**AP Statistics**

College Board 1-day workshop

December 12, 2016

Livonia, Michigan

Luke Wilcox

East Kentwood High School.

Kentwood, Michigan

[luke.wilcox@kentwoodps.org](mailto:luke.wilcox@kentwoodps.org)

Resources from today’s training are at www.thestatsmedic.com/livonia2016

Agenda for today:

1. Introductions
2. General Course Information and Resources
3. Activity 🡪 AP Exam Question
   * Hershey’s Kiss Activity 🡪 2016 AP Exam Question #5
   * M&Ms Activity 🡪 2016 AP Exam Question #2
4. Bonus Activities
   * (1) Can you guess my IQ? (Interpretation of r2)
   * (2) Does Beyonce write her own lyrics? (Sampling distribution, bias, variability)
   * (3) Which version of the Exam is harder? (Two sample t-test for means)
5. Additional Activities and Resources

Goals for today:

* Participants will be exposed to relevant resources and instructional strategies that can enhance the quality of their AP Statistics course.
* Participants will actively participate in activities that develop deeper understanding of statistical concepts.
* Participants will learn about the format, content, rubric, and grading of the AP Statistics Exam.
* Participants will have a better understanding of statistical inference.

General Course Information and Resources:

* AP Statistics Content Specifications:
  + <http://media.collegeboard.com/digitalServices/pdf/ap/ap-statistics-course-description.pdf>
* Available textbooks:
  + The Practice of Statistics 5e – Starnes, Tabor, Yates, Moore
  + Stats Modeling the World – Bock, Velleman, De Veaux
  + Introduction to Statistics and Data Analysis – Peck, Olsen, Devore
  + Statistics in Action – Watkins, Scheaffer, Cobb
* AP Course Audit
  + <http://www.collegeboard.com/html/apcourseaudit/courses/statistics.html>
* About the AP Exam and College Board Website Resources
  + <http://apcentral.collegeboard.com/apc/members/exam/exam_information/8357.html>
* Your College Board Accounts
  + <https://scores.collegeboard.org/pawra/home.action>
    - Online scores and instructional planning report
  + <https://account.collegeboard.org/login/login?appId=205>
    - Secured exams.
* AP Exam Questions by Topic – Josh Tabor

College Board Equity and Access

The College Board strongly encourages educators to make equitable access a guiding principle for their AP programs by giving all willing and academically prepared students the opportunity to participate in AP. We encourage educators to:

* + Eliminate barriers that restrict access to AP for students from ethnic, racial, and socioeconomic groups that have been traditionally underserved.
  + Make every effort to ensure their AP classes reflect the diversity of their student population.
  + Provide all students with access to academically challenging coursework **before** they enroll in AP classes

Only through a commitment to equitable preparation and access can true equity and excellence be achieved.

 **Which way will the Hershey Kiss land?**



When you toss a Hershey Kiss, it sometimes lands flat and sometimes lands on its side. What proportion of tosses will land flat?

Each group of four selects a random sample of 50 Hershey’s Kisses to bring back to their desks. Toss the 50 Kisses and then calculate the proportion that land flat. Let = the proportion of the Kisses that land flat. Do this 5 times. For each trial, record the number of Kisses that land flat and the proportion,  that land flat.

1. Trial 1: # that land flat = \_\_\_\_\_\_\_ =\_\_\_\_\_\_\_.

Trial 2: # that land flat = \_\_\_\_\_\_\_ =\_\_\_\_\_\_\_.

Trial 3: # that land flat = \_\_\_\_\_\_\_ =\_\_\_\_\_\_\_.

Trial 4: # that land flat = \_\_\_\_\_\_\_ =\_\_\_\_\_\_\_.

Trial 5: # that land flat = \_\_\_\_\_\_\_ =\_\_\_\_\_\_\_.

1. As a class, make a dotplot for the number that land flat and a dotplot for the  values.
2. For the sampling distribution of , answer the following:

Describe the shape:

Describe the center:

Describe the variability:

Now we will make one final attempt to estimate the proportion of Hershey’s Kisses that land flat. We will use this estimate to construct a **confidence interval**.

Toss all 50 Hershey’s kisses.

1. Based on this trial, what is your **point estimate** for the true proportion that land flat? \_\_\_\_\_\_\_\_
2. *Identify the population, parameter, sample and statistic.*

Population: Parameter:

Sample: Statistic:

1. Calculate the standard deviation of the sampling distribution of  using 
2. Was the sample a random sample? Why is this important?
3. Would it be appropriate to use a normal distribution to model the sampling distribution of ? Justify your answer.
4. In a normal distribution, 95% of the data lies within \_\_\_\_ standard deviations of the mean.
5. Calculate the value that is 2 X (standard deviation of the sampling distribution of ).

This value is called the **margin of error.**

1. Find the confidence interval using **point estimate +/- margin of error**.

**State, Plan, Do, Conclude**

**S**tatistics **P**roblems **D**emand **C**onsistency

A format for understanding inference and success on the AP Exam

**Confidence Intervals: A Four-Step Process**

**1. State:** What parameter do you want to estimate and at what confidence level?

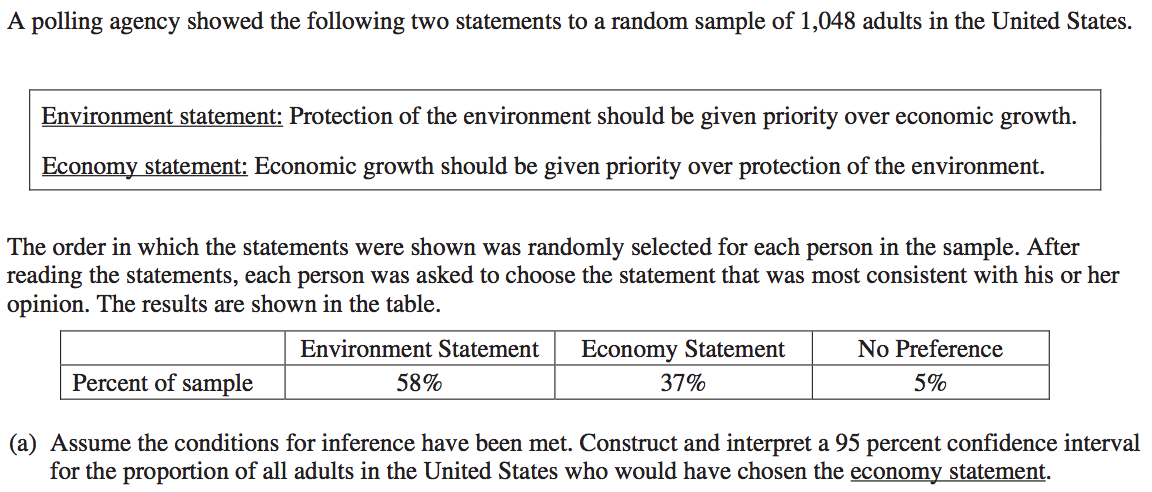
**2. Plan:** Identify the appropriate inference method and check conditions.

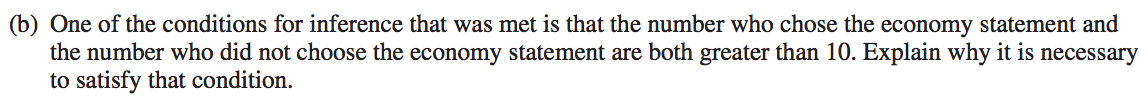
**3. Do:** Perform the calculations.

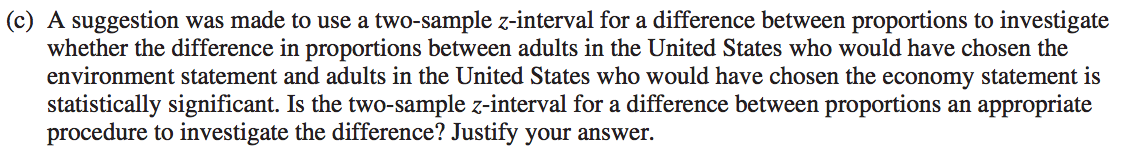
**4. Conclude:** Interpret your interval in the context of the problem.

Construct and interpret a 95% confidence interval for the true proportion of all Hershey’s Kisses that would land flat when tossed?

2016 AP Exam Question #5







**Which color M&M is the most common?**



The company that makes milk chocolate M&Ms claims the following distribution:

13% Brown, 14% Yellow, 20% Orange, 16% Green, 24% Blue, and 13% Red. Is this true?

1. Observed values: Brown:\_\_\_\_\_ Yellow:\_\_\_\_\_ Orange:\_\_\_\_\_ Green:\_\_\_\_\_ Blue:\_\_\_\_\_ Red:\_\_\_\_\_

Total number of M&Ms:\_\_\_\_\_\_\_

2. As a class, write down hypotheses for a significance test.

H0:

Ha:

3. Let’s suppose that M&Ms claimed distribution is correct. If they are correct, how many of each color would we expect to get in our sample.

Expected values: Brown:\_\_\_\_\_ Yellow:\_\_\_\_\_ Orange:\_\_\_\_\_ Green:\_\_\_\_\_ Blue:\_\_\_\_\_ Red:\_\_\_\_\_

Use the table to calculate the test statistic.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Observed | Expected | (Observed - Expected) | (Observed - Expected)2 |  |
| Brown |  |  |  |  |  |
| Yellow |  |  |  |  |  |
| Orange |  |  |  |  |  |
| Green |  |  |  |  |  |
| Blue |  |  |  |  |  |
| Red |  |  |  |  |  |

Add up all the numbers in the last column. This is our test statistic:\_\_\_\_\_\_\_\_\_

1. What value would we get for the test statistic if our sample was very close to what is expected? Explain.
2. What value would we get for the test statistic if our sample was very far from what is expected? Explain.

**State, Plan, Do, Conclude**

**S**tatistics **P**roblems **D**emand **C**onsistency

A format for understanding inference and success on the AP Exam

**Significance Tests: A Four-Step Process:**

**1. State:** Formulate hypotheses, define parameters.

**2. Plan:** Identify the type of significance test. Check conditions.

**3. Do:** Picture, general formula, specific formula, plug #s in, **test statistic, P-value.**

**4. Conclude:** Write a short novel.

Do the data provide convincing evidence that M&Ms are lying about the distribution of colors?

Chi Square Test of Homogeneity

Mr. Wilcox was interested in the Facebook habits of students at East Kentwood High School compared to students at Rockford High School. He took a random sample from each high school and recorded the results below:

Use Facebook EKHS Rockford Totals:

Not much 16 4

1+ per week 14 16

1+ per day 30 20

Totals

a. Find the marginal distribution of Facebook habits (in percents)

b. Find the conditional distribution of the Facebook habits for students from Rockford (in percents)

c. If the null hypothesis was that there is no difference between the Facebook habits of students at EKHS vs. students at Rockford, how would we calculate the expected counts?

Expected count (row 2, column 1) = Expected count (1+ per week for EKHS) =

Formulas:

Expected count =

df =

Do the data provide statistically significant evidence that the students at EKHS have different Facebook habits than the students at Rockford?

Chi Square Test of Association/Independence

Suppose we were interested in finding out what sports high school students play. In particular, we are interested in finding out if a student’s current math class has any relationship with the sport that they play. We took a random sample of 335 high school students who play 1 sport.

## Math Class Baseball Volleyball Football Basketball

Geometry 35 34 12 25

Algebra II 42 25 50 29

Pre-Calculus 11 21 30 21

Do these data show a statistically significant relationship between current math class and sport played?

Summary of the chi-square tests:

1. AGOF looks at \_\_\_\_\_ variable in \_\_\_\_\_ population.

Example:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

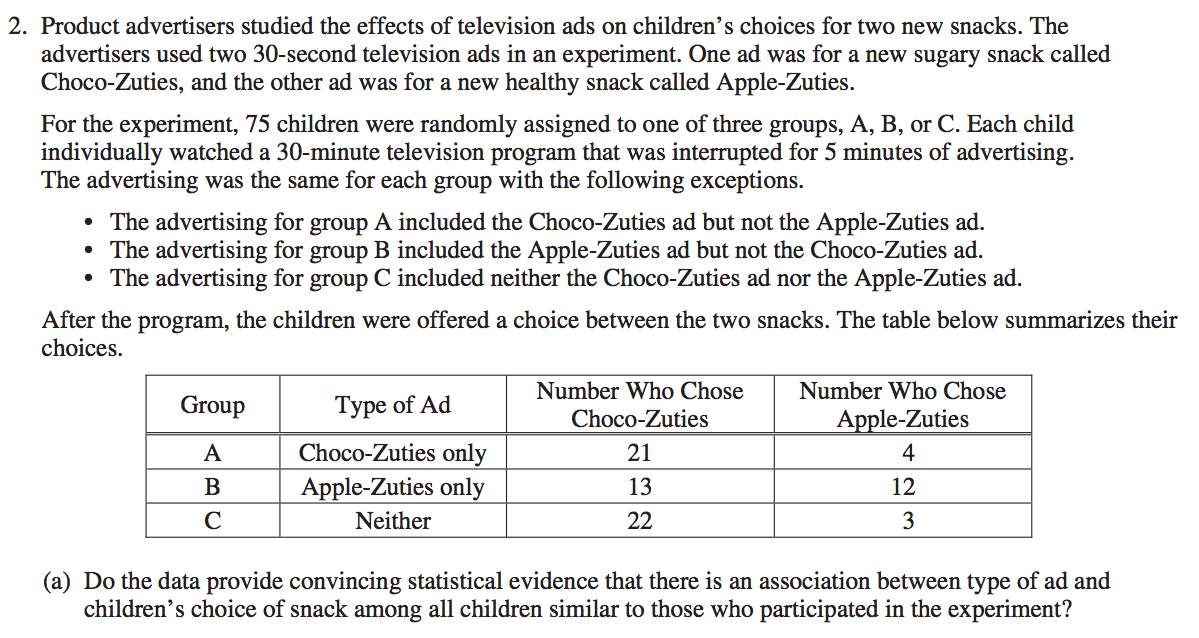
2.test of homogeneity looks at \_\_\_\_\_ variable in \_\_\_\_\_\_\_\_\_\_\_\_ populations.

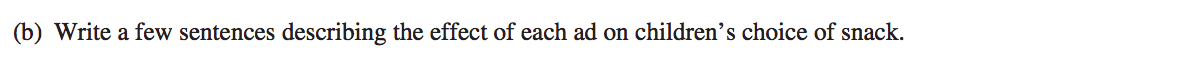
Example:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

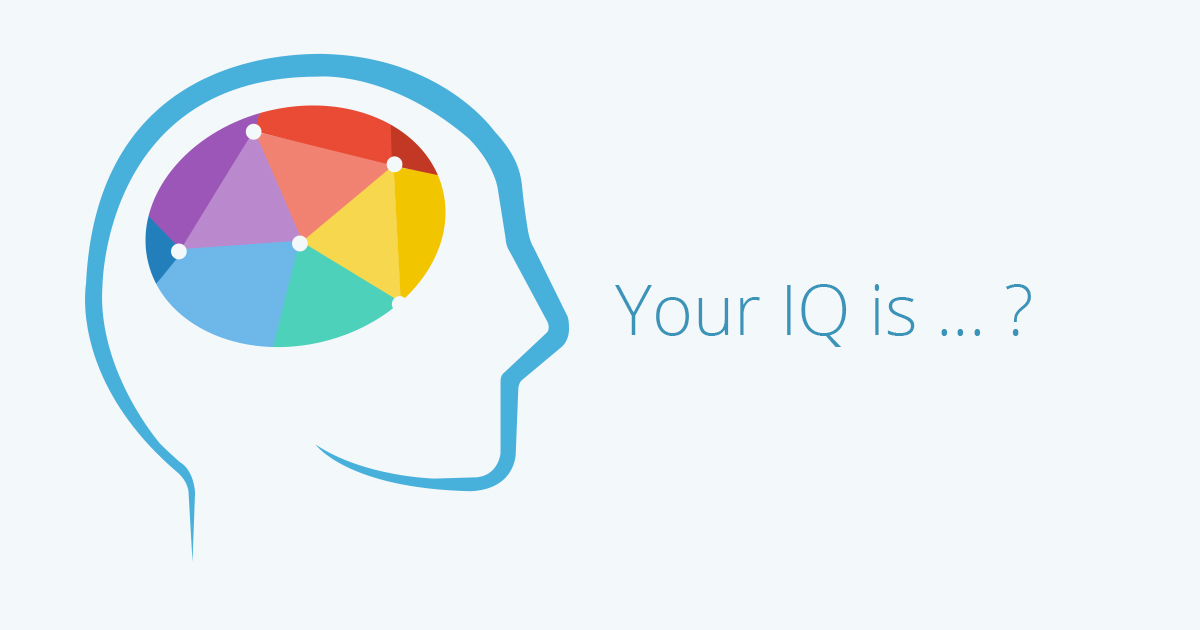
3. test of association looks at \_\_\_\_\_ variables in \_\_\_\_\_ population

Example:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

2016 AP Exam Question #2





**Can You Guess My IQ?**

As part of a new transcript at our school, the counselors

have decided to include an IQ score in addition to GPA.

Five students requested that the counselors update their transcripts for them.

Adam GPA = 1.8

Bernard GPA = 2.4

Christie GPA = 2.9

Deja GPA = 3.4

Eldin GPA = 3.8

Their IQ scores are 110, 85, 120, 95, 105 but they have been all mixed up and the counselors don’t know which IQ score goes with which GPA. The guidance counselors are forced to guess the IQ for each student (and they realize that higher GPA doesn’t always go with higher IQ).

Each counselor takes a different approach

**Counselor #1: The New Guy**

The New Guy is so nervous about being wrong, so he wants to play it safe with his guesses and minimize his error. He decides to find the average IQ and use it as his prediction for all five of the students:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| GPA | 1.8 | 2.4 | 2.9 | 3.4 | 3.8 |
| Predicted IQ |  |  |  |  |  |

**Counselor #2: The Veteran**

The Veteran noticed an equation written on the board in the AP Statistics room: . She uses this line of best fit to make her guesses.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| GPA | 1.8 | 2.4 | 2.9 | 3.4 | 3.8 |
| Predicted IQ |  |  |  |  |  |

**Counselor #3: The Truth Seeker**

Guidance counselor #3 pulled the five students out of class and found the truth.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| GPA | 1.8 | 2.4 | 2.9 | 3.4 | 3.8 |
| Actual IQ | 85 | 95 | 110 | 105 | 120 |

**Who made the better predictions?**

Now let’s see which counselor made better guesses:

**Counselor #1: The New Guy (used the mean IQ for every guess)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| GPA | 1.8 | 2.4 | 2.9 | 3.4 | 3.8 |
| Actual IQ | 85 | 95 | 110 | 105 | 120 |
| Predicted IQ |  |  |  |  |  |
| Error (Actual - Predicted) |  |  |  |  |  |
| Squared error |  |  |  |  |  |
|  |  |  |  |  |  |
| Sum of the squared errors: | |  |  |  |  |

**Counselor #2: The Veteran (used the line of best fit for every guess)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| GPA | 1.8 | 2.4 | 2.9 | 3.4 | 3.8 |
| Actual IQ | 85 | 95 | 110 | 105 | 120 |
| Guess IQ |  |  |  |  |  |
| Error (Actual - Predicted) |  |  |  |  |  |
| Squared error |  |  |  |  |  |
|  |  |  |  |  |  |
| Sum of the squared errors: | |  |  |  |  |

Who did better? Why?

What percentage of the errors by Guidance Counselor #1 were fixed (accounted for, explained by) because Guidance Counselor #2 used the line of best fit?

Now find the r2 value for the data using your calculator. r2 = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| GPA | 1.8 | 2.4 | 2.9 | 3.4 | 3.8 |
| Actual IQ | 85 | 95 | 110 | 105 | 120 |

Interpretation for r2:

** Does Beyonce write her own lyrics?**

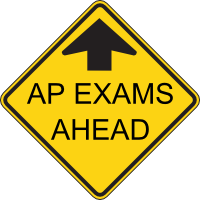
****

*"' “Crazy in Love' was really hard to write because there was so much going on … I mean, I had written —*

*what? —seven, eight number one songs with* ***Destiny's Child****, in a row." Beyonce, Vanity Fair.*

We can use statistics to help determine whether or not Beyonce wrote the song “Crazy in Love”. If we can find the average word length from the song, we can compare it to the average word length for songs that we know for sure were written by Beyonce.

1. Quickly circle a random sample of 5 words from the song. Write them below. How many letters in each word?
2. What is the average word length of your sample? \_\_\_\_\_\_\_\_.
3. Put your average on the dotplot on the white board at the front of the room. Copy the class dotplot below.
4. Find a new sample of 5 words using a random number generator. Put your average on the dotplot on the white board at the front of the room. Copy the class dotplot below.
5. How is the dotplot from #4 different than the dotplot for #3?
6. Take a random sample of 10 words from “Crazy in Love”. Find the average of your sample and put it on the dotplot at the front of the room. Copy the dotplot below.
7. What happens to the dotplot when we increase the sample size?
8. It is a well known fact that Beyonce wrote the lyrics for all of the Destiny’s child songs. The average word length for these songs is 3.64 letters. Based on your samples, do you have good evidence that Beyonce did not write the lyrics for “Crazy in Love”. Explain.

Which version of the Exam was harder?

Last year, East Kentwood High School had 30 students take the AP Statistics

Exam. We were informed later that the College Board gave two versions of the

Exam, which were randomly assigned to the students. Here are the results:

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Version A | | 3 | | 3 | 3 | 3 | 4 | 4 | 4 | 4 | 5 | 5 |
|  | | 5 | | 5 | 5 | 5 | 5 |  |  |  |  |  |
|  | |  | |  |  |  |  |  |  |  |  |  |
| Version B | | 2 | | 2 | 3 | 3 | 4 | 4 | 4 | 4 | 4 | 5 |
|  | | 5 | | 5 | 5 | 5 | 5 |  |  |  |  |  |
|  |  | |

What is the mean AP Exam score for students who took Version A () ? \_\_\_\_\_\_

What is the mean AP Exam score for students who took Version B () ? \_\_\_\_\_\_

What is the difference in means  - ? \_\_\_\_\_\_\_\_\_\_\_

Mr. Wilcox believes that Version B is unfair because it was a harder exam. Do you agree? Explain.

Suppose that the version of the test has no affect on the student’s AP score. Then a student would have the same AP score regardless of whether he/she was assigned to Version A or Version B. In that case, we could examine the results of repeated random assignments of the student to the two version.

Let’s see what would happen purely by chance if we randomly assign the 30 students to the two versions of the test many times, *assuming the version of the exam make does not affect the AP Exam score.*

**Activity:** Write each of the 30 AP Exam scores on a separate index card. Mix the cards well and then deal them face down into two piles of 15 cards each. Be sure to decide which pile is Version A and which is Version B *before* you look at the cards.

= mean of AP Exam scores for Version A = \_\_\_\_\_\_\_\_\_\_

= mean of AP Exam scores for Version B = \_\_\_\_\_\_\_\_\_\_

Difference in means =  - = \_\_\_\_\_\_\_\_\_\_

A negative difference in means would suggest that the mean AP Exam score for Version A was less than the mean AP Exam score for Version B.

Repeat the process four more times so that you have a total of 5 trials. Records your results in a table like this:

Mean AP Exam Mean AP Exam Difference in

Trial score Version A,  score Version B,  means ( - )

1

2

3

4

5

Make a class dotplot of the difference in means. Sketch below:

In what percent of the class’s trials did the difference in means equal or exceed 4.20 – 4.0 = 0.2? Is the result statistically significant? What conclusion can you draw about the two versions of the AP Exam?