AP Annual 2014 Conference

Results from the 2014 AP Statistics Exam

Jessica Utts, University of California, Irvine

Chief Reader, AP Statistics

jutts@uci.edu

CollegeBoard

The six free-response questions

- Question #1: Extracurricular activities on and off campus
  - Comparing bar graphs, conclusion from *p*-value
- Question #2: Selecting conference attendees
  - Probability, simulating probability
- Question #3: School attendance and funding
  - Normal probability, probability for independent events
- Question #4: Alumni incomes
  - Mean vs median, bias in surveys
- Question #5: Car purchase price for men vs women
  - Paired t-test
- Question #6 (investigative task): Fuel consumption rate
  - Residuals, choosing a new variable for prediction

Plan for each question

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

Question 1
Extracurricular activities

Comparing bar graphs, conclusion from p-value

AP Annual 2014 Conference

Question #1

An administrator at a large university is interested in determining whether the residential status of a student is associated with level of participation in extracurricular activities. Residential status is categorized as on campus for students living in university housing and off campus otherwise. A simple random sample of 100 students in the university was taken, and each student was asked the following two questions.

- Are you an on campus student or an off campus student?
- In how many extracurricular activities do you participate?

The responses of the 100 students are summarized in the frequency table shown [next slide].

Question 1, continued

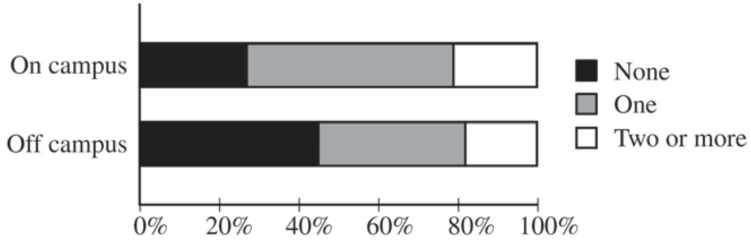
	Residential Status		
Level of Participation in Extracurricular Activities	On campus	Off campus	Total
No activities	9	30	39
One activity	17	25	42
Two or more activities	7	12	19
Total	33	67	100

(a) Calculate the proportion of on campus students in the sample who participate in at least one extracurricular activity and the proportion of off campus students in the sample who participate in at least one extracurricular activity.

- On campus proportion:
- Off campus proportion:

Question 1, continued

The responses of the 100 students are summarized in the segmented bar graph shown.



(b) Write a few sentences summarizing what the graph reveals about the association between residential status and level of participation in extracurricular activities among the 100 students in the sample.

(c) After verifying that the conditions for inference were satisfied, the administrator performed a chi-square test of the following hypotheses.

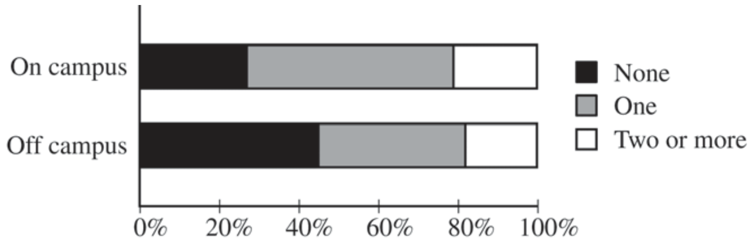
H_0 : There is no association between residential status and level of participation in extracurricular activities among the students at the university.

H_a : There is an association between residential status and level of participation in extracurricular activities among the students at the university.

The test resulted in a p-value of 0.23. Based on the p-value, what conclusion should the administrator make?

Question 1 Solution

a) On campus: $\frac{17+7}{33} \approx 0.727$; Off campus: $\frac{25+12}{67} \approx 0.552$



b) The graph reveals that on campus residents in this sample are more likely to participate in extra-curricular activities than off campus residents. The proportions who participate in *two or more extra-curricular activities* are similar between the two groups although slightly larger for on campus residents (on campus: 0.212, off campus: 0.179). On campus residents have a *greater* proportion who participate in *one activity* (on campus: 0.515, off campus 0.373) and a *smaller* proportion who participate in *no activities* (on campus: 0.273, off campus: 0.448) than off campus residents.

Question 1 Solution, continued

c) The *p* -value of 0.23 is larger than conventional significance levels such as 0.05 or 0.10. Therefore, the *p* -value indicates that the sample data do not provide strong enough evidence to conclude that participation in extra-curricular activities differs between on and off campus residents in the population of all students at the university.

Question 1 Common Student Errors

Part (a): Failure to show work, just giving decimal answer

Part (b), errors resulting from poor communication:

- Lack of comparative language, giving a statement about one or both groups without comparing.
- Comparative language that doesn't mention both groups
  - "larger than" or "smaller than" *what*?
- Using language implying counts instead of proportions
  - "more students …" instead of "larger proportion of students".

Part (c), language equivalent to <u>accepting</u> the null hypothesis:

- Accepting the null hypothesis as a stated decision.
- Giving a contextual conclusion equivalent to accepting the null. For example, "There is <u>no</u> association between …"

Question 1 Teaching Tips

- Remind students to always show their work. No "bald answers."
- Provide practice with *comparisons* based on visual displays.
- Differentiate between comparison of counts and proportions
- Teach the conceptual difference between "fail to reject" and "accept" the null hypothesis. (Never say "accept the null hypothesis" or the equivalent in context.)
- Always state the conclusion *in context*.
- Always link the *p*-value to a value of alpha to make a conclusion. It's not enough to say "the *p*-value is 0.04 and alpha is 0.05 so reject the null hypothesis." Instead, say "because the *p*-value of 0.04 is less than alpha of 0.05…"

Question 2
Selecting conference attendees

Probability, simulating probability

Question #2, parts (a) and (b)

Nine sales representatives, 6 men and 3 women, at a small company wanted to attend a national convention. There were only enough travel funds to send 3 people. The manager selected 3 people to attend and stated that the people were selected at random. The 3 people selected were women. There were concerns that no men were selected to attend the convention.

(a) Calculate the probability that randomly selecting 3 people from a group of 6 men and 3 women will result in selecting 3 women.

(b) Based on your answer to part (a), is there reason to doubt the manager's claim that the 3 people were selected at random? Explain.

Question 2 Solution, part (a)

(a) The probability that all 3 people selected are women can be calculated using the multiplication rule, as follows:

P(all three selected are women)

= $P(1^{st}$ is a woman$) \times P(2^{nd}$ is a woman$|1^{st}$ is a woman$) \times P(3^{rd}$ is a woman$|1^{st}$ and 2^{nd} are women$)$

$$= \frac{3}{9} \times \frac{2}{8} \times \frac{1}{7} \approx 0.012$$

Question 2 (a) Common Student Errors

The correct answer is based on sampling without replacement:

$$= \frac{3}{9} \times \frac{2}{8} \times \frac{1}{7} \approx 0.012$$

The most common error was computing the probability as if sampling with replacement:

$$\frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} = \frac{1}{27} \approx 0.037$$

Another common error:

$$\frac{1}{9} \times \frac{1}{8} \times \frac{1}{7} \approx 0.002$$

Question 2 (a) Teaching Tips

- Provide students with practice in situations in which the assumption of independent observations is not reasonable.
- It may be that students are exposed to probability computations primarily in abstract situations (*e.g.*, cards, dice) and do not see the connection to real-world situations. Teaching probability using such situations, even if contrived, may better help them understand how and why probabilities matter in real life.

Question 2 Solution, part (b)

(b) The probability calculated in part (a) does provide a reason to doubt the manager's claim that the selections were made at random. The calculation shows that there is only about a 1.2% chance that random selection would have resulted in three women being selected. The probability is small enough that it may cast doubt on the manager's claim that the selections were made at random.

Question 2 (b) Common Student Errors

- Computed the wrong probability, and then said it was low because that's what they thought was expected: "0.34 is a very low probability…" (really?)
- "The probability that he picked the three women at random is very low…" (wrong event)
- "There is reason to doubt his claim because the probability is 0.012." ("linkage" missing)
- "The probability is very low, so it shows the manager didn't really select them at random." (Too strong a conclusion)
- "The probability is very low, but there's still a chance it could happen, so there's no reason to doubt the manager's claim." (Why do we do statistics?)

Question 2 (b) Teaching Tips

- Students need to understand better the logical structure of “event would be unusual under Model X, so event occurrence should make us question Model X”.
  - Teach hypothesis testing not as a mechanical process, but as a logical structure. Play hypothesis testing “games” with them.
- Students need lots and lots of practice writing coherent logical arguments. This should not be done by having them “fill in templates”, but by having them write in words what they think they understand. Many, many times!

Question 2, part (c)

(c) An alternative to calculating the exact probability is to conduct a simulation to estimate the probability. A proposed simulation process is described below.

> Each trial in the simulation consists of rolling three fair, six-sided dice, one die for each of the convention attendees. For each die, rolling a 1, 2, 3, or 4 represents selecting a man; rolling a 5 or 6 represents selecting a woman. After 1,000 trials, the number of times the dice indicate selecting 3 women is recorded.

Does the proposed process correctly simulate the random selection of 3 women from a group of 9 people consisting of 6 men and 3 women? Explain why or why not.

Question 2 Solution, part (c)

(c) No, the process does not correctly simulate the random selection of 3 women from a group of 9 people of whom three are women and six are men. The selection of three people among nine is done *without* replacement, but the simulation with dice simulates selecting three people *with* replacement, since the three dice rolls in any given trial are independent of one another.

Question 2 (c) Common Student Errors

- Failure to recognize dependence structure in the selection of people—even when they got part (a) correct. “Yes, the simulation works, because $\frac{2}{6} = \frac{3}{9}$ ”
- “No, because a die only has six sides and there are nine people.”
- “No, because the dice don’t take into account sampling without replacement.” (Need to be more explicit in their communication; this typically occurred in responses that were strong overall)
- “No, because the trials are independent. You can’t select the same person more than once.” (Poor communication: what “trials” are we talking about?)

Question 2 (c) Teaching Tips

- Have students design and conduct their own simulations, and describe them in writing.
- Be sure students are exposed to simulations involving both dependence (as with sampling “from a hat”) and independence (as with dice rolls). Juxtapose them and contrast them explicitly.
- The so-called “10 percent rule” is a mystery for most students. Have them conduct simulations in situations when it is satisfied (so with and without replacement result are similar) and when it is not (so with and without replacement are different).

Question 3
School attendance and funding

Normal probability, probability for independent events

Question #3, Part (a)

Schools in a certain state receive funding based on the number of students who attend the school. To determine the number of students who attend a school, one school day is selected at random and the number of students in attendance that day is counted and used for funding purposes. The daily number of absences at High School A in the state is approximately normally distributed with mean of 120 students and standard deviation of 10.5 students.

(a) If more than 140 students are absent on the day the attendance count is taken for funding purposes, the school will lose some of its state funding in the subsequent year. Approximately what is the probability that High School A will lose some state funding?

Question 3, Part (a) Solution

Because the distribution of the daily number of absences is approximately normal with mean 120 students and standard deviation 10.5 students, the z-score for an absence total of 140 students is

$$z = \frac{140 - 120}{10.5} \approx 1.90.$$

The table of standard normal probabilities or a calculator reveals that the probability that 140 or fewer students are absent is 0.9713. So the probability that more than 140 students are absent and, therefore, the school will lose some state funding is 1 – 0.9713 = 0.0287

Question 3 (a) Common Student Errors

1. Lack of Communication
   - Not identifying that the Normal distribution is being used.
   - Not identifying the parameter values
   - Omitting "z" in a z-score equation
   - Inadequate sketch to get credit

2. Reversing x and μ in the *z*-score equation

3. Using "Table A" incorrectly looking at 1.95 instead of 1.905

Annual 2014 Conference

Question 3 (a) Teaching Tips

Remind students to:

1. To identify the type of distribution… to list and identify the parameters.

2. Label values when using calculator syntax.

3. Use the formula and show your work in a z-score calculation (**include the "z"**)

$$z = \frac{\text{value} - \text{mean}}{\text{S.D.}}$$

4. Sketch a normal curve and shade the area to help estimate the probability being computed.

CollegeBoard

Question 3, Part (b)

(b) The principals' association in the state suggests that instead of choosing one day at random, the state should choose 3 days at random. With the suggested plan, High School A would lose some of its state funding in the subsequent year if the mean number of students absent for the 3 days is greater than 140. Would High School A be more likely, less likely, or equally likely to lose funding using the suggested plan compared to the plan described in part (a)? Justify your choice.

Question 3 Part (b) Solution

High School A would be ***less likely*** to lose state funding. With a random sample of 3 days, the distribution of the sample mean number of students absent would have less variability than that of a single day. With less variability, the distribution of the sample mean would concentrate more narrowly around the mean of 120 students, resulting in a smaller probability that the mean number of students absent would exceed 140.

In particular, the standard deviation of the sample mean number of absences is $\frac{12.5}{\sqrt{3}} = 6.062$ so the *z*-score for a sample mean of 140 is

$$z = \frac{140 - 120}{6.062} \approx 3.30$$

The probability that High School A loses funding using the suggested plan would be 1 – 0.9995 = 0.0005 as determined from the table of standard normal probabilities or from a calculator, which is less than a probability of 0.0287 obtained for the plan described in part (a)

Question 3 (b) Common Student Errors

1. Lack of Communication
   - Not explicitly communicating that the sampling distribution of the mean is appropriate
   - Did not answer the question *AND/OR* justify
   - No comparison of probabilities

2. Answering $(.0287)^3$, because each day must have more than 140 absences for the mean to be > 140.

3. CLT does not apply for a sample size of n=3.

Question 3 (b) Teaching Tips

- Make sure the question is answered *AND* justified (explain why)

  "less likely" because ...

- Emphasize the difference between the distribution of a population and a sampling distribution

- Using a well-labeled sketch would be an effective method to communicate the difference in distributions (one day versus mean of 3 days)

Question 3, Part (c)

(c) A typical school week consists of the days Monday, Tuesday, Wednesday, Thursday, and Friday. The principal at High School A believes that the number of absences tends to be greater on Mondays and Fridays, and there is concern that the school will lose state funding if the attendance count occurs on a Monday or Friday. If one school day is chosen at random from each of 3 typical school weeks, what is the probability that none of the 3 days chosen is a Tuesday, Wednesday, or Thursday?

Question 3 Part (c) Solution

For any one typical school week, the probability is 2/5 = 0.4 that the day selected is not a Tuesday, Wednesday, or Thursday. Therefore, because the days are selected independently across the three weeks, the probability that none of the three days selected would be a Tuesday or Wednesday or Thursday is $(0.4)^3 = 0.064$.

This is equivalent to finding a binomial probability of 3 successes in 3 trials (weeks) when the probability of a success (no Tues, Wed, Thurs) is 0.4 on each trial.

Question 3 (c) Common Student Errors

1. P(None) = 1 – P(All)

2. Used the Binomial with $n = 5$ or 15.

Question 3 (c) Teaching Tips

1. Practice "easier" probability problems.

2. Recognize independence.

Remind students that:

3. Probabilities >1 are not a good thing.

4. Make sure the answer you get makes sense.

Question 4
Alumni incomes

Mean vs median, bias in surveys

Question #4

As part of its twenty-fifth reunion celebration, the class of 1988 (students who graduated in 1988) at a state university held a reception on campus. In an informal survey, the director of alumni development asked 50 of the attendees about their incomes. The director computed the mean income of the 50 attendees to be $189,952. In a news release, the director announced, "The members of our class of 1988 enjoyed resounding success. Last year's mean income of its members was $189,952!"

(a) What would be a statistical advantage of using the median of the reported incomes, rather than the mean, as the estimate of the typical income?

Question 4, continued

(b) The director felt the members who attended the reception may be different from the class as a whole. A more detailed survey of the class was planned to find a better estimate of the income as well as other facts about the alumni. The staff developed two methods based on the available funds to carry out the survey.

Method 1: Send out an e-mail to all 6,826 members of the class asking them to complete an online form. The staff estimates that at least 600 members will respond.

Method 2: Select a simple random sample of members of the class and contact the selected members directly by phone. Follow up to ensure that all responses are obtained. Because method 2 will require more time than method 1, the staff estimates that only 100 members of the class could be contacted using method 2.

Which of the two methods would you select for estimating the average yearly income of all 6,826 members of the class of 1988? Explain your reasoning by comparing the two methods and the effect of each method on the estimate.

Question 4 Solution

(a) The median would be less affected by skewness and outliers than the mean. With a variable such as income, a small number of very large incomes could dramatically increase the mean but not the median. Therefore, the median would be more likely than the mean to provide a better estimate of a typical income value.

(b) Method 2 is better than Method 1. Method 1 could be biased because it seems plausible that class members with larger incomes would be more likely to return the form than class members with smaller incomes. The mean income in this sample is likely to overestimate the mean income of all class members. With Method 2, despite the smaller sample size, the random selection is likely to result in a sample that is more representative of the entire class and also to produce an unbiased estimate of mean yearly income of all class members.

Question 4 (a) Common Student Errors

- Many students provided only a generic statement describing how outliers and skewness affect the mean or don't affect the median *without* referencing the distribution of incomes. Using the variable name (income) wasn't enough to earn credit. It had to be clear that the student was relating a statement about the effect of skewness or outliers to the fact that the distribution of incomes may be skewed or have outliers.
- Some students didn't describe the *effect* of skewness or outliers on the mean/median. Instead, they included statements such as "use the median when there are outliers" without explaining *why* they should use the median.
- Many students are sloppy with their language, saying things such as "outliers will skew the mean" or "means are biased when there is skewness."

Question 4 (a) Teaching Tips

- Responses should address the context of the problem in a meaningful way.
- Always justify your choice with a reason.
- Don’t use statistical terms (bias, skew, and so on) unless you use them correctly.

Question 4 (b) Common Student Errors

- Using incorrect names for describing the problem with nonresponse in Method 2, such as “response bias” Naming but not describing the bias was another error.
- Choosing Method 2 and providing good reasons not to choose Method 1, but never providing any evidence that Method 2 was better.
- Some students didn’t explicitly choose a method
- Some students chose Method 1, ignoring nonresponse bias, focusing on sample size.
- Many students discussed possible problems with Method 2, including potential nonresponse. However, the question said that all responses would be obtained in Method 2.

Question 4 (b) Common Student Errors (cont.)

- Many, many students did not address the effect of the bias on the estimated mean, even though this was specifically asked for in the question.
- When trying to address the effect on the estimate, many students used ambiguous phrases such as “the results will be more accurate. It was unclear what students meant by “more accurate” other than “the results will be better.”
- Some students discussed the conditions for inference when choosing Method 1 or Method 2. Although it is reasonable to consider the conditions for inference when choosing a data collection method, many students made incorrect statements in the process, such as “large samples mean that the sample data will be normal” or “a SRS means that we can apply the Central Limit Theorem.”

Question 4 (b) Teaching Tips

- Worry less about memorizing labels and more about describing the concept.
- Always gives reasons why you are choosing what you're choosing, and reasons why you are not choosing what you are not choosing.
- Answer the question, and always in context. Don't over-think it.
- Remind students that sample size is important, but it isn't the only concern. Using a biased sampling method with a larger sample size makes it *less* likely that the parameter value is within the margin of error of the estimate. Also, make sure students understand the difference between accuracy and precision.
- Don't allow your students to use phrases with statistical meaning in a casual way.
- Make sure students understand when they should address the conditions for inference and why these conditions need to be checked.

Question 5
Car purchase price for men vs women
Paired t-test

Question #5

A researcher conducted a study to investigate whether local car dealers tend to charge women more than men for the same car model. Using information from the county tax collector's records, the researcher randomly selected one man and one woman from among everyone who had purchased the same model of an identically equipped car from the same dealer. The process was repeated for a total of 8 randomly selected car models. The purchase prices and the differences (woman – man) are shown in the table *on the next page*. Summary statistics are also shown.

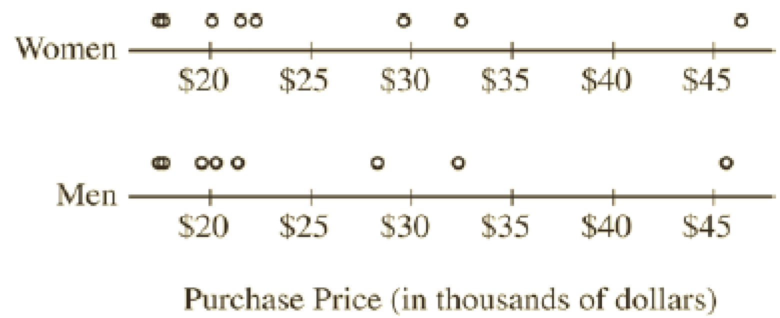
Do the data provide convincing evidence that, on average, women pay more than men in the county for the same car model?

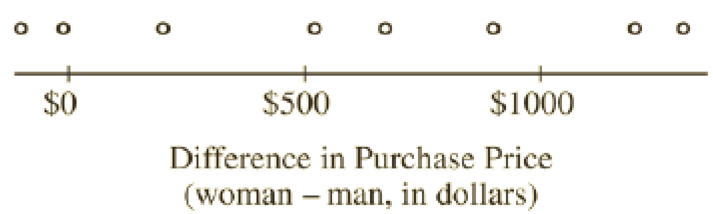
Question 5, continued

Car model	1	2	3	4	5	6	7	8
Women	$20,100	$17,400	$22,300	$32,500	$17,710	$21,500	$29,600	$46,300
Men	$19,580	$17,500	$21,400	$32,300	$17,720	$20,300	$28,300	$45,630
Difference	$520	–$100	$900	$200	–$10	$1,200	$1,300	$670

	Mean	Standard Deviation
Women	$25,926.25	$9,846.61
Men	$25,341.25	$9,728.60
Difference	$585.00	$530.71

Dotplots of the data and the differences are shown below.





Question 5 Solution

Notice that the *data are paired* because each car model in the study includes one purchased by a man and one purchased by a woman.

Step 1: States a correct pair of hypotheses

Let μ_{Diff} represent the population mean difference in purchase price (woman – man) for identically equipped cars of the same model, sold to both men and women by the same dealer, in this county. The hypotheses to be tested are H_0: $\mu_{Diff} = 0$ versus

H_a: $\mu_{Diff} > 0$.

Step 2: Identifies a correct test procedure (by name or by formula) and checks appropriate conditions.

- Paired *t*-test
- Randomly selected sample (explain)
- Small *n*, so check dotplot for extreme skew and/or outliers – OK

Question 5 Solution, continued

Step 3: Correct mechanics, including the value of the test statistic and p-value (or rejection region)

$$t = \frac{585 - 0}{\frac{530.71}{\sqrt{8}}} \approx 3.12$$

The p-value, based on a t-distribution with df = 8 – 1 = 7 is 0.008.

Step 4: States a correct conclusion in the context of the study, using the result of the statistical test.

Because the *p*-value is very small (for instance, smaller than α = 0.05), we reject the null hypothesis. The data provide convincing evidence that, on average, women pay more than men in the county for the same car model.

Question 5 Common Student Errors

- Not doing a test at all—just doing descriptive analysis
- Doing a two-sample t-test
- Defining parameter incorrectly
- Verifying conditions: It's not clear that students understand what they are trying to verify, or why
- Using z procedures with unknown σ
- Incorrect formulas for test statistic; improper notation
- Accepting or retaining the null; proving the alternative
- Adding a p-value interpretation to the conclusion and doing it wrong

Question 5 Teaching Tips

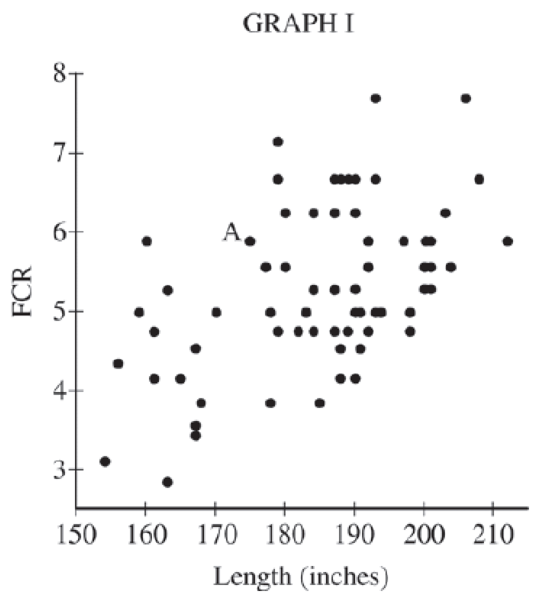
- Practice identifying the appropriate inference procedure for a given research question. Stress that design determines analysis.
- Advice to students: Don't waste time typing data into your calculator when summary statistics are provided
- Have students use the same inference framework every time
- Don't teach inference for a mean using *z* procedures
- Provide extra information on some questions so that students have practice filtering out what's important/relevant
- Stress that there are only two decisions one can make about the null hypothesis based on the data: reject H_0 or fail to reject H_0.
- Insist on a conclusion *in context* that addresses strength of evidence for H_a and that *includes the parameter(s) of interest*.

Question 6
Fuel consumption rate

Residuals, choosing a new variable for prediction

Question #6, parts (a) and (b)

Jamal is researching the characteristics of a car that might be useful in predicting the fuel consumption rate (FCR); that is, the number of gallons of gasoline that the car requires to travel 100 miles under conditions of typical city driving. The length of a car is one explanatory variable that can be used to predict FCR. Graph I is a scatter plot showing the lengths of 66 cars plotted with the corresponding FCR. One point on the graph is labeled A.



Jamal examined the scatterplot and determined that a linear model would be a reasonable way to express the relationship between FCR and length. A computer output from a linear regression is shown below. [*To the right.*]

Linear Fit

FCR = −1.595789 + 0.0372614 * Length

Summary of Fit

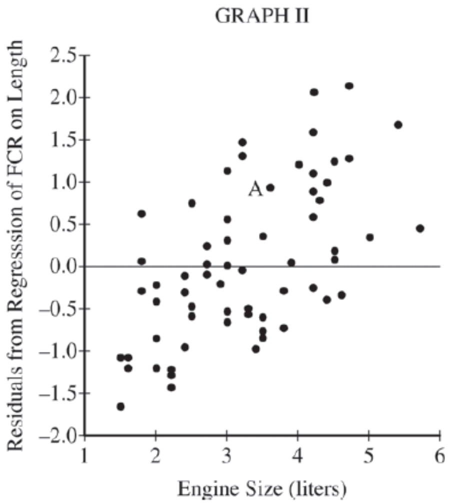
RSquare	0.250401
Root Mean Square Error	0.902382
Observations	66

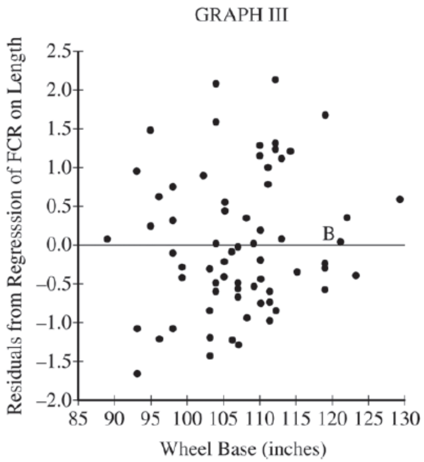
Question 6ab, continued

(a) The point on the graph labeled A represents one car of length 175 inches and an FCR of 5.88. Calculate and interpret the residual for the car relative to the least squares regression line.

Jamal knows that it is possible to predict a response variable using more than one explanatory variable. He wants to see if he can improve the original model of predicting FCR from length by including a second explanatory variable in addition to length. He is considering including engine size, in liters, or wheel base (the length between axles), in inches. Graph II is a scatterplot showing the engine size of the 66 cars plotted with the corresponding residuals from the regression of FCR on length. Graph III is a scatterplot showing the wheel base of the 66 cars plotted with the corresponding residuals from the regression of FCR on length.

Question #6ab, continued





(b) In graph II, the point labeled A corresponds to the same car whose point was labeled A in graph I. The measurements for the car represented by point A are given below.

FCR	Length (inches)	Engine Size (liters)	Wheel Base (inches)
5.88	175	3.6	93

(i) Circle the point on graph III that corresponds to the car represented by point A on graphs I and II.
(ii) There is a point on graph III labeled B. It is very close to the horizontal line at 0. What does that indicate about the FCR of the car represented by point B?

Question 6ab Solution

(a) For a car with length 175 inches, the predicted value for the car's FCR, based on the least squares regression line, is

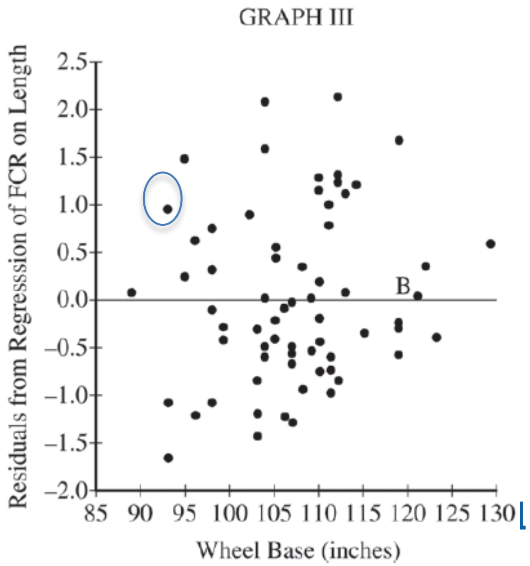
Predicted FCR = -1.595789 + 0.0372614 (175) ≈ 4.92 gallons per 100 miles

The actual FCR for the car is 5.88, so the residual is 5.88 – 4.92 = 0.96.

The residual value means that the car's FCR is 0.96 gallons per 100 miles greater than would be predicted for a car of its length.

(b) (i) See circled point.

(b) (ii) Point B corresponds to a car with an actual FCR that is very close to the FCR that would be predicted for a car with its length.



Question 6ab, Common Student Errors

Part (a) Successes and Mistakes:

- Many students successfully calculated and interpreted the residual value.
- Context was seldom missing in the interpretation.
- "***away from the predicted value***" was not good enough – need direction.
- Magnitude (.96) was missing in some students' interpretation.

Part (b) Successes and Mistakes:

- Almost all students circled the point correctly.
- It was important that the response described the FCR of the car corresponding to the point B being close the predicted FCR using the regression of FCR based on length (***not wheel base – a major mistake***).

Question 6, parts c, d

(c) Write a few sentences to compare the association between the variables in graph II with the association between the variables in graph III.

(d) Jamal wants to predict FCR using length and one of the other variables, engine size or wheel base. Based on your response to part (c), which variable, engine size or wheel base, should Jamal use in addition to length if he wants to improve the prediction? Explain why you chose that variable.

Question 6 Solution, parts c, d

(c) Graph II reveals an association that is positive and moderately linear. In contrast, there is a weak positive, or no, linear association in graph III. The association between engine size and residual (from predicting FCR based on length) is stronger than the association between wheel base and residual (from predicting FCR based on length).

(d) Engine size is a better choice than wheel base for including with length in a regression model for predicting FCR. The stronger association between engine size and residual (from predicting FCR based on length) indicates that engine size is more useful than wheel base for reducing variability in FCR values that remains unexplained (as indicated by residuals) after predicting FCR based on length.

Question 6cd, Common Student Errors

Part (c) Successes and Mistakes:

- Many students misinterpreted the two graphs as residual plots from the regression of FCR based on engine size and wheel base instead of the residuals of FCR based on length plotted against variables engine size and wheel base.
- This misinterpretation invariably led to the incorrect choice of wheel base in part (d).
- Scoring on part (c) was focused on the comparisons of the associations in the graphs (as directed by the stem of question). However, an incorrect description of the variables and/or the graphs was unacceptable.
- For graph II, it was necessary to use the word "linear" when describing the form. "Correlation" is not synonymous with "linear"

Question 6 Common Student Errors, cont.

Part (d) Successes and Mistakes:

- Misinterpretation of the graphs as residual plots invariably led to the incorrect choice of wheel base in part (d).
- It was difficult for students to justify the choice with the correct description of the associations of the appropriate variables, and an incorrect description of the variables and/or the graphs was unacceptable in the justification of the choice.
- Most justifications of the choice engine size did not include the reduction of variability remaining in the residuals from the regression of FCR based on length.
- This component was the main thrust of the investigative task and therefore played a large role in holistic grading.

Question 6 Teaching Tips

- In general, read the question carefully. Especially in the investigative task (Question 6) don’t assume you are seeing something familiar (in this case, a residual plot).
- For the investigative task, remind students that they are probably going to be asked to do something standard, and then to think about something they have not seen before.
- Practice by going over previous investigative tasks, and pointing out where the “investigative” part starts.

Score Summary

Q #	1	2	3	4	5	6
Mean score (Possible 0 to 4)	1.96	1.61	1.43	1.79	1.10	1.29

Exam Score	1	2	3	4	5
%	22.7%	17.9%	24.46%	20.91%	14.0%

Final Notes:

- Consider becoming an AP Statistics Reader!

  http://apcentral.collegeboard.com
- Questions/Comments: jutts@uci.edu