Teaching Statistical Inference With Multiple Decks of Cards

Barbara C. Beliveau
Assistant Professor of Economics
St. Mary’s College of Maryland
bcbeliveau@smcm.edu

INTRODUCTION

Students frequently have difficulty understanding the random nature of statistical inference. In particular, they struggle with the idea that classical statistical analysis does not prove hypotheses, but merely allows us to reject hypotheses that are deemed to be sufficiently unlikely. A simple classroom experiment with playing cards has been helpful in enhancing student comprehension of this concept.

CLASSROOM PREPARATION

Preparation begins with an informal discussion of probabilities and randomness. Students will have been encoleed read a chapter on statistical inference and view a clip from a performance of a movie adaptation of the play “Rosencrantz and Guildenstern Are Dead” by Tom Stoppard. In this scene a coin repeatedly comes up heads and the characters discuss the idea of a fair coin.

EXPERIMENTAL PROTOCOL

The experiment begins by introducing the practice of casinos using multiple decks of cards at the blackjack tables as a way to keep players from gaining information about the cards remaining in play.

Students are told that they will be given a sample of cards from an altered set of decks. The alteration may be fairly easy to spot (all red cards removed) or more subtle (removing half of the black deuces). Their task is to determine how the sleeve of 10 decks may have been altered.

Round 1: Student must hypothesize 3 alterations that are consistent with their sample. These are listed on the work sheet.

Round 2: Students receive 5 additional cards and determine if this sample allows them to exclude any of their previous hypotheses. If all three are excluded, they must formulate at least one additional hypothesis.

Students are asked if any of them have data that would disprove any of these hypotheses. This typically results in change? Explain.

PRE-EXPERIMENT QUESTION & RESULTS

After the initial class discussion, students are asked to answer the following question.

According to the site BetUS.com the following data describes the coin toss at the beginning of the first 43 Super Bowls. Each coin is unique, and produced by the same company.

Heads has landed 22 times, Tails 21

Does this indicate that the coins used are fair or not? Explain.

CLASSROOM DISCUSSION

Students are asked to recall their hypotheses about the sleeves of cards. Usually there are several competing possibilities.

Students are then asked if any of them have data that would disprove any of these hypotheses. This typically results in either all the hypotheses being eliminated or a reduced number of surviving hypotheses.

The following discussion questions are considered:

1. Can you be sure the sleeve hasn’t been altered in some other way?
2. Can you be sure the sleeve has been altered?
3. If you had a sample that included all 52 cards in a standard deck, would this prove that the sleeve was unaltered?

POST- EXPERIMENT EVALUATION

The following question is included on a quiz or test in the week following the experiment.

According to the site BetUS.com the following data describe the coin toss at the beginning of the first 43 Super Bowls.

Heads has landed 22 times, Tails 21

The NFC has won the coin toss 12 straight years

The winner of the coin toss has a losing record in the Super Bowl 20-23

(a) Based on this information, assess the fairness of the coin used in the tosses.

(b) If you knew that the last 10 tosses contained 8 heads and 2 tails, would your answer change? Explain.

All 33 students answered part (a) by saying the coins used seemed to be “fair” in that Heads and Tails were equally likely.

Most students, 25 of the 33, said that if the same coin was used the fact that the last 10 tosses seemed biased didn’t negate the experience with the larger sample.

COMMENTS

In one section, students spontaneously began to collaborate and eliminate possible hypotheses about how the deck may have been altered.

The idea of a Fair sleeve of cards became a reference point throughout the class. It was used in demonstrations of the SAMPLING and HYPOTHESIS TESTING.

SAMPLING: Students selected 5-card samples from the sleeve and computed the average value of the cards in their sample. They then reported their results and the resulting sampling distribution was graphed. The contrast between the uniform population frequency distribution and the empirical sampling distribution served to make the idea concrete and comprehensible.

HYPOTHESIS TESTING: After tampering with a deck, students selected 5-card samples and tried to determine if the deck had been altered. Although the tampering was subtle (some 2’s and 4’s had been removed) it was detectable when all the student data was combined.

EXTENSION: Discussion of how to determine if the deck had been altered in a mean-preserving way.