Permutation Tests: Comparing Proportions

Math 107, Exploration 3

1 Draw inferences beyond the data

Exploratory data analysis reveals that the proportion of subjects in the lithium group who relapsed is greater than the proportion who relapsed in the control group, but does this provide convincing evidence of a genuine difference in the long-run proportions (i.e., probabilities)? To determine whether the results provide convincing enough evidence that desipramine is a more effective treatment of cocaine addiction, we will use a permutation test to determine whether these results are typical or surprising for what we could find if the treatments were equally as effective.

The key to creating the randomization distribution for our permutation test is to assume that if the treatments are equally effective (null hypothesis), then **the 28 subjects that relapsed would have relapsed regardless of their assigned treatment**. Similarly, we must assume that the 20 subjects that did not relapsed would not have relapsed regardless of their assigned treatment. In other words, the randomization distribution assumes the null hypothesis is true.

1. We cannot use coins to conduct this simulation analysis since we have two variables to consider: the treatment group, and whether or not a subject relapsed. How might we compute a randomization sample for these data? (Try to think about some tactile way to simulate, don't just say that you would use R.)

2. Use R to perform the process you outlined above once.
   (a) Calculate the difference of these proportions.

   (b) Is your simulated statistic at least as "extreme" as the observed statistic from the reported study?

3. Use R to perform this process a large number of times (say, 1,000) in order to create a randomization distribution to assess the original result.

4. Look closely at the randomization distribution for the difference in proportions that you produced in R.
   (a) Is the randomization distribution centered around 0? Explain why this makes sense (*Hint*: think about the choice of statistic and about the null hypothesis.)

   (b) Is the observed value of the sample statistic out in the tail of the randomization distribution, or not so much? In other words, does the observed result appear to be typical or surprising when the null hypothesis is true?

(c) As you know, we can assess the strength of evidence against the null hypothesis by calculating a p-value. To calculate the p-value from this distribution, you will count the number of simulations that produced a

simulated difference of proportions equal to ______________ or ______________.

(d) In R, calculate the p-value using the randomization distribution you created.

Guidelines for evaluating strength of evidence from *p*-values:

$0.10 \leq$	*p*-value		not much evidence against the null hypothesis; null is plausible
$0.05 \leq$	*p*-value	< 0.10	little evidence against the null hypothesis
$0.01 \leq$	*p*-value	< 0.05	evidence against the null hypothesis
	p-value	< 0.01	strong evidence against the null hypothesis
Summary: the smaller the *p*-value, the stronger the evidence against the null hypothesis.			

5. Is the approximate *p*-value from your analysis small enough to provide much evidence against the null hypothesis? If so, how strong is the evidence?

2 Formulate conclusions

6. Do you consider the observed result to be statistically significant? Recall that this means that the observed result is unlikely to have occurred by chance alone.

7. How broadly are you willing to generalize your conclusions? Would you be willing to generalize your conclusions to cocaine addicts beyond the subjects of this study? Explain your reasoning.