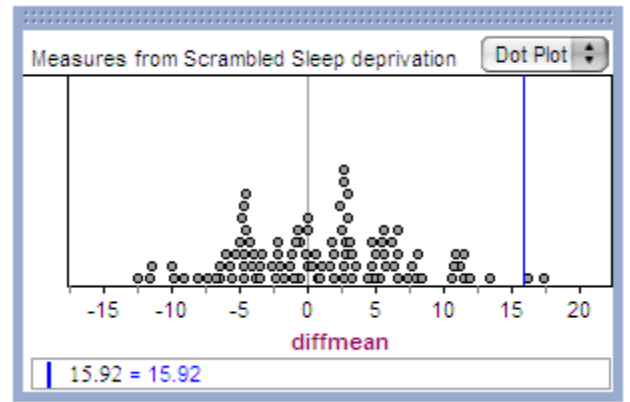
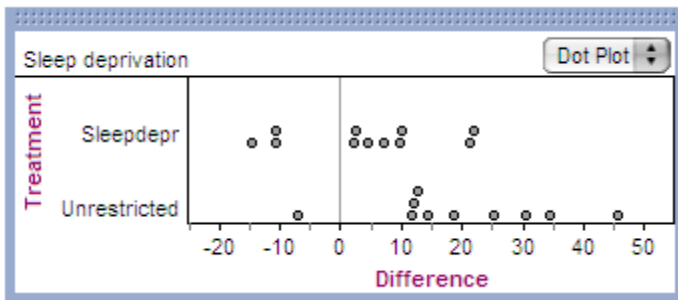


# Making Sense of Statistical Inference



Daren Starnes  
Mathematics Department Chair  
The Lawrenceville School  
dstarnes@lawrenceville.org

2013 NCTM Annual Meeting  
Denver, CO

## Pre-test: Choosing the correct inference method

- a.  $t$  test
- b.  $t$  interval
- c. two-sample  $t$  test
- d. two-sample  $t$  interval
- e. one proportion  $z$  test
- f. one proportion  $z$  interval
- g. two proportion  $z$  test
- h. two proportion  $z$  interval
- i. chi-square test for goodness-of-fit
- j. chi-square test for association/independence
- k. chi-square test for homogeneity
- l. linear regression  $t$  interval for slope
- m. linear regression  $t$  test for slope

- \_\_\_\_\_ 1. Which brand of AA batteries last longer—Duracell or Eveready?
- \_\_\_\_\_ 2. According to a recent survey, a typical teenager has 38 contacts stored in his/her cellphone. Is this true at your school?
- \_\_\_\_\_ 3. What percent of students at your school have a MySpace page?
- \_\_\_\_\_ 4. Is there a relationship between the age of a student's car and the mileage reading on the odometer at a large university?
- \_\_\_\_\_ 5. Is there a relationship between students' favorite academic subject and preferred type of music at a large high school?
- \_\_\_\_\_ 6. Who is more likely to own an iPod—middle school girls or middle school boys?
- \_\_\_\_\_ 7. How long do teens typically spend brushing their teeth?
- \_\_\_\_\_ 8. Are the colors equally distributed in Fruit Loops?
- \_\_\_\_\_ 9. Which brand of razor gives a closer shave? To answer this question, researchers recruited 25 men to shave one side of their face with Razor A and the other side with Razor B.
- \_\_\_\_\_ 10. How much more effective is exercise and drug treatment than drug treatment alone at reducing the incidence of heart attacks among men aged 65 and older?

For more practice identifying the correct inference method, try Larry Green's Web site at Lake Tahoe Community College:

<http://www.ltcconline.net/greenl/java/Statistics/catStatProb/categorizingStatProblems12.html>

*What other obstacles prevent students from succeeding on inference questions?*

# Critical Issues for Helping Students Perform Inference Successfully

## 1. Deciding which inference method to choose

- Inference about what?
- How were the data produced?
- Estimating or testing a claim?

## 2. Checking conditions—which ones and why?

- Random sampling/random assignment
  - Random selection allows us to generalize from sample to population
  - Random assignment allows cause-and-effect conclusions
  - Both permit the use of probability distributions as models in inference
- Normality or large sample size
  - For the desired probability distribution to give accurate results
- Independence (of measurements/samples)
  - So that the standard deviation of the statistic is approximately correct

## 3. Communicating effectively

- Using notation and statistical terminology correctly
  - Distinguishing among samples, populations, and sampling distributions
- Stating technically correct conclusions in context
  - Connecting sampling distributions (what happens if we repeat the randomization process many times) and inference (conclusions based on a single observational study or experiment).
  - *Confidence interval*: set of plausible values for the parameter
  - *Significance test*: is it believable that the observed difference occurred just by random chance?

## 4. Using technology as a tool

- Simulations to understand the underlying logic of inference
- Calculations for inference methods

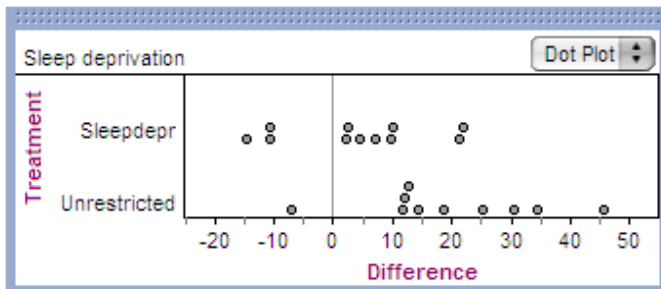
## Does sleep deprivation linger?

(from Rossman, Cobb, Chance, and Holcomb's NSF project shared at JMM 2008 in San Diego)

Researchers have established that sleep deprivation has a harmful effect on visual learning. But do these effects linger for several days, or can a person “make up” for sleep deprivation by getting a full night's sleep on subsequent nights? A recent study (Stickgold, James, and Hobson, 2000) investigated this question by randomly assigning 21 subjects (volunteers between the ages of 18 and 25) to one of two groups: one group was deprived of sleep on the night following training and pre-testing with a visual discrimination task, and the other group was permitted unrestricted sleep on that first night. Both groups were then allowed as much sleep as they wanted on the following two nights. All subjects were then re-tested on the third day. Subjects' performance on the test was recorded as the minimum time (in milliseconds) between stimuli appearing on a computer screen for which they could accurately report what they had seen on the screen. The sorted data and dotplots presented here are the improvements in those reporting times between the pre-test and post-test (a negative value indicates a decrease in performance):

Sleep deprivation ( $n = 11$ ):  $-14.7, -10.7, -10.7, 2.2, 2.4, 4.5, 7.2, 9.6, 10.0, 21.3, 21.8$

Unrestricted sleep ( $n = 10$ ):  $-7.0, 11.6, 12.1, 12.6, 14.5, 18.6, 25.2, 30.5, 34.5, 45.6$



Sleep deprivation		Treatment		Row Summary
		Sleepdepr	Unrestricted	
Difference		11	10	21
		3.9	19.82	11.481

S1 = **count** ( )

S2 = **mean** ( )

*Question:* Is it **plausible (believable)** that there's really no harmful effect of sleep deprivation, and random chance alone produced the observed differences between these two groups?

### Simulation of the random assignment

- If no difference in treatment effects, then values will be the same as in the original study.
- Write each of the 21 data values on a separate card.
- Place all of the cards (subjects) in a bag.
- How large a difference in group means with different random assignments?
- Mix your cards and deal two groups—one with 10 cards (unrestricted sleep) and one with 11 cards (sleep deprived).
- Calculate the difference in mean time improvement for the two groups (unrestricted – sleep).

## 2012 AP<sup>®</sup> STATISTICS FREE-RESPONSE QUESTIONS

4. A survey organization conducted telephone interviews in December 2008 in which 1,009 randomly selected adults in the United States responded to the following question.

At the present time, do you think television commercials are an effective way to promote a new product?

Of the 1,009 adults surveyed, 676 responded “yes.” In December 2007, 622 of 1,020 randomly selected adults in the United States had responded “yes” to the same question. Do the data provide convincing evidence that the proportion of adults in the United States who would respond “yes” to the question changed from December 2007 to December 2008 ?