What do you call a chicken
who is staring at lettuce?

What do you call a chicken
who is staring at lettuce?

Chicken Caesar Salad

TODAY • Ch. 12 Review Materials (12.1 only)
Quiz Wed.

Start 2020 AP Exam Review

→ Overview of Exam

→ Notation review

→ COMMON INTERPRETATIONS

Between now and the 12.1 quiz on Wednesday:

→ Ch. 12 Review Exercises (p. 821..)
1-4 only

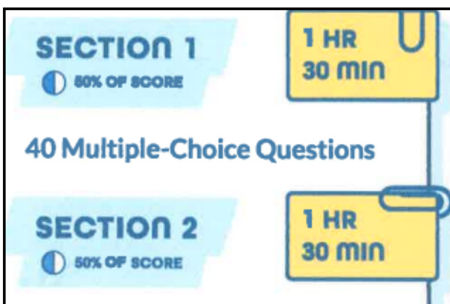
→ Ch. 12 AP Practice Test (p. 823....)
1, 3-8, 11 only

→ FRAPPY! ← Gets a bit into 12.2

→ LCQ 12.1 (all quizzes & solutions remain in class)

→ Strive for a 5

2020
AP Exam Review



6 Free-Response Questions

Part A | 5 problems:

- 1 multipart question with a primary focus on collecting data
- 1 multipart question with a primary focus on exploring data
- 1 multipart question with a primary focus on probability and sampling distributions
- 1 question with a primary focus on inference
- 1 question that combines 2 or more skill categories

Part B | 1 problem:

- 1 investigative task that assesses multiple skill categories and content areas, focusing on the application of skills and content in new contexts or in non-routine ways

Probably during an April review time

EXAM TOPICS

Unit 1: Exploring One-Variable Data

Unit 2: Exploring Two-Variable Data

Unit 3: Collecting Data

Unit 4: Probability, Random Variables, and Probability Distributions

Unit 5: Sampling Distributions

Lots of Inference Review next 5 days

Inference Test next Mon.

Unit 6: Inference for Categorical Data: Proportions

Unit 7: Inference for Quantitative Data: Means

Unit 8: Inference for Categorical Data: Chi-Square

Unit 9: Inference for Quantitative Data: Slopes

Unit 1: Exploring One-Variable Data

Unit 2: Exploring Two-Variable Data

Unit 3: Collecting Data

Unit 4: Probability, Random Variables, and Probability Distributions

Unit 5: Sampling Distributions

Already completed PPC's

Unit 6: Inference for Categorical Data: Proportions

Unit 7: Inference for Quantitative Data: Means

Unit 8: Inference for Categorical Data: Chi-Square

Unit 9: Inference for Quantitative Data: Slopes

PPC's available from now till May 15

Notation
Review

[mini quiz later this week]

Statistics Notation

a b n p q r s t x y z α β χ μ σ E F H P

Notation is an important part of communication in mathematics. Using the correct notation for statistical concepts is essential. BE CAREFUL! In statistics, unlike algebra, you are NOT free to substitute another letter in place of standard notation. Each of the above letters has a specific meaning in statistics. Also remember that "hats" and "bars" change those meanings. For example, y, $\hat{y}$, and $\overline{y}$ each have a very different meaning. Also, capitalizing a letter can change its meaning.

Work on for next 10 min.

First Semester Concepts:

1. Identify the letter used for the mean of a population.
2. Identify the letter used for the mean of a sample.
3. Identify the letter used for the standard deviation of a population.
4. Identify the letter used for the standard deviation of a sample.
5. Explain the difference between x_2 and x_i.

d March 01, 2020

First Semester Concepts:

1. Identify the letter used for the mean of a population. μ
2. Identify the letter used for the mean of a sample. $\bar{x}$
3. Identify the letter used for the standard deviation of a population. σ
4. Identify the letter used for the standard deviation of a sample. s s_x s_y
5. Explain the difference between x_2 and x_i.

x_2 represents the 2nd observed x-value while x_i represents all possible x-values.

6. Identify the letter that represents the standard normal variable.
7. Which letter represents the slope of the least-squares regression line?
8. Which letter represents the y-intercept of the least-squares regression line?
9. Explain the difference between y, $\hat{y}$, and $\bar{y}$.

$\hat{y} = a + bx$

6. Identify the letter that represents the standard normal variable. Z
7. Which letter represents the slope of the least-squares regression line? b
8. Which letter represents the y-intercept of the least-squares regression line? a
9. Explain the difference between y, $\hat{y}$, and $\bar{y}$.

y observed y-values

$\hat{y}$ predicted y-values

$\bar{y}$ average y-value

10. Explain the difference between y_2 and $\hat{y}_2$.

2nd observed y-value

2nd predicted y-value

11. Identify the letter used for correlation. r
12. Which letters are most commonly used for random variables? X and Y
13. Explain the difference between the variables x and X.

X is a random variable and x is not.

14. Identify the letter that represents the number of observations in a sample. n
15. Identify the letter used for the probability of an event. P

Notation Quiz (Practice)

- due by Thursday
- small quiz on notation on Thursday

Common Interpretations (prior to inference)

fill out while we review

Common Interpretations Prior to Inference (Notes to be filled out)

_____________________________ $\mu = 6$

The average number of cheeseburgers eaten by the sample of JCS students is 6 cheeseburgers.

_____________________________ $IQR = 5$

The middle 50% of the number of cheeseburgers eaten by the sample of JCS students are within 5 cheeseburgers of each other.

_____________________________ $\sigma = 2.74$

The number of cheeseburgers students eat per month typically varies from the mean by about 2.74 cheeseburgers.

Common Interpretations Prior to Inference (Notes to be filled out)

Interpret the mean $\mu = 6$ cheeseburgers

The average number of cheeseburgers eaten by the sample of JCS students is 6 cheeseburgers.

$Q_3 - Q_1$

Interp. the IQR $IQR = 5$ cheeseburgers

The middle 50% of the number of cheeseburgers eaten by the sample of JCS students are within 5 cheeseburgers of each other.

Pop. Std Deviation $\sigma = 2.74$

The number of cheeseburgers students eat per month typically varies from the mean (6) by about 2.74 cheeseburgers.

Classifying Outliers:

________________ $= Q3 + 1.5(IQR)$ ________________ $= Q1 - 1.5(IQR)$

Interpret ____________ 80^{th} percentile

80% of the students in the US scored 27 or lower on the ACT Exam.

Classifying Outliers:

Upper fence $= Q3 + 1.5(IQR)$ Lower Fence $= Q1 - 1.5(IQR)$

Interpret Percentiles 80^{th} percentile

80% of the students in the US scored 27 or lower on the ACT Exam.

Interpret ________________ $z = \frac{x - \mu}{\sigma} = -1.76$

Einstein scored 1.76 standard deviations below the mean on the statistics test.

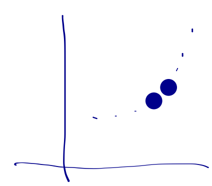
Interpret ________________ $r = -0.85$

There is a strong, negative, linear relationship between time spent studying for the AP Exam and score on the AP Exam. (Assuming a scatterplot indicates the relationship IS in fact linear)

Interpret r^2, the ____________________________ $r^2 = 0.72$

72% of the variation in AP Exam scores (y) are explained by the linear model that uses time spent studying (x) to predict AP Exam scores (y).

Interpret Z-Scores $z = \frac{x - \mu}{\sigma} = -1.76$



Einstein scored 1.76 standard deviations below the mean on the statistics test.

Interpret correlation coef. $r = -0.85$

There is a strong, negative, linear relationship between time spent studying for the AP Exam and score on the AP Exam. (Assuming a scatterplot indicates the relationship IS in fact linear)

Interpret r^2, the coeff. of determination $r^2 = 0.72$

72% of the variation in AP Exam scores (y) are explained by the linear model that uses time spent studying (x) to predict AP Exam scores (y).

Interpret the residual $resid = y_{actual} - y_{predicted} = -2.6$

The predicted height of this student is 2.6 inches larger than the actual height of this student when using a linear model that uses shoe size (x) to predict height (y).

Interpret the Slope $b = -2.3$

For every increase of 1 unit of x (in context), the predicted value of y (in context) will decrease by 2.3 units.

S_x S_y

Interpret the y-intercept $a = 40.1$

If a student studies zero hours (x), their predicted score on the test (y) is 40.1 points.

S_b

Interpret s, stand. deviation of the residuals *s = 1.235 wins*

When using a linear model with points per game (x) to predict wins (y), our win predictions with typically be off by 1.2 wins.

Notes and Tips for Inference

____

*Always check if there is random sampling and you don't know if there was replacement. Don't check the 10% condition on an experiment (Not sure but it could **cost you a point**!)*

____

*This is a check for normality of the sampling distribution for z and t tests (not χ^2). You should state the word **normal** for the check, but do NOT state that the sample data is normal (it isn't). Rather state that the sample data is **symmetric with no outliers**.*

Notes and Tips for Inference

10% Condition

*Always check if there is random sampling and you don't know if there was replacement. Don't check the 10% condition on an experiment (Not sure but it could **cost you a point**!)*

Normal Check

mean CLT (large sample)
proportion large counts

*This is a check for normality of the sampling distribution for z and t tests (not χ^2). You should state the word **normal** for the check, but do NOT state that the sample data is normal (it isn't). Rather state that the sample data is **symmetric with no outliers**.*

Name the Test vs. Formula

*You must give the formal name of the test **OR** the formula for the statistic/conf. interval. You do not need to do both. Give the name as it is easier and avoid the formula as it is easy to mess up details/notation.*

____

*Explicitly comparing the p-value to the alpha value is **ALWAYS** required. Don't forget to write that step.*

Name the Test vs. Formula

You must give the formal name of the test ***OR*** *the formula for the statistic/conf. interval. You do not need to do both. Give the name as it is easier and avoid the formula as it is easy to mess up details/notation.*

Don't forget linkage

Explicitly comparing the p-value to the alpha value is ***ALWAYS*** *required. Don't forget to write that step.*

Wording Conclusions

1. *Never reference the sample in your parameter, hypotheses, or your conclusion. Don't use wording that implies you are talking about the sample!*

2. *You MUST MUST MUST reference the parameter in your conclusion. Have the word MEAN, PROPORTION, or SLOPE in your conclusions for your z and t procedures. (chi-sq you word differently).*

Good

We are 94% confident that the interval from 44.1% to 49.9% captures the true population proportion of U.S. women that would say they don't get enough time to themselves.

We are 94% confident that the interval from 44.1% to 49.9% captures the true population proportion of U.S. women that said they don't get enough time to themselves.

(this references the women in the sample)

Good

We are 94% confident that the interval from 44.1% to 49.9% captures the true population proportion of U.S. women that would say they don't get enough time to themselves.

Bad

We are 94% confident that the interval from 44.1% to 49.9% captures the true population proportion of U.S. women that said they don't get enough time to themselves.

(this references the women in the sample)

Also Bad

We are 94% confident that the interval from 44.1% to 49.9% captures the true number of U.S. women that would say they don't get enough time to themselves.

(this doesn't use the word proportion)

(0.25,0.32)

We are 95% confident that the interval from 25% to 32% captures the (parameter in context) true proportion of people in the population that scored a 5 on the AP Exam.

Also Bad

We are 94% confident that the interval from 44.1% to 49.9% captures the true number of U.S. women that would say they don't get enough time to themselves.

(this doesn't use the word proportion)

Interp. Confidence Interval (0.25,0.32)

We are 95% confident that the interval from 25% to 32% captures the (parameter in context) true proportion of people in the population that scored a 5 on the AP Exam.

Interp. Confidence Level 95%

If we take many sample of size 800 (n) from this populations, about 95% of them will result in an interval that captures the (parameter in context) true proportion of people in the population that scored a 5 on the AP Exam.

$s_{\bar{x}}$ or $s_{\hat{p}}$ or s_b or $s_{\widehat{p_1}-\widehat{p_2}}$ etc ...

When taking samples of size 800 and calculating the sample proportion, these sample proportions will typically be 0.027 away from the true proportion of people in the population that scored a 5 on the AP Exam.

Interp. Confidence Level 95%

If we take many sample of size 800 (n) from this populations, about 95% of them will result in an interval that captures the (parameter in context) true proportion of people in the population that scored a 5 on the AP Exam.

Inter. std error of a statistic $s_{\bar{x}}$ or $s_{\hat{p}}$ or s_b or $s_{\widehat{p_1}-\widehat{p_2}}$ etc ...

When taking samples of size 800 and calculating the sample proportion, these sample proportions will typically be 0.027 away from the true proportion of people in the population that scored a 5 on the AP Exam.

$H_o: p = 0.80$ $H_a: p \neq 0.80$ $p =$ true population proportion of free throws that Rickey Bobby makes

______________________________ $p - value = 0.0075$

$p - value = 0.0075 < 0.05 = \alpha$, reject the H_o

We have convincing evidence that the (H_a with parameter with context) true population proportion of free throws that Ricky Bobby makes is not equal to 80%.

______________________________ $p - value = 0.0075$

Assuming the true proportion of the shots Ricky Bobby makes is 0.80 (H_o is true) there is a 0.0075 (p-value) probability of getting a sample proportion of 0.64 (sample statistic we got) or one that is more extreme

$H_o: p = 0.80$ $H_a: p \neq 0.80$ $p =$ true population proportion of free throws that Rickey Bobby makes

Conclusion of a Signif. Test $p - value = 0.0075$

$p - value = 0.0075 < 0.05 = \alpha$, reject the H_o

We have convincing evidence that the (H_a with parameter with context) true population proportion of free throws that Ricky Bobby makes is not equal to 80%.

______________________________ $p - value = 0.0075$

Assuming the true proportion of the shots Ricky Bobby makes is 0.80 (H_o is true) there is a 0.0075 (p-value) probability of getting a sample proportion of 0.64 (sample statistic we got) or one that is more extreme

$H_o: p = 0.80$ $H_a: p \neq 0.80$ $p =$ true population proportion of free throws that Rickey Bobby makes

Conclusion of a Signif. Test $p-value = 0.0075$

$p-value = 0.0075 < 0.05 = \alpha$, reject the H_o

We have convincing evidence that the (H_a with parameter with context) true population proportion of free throws that Ricky Bobby makes is not equal to 80%.

Inter. P-Value $p-value = 0.0075$

Assuming the true proportion of the shots Ricky Bobby makes is 0.80 (H_o is true) there is a 0.0075 (p-value) probability of getting a sample proportion of 0.64 (sample statistic we got) or one that is more extreme

Type I Error: Ho is really true but we reject Type II Error: ____________________

Interpret ___________: *We conclude that Ricky Bobby lied about making 80% of his free throws; when in reality he really DOES make 80%.*

Interpret ___________: *We do not have enough evidence to conclude Ricky Bobby makes less than 80% of his free throws; when in reality he does not make than 80%.*

Interpret ___________: $power = 0.35$

Given that Rickey Bobby really doesn't make 80% of his shots, when sampling 100 shots, there is a 35% chance that we will correctly reject the null and find enough evidence to conclude that he does not make 80% of his free throws.

$P(Type\ I\ error) =$ P(reject Ho | Ho is true) = α = signif. level

$P(Type\ II\ error) =$ ______

$Confidence\ Level =$ ______

$Power =$ ______

Type I Error: Ho is really true but we reject Type II Error: Ho is really false but we don't reject

Interpret ______: *We conclude that Ricky Bobby lied about making 80% of his free throws; when in reality he really DOES make 80%.*

Interpret ______: *We do not have enough evidence to conclude Ricky Bobby makes less than 80% of his free throws; when in reality he does not make than 80%.*

Interpret ______: $power = 0.35$

Given that Rickey Bobby really doesn't make 80% of his shots, when sampling 100 shots, there is a 35% chance that we will correctly reject the null and find enough evidence to conclude that he does not make 80% of his free throws.

$P(\text{Type I error}) =$ P(reject Ho | Ho is true) = α = signif. level

$P(\text{Type II error}) =$ P(fail to reject Ho | Ho is false) = β

$\text{Confidence Level} =$ P(fail to reject Ho | Ho is true) = $1 - \alpha$

$\text{Power} =$ ______________________

Type I Error: Ho is really true but we reject

Type II Error: Ho is really false but we don't reject

Interpret Type I error**:** *We conclude that Ricky Bobby lied about making 80% of his free throws; when in reality he really DOES make 80%.*

Interpret Type II error**:** *We do not have enough evidence to conclude Ricky Bobby makes less than 80% of his free throws; when in reality he does not make than 80%.*

Interpret ____________**:** $power = 0.35$

Given that Rickey Bobby really doesn't make 80% of his shots, when sampling 100 shots, there is a 35% chance that we will correctly reject the null and find enough evidence to conclude that he does not make 80% of his free throws.

Type I Error: Ho is really true but we reject Type II Error: Ho is really false but we don't reject

Interpret Type I error: *We conclude that Ricky Bobby lied about making 80% of his free throws; when in reality he really DOES make 80%.*

Interpret Type II error: *We do not have enough evidence to conclude Ricky Bobby makes less than 80% of his free throws; when in reality he does not make than 80%.*

Interpret Power: $power = 0.35$

Given that Rickey Bobby really doesn't make 80% of his shots, when sampling 100 shots, there is a 35% chance that we will correctly reject the null and find enough evidence to conclude that he does not make 80% of his free throws.

$P(Type\ I\ error) =$ P(reject Ho | Ho is true) = α = signif. level

$P(Type\ II\ error) =$ P(fail to reject Ho | Ho is false) = β

$Confidence\ Level =$ P(fail to reject Ho | Ho is true) = $1 - \alpha$

$Power =$ P(correctly rejecting a false Ho) = P(reject Ho | Ho is false)

① finish the AP Stats Notation Quiz (Practice) by Thursday

↓

Mini quiz on it

② Prepare for The Ch. 12 Quiz on Wednesday