AP Annual 2015 Conference

Results from the 2015 AP Statistics Exam

Jessica Utts, University of California, Irvine

Chief Reader, AP Statistics

jutts@uci.edu

CollegeBoard

The six free-response questions

- Question #1: Accountant salaries five years after hire
  - Compare boxplots; pros and cons of working for each company
- Question #2: Restaurant discounts for 20% of bills
  - Use confidence interval to test value; effect of change in *n*
- Question #3: Automated teller machines at the mall
  - Probability and expected value for discrete random variable
- Question #4: Aspirin and colon cancer
  - Two sample z-test for comparing proportions
- Question #5: Heights and arm spans
  - Interpret scatter plots and lines; regression prediction
- Question #6 (investigative task): Tortilla diameters
  - Compare sampling methods; sampling distributions

Plan for each question

- State question
- Present solution
- Describe common student errors
- Suggest teaching tips
- Report average score (all at the end)

Question 1
Accountant salaries five years after hire
Compare boxplots; pros and cons of working for each company

Question #1

Two large corporations, A and B, hire many new college graduates as accountants at entry-level positions. In 2009 the starting salary for an entry-level accountant position was $36,000 a year at both corporations. At each corporation, data were collected from 30 employees who were hired in 2009 as entry-level accountants and were still employed at the corporation five years later. The yearly salaries of the 60 employees in 2014 are summarized in the boxplots below.

Question 1, continued

(a) Write a few sentences comparing the distributions of the yearly salaries at the two corporations.

(b) Suppose both corporations offered you a job for $36,000 a year as an entry-level accountant.

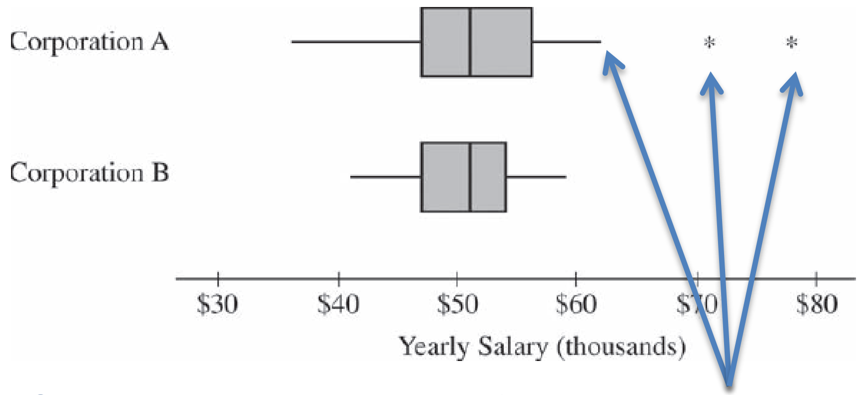
- (i) Based on the boxplots, give one reason why you might choose to accept the job at corporation A.
- (ii) Based on the boxplots, give one reason why you might choose to accept the job at corporation B.

Question 1(a) Solution

- The median salary is the same for both companies.
- The range and interquartile range of the salaries are much larger for Corporation A than for Corporation B.
- Corporation A has two outlier salaries at the high end, while Corporation B has no outliers.

Question 1(b) Solution

Q: Based on the boxplots, give one reason why you might choose to accept the job at Corporation A.

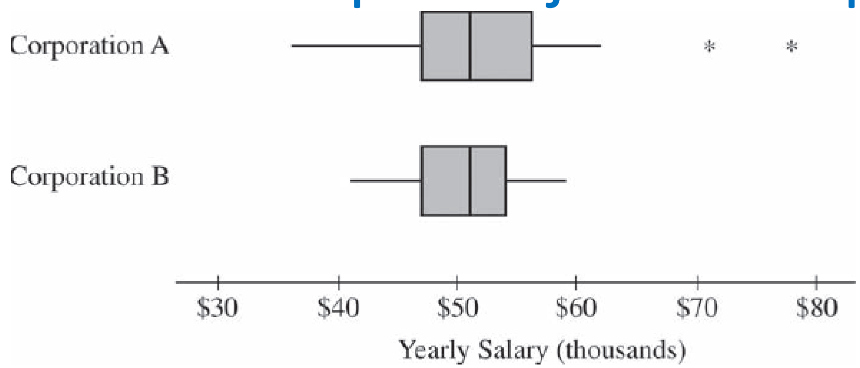


Five years after starting, at least 3 out of 30 (10%) of the salaries at Corporation A are higher than the maximum salary at Corporation B. If I accept the offer from Corporation A, I might be able to make a higher salary at A than at B.

Question 1(b) Solution, continued

AP Annual 2015 Conference

Q: Based on the boxplots, give one reason why you might choose to accept the job at Corporation B.



Five years after starting, the minimum salary at Corporation B is higher than at Corporation A. In fact, at Corporation A, it looks like some people are still making the starting salary of $36,000 and never got a raise in the 5 years since they were hired. So if I work at Corporation A, I might never get a raise.

CollegeBoard

Question 1 Common Student Errors

- Weak communication or incomplete answer (e.g. comparing spread but not mentioning median or outliers)
- Describing the boxplots but not using comparison words.
- Not understanding the limitations of what can be determined from boxplots.
- From boxplots you can tell:
  - Range and IQR
  - Outliers
  - Medians
- From boxplots you cannot tell:
  - Complete shape information.
  - Which data set has higher mean or standard deviation

Question 1 Teaching Tips

- Revisit material from earlier in the course closer to the exam
- Provide practice with *comparisons* based on visual displays, including language indicating comparison. If you explain why one choice is good, explain why the other is not as good.
- Always include *context*.
- Discuss what can be learned from each type of visual display, and what cannot.

Question 2
Restaurant discounts for 20% of bills

Use confidence interval to test value;

effect of change in n

Question #2, stem

AP Annual 2015 Conference

To increase business, the owner of a restaurant is running a promotion in which a customer's bill can be randomly selected to receive a discount. When a customer's bill is printed, a program in the cash register randomly determines whether the customer will receive a discount on the bill. **The program was written to generate a discount with a probability of 0.2, that is, giving 20 percent of the bills a discount in the long run**. However, the owner is concerned that the program has a mistake that results in the program not generating the intended long-run proportion of 0.2.

The owner selected a random sample of bills and found that only 15 percent of them received discounts. **A confidence interval for p, the proportion of bills that will receive a discount in the long run, is 0.15±0.06**.

All conditions for inference were met.

Question 2, part a(i)

Consider the confidence interval 0.15 ± 0.06

(i) Does the confidence interval provide convincing statistical evidence that the program is not working as intended? Justify your answer.

Solution

No. The confidence interval is 0.09 to 0.21, which includes the value of 0.20. Therefore, it is *plausible* that the computer program is generating discounts with probability 0.20. So the confidence interval does not provide convincing statistical evidence that the program is <u>not</u> working as intended.

Question 2, part a(ii)

Consider the confidence interval 0.15 ± 0.06

(ii) Does the confidence interval provide convincing statistical evidence that the program generates the discount with a probability of 0.2 ? Justify your answer.

Solution

No. The confidence interval includes values from 0.09 to 0.21, so any value in that interval is a plausible value for the probability that the computer is using to generate discounts.

Question 2, part (b)

A second random sample of bills was taken that was four times the size of the original sample. In the second sample 15 percent of the bills received the discount.

(b) Determine the value of the margin of error based on the second sample of bills that would be used to compute an interval for *p* with the same confidence level as that of the original interval.

Solution

The margin of error for a confidence interval for a proportion includes the square root of the sample size in the denominator. Therefore, when the sample size is multiplied by 4, the margin of error is divided by two. So the new margin of error is 0.03.

Question 2, part (c)

(c) Based on the margin of error in part (b) that was obtained from the second sample, what do you conclude about whether the program is working as intended? Justify your answer.

Solution

Using the new margin of error of 0.03, the confidence interval for *p* obtained from the second sample is 0.15 ± 0.03, or 0.12 to 0.18. This interval does not include 0.20, so there is convincing evidence that the computer program has the mistake described and is not generating discounts with probability 0.20.

Question 2 Common Student Errors

- Not knowing how to use a confidence interval to make a conclusion.
- Thinking that because 0.2 is in the interval there *is* evidence that the program is working (equivalent to accepting null hypothesis).
- Recognizing that 0.2 is a plausible value because it is in the confidence interval, but not recognizing that there are other plausible values (anything in the interval).
- Not recognizing the relationship between the new n and new margin of error (divide by square root of 4).
- Obtaining a new margin of error > .06 and not recognizing that a larger sample would mean a smaller margin of error.
- Giving .03 with no work shown, or not giving a value.
- Not knowing how to use the new margin of error to make a conclusion in part (c).

Question 2 Teaching Tips

- Teach students that a confidence interval provides a range of plausible values for the population parameter.
- Explain that there is an inverse square root relationship between sample size and margin of error. (This is one situation for which it's useful to discuss a formula.)
- Teach students that it is important to justify conclusions and calculations with a relevant explanation or formula.
- Instruct students to make sure they read the question carefully and provide an answer to the question that is asked. For example, if the questions states, "Determine the value..." then a value should be provided.

Question 3
Automated teller machines at the mall
Probability and expected value for a discrete random variable

Question #3, stem

A shopping mall has three automated teller machines (ATMs). Because the machines receive heavy use, they sometimes stop working and need to be repaired. Let the random variable X represent the number of ATMs that are working when the mall opens on a randomly selected day. The table shows the probability distribution of X.

Number of ATMs working when the mall opens	0	1	2	3
Probability	0.15	0.21	0.40	0.24

Question 3, Part (a)

Number of ATMs working when the mall opens	0	1	2	3
Probability	0.15	0.21	0.40	0.24

(a) What is the probability that at least one ATM is working when the mall opens?

Solution:

The probability that at least one ATM is working when the mall opens is

$$P(X \geq 1) = 0.21 + 0.40 + 0.24 = 0.85.$$

Question 3, Part (b)

Number of ATMs working when the mall opens	0	1	2	3
Probability	0.15	0.21	0.40	0.24

(b) What is the expected value of the number of ATMs that are working when the mall opens?

Solution:

The expected value of the number of ATMs that are working when the mall opens is

$$E(X) = 0(0.15) + 1(0.21) + 2(0.40) + 3(0.24) = 1.73$$

Question 3, Part (c)

Number of ATMs working when the mall opens	0	1	2	3
Probability	0.15	0.21	0.40	0.24

(c) What is the probability that all three ATMs are working when the mall opens, given that at least one ATM is working?

Solution:

The probability that all three ATMs are working when the mall opens, given that at least one ATM is working is

$$P(X = 3 \mid X \geq 1) = \frac{P(X = 3 \text{ and } X \geq 1)}{P(X \geq 1)} = \frac{P(X = 3)}{P(X \geq 1)} = \frac{0.24}{0.85} \approx 0.282$$

Question 3, Part (d)

(d) Given that at least one ATM is working when the mall opens, would the expected value of the number of ATMs that are working be less than, equal to, or greater than the expected value from part (b) ? Explain.

Solution:

Given the information that at least one ATM is working, the expected value of the number of working ATMs is greater than the expected value with no additional information. By eliminating the possibility of 0 working ATMs (the smallest possible number without the additional information), the probabilities for 1, 2, and 3 working ATMs all increase proportionally, so the expected value must increase.

Question 3 Common Student Errors

- Answering a question about $X \leq 1$ or $X > 1$ instead of $X \geq 1$.
- Rounding the expected value to 2 or saying the expected value was approximately 2, suggesting that they thought the mean of a discrete random variable has to be a whole number.
- Assuming events were independent when they were not. They tried to calculate by multiplying $P(X = 3)$ and $P(X \geq 1)$.
- Not showing their work when calculating probabilities or expected values.
- Using incorrect notation. For example, $P(0.24)$ instead of $P(3) = 0.24$
- Some students did not seem to know that appropriate formulas were provided on the formula sheet.

Question 3 Teaching Tips

- On computational questions involving probability and random variables, don't give credit for correct answers with no supporting work.
  - Students should show the arithmetic they are performing, even if they use a calculator to do the arithmetic.
  - Writing a generic formula is *not* sufficient for showing work, such as $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$ or $E(X) = \sum x_i p_i$
  - Calculator commands like 1-VarStats L1,L2 are *not* sufficient for showing work.
- Explain the meaning of expected value by presenting it as the long-run average if a chance process is repeated many, many times.

Question 3, More Teaching Tips

- Give students practice with using the formula sheet on assessments throughout the year.
- Stress that the multiplication rule for independent events can only be used when two events are independent.
- Provide opportunities for students to explain statistical concepts in words.
- Emphasize that statistics has its own very precise language, and that careful communication matters.

Question 4
Aspirin and colon cancer

Two sample z-test for comparing proportions

Question #4

A researcher conducted a medical study to investigate whether taking a low-dose aspirin reduces the chance of developing colon cancer. As part of the study, 1000 adult volunteers were randomly assigned to one of two groups. Half of the volunteers were assigned to the experimental group that took a low-dose aspirin each day, and the other half were assigned to the control group that took a placebo each day. At the end of six years, 15 of the people who took the low-dose aspirin had developed colon cancer and 26 of the people who took the placebo had developed colon cancer. At the significance level $\alpha = 0.05$, do the data provide convincing statistical evidence that taking a low-dose aspirin each day would reduce the chance of developing colon cancer among all people similar to the volunteers?

Question 4, Solution

Step 1: Hypotheses

Let p_{asp} represent the population proportion of adults similar to those in the study who would have developed colon cancer within the six years of the study if they had taken a low-dose aspirin each day.

Similarly, let p_{plac} represent the population proportion of adults similar to those in the study who would have developed colon cancer within the six years of the study if they had taken a placebo each day.

The hypotheses to be tested are:

$H_0: p_{asp} = p_{plac}$ versus $H_a: p_{asp} < p_{plac}$

or equivalently,

$H_0: p_{asp} - p_{plac} = 0$ versus $H_a: p_{asp} - p_{plac} < 0$

Question 4 Solution, continued

Step 2: Identify test by name or formula and check conditions.

- The appropriate procedure is a two-sample *z*-test for comparing proportions.
- This is a randomized experiment, so the first condition is that subjects were randomly assigned to treatment groups. This condition is met because we are told that the subjects were randomly assigned to take low-dose aspirin or placebo.
- The second condition is that the sample sizes are large, relative to the proportions involved. This condition is satisfied because all sample counts (15 with colon cancer in aspirin group, 26 with colon cancer in placebo group, 500 – 15 = 485 cancer-free in aspirin group, 500 – 26 = 474 cancer-free in placebo group) are large enough.

Question 4 Solution, continued

Step 3: Appropriate test statistic and *p*-value

The sample proportions who developed colon cancer are

$$\hat{p}_{asp} = \frac{15}{500} = 0.030 \qquad \hat{p}_{plac} = \frac{26}{500} = 0.052 \qquad \hat{p}_{combined} = \frac{15+26}{500+500} = 0.041.$$

$$z = \frac{0.030 - 0.052}{\sqrt{0.041(1-0.041)\left(\frac{1}{500}+\frac{1}{500}\right)}} = -1.75$$

The *p*-value is $P(Z \leq -1.75) = 0.0401$ (0.0397 from calculator), where Z has a standard normal distribution.

Question 4 Solution, continued

Step 4: Conclusion in context

Because the *p*-value is less than the given significance level of $\alpha = 0.05$, we reject the null hypothesis and conclude that the data provide convincing statistical evidence that the proportion of all adults similar to the volunteers who <u>would develop</u> colon cancer if given low-dose aspirin every day is smaller than the proportion of all adults similar to the volunteers who <u>would develop</u> colon cancer if given a placebo every day.

Question 4 Common Student Errors

- Trouble defining the parameters appropriately. Common errors were:
  - Using subscripts that do not clearly convey which group is associated with which parameter and with no explanation of which is which.
  - Defining the parameter symbol as the group rather than as a population proportion associated with the group, e.g., p_1 = placebo group.
  - Defining symbols that refer to (or imply reference to) the *sample* rather than to a *population* proportion: e.g., "p_1 is the proportion of adults who *took* low-dose aspirin daily and then *developed* cancer."
- Trouble checking the appropriate conditions for the test. For instance:
  - Students incorrectly stated that the randomness condition was satisfied because a simple random sample was chosen, rather than because of random assignment.
  - Students incorrectly stated that the normality condition was satisfied because both groups were larger than 30.

Question 4 More Common Student Errors

- Not reporting the value of the test statistic – reporting only the *p*-value.
- Using the formula for the standard error of the difference in sample proportions as the *z* statistic.
- Not providing an explicit conclusion about the research question, but simply restating a rejection of the null hypothesis in context.
- Omitting explicit justification for a decision or conclusion by failing to compare the *p*-value to the given alpha.

Question 4 Teaching Tips

- Teach students the importance of clearly defining parameters used in hypotheses. Some important factors are:
  - Making sure subscripts are defined. It is not sufficient to use subscripts of 1 and 2 without describing what they mean.
  - Making sure the parameters are explicitly defined to be about the population(s) and not the sample(s). Give students examples of definitions contrasting descriptions of sample quantities (not valid population parameters) to definitions that describe population quantities (parameters). For instance, “the proportion of adult volunteers who took aspirin and then developed colon cancer” refers to a sample quantity, but “the proportion of all adults similar to the volunteers who would have developed colon cancer if they had taken a daily aspirin” refers to a population parameter.

Question 4 More Teaching Tips

- Emphasize distinction between random *samples* and random *assignment*
- Avoid use of abbreviations such as “SRS”
- Remind students to include a test statistic, not just a *p*-value.
- Teach students that using technology is fine, but they need to report enough information from their calculator to justify their response.
- Teach students that a *decision* (reject or fail to reject the null hypothesis) is not enough. They must also include a *conclusion*, which is an answer to the scientific question asked, in context.
- Teach students to justify their *conclusion* by using statistical information:
  - providing a decision to reject or fail to reject the null hypothesis;
  - justifying that decision by making an explicit comparison of the *p*-value to the significance level (when it is provided)
  - stating a conclusion in the context of the problem.

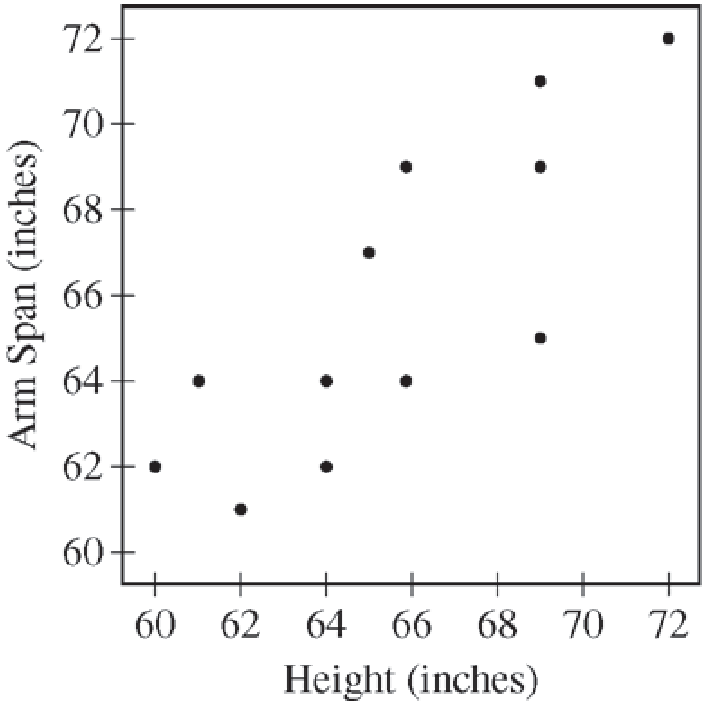
Question 5
Heights and arm spans

Interpret scatter plots and lines; regression prediction

Question #5

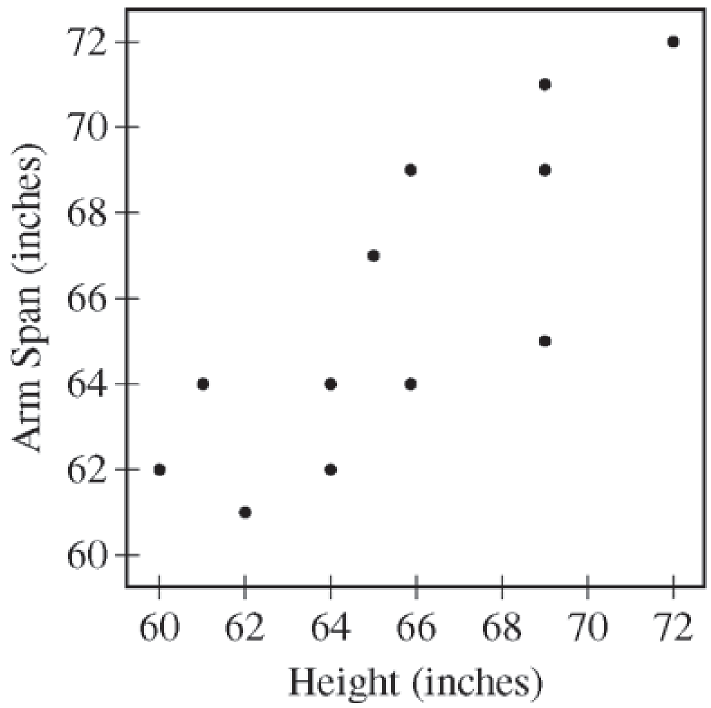
A student measured the heights and the arm spans, rounded to the nearest inch, of each person in a random sample of 12 seniors at a high school. A scatterplot of arm span versus height for the 12 seniors is shown.

(a) Based on the scatterplot, describe the relationship between arm span and height for the sample of 12 seniors.



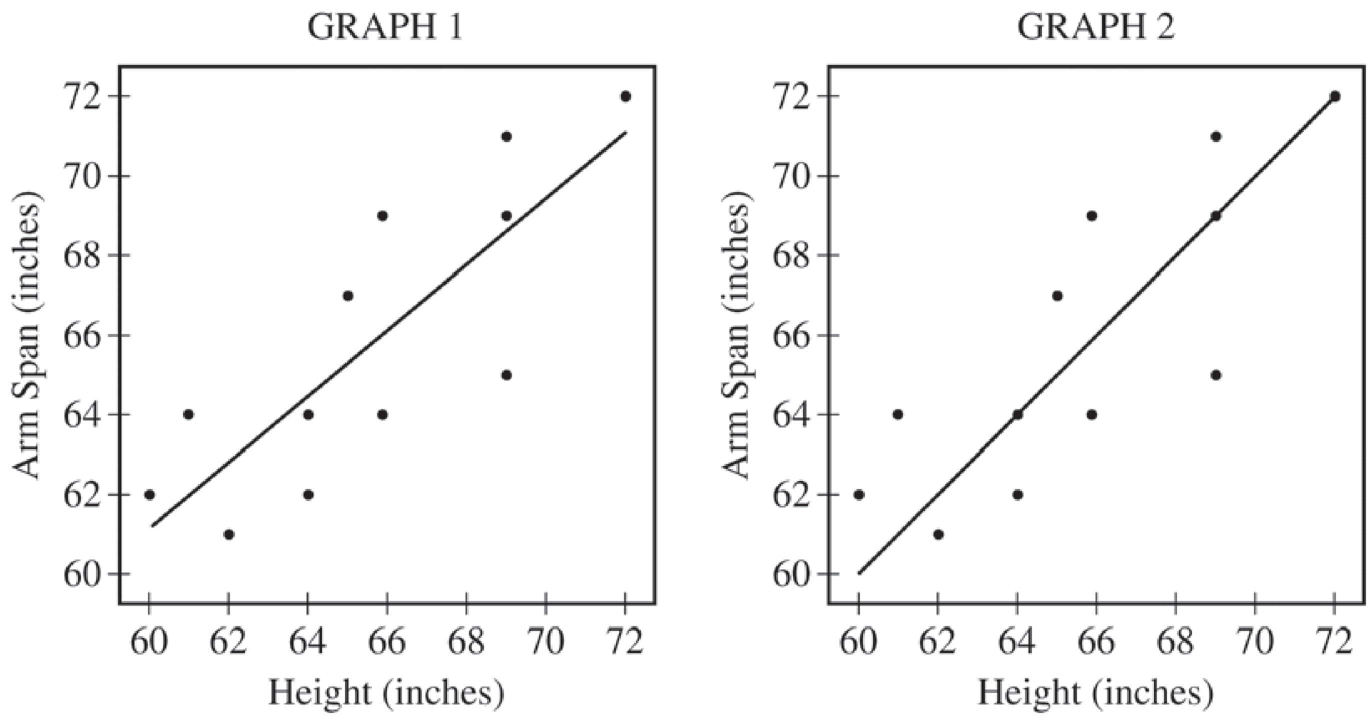
Question #5a, solution

There is a moderately strong, positive, linear relationship between height and arm span, so that taller students tend to have longer arm spans.



Question #5, continued

Let x represent height, in inches, and let y represent arm span, in inches. Two scatterplots of the same data are shown below. Graph 1 shows the data with the least squares regression line $\hat{y} = 11.74 + 0.8247x$ and graph 2 shows the data with the line $y = x$.



Question #5b

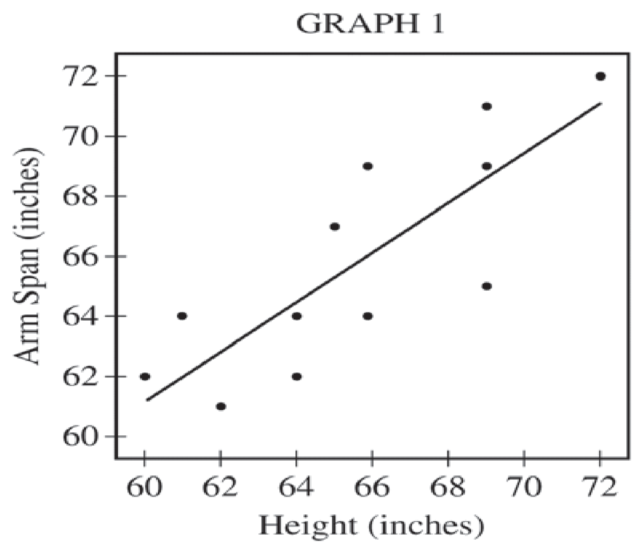
(b) The criteria described in the table below can be used to classify people into one of three body shape categories: square, tall rectangle, or short rectangle.

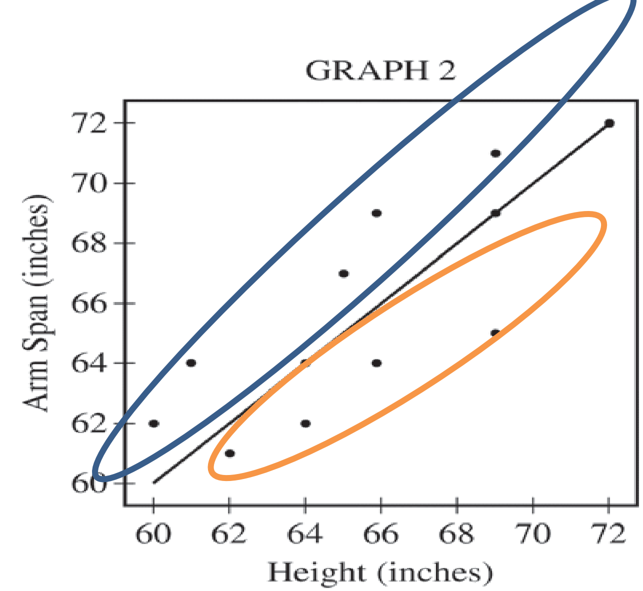
Square	Tall Rectangle	Short Rectangle
Arm span is equal to height.	Arm span is less than height.	Arm span is greater than height.

(i) For which graph, 1 or 2, is the line helpful in classifying a student’s body shape as square, tall rectangle, or short rectangle? Explain.

(ii) Complete the table of classifications for the 12 seniors.

Classification	Square	Tall Rectangle	Short Rectangle
Frequency			





Part b, solution:

(i) The line in Graph 2 is the one that is helpful. For each student, the graph illustrates whether arm span is equal to height (square = points on the line), arm span is less than height (tall rectangle = points below the line), or arm span is greater than height (short rectangle = points above the line).

(ii) Complete the table of classifications for the 12 seniors.

Classification	Square	Tall Rectangle	Short Rectangle
Frequency	3	4	5

Question #5, part c

Let x represent height, in inches, and let y represent arm span, in inches. Two scatterplots of the same data are shown below. Graph 1 shows the data with the least squares regression line $\hat{y} = 11.74 + 0.8247x$ and graph 2 shows the data with the line $y = x$.

(c) Using the best model for prediction, calculate the predicted arm span for a senior with height 61 inches.

Solution:

The predicted arm span is

$\hat{y} = 11.74 + 0.8247x = 11.74 + 0.8247(61) = 62.05$ inches

Question 5 Common Student Errors

Part a:

- Failure to use the word “linear”
- Use of “correlation” instead of “linear relationship”
- Failure to include context

Part b-i:

- Saying that the y=x line is more helpful without explaining how the y=x line divides the graph into three regions corresponding to the three body shape categories.
- Reversing the position of the “tall rectangle” and “short rectangle” categories relative to the y=x line.
- Choosing Graph 2 for “squares” and Graph 1 for “rectangles”
- Referring to the y=x line as a “regression line”.

Question 5, Common Errors

Part b-ii:

- Reporting proportions or relative frequencies, instead of frequencies.
- Reversing counts for “short rectangle” and “tall rectangle categories”.
- Not using Graph 2 as an aid to count, even when selected in part (b-i).
- Using Graph 1 as a counting aid.

Part c:

- Not using the given least square formula to predict arm span
  - Estimating from the graph
  - Computing another formula
  - Selecting a point on the plot
- Failing to show the formula with 61 inserted for height
- Failing to report units of measurement (inches)
- Not checking for “reasonableness” of the prediction

Question 5 Teaching Tips

AP Annual 2015 Conference

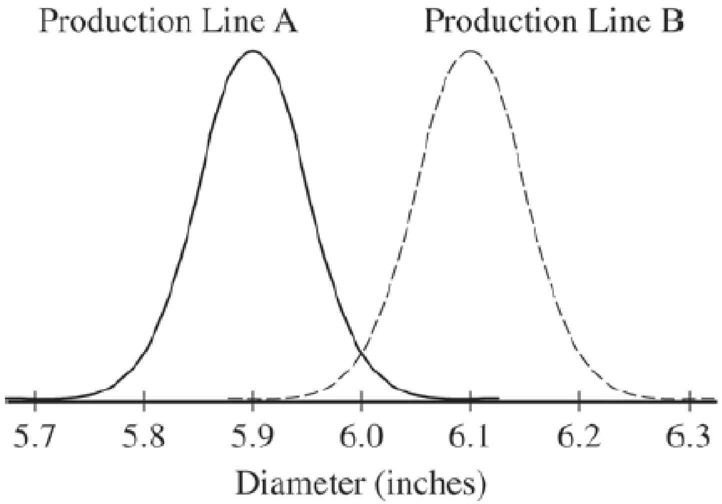
- Encourage clear handwriting.
- Give students many types of scatterplots to describe.
- For practice with scatterplots use bullets (direction, form, strength) and have students fill in a description of each.
- Do not accept answers without context.
- Use “calculate” as an instruction in student assignments and expect work to be shown.
- Give students problems in which a formula they are expected to use is presented (e.g. for a problem with data on a scatterplot, present the regression line needed to make a prediction).
- Always have students report units of measurements

Question 6
Tortilla diameters

Compare sampling methods; sampling distributions

Question 6

Corn tortillas are made at a large facility that produces 100,000 tortillas per day on each of its two production lines. The distribution of the diameters of the tortillas produced on production line A is approximately normal with mean 5.9 inches, and the distribution of the diameters of the tortillas produced on production line B is approximately normal with mean 6.1 inches. The figure below shows the distributions of diameters for the two production lines.



Question 6, continued

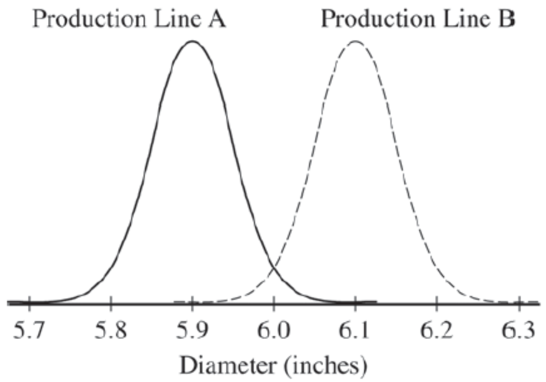
The tortillas produced at the factory are advertised as having a diameter of 6 inches. For the purpose of quality control, a sample of 200 tortillas is selected and the diameters are measured. From the sample of 200 tortillas, the manager of the facility wants to estimate the mean diameter, in inches, of the 200,000 tortillas produced on a given day. Two sampling methods have been proposed.

Method 1: Take a random sample of 200 tortillas from the 200,000 tortillas produced on a given day. Measure the diameter of each selected tortilla.

Method 2: Randomly select one of the two production lines on a given day. Take a random sample of 200 tortillas from the 100,000 tortillas produced by the selected production line. Measure the diameter of each selected tortilla.

Question 6, Basic Idea of IT

Method 1:
Simple random sample
Method 2:
Randomly choose a production line; entire sample from that line.



Variability in the collection of 200 *individual* tortillas in one day:

- Larger with Method 1, because they come from both lines.
- Smaller with Method 2, concentrated around 5.9 or 6.1.

Variability in the *sample means* from day to day:

- Smaller with Method 1, because always centered around 6.0
- Larger with Method 2, sometimes close to 5.9 and sometimes close to 6.1.

Question 6, part a

(a) Will a sample obtained using Method 2 be representative of the population of all tortillas made that day, with respect to the diameters of the tortillas? Explain why or why not.

Solution: No, a sample obtained using Method 2 will <u>not</u> be representative of all tortillas made that day. The sample obtained using Method 2 will only represent the tortillas from one production line, not from the entire population, because the distributions of tortilla diameters for the two production lines are different.

Question 6a Common Errors + Teaching Tips

AP Annual 2015 Conference

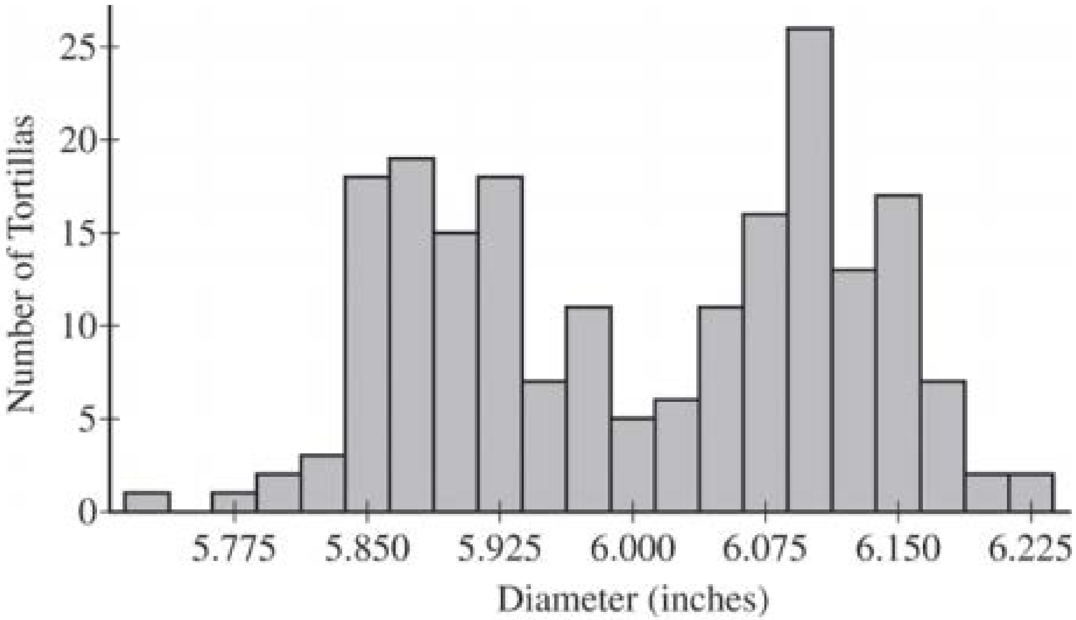
Most students correctly said that the sample would not be representative of all tortillas made that day and gave an adequate justification (e.g., only one line was selected). However, many of these students didn't give a *complete justification* for why selecting from only one line wouldn't be representative. It would have been better if students said something like "because the lines produce tortillas with different mean diameters, selecting from only one line won't produce a representative sample."

- TIP: Require that students provide complete explanations that don't require a reader to finish the argument.

CollegeBoard

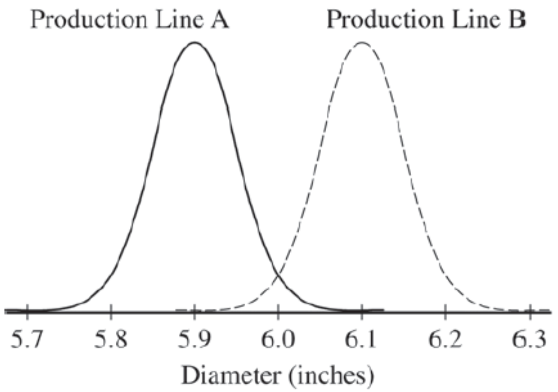
Question 6, part b

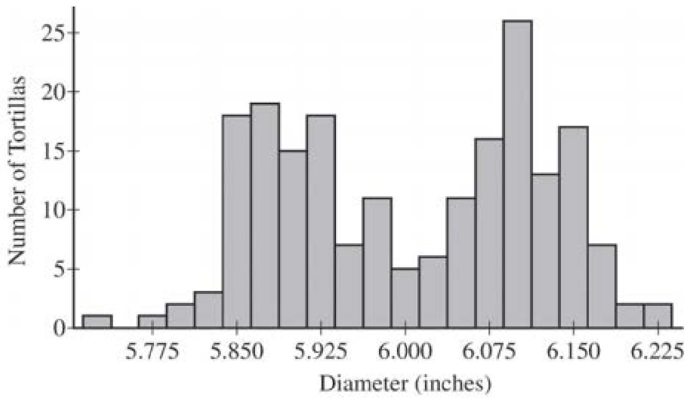
The figure below is a histogram of 200 diameters obtained by using one of the two sampling methods described. Considering the shape of the histogram, explain which method, Method 1 or Method 2, was most likely used to obtain such a sample.



Part b, solution

Method 1 was used to select this sample. The bimodal shape in the histogram of sample data indicates that tortillas were selected from both production lines, which is what would happen using Method 1. Method 2 would be likely to produce a unimodal distribution centered at either 5.9 or 6.1.





Question 6b Common Errors + Teaching Tips

Most students correctly said that the sample came from Method 1 and gave an adequate justification (e.g., the histogram is bimodal). However, many of these students didn't give a *complete justification* that also referred to the population. It would have been better if students said something like "because the histogram is bimodal which is what I would expect when sampling from two production lines that have different means."

- TIP: Require that students provide complete explanations that don't require a reader to finish the argument.

In the stem of part (b), students were told to consider the *shape* of the histogram, but some students focused on center or variability instead or simply restated that Method 1 uses tortillas from both lines.

- TIP: Do what the question asks.

Question 6, part c

Which of the two sampling methods, Method 1 or Method 2, will result in less variability in the diameters of the 200 tortillas in the sample on a given day? Explain.

Solution: Method 2 would result in less variability in the sample of 200 tortillas on a given day because the sample comes from only one production line. Because the distributions of tortilla diameters are not the same for the two production lines, selecting tortillas from both lines (as in Method 1) would result in more variable sample data.

Question 6c Common Errors + Teaching Tips

Most students correctly said that Method 2 will result in less variability in diameters on a given day and gave an adequate justification (e.g., the sample comes from only one production line). However, many of these students didn't give a *complete justification* for why selecting from only one line would result in less variability. It would have been better if students said something like "because the production lines have different means, using a sample from both production lines would likely result in more variable diameters" or "the tortilla diameters would have a range of about 0.6 inches when selecting from both lines but only about 0.4 inches when selecting from one line only."

- TIP: Require that students provide complete explanations that don't require a reader to finish the argument.
- TIP: Consider using a numerical justification when appropriate.

Question 6, part d

Each day, the distribution of the 200,000 tortillas made that day has mean diameter 6 inches with standard deviation 0.11 inches.

(d) For samples of size 200 taken from one day's production, describe the sampling distribution of the sample mean diameter for samples that are obtained using Method 1.

Solution:

The sampling distribution of the sample mean diameter for Method 1 would be approximately normal with mean 6 inches and standard deviation $\frac{0.11}{\sqrt{200}} = 0.0078$ inches.

Question 6d Common Errors

- Didn't describe *all three characteristics* of the sampling distribution (shape, center, variability).
- Unable to identify the shape of the sampling distribution of the sample mean as *approximately normal*. Some repeated the population shape (bimodal) but most didn't describe the shape at all. Among the students who were able to identify the shape, few were able to give a justification for why the shape is approximately normal.
- Many students did not remember to divide by $\sqrt{200}$ when calculating the standard deviation of the sampling distribution of the sample mean. Some of these students repeated the population standard deviation (0.11 inch) and many didn't include the standard deviation at all.
- Many students used *incorrect notation* or *sloppy language*, such as $\bar{x} = 6$ instead of $\mu_{\bar{x}} = 6$. Also, some students stated that the shape is "normal" instead of "approximately normal."

Question 6d Teaching Tips

- Give students practice with finding the sampling distribution of a mean when the original population has various, non-normal shapes.
- Remind students of the details of the Central Limit Theorem.
- Use the terminology "standard deviation of …." and fill in the blank, even when talking about the "standard deviation of the original population."

Question 6, part e

(e) Suppose that one of the two sampling methods will be selected and used every day for one year (365 days). The sample mean of the 200 diameters will be recorded each day. Which of the two methods will result in less variability in the distribution of the 365 sample means? Explain.

Solution: Method 1 would result in less variability in the sample means over the 365 days. With Method 2, roughly half of the sample means will be around 5.9 inches and the other half will be around 6.1 inches. With Method 1, however, the sample means will all be very close to 6 inches.

Question 6e, Common errors

Although most students answered “Method 1,” few students were able to describe the sampling distribution of the 365 sample means for Method 2 as having roughly half the sample means around 5.9 inches and the other half of the sample means around 6.1 inches. In some cases, it wasn’t clear that the student was describing *more than one* mean. In other cases, students didn’t imply that the sample means *will vary* from the population means by correctly using phrases such as “the sample means will cluster *around* 5.9 or *around* 6.1.”

Question 6e, Teaching tips

- Give students practice thinking about sampling distributions in *unfamiliar contexts*. For example, to estimate the standard deviation of the sampling distribution of the sample median for samples of size 5 from a population of 1000 students, explain how to simulate samples of size 5, record the median for each sample, do this many times, and calculate the standard deviation of the simulated distribution of sample medians.
- Make sure students are aware that the purpose of Question 6 (the investigative task) is to assess their ability to *integrate statistical ideas* and apply them in a new context or in a non-routine way. Prepare students by showing them examples of previous investigative tasks or asking similar questions on assessments.

Question 6, part f

A government inspector will visit the facility on June 22 to observe the sampling and to determine if the factory is in compliance with the advertised mean diameter of 6 inches. The manager knows that, with both sampling methods, the sample mean is an unbiased estimator of the population mean. However, the manager is unsure which method is more likely to produce a sample mean that is close to 6 inches on the day of sampling. Based on your previous answers, which of the two sampling methods, Method 1 or Method 2, is more likely to produce a sample mean close to 6 inches? Explain.

Solution: Method 1 is more likely to produce a daily sample mean close to 6 inches. Because the sample mean is an unbiased estimator for both methods, the manager should pick the method that would result in less variability in the sampling distribution of the sample mean. Based on the answer to part (e), Method 1 results in less variability.

Question 6f Common Errors + Teaching Tips

- Many students answered "Method 1" and gave an adequate justification based on what could happen on a single day. However, some of these students didn't indicate that the sample mean will vary from the population mean, for instance by using phrases such as "the sample mean will be *around* 5.9 or *around* 6.1" when discussing Method 2.
  - TIP: Make sure students are conscious of sampling variability and that estimates from samples are very rarely exactly correct.
- Many students did not see the connection between part (f) and part (e).
  - TIP: Remind students that the parts are often connected on free-response questions, *especially* the investigative task.

Score Summary

Q #	1 Salaries	2 Discount	3 ATMs	4 Aspirin	5 Rectangle	6 Tortillas
Mean score (Possible 0 to 4)	2.24	1.23	1.70	1.52	2.07	1.08

Exam Score	1	2	3	4	5
%	23.8%	18.9%	25.2%	18.9%	13.2%

Final Notes:

- Consider becoming an AP Statistics Reader!

  http://apcentral.collegeboard.com

- Questions/Comments: jutts@uci.edu