Statistics for Common Core:

From Randomness to Inference

Daren Starnes

The Lawrenceville School

dstarnes@lawrenceville.org

CMC South 2013

Common Core State Standards in Mathematics

Making Inferences and Justifying Conclusions S-IC

**Understand and evaluate random processes underlying statistical experiments**

1. Understand statistics as a process for making inferences about population parameters based on a random sample from that population.

2. Decide if a specified model is consistent with results from a given data-generating process, e.g., using simulation. *For example, a model* *says a spinning coin falls heads up with probability 0.5. Would a result of 5* *tails in a row cause you to question the model?*

**Make inferences and justify conclusions from sample surveys, experiments, and observational studies**

3. Recognize the purposes of and differences among sample surveys, experiments, and observational studies; explain how randomization relates to each.

4. Use data from a sample survey to estimate a population mean or proportion; develop a margin of error through the use of simulation models for random sampling.

5. Use data from a randomized experiment to compare two treatments; use simulations to decide if differences between parameters are significant.

6. Evaluate reports based on data.

**Activity #1: Red, White, and You Inference for Sampling**

Goal: To estimate the true proportion *p* of red beads in the container with 95% confidence.

I. How will we collect the data? What potential pitfalls do we need to avoid?

II. Data collection and analysis

III. Making an inference

* What is our “best guess” at the value of *p* based on the data? This is our *point estimate*.
* How close does the sample result tend to be to the true population value in random samples of this size? To find out, let’s do a simulation.
* Conclusion:

**Activity #2: Do subliminal messages work? Inference for experiments**

From *The Practice of Statistics,* 5th edition, by Starnes, Tabor, Yates, and Moore, W. H. Freeman and Co., in publication

A “subliminal” message is below our threshold of awareness but may nonetheless influence us. Can subliminal messages help students learn math? A group of 18 students who had failed the mathematics part of the City University of New York Skills Assessment Test agreed to participate in a study to find out. All received a daily subliminal message, flashed on a screen too rapidly to be consciously read. The treatment group of 10 students (assigned at random) was exposed to “Each day I am getting better in math.” The control group of 8 students was exposed to a neutral message, “People are walking on the street.” All 18 students participated in a summer program designed to raise their math skills, and all took the assessment test again at the end of the program. The table below gives data on the subjects’ scores before and after the program.

 Mathematics skills scores before and after a subliminal message

 Treatment Group Control Group

 Pre-test Post-test Pre-test Post-test

 18 24 18 29

 18 25 24 29

 21 33 20 24

 18 29 18 26

 18 33 24 38

 20 36 22 27

 23 34 15 22

 23 36 19 31

 21 34

 17 27

*Source:* Data provided by Warren Page, New York City Technical College, from a study done by John Hudesman.

Do the data provide convincing evidence that the subliminal message about getting better at math results in larger increases in scores, on average, than the neutral message?

Simulation of the random assignment:

* Write the 18 difference values on identical slips of paper.
* Place all 18 slips in a bag and mix them well.
* If there’s no treatment effect, then pretest scores, posttest scores, and difference values same as in original study.
* How large a difference in group means with different random assignments?
* Pull out slips one at a time until you have 10 to represent the Treatment group. The remaining 8 slips represent the Control group.
* Calculate the difference in mean improvement scores for the two groups (Treatment – Control). Record this value in the table below.
* Repeat this process two more times.

|  |  |  |  |
| --- | --- | --- | --- |
| Trial | 1 | 2 | 3 |
| Difference in mean improvement scores (T – C) |  |  |  |

Combine results: How often did a difference as large as or larger than the one observed in the study ($\overbar{x}\_{T}-\overbar{x}\_{C}=3.15$) occur just due to the chance involved in the random assignment?