

Results of in-class survey for Fisher's Exact Test Activity:

There were two versions of the survey given in class, both of which are shown below, followed by information about the results of the survey.

For each question, check the one response that best captures your reaction to the scenario presented.

Version 1:

1. Suppose that you have decided to see a play for which the admission charge is \$20 per ticket. As you prepare to purchase the ticket, you discover that you have lost a \$20 bill. Would you still pay \$20 for a ticket to see the play?

Yes _____ No _____

2. Suppose that the United States is preparing for the outbreak of an unusual disease which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows:

- If Program A is adopted, 200 people will be saved.
- If Program B is adopted, there is a 1/3 probability that 600 people will be saved and a 2/3 probability that nobody will be saved.

Which of the two programs would you favor? Program A _____ Program B _____

Version 2:

1. Suppose that you have decided to see a play and paid the admission price of \$20 per ticket. As you prepare to enter the theater, you discover that you have lost the ticket. The seat was not marked, and the ticket can not be recovered. Would you pay \$20 for another ticket?

Yes _____ No _____

2. Suppose that the United States is preparing for the outbreak of an unusual disease which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows:

- If Program A is adopted, 400 people will die.
- If Program B is adopted, there is a 1/3 probability that nobody will die and a 2/3 probability that 600 people will die.

Which of the two programs would you favor? Program A _____ Program B _____

Results:

Question 1	Lost \$20	Lost ticket
Yes	13	8
No	2	6

Question 2	Save	Die
Program A	8	2
Program B	7	12

Before doing the survey, I conjectured that people who lost \$20 would be more likely to pay another \$20 for a ticket than those who had lost the ticket. I also conjectured that people would be more likely to select Program A in question 2 if was presented as *saving* 200 people rather than allowing 400 to *die* (note that the actual effect of Program A is identical either way – the only difference is how it was presented). I made this conjecture because I expected that people would be more likely to gamble when faced with a *guarantee* of a large number of people *dying*, and would be less likely to gamble if faced with a guarantee that a number of people would be saved.

Exercise 9 from Chapter 3 asks you to apply Fisher’s Exact Test to these results. I will show you how this works for question 1 from the survey and ask you to run the test for question 2 (regarding the disease). Our goal is to determine whether the results of this survey provide strong evidence in favor of my conjectures. The results of the survey on question 1 certainly match my conjecture: approximately 87% of those with the question that said they had lost \$20 said they would buy a ticket, while only about 57% of those who were told they lost a ticket said they would buy another one. But the critical question is whether this survey really provides good evidence that this pattern would continue to hold up if I asked the question to more people, or was it just random luck of the draw in this particular sample that created the difference? We answer this question through the use of Fisher’s Exact Test.

We begin by assuming that the wording of the question made absolutely no difference in the answers given. That is, we *assume* that the 21 people that answered “yes” to question 1 would have answered “yes” regardless of which version of the question they received (that is, it doesn’t matter whether they lost \$20 cash or the \$20 ticket, either way they would pay for a ticket). We also assume that the 8 people who said “no” would have said “no” regardless of which version of the question they got. In other words, we assume that it was just luck of the draw that caused 13 of the 21 people that said “yes” to receive the first version of the survey (the “lost \$20” version). This is our *null hypothesis*. The *alternative hypothesis* is that the wording of the question *does* make a difference (as described in my conjecture above).

We now ask *how likely* it is (i.e., what is the *probability*) that we would get a result as extreme or more extreme than the result we actually got on the class survey *assuming* that the wording actually made no difference. To do this, we will use hypergeometric probabilities. We are dealing with a population of $N = 29$ people. Of these, $M = 21$ would say “yes” to buying a ticket (treat this as a “success”) and $N - M = 8$ would say “no” regardless of which version of the survey they received. We then randomly select $n = 15$ of the people to receive Version 1 of the questionnaire (the “lost \$20” version). Let X equal the number of “yes” answers among these 15 people. Then X has the hypergeometric distribution. Finally, we ask how likely it is that we would get a result as extreme *or more extreme* than the result we actually got when we did the survey with the class. That is, we want to know the probability that 13 or more of the 15 people receiving Version 1 would say “yes” if we were just randomly choosing 15 out of the 29 people. This probability, called the *p-value* for our test, is given by

$$P(X \geq 13) = P(X = 13) + P(X = 14) + P(X = 15)$$

$$= \frac{\binom{21}{13} \binom{8}{2}}{\binom{29}{15}} + \frac{\binom{21}{14} \binom{8}{1}}{\binom{29}{15}} + \frac{\binom{21}{15} \binom{8}{0}}{\binom{29}{15}} \approx 0.0735 + 0.0120 + 0.0007 = 0.0862$$

This p-value is fairly small, but not extremely small. As a result, we conclude that this data provides some evidence to support my conjecture, but the evidence is not extremely strong. Even if the wording of the question made no difference, there is still approximately an 8.6% chance (p-value = 0.0862) that we would wind up with a difference in responses as large as I got from this survey just because of the random luck of who got which version of the survey. This probability is small (so it is unlikely that this was purely random luck, but it is still certainly possible), so most likely the wording of the question did have an effect. To obtain more convincing evidence, a larger sample would be needed.