nearly 40% of the time in any case.
(d) Using Model 3, sketch the estimated regression model for contact wearers and the estimated regression model for noncontact wearers on the grid below.

![Graph showing the estimated regression models for contact and non-contact wearers.]

(e) In the context of this study, provide an interpretation of the estimated coefficients for Model 3.

Contact wearers overestimate the distance more.

\[
\text{model } \equiv y = (\text{actual distance}) + 0.05 \times (\text{actual distance}) + 0.12 \times (\text{contact}) \times (\text{actual distance})
\]

Everyone overestimates by 5% of the actual distance on average; contact wearers overestimate by an additional 12% of the actual distance.

Student Sample C: Score = 2

(a) Provide an interpretation in context for the estimated slope in Model 1.

For every unit farther apart the objects are placed, the subject will estimate that the objects are an additional 1.080 units apart.
(b) Explain why the researcher might prefer Model 2 to Model 1 in this context.

The researcher might prefer Model 2 because the y-intercept is 0. It contains only one estimated value and therefore has less variability. The subject doesn't start off being automatically wrong when the researcher computes the expected values when the actual distance is zero.

(c) Using Model 2, test the researcher’s hypothesis that in dim light participants overestimate the distance, with the overestimate increasing as the actual distance increases. (Assume appropriate conditions for inference are met.)

Linear Regression t-test \( \beta \) - the slope between the actual distance between objects, and the perceived distance in dim light.

\[ H_0: \beta = 1 \] : There is no difference between the actual distance and the perceived distance in dim light.

\[ H_a: \beta > 1 \] : As the actual distance between the objects increases, the distance perceived by the participant increases more.

All conditions for inference are met (given).

\[ b_1 = 1 \] proposed value
\[ b_2 = 1.102 \] model value
\[ s_e = 0.393 \] standard error
\[ \alpha = 0.05 \]

\[ t = \frac{b_2 - b_1}{s_e} = \frac{1.102 - 1}{0.393} = 0.259 \]

\( t \)-distribution
\[ t^* = 1.684 \]
\[ P(t > 1.684) = 0.05 \]

\[ P(t > 0.259) = 0.398 > 0.05 \]

Fail to reject \( H_0 \).

There is not sufficient evidence to reject \( H_0 \).

The slope between the actual distance and perceived distance is equal to 1. If \( H_0 \) were true, we would get results this extreme 39.8% of the time. This is not significant at the 0.05 level.

GO ON TO THE NEXT PAGE.
(d) Using Model 3, sketch the estimated regression model for contact wearers and the estimated regression model for noncontact wearers on the grid below.

![Graph showing perceived distance vs. actual distance.](image)

(e) In the context of this study, provide an interpretation of the estimated coefficients for Model 3.

Every participant overestimated the actual distance by 1.05 units. For every foot the actual distance increased, the perceived distance went up 1.05 feet. If someone wore contacts, their perceived distance went up an additional 0.12 feet on top of the 1.05 feet for every 1 foot increase in the actual distance.