Results from the 2017 AP Statistics Exam

Jessica Utts,
University of California, Irvine
Chief Reader, AP Statistics
jutts@uci.edu

July 28, 2017

CollegeBoard AP

AP Annual Conference 2017

Plan for each of the six questions

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

Question #1: Regression for length and weight of gray wolves

Explain positive, linear, strong; interpret slope; find weight given length & residual

Question 1

Researchers studying a pack of gray wolves in North America collected data on the length *x*, in meters, from nose to tip of tail, and weight *y*, in kilograms, of the wolves. A scatterplot of weight versus length revealed a relationship between the two variables described as positive, linear, and strong.

Question 1 – Part (a)

a) For the situation described above, explain what is meant by each of the following words.

i. Positive:

ii. Linear:

iii. Strong:

Solution Part (a):

In the context of a scatter plot of

y = weight and x = length:

A <u>positive</u> relationship means that wolves with higher values of length also tend to have higher weights.

A <u>linear</u> relationship means that when length increases by one meter, weight tends to change by a constant amount, on average.

A <u>strong</u> relationship means that the data points fall close to a line (or curve).

Common Student Errors, Q1(a)

- Defining a positive relationship by simply saying that there is a positive correlation.
- Using "correlation" to define a linear relationship. A correlation coefficient is a measure of the strength of a linear relationship, but it does not by itself explain the meaning of a linear relationship. It is more appropriate to use correlation to discuss a strong relationship.
- Indicating that a relationship is strong when points in the scatterplot are close together or not too scattered. The response should indicate that a relationship is strong when the points in the scatterplot are close to a line, or more generally, a curve.

Question 1 – Part (b)

The data collected from the wolves were used to create the least-squares equation $\hat{y}$ = -16.46 + 35.02x

b) Interpret the meaning of the slope of the least-squares regression line in context

Solution Part (b):

The slope of 35.02 indicates that two wolves that differ by one meter in length are predicted to differ by 35.02 kilograms in weight, with the longer wolf having the higher weight.

Common Student Errors, Q1(b)

- Implying that the slope of a least squares regression line corresponds to an exact relationship between changes in observed values of y as x changes.
- Failure to link the increase in the predicted response to an increase of a specific size in the explanatory variable. For instance, an unacceptable response is "For any change in length, the predicted weight increases by 35.02 kg."

Question 1 – Part (c)

[The data collected from the wolves were used to create the least-squares equation $\hat{y}$ = -16.46 + 35.02x]

c) One wolf in the pack with a length of 1.4 meters had a residual of -9.67 kilograms. What was the weight of the wolf?

Solution Part (c):

In general, residual = actual weight – predicted weight, or equivalently, actual weight = predicted weight + residual. In this situation:

Predicted weight = –16. 46 + 35.02(1.4) = 32.568 kg

Actual weight = predicted weight + residual

So y = 32.568 – 9.67 = 22.9 kilograms.

Common Student Errors, Q1(c)

- In calculating the actual weight from the regression equation for a specific x when a residual is given, many responses stopped after computing the predicted value. They most likely forgot the question!
- Some responses incorrectly replaced the residual with the intercept in the calculation of the actual weight, or the intercept with the residual in the calculation of the predicted weight.

Q1 Teaching tips - Encourage students to:

- Read the problem
- Use better handwriting
- Focus on the question, keep thoughts clear and concise
- Use graphs for clarification (in part a), but include words
- Complete calculations—follow problem to the end with work
- Avoid sloppy terminology
  - (“the scatterplot is straight” vs. “the points on the scatterplot…”)
- Distinguish between statistical use of words like correlation and ordinary English use. For example, “r close to 1” or “correlation coefficient” clearly indicate statistical use.

Q1 Teaching tips, continued

- Things to reinforce
  - Don't use deterministic wording in slope interpretation (include "predicted" or "estimated")
  - Slope is for a 1 unit change in x, not just a "change"
  - "Strength" is a measure of how closely the data follow a pattern (line or curve), not how close the points are to each other
  - Small residuals can be an indication of the strength of the model
- Use "backwards" questions in class
  - Write the regression equation associated with the following interpretation…
  - Explain what is meant when you have high correlation…

Question #2:
Asking for a water cup but filling it with soda

Construct and interpret a confidence interval for a proportion; use it to get a confidence interval for the cost to the restaurant

Question 2 Part (a):

The manager of a local fast-food restaurant is concerned about customers who ask for a water cup when placing an order but fill the cup with a soft drink from the beverage fountain instead of filling the cup with water. The manager selected a random sample of 80 customers who asked for a water cup when placing an order and found that 23 of those customers filled the cup with a soft drink from the beverage fountain.

(a) Construct and interpret a 95 percent confidence interval for the proportion of all customers who, having asked for a water cup when placing an order, will fill the cup with a soft drink from the beverage fountain.

Solution Part (a) (Step 1 of 3):

Step 1: Identify the appropriate confidence interval by name or formula and check appropriate conditions.

The appropriate procedure is a one-sample *z*-interval for a population proportion *p*. In this case, the population is all customers of this restaurant who ask for a water cup, and *p* is the proportion of that population that will fill the cup with a soft drink. The appropriate formula is

$$\hat{p} \pm z^{*} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Solution Part (a) (Step 1 of 3, continued):

Step 1:

Conditions:

1. Random sample
2. Large sample (number of successes $n\hat{p} \geq 10$ and number of failures $n(1-\hat{p}) \geq 10$)

For condition 1, the stem of the problem states that a random sample of customers who asked for a water cup was used.

For condition 2, the number of "successes" (filled cup with soft drink) is 23 and the number of "failures" is 57, both of which are greater than 10.

Solution Part (a) (Step 2 of 3):

Step 2: Correct mechanics

The sample proportion is $\hat{p} = \frac{23}{80} = 0.2875$. The confidence interval is:

$$0.2875 \pm 1.96\sqrt{\frac{0.2875\,(1-0.2875)}{80}}$$

= 0.2875 ± 1.96(0.0506)
= 0.2875 ± 0.0992
= 0.1883 to 0.3867

Solution Part (a) (Step 3 of 3):

Step 3: Interpretation

We can be 95% confident that in the population of all customers at this fast-food restaurant who ask for a water cup, the proportion that will fill it with a soft drink is between 0.1883 and 0.3867.

Common Student Errors, Q2(a)

- Identification of procedure missing or incorrect
- Not checking conditions
- Omitting the large sample condition or verifying only one of the two inequalities
- Mislabeling conditions, e.g. “Independence” for the sample size condition
- Vague reference to a normal distribution
- Stating that the sample or population has a normal distribution (for a categorical variable)
- Inappropriate large sample condition: $n \geq 30$
- Using incorrect critical value (wrong z^* or a t^*) or using t and showing df
- Using an incorrect formula for the standard error of a sample proportion

Common Student Errors, Q2(a), continued

- Supporting work that doesn’t match the confidence interval given
- Using incorrect or nonstandard notation without defining it
- Calculating an unreasonable interval for a proportion—not realizing that a proportion must be between 0 and 1
- Not recognizing that the population of interest is all customers who ask for a water cup at this fast-food restaurant
- Interpreting confidence level instead of confidence interval, or attempting to blend the interpretation of the interval with the interpretation of the confidence level in a single sentence, or providing interpretations of both confidence interval and confidence level
- Saying “mean” in the interpretation of the interval

Question 2 Part (b):

(b) The manager estimates that each customer who asks for a water cup but fills it with a soft drink costs the restaurant \$0.25. Suppose that in the month of June 3000 customers ask for a water cup when placing an order. Use the confidence interval constructed in part (a) to give an interval estimate for the cost to the restaurant for the month of June from the customers who ask for a water cup but fill the cup with a soft drink.

Solution Part (b):

Using the confidence interval in Part (a), a 95% interval estimate for the number of customers in June who ask for a water cup but then filled it with a soft drink is 3000×0.1883 to 3000×0.3867, or 565 to 1160. At a cost of \$0.25 per customer, a 95% interval estimate for the cost to the restaurant in June is \$141.25 to \$290.00.

Common Student Errors, Q2(b)

- Calculating a single value (point estimate) in part (b) rather than an interval
- Not using the interval from part (a) as directed
- Not showing work in part (b)

Q2 Teaching tips

- In inference questions, ask students to identify the population and parameter of interest. Encourage students to use the language in the stem of the question when defining the parameter.
- Discuss <u>why</u> each condition is being checked for an inference procedure as well as how to check the condition. Use applets and hands-on activities to demonstrate what happens when each condition isn’t met.
- Insist on proper notation throughout the course—refer students to the formula sheet

Q2 Teaching tips, continued

- Emphasize the difference between interpreting a confidence interval and a confidence level. Use hands-on activities and applets involving repeated sampling to illustrate the idea of a confidence level interpretation.
- Give students ample practice in distinguishing categorical variables/data (proportions) from quantitative variables/data (means).
- Have students make a summary chart of inference procedures with the appropriate names, conditions, and formulas for each.

Question #3:

Melons from organic and nonorganic farms

Calculate a normal probability, a weighted probability and a conditional probability

Question 3

A grocery store purchases melons from two distributors, J and K. Distributor J provides melons from organic farms. The distribution of the diameters of the melons from Distributor J is approximately normal with mean 133 millimeters (mm) and standard deviation 5 mm.

(a) For a melon selected at random from Distributor J, what is the probability that the melon will have a diameter greater than 137 mm?

Solution – Part (a)

Let X denote the diameter of a randomly selected melon from Distributor J. X has an approximately normal distribution with mean 133 mm and standard deviation 5 mm.

The z-score for a diameter of 137 mm is

$$z = \frac{137 - 133}{5} = \frac{4}{5} = 0.8.$$

Therefore $P(X > 137) = P(Z > 0.8) = 1 - 0.7881 = 0.2119$

Common Student Errors, Q3(a)

- Thinking that a normal probability problem can be solved using the t-distribution
- Thinking that a normal probability problem always involves using x-bar.
- Thinking that normalcdf(137, 1000000, 133, 5) “shows work;” i.e. identifies parameters and boundary conditions.
- Not knowing the difference between notation for parameters and statistics
- Thinking that you need to adjust 137 as if the normal random variable is discrete and using 137.5 or 138 in place of 137
- Trying to use the Empirical Rule and interpolate probabilities for z values between two whole numbers (0, 1, 2, 3)

Question 3, Part (b)

Distributor K provides melons from nonorganic farms. The probability is 0.8413 that a melon selected at random from Distributor K will have a diameter greater than 137 mm. For all melons at the grocery store, 70 percent of the melons are provided by Distributor J and 30 percent are provided by Distributor K.

(b) For a melon selected at random from the grocery store, what is the probability that the melon will have a diameter greater than 137 mm?

Solution – Part (b)

- Define events:

  J: melon is from Distributor J

  K: melon is from Distributor K

  G: melon diameter is greater than 137 mm

For a randomly selected melon from the grocery store,

$$P(G) = P(G \mid J) \times P(J) + P(G \mid K) \times P(K) = (0.2119)(0.7) + (0.8413)(0.3)$$

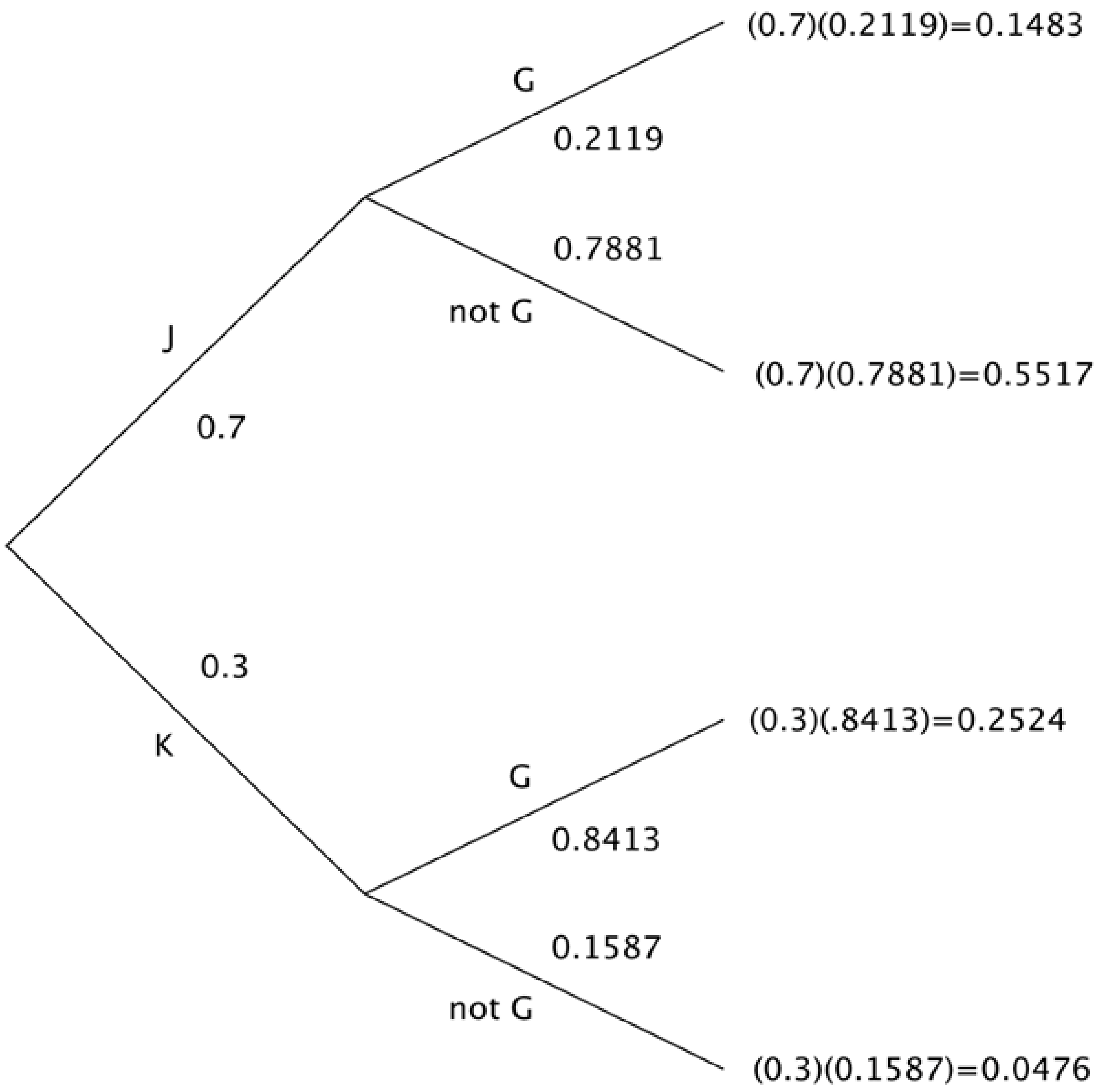
which is 0.1483 +0.2524 = 0.4007

Solution to Part (b) using a tree diagram

$$P(G) = P(G \text{ and } J) + P(G \text{ and } K)$$

$= 0.1483 + 0.2524$

$= 0.4007$



Common Student Errors, Q3(b)

- Not being able to properly create a tree diagram, such as putting joint instead of conditional probabilities on the 2nd set of branches.
- Not recognizing that two events are mutually exclusive, particularly the ends of the branches of a tree diagram.
- Not being able to find the appropriate probabilities from a tree diagram. For example, thinking the conditional branch on the tree is actually the intersection probability.
- Generally not knowing how much work is needed to justify probability calculations. It should be clear where each number in the response came from.

Question 3, Part (c)

(c) Given that a melon selected at random from the grocery store has a diameter greater than 137 mm, what is the probability that the melon will be from Distributor J?

Solution, Part (c)

Using the events defined in part (b), the requested probability is

$$P(J \mid G) = \frac{P(J \text{ and } G)}{P(G)} = \frac{P(G \mid J)P(J)}{P(G)}$$

which is $$\frac{(0.2119)(0.7)}{0.4007} = \frac{0.1483}{0.4007} = 0.3701$$

Common Student Errors, Q3(c)

- Not being able to distinguish between a question that is asking for a conditional probability and one that is asking for an intersection probability; i.e, ignoring the "given" in a question. In this context, computing

  P(Greater than 137mm *and* Distributor J)

  rather than

  P(Distributor J | Greater than 137mm)

- Not recognizing that a calculation from earlier work could be used in subsequent calculations; in this case, not recognizing that P(G) from Part (b) can be used as the denominator of Part (c).

Q3 Teaching tips

- In solving problems, model good behavior by showing all work/steps in a probability problem (and inference) all the time.
- Emphasize that the Empirical Rule gives only approximations of normal values and only for 1, 2, and 3 standard deviations from the mean; any sort of interpolation will give incorrect answers.
- When introducing Student's t distribution, emphasize that the only time the z distribution and t distribution are the same is when the t distribution is based on an infinite number of degrees of freedom.
- In general, it is not a good idea to use "calculator speak" in answering any question.

Q3 Teaching tips

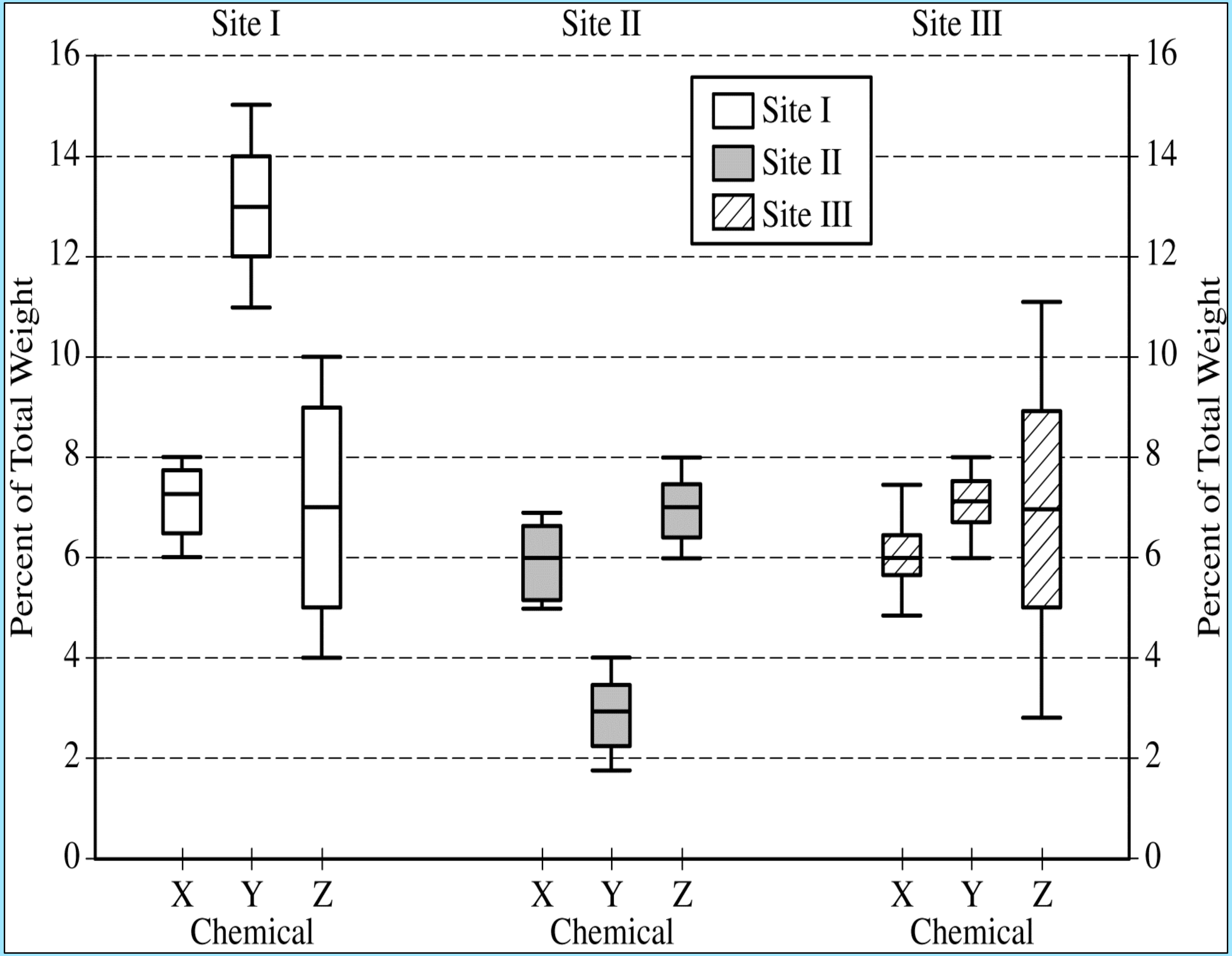
- Show the students problems where there are multiple parts and the answers for the later problems depend on the results from the earlier parts.
- If any continuous approximation of a discrete random variable is taught, explain as much as possible the reason for any adjustment or continuity correction, and that they apply to discrete variables only.
- Give students lots of practice with creating and using tree diagrams to solve word problems. Try "working backwards" by giving a tree diagram and asking them to answer probability problems using it.
- Give lots of examples distinguishing conditional and joint probabilities. Have students practice setting up probability statements for word problems, even if they don't solve them.

Question #4: Chemical analysis of clay in pottery to assess origin

Compare boxplots; answer contextual questions based on information in boxplots

Question 4

The chemicals in clay used to make pottery can differ depending on the geographical region where the clay originated. Sometimes, archaeologists use a chemical analysis of clay to help identify where a piece of pottery originated. Such an analysis measures the amount of a chemical in the clay as a percent of the total weight of the piece of pottery. The boxplots below summarize analyses done for three chemicals — X, Y, and Z — on pieces of pottery that originated at one of three sites: I, II, or III.


Site I
Site II
Site III
16
14
12
10
8
6
4
2
0
Percent of Total Weight
Site I
Site II
Site III
X
Y
Z
Chemical
X
Y
Z
Chemical
X
Y
Z
Chemical

Question 4, Part (a)

(a) For chemical Z, describe how the percents found in the pieces of pottery are similar and how they differ among the three sites.

SOLUTION to Part (a)

The median value for the percent of chemical Z in the pottery pieces is similar for all three sites, at about 7%. The ranges for the percent of chemical Z are much different for the three sites, with the smallest range being about 2% (from 6% to 8%) at Site II, a much higher range of about 6% (from about 4% to 10%) at Site I and the largest range of about 8% (from about 3% to 11%) at Site III.

Common Student Errors, Q4(a)

- Some students stated that symmetric boxplots indicate that the distribution is normal.
- Some students described many attributes of the boxplots for Chemical Z at each site but never clearly stated what is similar and what is different. This is a “Laundry List.”
- Some students referred to the interval from the minimum to the maximum as the range.
- Some students compared Chemical Z to Chemicals X and Y *within* each site rather than comparing Chemical Z to itself *across* sites.

Question 4, Part (b-i)

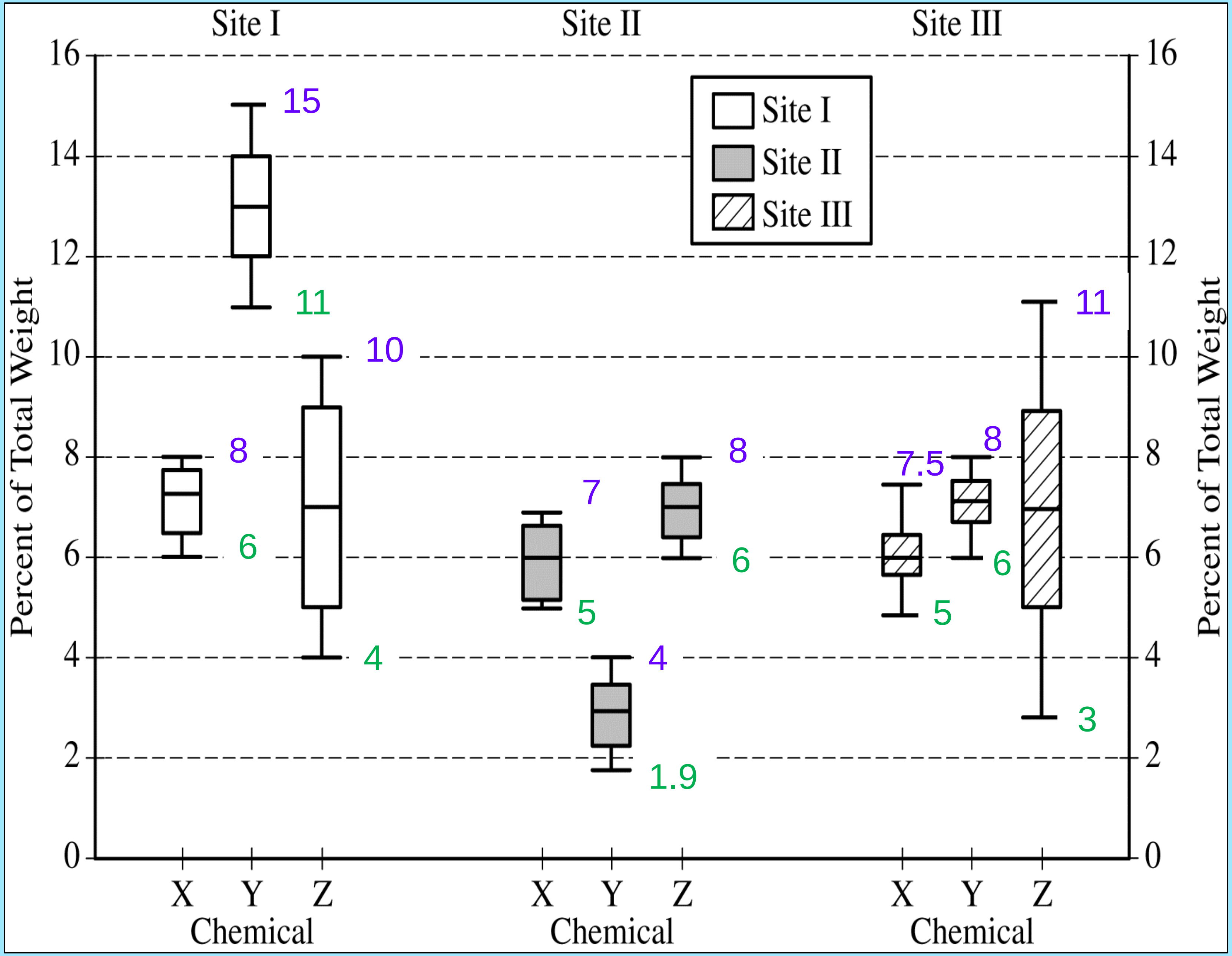
Consider a piece of pottery known to have originated at one of the three sites, but the actual site is not known.

(i) Suppose an analysis of the clay reveals that the sum of the percents of the three chemicals X, Y, and Z is 20.5%. Based on the boxplots, which site—I, II, or III—is the most likely site where the piece of pottery originated? Justify your choice.

Question 4 b-i Solution (sum of 3 chemicals)

The piece most likely originated at Site III. Although values outside of the range of data observed in the samples would be possible, using the available data results in approximate minimum and maximum sums of the percents for the 3 chemicals as shown in the table below. The only site that includes 20.5 between the sums of the minimum and maximum values is Site III.

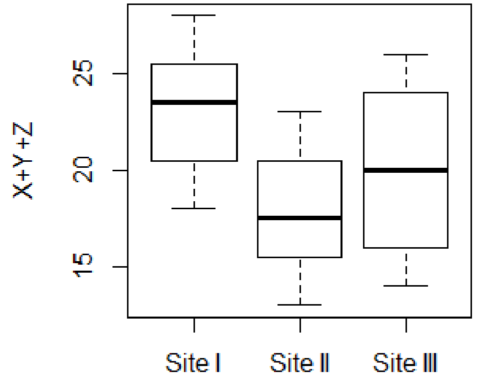
	Site I		Site II		Site III	
Chemical	Min	Max	Min	Max	Min	Max
X	6	8	5	7	5	7.5
Y	11	15	1.9	4	6	8
Z	4	10	6	8	3	11
Sum	21	33	12.9	19	14	26.5

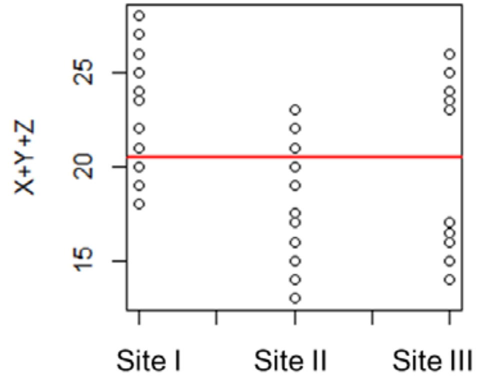


Site I
Site II
Site III
Site I
Site II
Site III
16
14
12
10
8
6
4
2
0
Percent of Total Weight
15
11
10
8
6
4
7
5
4
1.9
8
6
7.5
8
5
6
11
3
X Y Z
Chemical
X Y Z
Chemical
X Y Z
Chemical

Common Student Errors, Q4(b-i)

- Some students chose site III based on the sums of the medians instead of the sums of the minimums and maximums. An example of why that's not enough – sum of medians is 20.5; Site III is *least* likely:





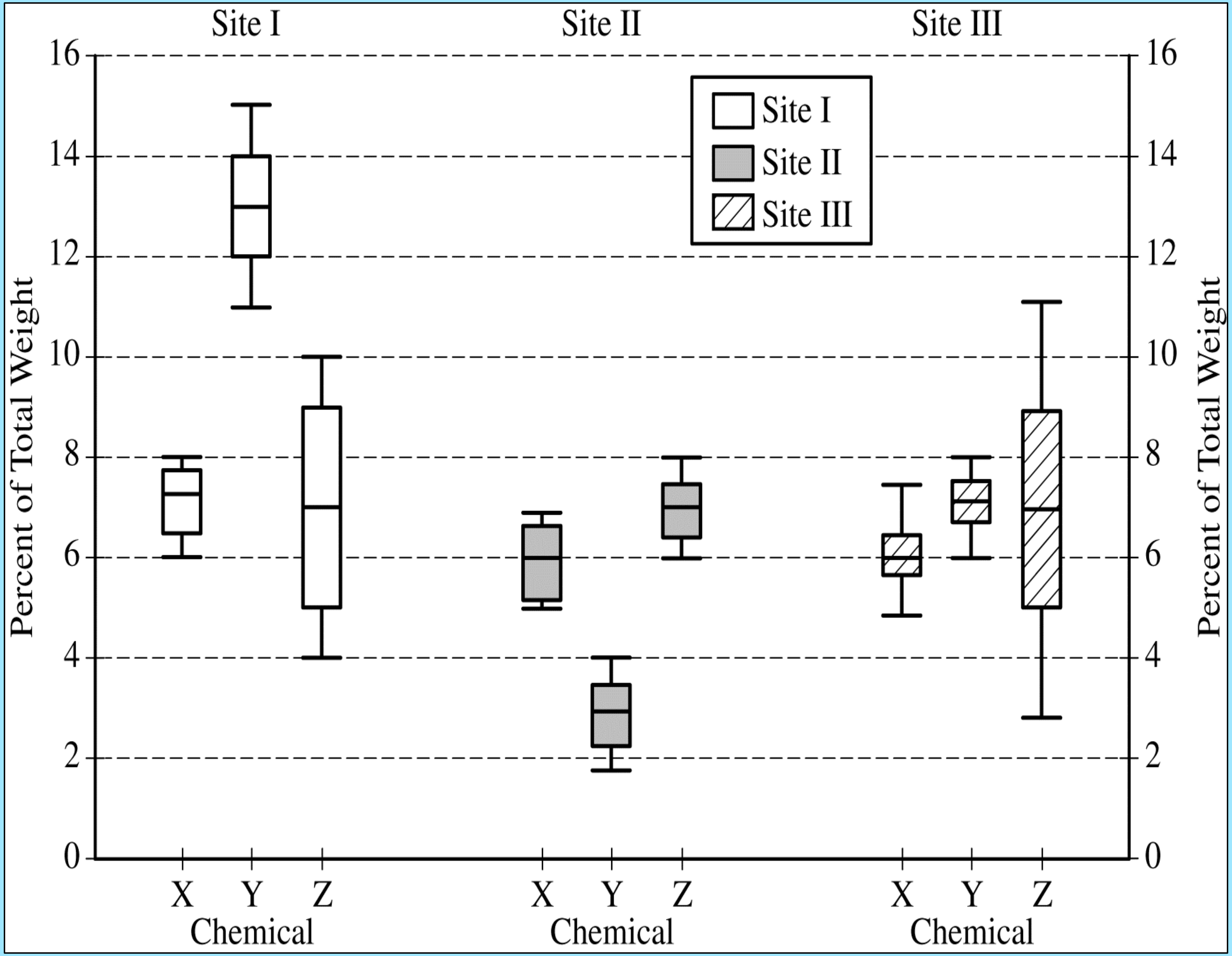
- Some students correctly chose site III but did not state why sites I and II were not the best choices.

Question 4, Part (b-ii)

[Consider a piece of pottery known to have originated at one of the three sites, but the actual site is not known.]

(ii) Suppose only one chemical could be analyzed in the piece of pottery. Which chemical - X, Y, or Z - would be the most useful in identifying the site where the piece of pottery originated? Justify your choice

SOLUTION (See next slide for boxplot picture):
Chemical Y would be most useful, because the distribution of the percentages of total weights at the three sites do not overlap. The distributions of Chemicals X and Z have substantial overlap.


Site I
Site II
Site III
Site I
Site II
Site III
Percent of Total Weight
Percent of Total Weight
0
2
4
6
8
10
12
14
16
X
Y
Z
Chemical
X
Y
Z
Chemical
X
Y
Z
Chemical

Common Student Errors, Q4(b-ii)

- Some students had difficulty clearly explaining that the boxplots do not overlap. In order to do this, both a difference in location and small variability had to be addressed. Some students described the boxplots as having different means, different variability or simply said that they have different boxplots. None of these descriptors indicate “no overlap”.

Q4 Teaching tips

- Read the question carefully and answer (only) the question asked. If the question asks for a similarity and a difference, clearly organize and state what is similar *and* what is different.
- When asked to identify one choice among three, complete justification includes reasoning for that particular choice and <u>also</u> rationale for not choosing the other options.
- Use clear communication within each part of the question. Do not assume that the reader will look back. Instead, incorporate previous work into the answer.

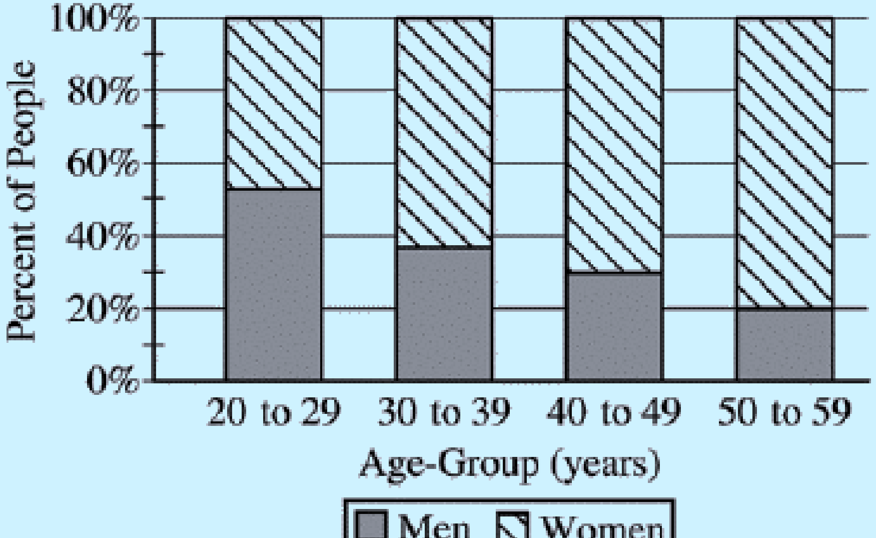
Question #5:
Gender and age at diagnosis of schizophrenia

Chi-square test for independence

Question 5

The table and the bar chart below summarize the age at diagnosis, in years, for a random sample of 207 men and women currently being treated for schizophrenia.

	Age-Group (years)				
	20 to 29	30 to 39	40 to 49	50 to 59	Total
Women	46	40	21	12	119
Men	53	23	9	3	88
Total	99	63	30	15	207



Question 5

Do the data provide convincing statistical evidence of an association between age-group and gender in the diagnosis of schizophrenia?

SOLUTION, Step 1

Step 1: State a correct pair of hypotheses.

The null hypothesis is that age group at diagnosis and gender are independent (that is, they are not associated) for the population of people currently being treated for schizophrenia.

The alternative hypothesis is that age group at diagnosis and gender are not independent for the population of people currently being treated for schizophrenia.

SOLUTION, Step 2:

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

The appropriate test is a chi-square test of independence.

The conditions for this test are satisfied because:

1. The question states that the sample was randomly selected.
2. The expected counts for all 8 cells of the table are at least 5, as seen in the following table, with expected counts shown below observed counts:

	Age at Diagnosis				
	20 to 29	30 to 39	40 to 49	50 to 59	Total
Women	46	40	21	12	119
	56.91	36.22	17.25	8.62	
Men	53	23	9	3	88
	42.09	26.78	12.75	6.38	

SOLUTION, Step 3:

Step 3: Find the value of the test statistic and the p-value

The test statistic is calculated as $\chi^2 = \sum \frac{(O-E)^2}{E}$, or:

$\chi^2 = 2.093 + 0.395 + 0.817 + 1.322 + 2.830 + 0.534 + 1.105 + 1.788$
$= 10.884$

The p-value is $P\left(\chi^2 \geq 10.884\right) = 0.012$, based on (4-1)x(2-1)= 3 degrees of freedom.

SOLUTION, Step 4:

Step 4: State the conclusion in context, with linkage to p-value

Because the p-value is very small (for instance much smaller than $\alpha = 0.05$), we would reject the null hypothesis and conclude that the sample data provide strong evidence that there is an association between age group at diagnosis and gender for the population currently being treated for schizophrenia.

Common Student Errors, Q5

- Using the idea of sufficient evidence (given in the stem of the problem) to state the hypotheses: “H0: there is sufficient evidence of no association” and “Ha: there is sufficient evidence of an association.”
- Listing incorrect conditions like n > 30, or “both samples independent”
- Stating the condition “expected counts > 5” but not verifying by computing them.
- Writing statistical conclusions as definitive statements “we conclude …” or “we prove”
- Stating conclusions about the sample data using the bar graph, but not carrying out any inference.

Q5 Teaching tips

- Remind students that the null hypothesis is about the population and not the sample. Also, it may be helpful to point out that no association (independence) is synonymous with status quo/no change similar to a null hypothesis of $\mu_1=\mu_2$.
- Remind students that it is important for them to name the test they are performing. Help students distinguish between chi-square test for independence and chi-square test for homogeneity. The distinction has to do with the sampling method, whether there was one sample taken or multiple samples taken.
- Help students understand not only the correct conditions needed for hypothesis tests but also emphasize why those conditions are necessary.
- Remind students to label their work well and clearly communicate what they are doing.

Q5 Teaching tips, continued

- Emphasize how values obtained from the calculator were calculated. Ideally, students should be familiar with how to calculate values by formula and calculator.
- Emphasize that statistical conclusions are not definitive. There should be some level of uncertainty.
- Highlight the importance of good communication; students should state why they are making the decisions that they are making, and include both a decision (reject null or not?) and a conclusion (use alternative to answer the scientific question).
- Emphasize to students the difference between evidence from a sample (which can be obtained from the graph) and convincing statistical evidence (which can be obtained from a hypothesis test).

Question #6:

Coin vs chip method for randomizing

Calculate probabilities for each method; decide which is best in a given situation

Question 6

6. Consider an experiment in which two men and two women will be randomly assigned to either a **treatment group** or a **control group** in such a way that each group has two people. The people are identified as Man 1, Man 2, Woman 1, and Woman 2. The six possible arrangements are shown below.

Arrangement A	
Treatment	Control
Man 1 Man 2	Woman 1 Woman 2

Arrangement B	
Treatment	Control
Man 1 Woman 1	Man 2 Woman 2

Arrangement C	
Treatment	Control
Man 1 Woman 2	Man 2 Woman 1

Arrangement D	
Treatment	Control
Woman 1 Woman 2	Man 1 Man 2

Arrangement E	
Treatment	Control
Man 2 Woman 2	Man 1 Woman 1

Arrangement F	
Treatment	Control
Man 2 Woman 1	Man 1 Woman 2

Question continued

Two possible methods of assignment are being considered: the sequential coin flip method, as described in part (a), and the chip method, as described in part (b).

For each method, the **order of the assignment** will be Man 1, Man 2, Woman 1, Woman 2.

part (a)

a) For the sequential **coin flip method**, a fair coin is flipped until one group has two people. An outcome of tails assigns the person to the treatment group, and an outcome of heads assigns the person to the control group. **As soon as one group has two people, the remaining people are automatically assigned to the other group.**

i) **Complete the table** below by calculating the probability of each arrangement occurring if the sequential coin flip method is used.

Arrangement	A	B	C	D	E	F
Probability						

SOLUTION, Part (a-i)

(i) Let T (tail) represent being assigned to the **treatment group** and H (head) represent being assigned to the **control group** for each coin flip. The process stops when either the treatment group or the control group has two members. The outcomes and their probabilities are as follows.

Arrangement	A	B	C	D	E	F
Coin outcomes	TT	THT	THH	HH	HTH	HTT
Calculation	(1/2)(1/2)	(1/2)(1/2)(1/2)	(1/2)(1/2)(1/2)	(1/2)(1/2)	(1/2)(1/2)(1/2)	(1/2)(1/2)(1/2)
Probability	1/4	1/8	1/8	1/4	1/8	1/8

Arrangement A	
Treatment	Control
Man 1 Man 2	Woman 1 Woman 2

Arrangement B	
Treatment	Control
Man 1 Woman 1	Man 2 Woman 2

Arrangement C	
Treatment	Control
Man 1 Woman 2	Man 2 Woman 1

Arrangement D	
Treatment	Control
Woman 1 Woman 2	Man 1 Man 2

Arrangement E	
Treatment	Control
Man 2 Woman 2	Man 1 Woman 1

Arrangement F	
Treatment	Control
Man 2 Woman 1	Man 1 Woman 2

Question Part (a-ii)

ii) For the sequential coin flip, **what is the probability** that Man 1 and Man 2 are assigned to the same group?

SOLUTION Part (a-ii)

ii) Man 1 and Man 2 are assigned to the same group for **arrangements A and D**, so the probability is

$$P(A) + P(D) = 1/4 + 1/4 = \mathbf{1/2}$$

part (b)

b) For the **chip method**, two chips are marked "treatment" and two chips are marked "control." Each person selects one chip without replacement.

i) **Complete the table** below by calculating the probability of each arrangement occurring if the chip method is used.

Arrangement	A	B	C	D	E	F
Probability						

SOLUTION, Part (b-i)

(i) Let T represent being assigned to the treatment group and C represent being assigned to the control group for each chip drawn. The process stops when either the treatment group or the control group has two members. **The probabilities differ from the coin flip method because chips are drawn *without* replacement.** The outcomes and their probabilities are as follows.

Arrangement	A	B	C	D	E	F
Chip outcomes	TT	TCT	TCC	CC	CTC	CTT
Calculation	(2/4)(1/3)	(2/4)(2/3)(1/2)	(2/4)(2/3)(1/2)	(2/4)(1/3)	(2/4)(2/3)(1/2)	(2/4)(2/3)(1/2)
Probability	1/6	1/6	1/6	1/6	1/6	1/6

Arrangement A	
Treatment	Control
Man 1 Man 2	Woman 1 Woman 2

Arrangement B	
Treatment	Control
Man 1 Woman 1	Man 2 Woman 2

Arrangement C	
Treatment	Control
Man 1 Woman 2	Man 2 Woman 1

Arrangement D	
Treatment	Control
Woman 1 Woman 2	Man 1 Man 2

Arrangement E	
Treatment	Control
Man 2 Woman 2	Man 1 Woman 1

Arrangement F	
Treatment	Control
Man 2 Woman 1	Man 1 Woman 2

Part (b)

(ii) For the chip method, what is the probability that Man 1 and Man 2 are assigned to the same group?

SOLUTION:

(ii) Man 1 and Man 2 are assigned to the same group for **arrangements A and D**, so the probability is

P(A) + P(D) = 1/6 + 1/6 = **1/3**

Common Errors and Teaching Tips: parts (a) and (b)

- In part (i), although many students listed the correct probabilities in the table, the justification was often missing or incomplete. Good justifications included calculations that were labeled, such as P(A) = (1/2)(1/2) = 1/4. Incomplete justifications were missing labels or proximity to the corresponding arrangements in the table.
  - TIP: For calculations, students should always provide justification that is easy to follow.
- In part (ii), students failed to include BOTH Arrangement A and Arrangement D when asked for the probability that both men end up in the same treatment group.
  - TIP: Make sure students read the question carefully, including information in the initial stem of the question.

Common Errors and Teaching Tips: parts (a) and (b)

- In part (ii), when attempting to find the combined probability of arrangements A and D, some students incorrectly multiplied the two probabilities rather than adding them.
  - TIP: Make sure students understand the difference between P(A or D) and P(A and D).
- In part (ii), when attempting to find the combined probability of arrangements A and D, some students incorrectly subtracted $P(A)\times P(D)$ from $P(A) + P(D)$.
  - TIP: When using the addition rule, remind students that $P(A \text{ and } D) = 0$ when events are mutually exclusive.

Part (c)

Sixteen participants consisting of **10 students and 6 teachers** at an elementary school will be used for an experiment to determine lunch preference for the school population of students and teachers. As the participants enter the school cafeteria for lunch, they will be **randomly assigned** to receive one of two lunches so that 8 will receive a salad, and 8 will receive a grilled cheese sandwich. The **students will enter the cafeteria first**, and the teachers will enter next. Which method, the sequential coin flip method or the chip method, should be used to assign the treatments? Justify your choice.

SOLUTION, Part (c)

Use the **chip method** (from part b). The chip method gives **equal probability** to all possible arrangements, but the **coin method does not**, as shown in the tables from parts (a-i) and (b-i). Furthermore, the coin method is more likely to result in **imbalanced treatment groups** with regard to students and teachers, based on the probabilities in parts (a-ii) and (b-ii). If **food preferences** for teachers are **different** than for students, this imbalance is a problem. For example, if one treatment group consists entirely of students, it would be impossible to know if a difference in the response variable is due to the treatment (type of meal) or the role of the person at the school (teacher or student).

Common Errors and Teaching Tips: part (c)

- Although many students said it would be good to have equally likely arrangements or to avoid imbalanced treatment groups, almost no students explained *why* this is important. Having one treatment group with mostly students and the other treatment group with mostly teachers is a problem only if teachers and students have different food preferences.
  - TIP: Suggest that students ask “so what?” or “why does this matter?” at the end of an explanation, especially when choosing a data collection method. Correctly addressing why the choice matters is often the difference between a substantial response and a complete response.

Common Errors and Teaching Tips: part (c)

- Some students included only a benefit of the chip method or only a drawback of the coin method, but not both.
  - TIP: When asked to make a choice between options, make sure students explain why they are choosing what they are choosing, and why they are not choosing what they are not choosing.
- Some students never explicitly made a choice, even though they included correct comments about the chip and coin methods.
  - TIP: Make sure students answer the question!

Question averages

1. 1.72 (wolves, regression)
2. 2.22 (water cups, C.I.)
3. 1.72 (melons, probability)
4. 1.71 (pottery, boxplots)
5. 1.51 (schizophrenia, chi-sq.)
6. 0.99 (coins/chips, random assign)

Thank You.

A pdf version of these slides will be posted on my homepage within the next week. Scroll to “Representative presentations.”

http://www.ics.uci.edu/~jutts/ (or search my name to easily find it)

Presentations from 2014 to 2016 are there already.