

Friendly Observer

In a study published in the *Journal of Personality and Social Psychology* (Butler and Baumeister, 1998), researchers investigated a conjecture that having an observer with a vested interest would decrease subjects' performance on a skill-based task. Subjects were given time to practice playing a video game that required them to navigate an obstacle course as quickly as possible. They were then told to play the game on final time with an observer present. Subjects were randomly assigned to one of two groups:

- Group A was told that the participant and the observer would each win \$3 if the participant beat a certain threshold.
- Group B was told that only that the participant would win the prize if the threshold was beaten.

The goal of this data analysis is to determine whether or not there is an effect from the observer on the performance. That is, like the  $\chi^2$  test, our hypotheses are:

$H_o$  : there is no association between the two variables

$H_a$  : there is an association between the two variables

The data from the 24 subjects is given below:

	A: shares prize	B: no sharing	Total
Beat threshold	3	8	11
Did not beat threshold	9	4	13
Total	12	12	24

Recall:

p-value = probability of seeing your data or more extreme if the null hypothesis is true.

In keeping with the reasoning of hypothesis tests, we want to ask how likely the sample results would have been if in fact the observer's incentive had no effect on the subject's performance. One way to analyze this question is to assume that those 11 successes would have been successes regardless of which group the subject had been assigned to. We can then *simulate* the process of assigning subjects at random to the two groups, observing how often we obtain a sample result at least as extreme (3 or fewer successes assigned to A) in the actual sample.

Red cards are successes and black cards are failures. Shuffle the cards well, and randomly assign 12 people (cards) to the A group and 12 people (cards) to the B group. How many of the 12 in the A group are successes? Is your result as extreme as the actual sample?

Repeat this experiment 5 times and record your results in the table:

Repetition #	1	2	3	4	5
"successes" assigned to group A					
as extreme as actual study?					