AP Stats Webinar

Putting It (Inference) All Together

Monday, 21 Mar 2016, 7:30-8:30PM EDT

When you first log-on to the webinar you should hear music. If you do not hear music, you will need to make adjustments so you have access to sound for the webinar.

Try the following:

1. At the top of your screen, click on the **communicate** tab. Select **call using computer**.
2. If this still does not work, you may need to log-off and log back on using a different browser (**not Internet Explorer**).

AP Stats Webinar

Putting It (Inference) All Together

Monday, 21 Mar 2016, 7:30-8:30PM EDT

If you are still having sound issues, please use the "chat" feature to ask for help. Be sure to select "All Panelists" when asking your questions, so our moderator can view your comments and help.

Putting It (Inference) All Together

Doug Tyson

Central York High School

tyson.doug@gmail.com

Files at MrTysonStats.com

Brief Resume' (My Apologies)

- Member of the ASA/NCTM Joint Committee on Curriculum in Statistics (2016-2019)
- AP Statistics Exam Reader and Table Leader
- ASA National Poster Competition judge (2011-Present)
- Contributor and supplements author for *The Practice of Statistics*, 5th edition (2013)
- Presenter of two-day *AP Statistics Teacher Workshop* (2015)
- Co-author of the College Board Curriculum Module on Random Sampling and Random Assignment (2012)
- Co-leader of *Experienced AP Statistics Teacher* workshop (2014)
- Co-leader of *Simulation Based Inference* workshop (2015)
- Various workshops and sessions at NCTM regional and annual conferences (2011-2015)

An Overview of Inference Procedures

(This content is non-negotiable.)

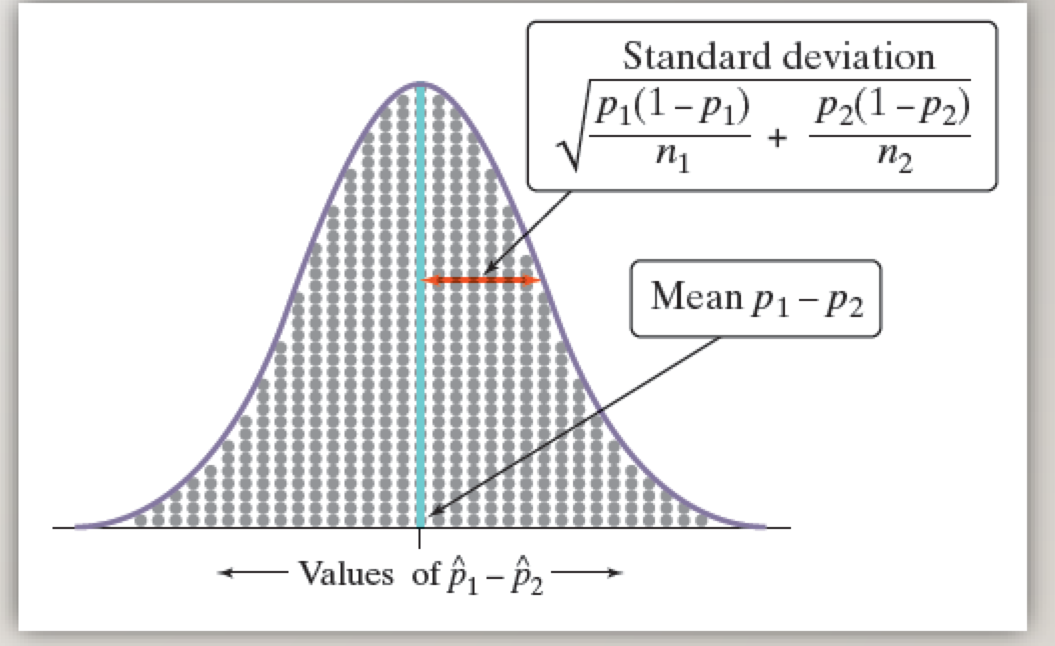
Inference for Proportions

- One-sample z interval for p
  - Calculator 1-PropZInt
- One-sample z test for p
  - Calculator 1-PropZTest
- Two-sample z interval for $p_1 - p_2$
  - Calculator 2-PropZInt
- Two-sample z test for $p_1 - p_2$
  - Calculator 2-PropZTest

Conditions for Inference about Proportions

- **Random:** data from a random sample or randomized experiment
  - ***10%:*** *n* < 0.10*N* when sampling without replacement (ignore otherwise)
- **Large Counts**:
  - Observed counts all at least 10 ($n\hat{p} \geq 10$ AND $n(1-\hat{p}) \geq 10$.
  - For 1-sample *z*-test for *p*, use expected counts.

DOUBLE THESE FOR 2 PROPORTIONS!



The Practice of Statistics, 5th edition by Starnes, Tabor, et al.

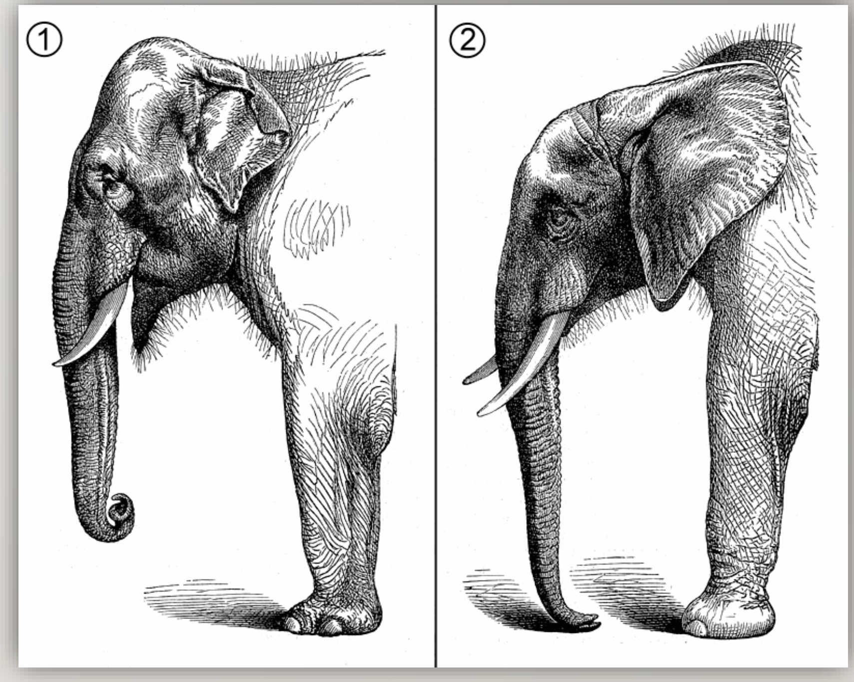
Student Difficulties with Inference about Proportions

- Watch out for the difference between $\hat{p}$ and p_0 (observed counts and expected counts)!
- When testing a difference in proportions, use the pooled/combined proportion of successes for the standard error:

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\frac{\hat{p}_C(1-\hat{p}_C)}{n_1} + \frac{\hat{p}_C(1-\hat{p}_C)}{n_1}}}, \text{ where } \hat{p}_c = \frac{X_1 + X_2}{n_1 + n_2}$$

Inference for Means

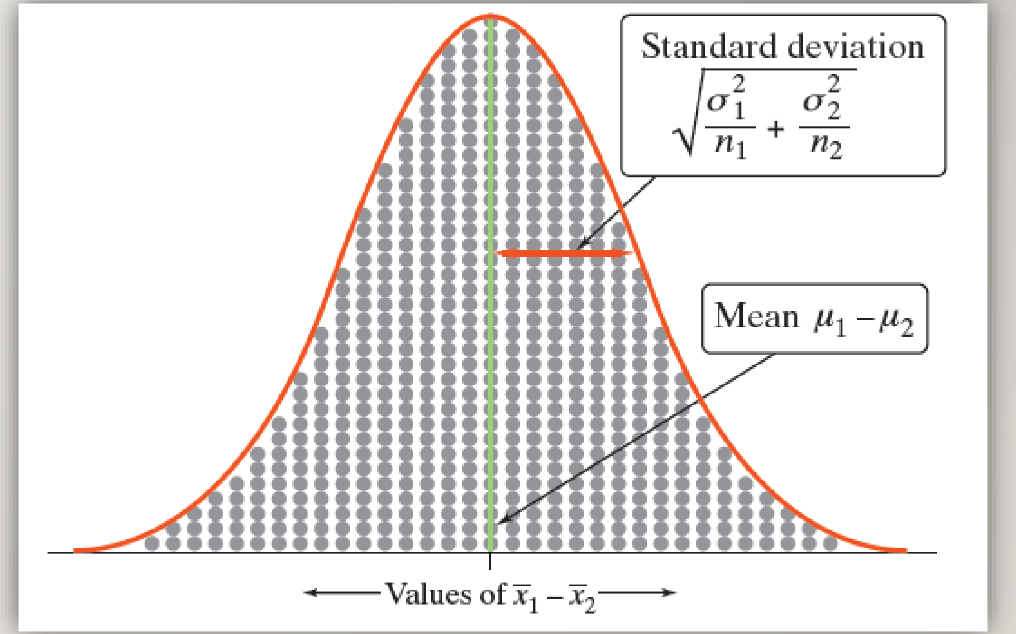
- One-sample t interval for μ
  - Calculator T-Interval
- One-sample t test for μ (also paired t test)
  - Calculator T-Test
- Two-sample t interval for $\mu_1 - \mu_2$
  - Calculator 2-SampTInt
- Two-sample t test for $\mu_1 - \mu_2$
  - Calculator 2-SampTTest



Conditions for Inference about Means

- **Random:** data from a random sample or randomized experiment
  - ***10%:*** *n* < 0.10*N* when sampling without replacement (ignore otherwise)
- **Normal/Large Sample**:
  - Normal population, OR $n \geq 30$, OR graph sample data and look for absence of heavy skewness and outliers.

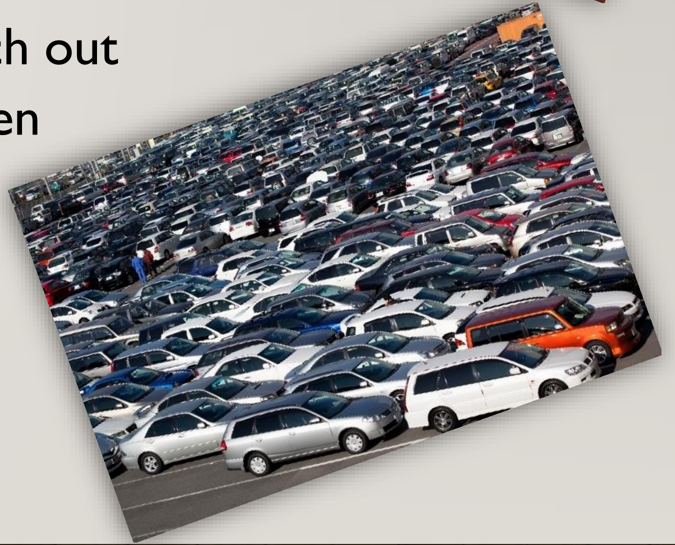
DOUBLE THESE FOR 2 MEANS!



The Practice of Statistics, 5th edition by Starnes, Tabor, et al.

Student Difficulties with Inference about Means

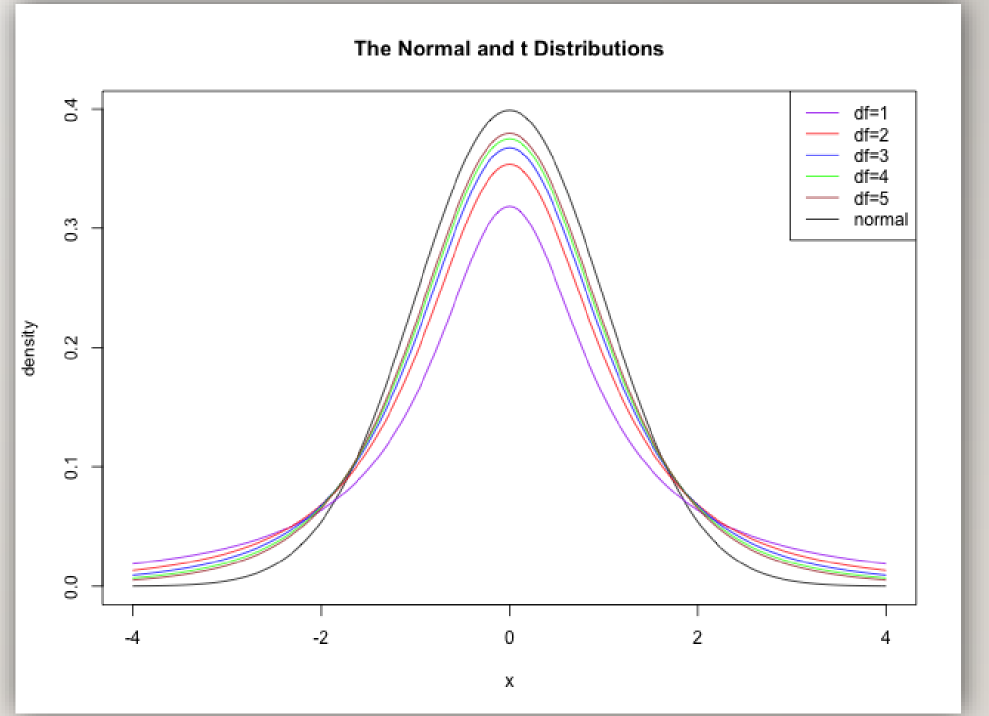
- The population standard deviation, σ, is almost never known. Don't write formulas with σ in them.
- We cheat on the rule above when determining sample size needed to produce a certain margin of error.
- Don't state hypotheses/conclusions about *the data*. In particular watch out for past tense. "The true difference in mean times for men and women who rode the bicycles…" is about the sample, not the population.



- Watch out for paired data (mean difference) versus two-sample (difference in means).
  - 2010 APFRQ #5 (Fish) = difference in means
  - 2014 APFRQ #5 (Cars) = mean difference

Degrees of Freedom

- For one sample: df = $n - 1$
- For two-sample:
  - Option 1 (preferred) – Use the *df* reported by technology
  - Option 2 (conservative) – Use the *df* that is the minimum of $n_1 - 1$ and $n_2 - 1$

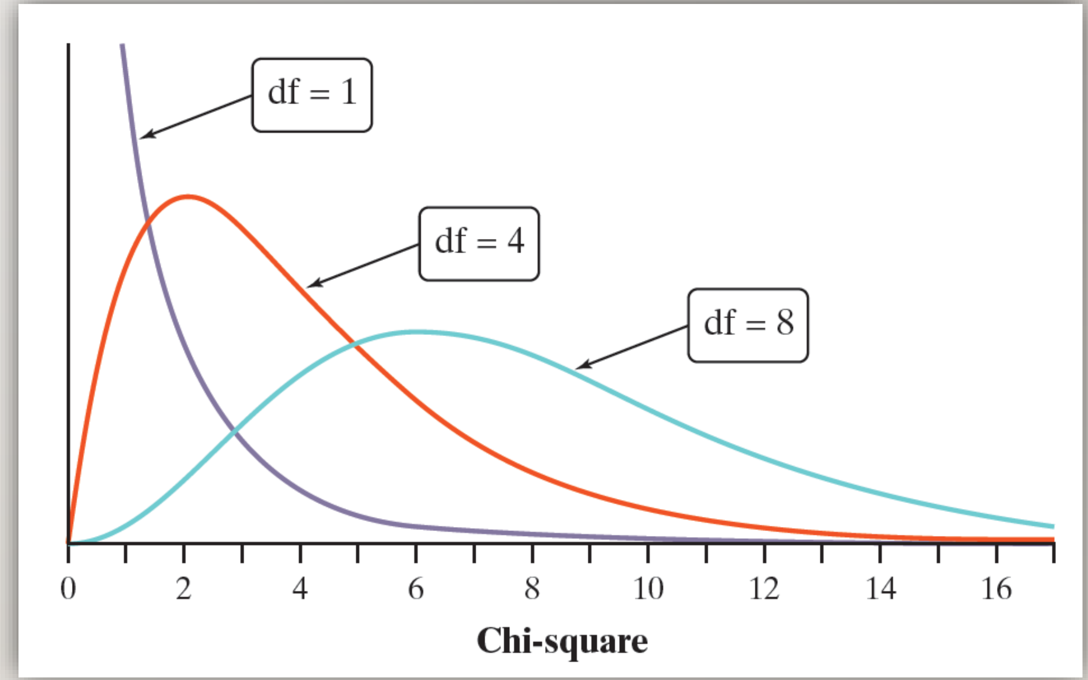


Inference for Distributions of Categorical Variables (Chi-square)

- χ^2 Goodness-of-fit test
  - Calculator χ^2GOF-Test
- χ^2 Test for Homogeneity
  - Calculator χ^2-Test
- χ^2 Test for Independence
  - Calculator χ^2-Test

Conditions for Chi-square tests

- **Random:** data from a random sample or randomized experiment
  - ***10%:*** $n < 0.10N$ when sampling without replacement (ignore otherwise)
- **Large Counts**:
  - All *expected* counts are at least 5.



The Practice of Statistics, 5th edition by Starnes, Tabor, et al.

Student Difficulties with Chi-square tests

- The Large Counts condition is about expected counts.
- You must show the expected counts to convince your reader that you really know that they are at least 5.
- Know the formula for expected counts in a two-way table:

$$expected\ count = \frac{row\ total\ \times column\ total}{n}$$

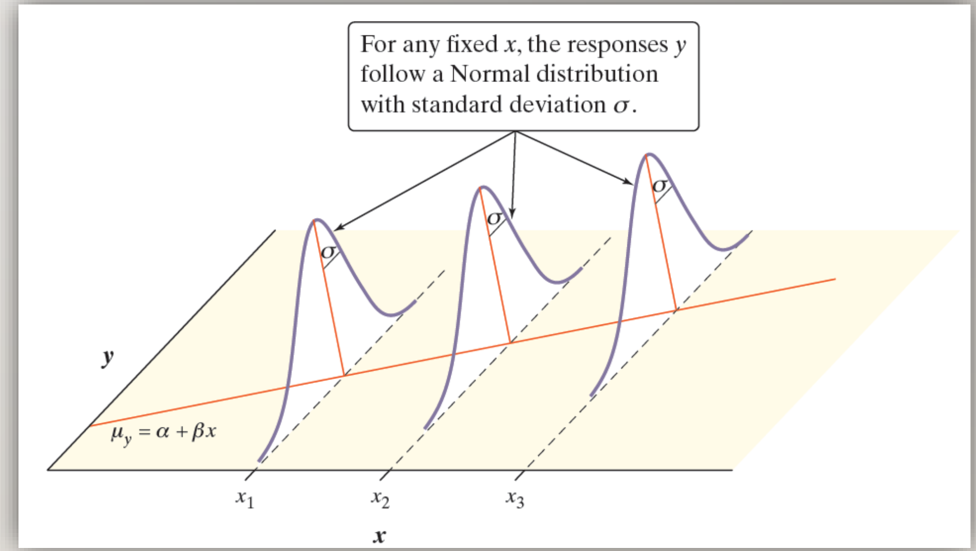
- In an FRQ:
  - Let the question set up your hypotheses
  - Call it a "chi-square two-way table test"

Inference about Relationships between Quantitative Variables (Slope)

- One-sample *t* interval for β
  - Calculator LinRegTInt
- One-sample *t* test for β
  - Calculator LinRegTTest

Conditions for Inference about Slope

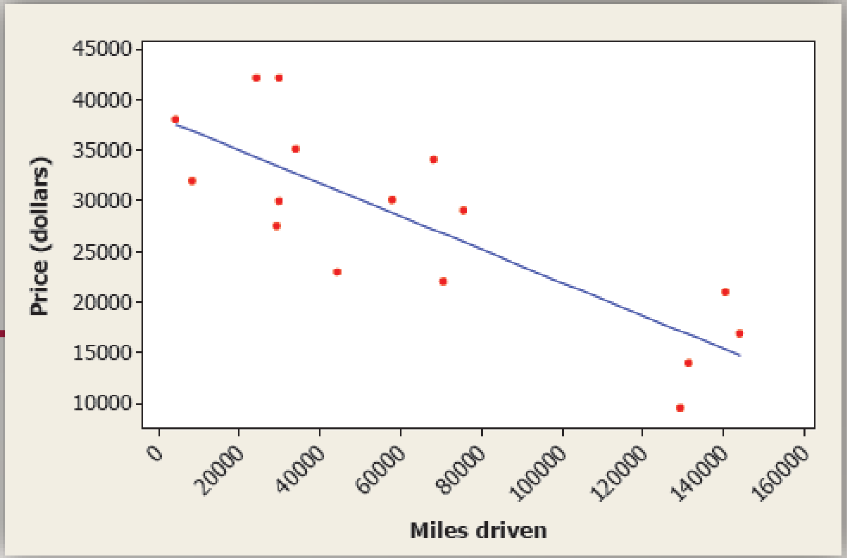
- **Linear:** true relationship between variables is linear
- **Independent (observations):** $n < 0.10N$ when sampling without replacement (ignore otherwise)
- **Normal:** the distribution of the response variable y is Normal for each value of explanatory variable x
- **Equal SD:** constant σ around the regression line for all values of the explanatory variable x
- **Random:** data from a random sample or randomized experiment

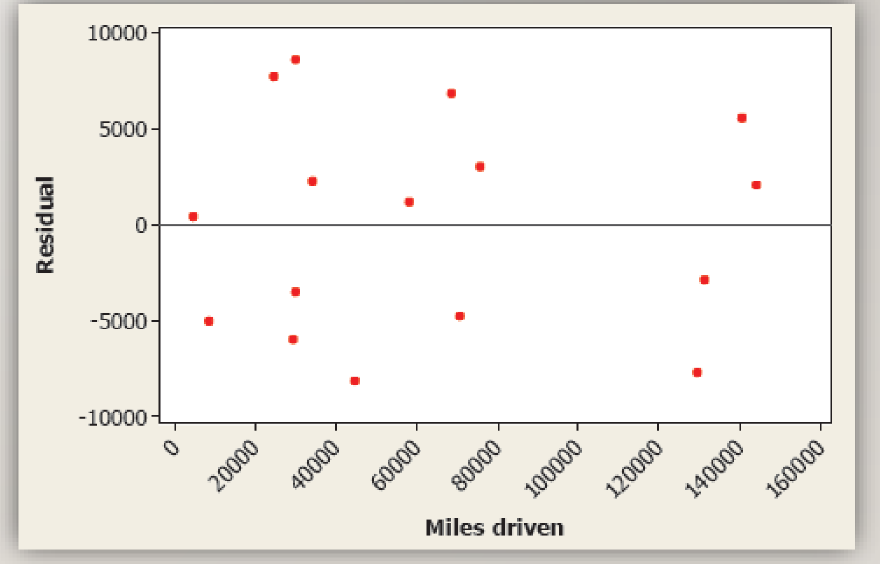


The Practice of Statistics, 5th edition by Starnes, Tabor, et al.

Student Difficulties with Inference about Slope

- df = $n - 2$ because we are estimating two parameters α, β
- Don't get hung up on conditions – rarely (never?) tested to date.
- Be able to calculate residuals (*observed y – predicted y*) and interpret residual plots.



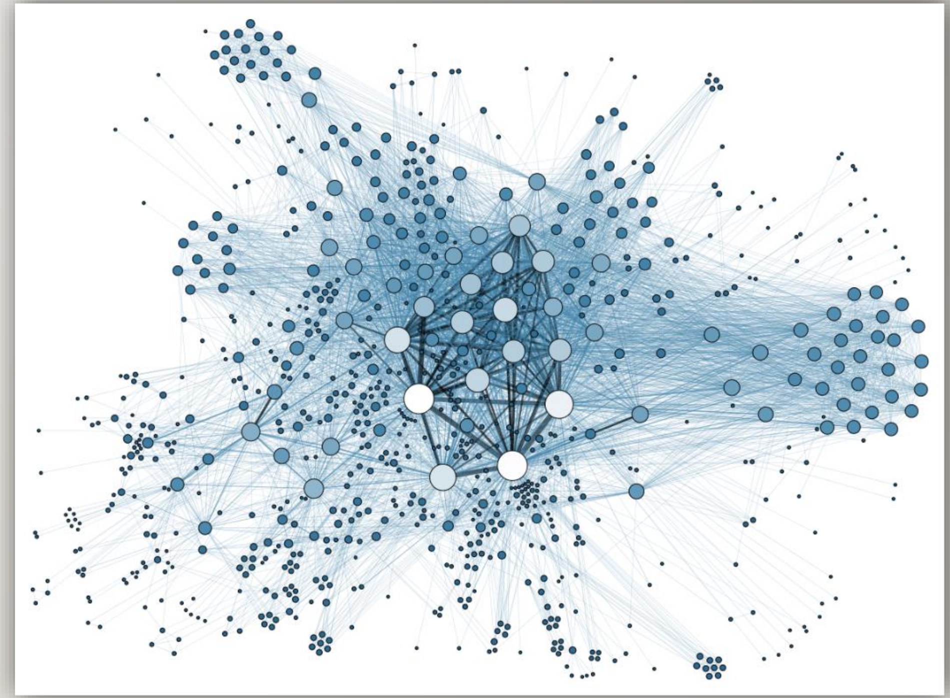


The Practice of Statistics, 5th edition by Starnes, Tabor, et al.

Teaching Frameworks & Ideas

Scope of Inference

- How is randomness used?
  - Random sampling from some larger population
  - Random assignment of subjects to treatments (or vice-versa)
- What conclusions can be drawn?
  - Random sampling allows inference about a larger population/association
  - Random assignment allows for inference about cause-and-effect



Teaching Inference – The Four-step Process

(*The Practice of Statistics*, 5th edition by Starnes, Tabor, et al.)

INTERVALS

1. STATE – define parameters and state confidence level
2. PLAN – give name (or formula) of procedure and check conditions
3. DO – perform calculations (give interval from calculator)
4. CONCLUDE – interpret interval

TESTS

1. STATE – hypotheses and significance level
2. PLAN – give name (or formula) of procedure and check conditions
3. DO – perform calculations (give test stat and *P*-value from calculator)
4. CONCLUDE – write conclusion

Intervals vs. Tests
Considerations Beyond the Four-step Process

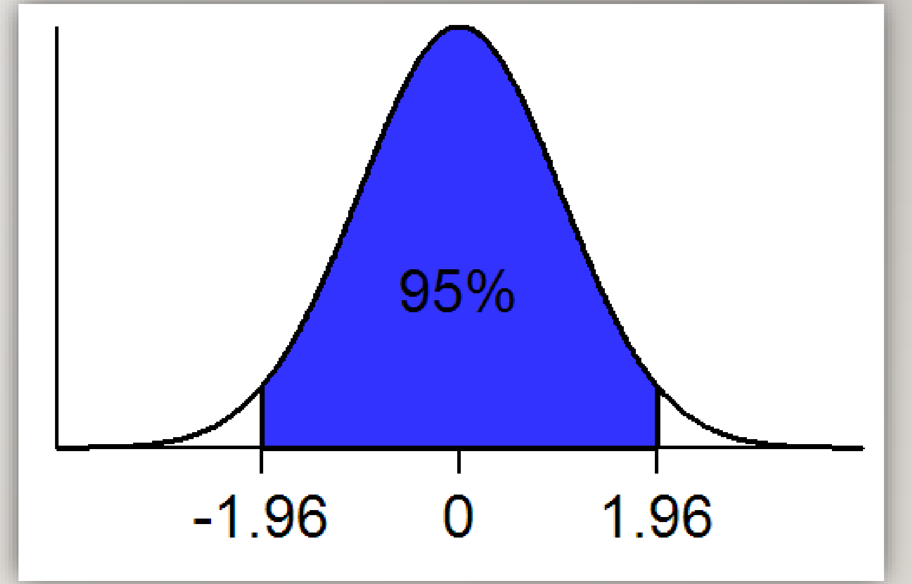
- Intervals
  - Interpret the confidence level
  - Know what factors affect the width of the interval (margin of error)
  - Does the interval include 0? (test decisions from an interval)
- Tests
  - Describe Type I & II errors (and consequences)
  - Interpret the *P*-value
  - Know what factors affect power of a test

Tips for The Exam and Formulas

- AP Exam FRQs – do not worry about writing out two-sample formulas on the AP Exam
- Why do I require the formulas for my own tests?
- AP Exam MCQs – it is possible the questions about two-sample formulas will be on the MCQ part of the AP Exam
- Remind students that variances add for independent random variables!

The Logic of a Test

- What did we expect?
- Did we observe what we expected?
- What are some reasons we didn't get what we expected?
- Can we eliminate chance as a reasonable explanation?
  - The purpose of a statistical test is to remove random chance as a reasonable explanation for the observed data



AP Exam Questions

2015 APFRQ #4 – Aspirin and Colon Cancer

2015 AP® STATISTICS FREE-RESPONSE QUESTIONS

4. A researcher conducted a medical study to investigate whether taking a low-dose aspirin reduces the chance of developing colon cancer. As part of the study, 1,000 adult volunteers were randomly assigned to one of two groups. Half of the volunteers were assigned to the experimental group that took a low-dose aspirin each day, and the other half were assigned to the control group that took a placebo each day. At the end of six years, 15 of the people who took the low-dose aspirin had developed colon cancer and 26 of the people who took the placebo had developed colon cancer. At the significance level $\alpha = 0.05$, do the data provide convincing statistical evidence that taking a low-dose aspirin each day would reduce the chance of developing colon cancer among all people similar to the volunteers?

2015 APFRQ #4 – Aspirin and Colon Cancer

What were common student errors or omissions?

- Many students had trouble defining the parameters appropriately. Some common errors were:
  - Using subscripts that do not clearly convey which group is associated with which parameter (for example, p_1 and p_2), and with no explanation of which is which.
  - Defining the parameter symbol as the group rather than as a population proportion associated with the group, such as, p_2 = placebo group.
  - Defining symbols that refer to (or imply reference to) the sample rather than to a population proportion, such as " p_1 is the proportion of adults who *took* low-dose aspirin daily and then *developed* cancer."
- Many students had trouble checking the appropriate conditions for the test. For instance:
  - Students incorrectly stated that the randomness condition was satisfied because a simple random sample was chosen, rather than because of random assignment.
  - Students incorrectly stated that the normality condition was satisfied because both groups were larger than 30.

2015 APFRQ #4 – Aspirin and Colon Cancer

- Some students had problems with the computing or stating the test statistics. Common errors included:
  - Not reporting the value of the test statistic, but reporting only the p-value.
  - Using the formula for the standard error of the difference in sample proportions as the z-statistic.
  - Calculating the z-statistic "by hand" from the formula, plugging in the numbers, simplifying the expression incorrectly, and getting the wrong value.

- Some students had trouble making an appropriate conclusion. Common mistakes included:
  - Not providing an explicit conclusion about the research question, but simply restating a rejection of the null hypothesis in context.
  - Omitting explicit justification for a decision or conclusion by failing to compare the p-value to the given $\alpha = 0.05$, such as, "at $\alpha = 0.05$ we reject the null hypothesis."
  - Failure to make any use of the significance level provided in the problem $(\alpha = 0.05)$, such as, "since the p-value is low …".
  - Overstating the conclusion to imply that the alternative hypothesis has been proven, such as, "since the p-value is less than $\alpha = 0.05$, we know that taking a low dose aspirin every day will reduce the chance of getting colon cancer."

Multiple Choice Questions (2008)

8. Sophomore, junior, and senior students at a high school will be surveyed regarding a potential increase in the extracurricular student activities fee. There are three possible responses to the survey question—agree with the increase, do not agree with the increase, or no opinion. A chi-square test will be conducted to determine whether the response to this question is independent of the class in which the student is a member. How many degrees of freedom should the chi-square test have?

   (A) 9

   (B) 6

   (C) 4

   (D) 2

   (E) 1

Multiple Choice Questions (2008)

20. The National Honor Society at Central High School plans to sample a random group of 100 seniors from all high schools in the state in which Central High School is located to determine the average number of hours per week spent on homework. A 95 percent confidence interval for the mean number of hours spent on homework will then be constructed using the sample data. Before selecting the sample, the National Honor Society decides that it wants to decrease the margin of error. Which of the following is the best way to decrease the margin of error?

    (A) Increase the confidence level to 99%

    (B) Use the population standard deviation

    (C) Use the sample standard deviation

    (D) Increase the sample size

    (E) Decrease the sample size

Multiple Choice Questions (2008)

25. In a large school district, 16 of 85 randomly selected high school seniors play a varsity sport. In the same district, 19 of 67 randomly selected high school juniors play a varsity sport. A 95 percent confidence interval for the difference between the proportion of high school seniors who play a varsity sport in the school district and high school juniors who play a varsity sport in the school district is to be calculated. What is the standard error of the difference?

(A) 0.0347

(B) 0.0695

(C) 0.1362

(D) 0.9800

(E) 1.6900

Multiple Choice Questions (2008)

35. To determine whether employees at Site X have higher salaries, on average, than employees at Site Y of the same company do, independent random samples of salaries were obtained for the two groups. The data are summarized below.

	Site X	Site Y
Mean	$61,234	$60,529
Standard Deviation	$4,352	$3,456
n	235	183

Based on the data, which of the following statements is true?

(A) At the 5% significance level, employees at Site Y have a significantly higher mean salary than employees at Site X do.

(B) At the 1% significance level, employees at Site Y have a significantly higher mean salary than employees at Site X do.

(C) At the 5% significance level, employees at Site X have a significantly higher mean salary than employees at Site Y do.

(D) At the 1% significance level, employees at Site X have a significantly higher mean salary than employees at Site Y do.

(E) At the 10% significance level, there is no significant difference in salaries between the employees at the two sites.

Contact Me

Doug Tyson

Email: tyson.doug@gmail.com

FB: tyson.doug

Twitter: @tyson_doug

LI: tyson.doug

URL: MrTysonStats.com