Hypothesis Testing Applications

YOUR NAME

16 July, 2020

Exercises

1. Repeat the analysis of the yawning data from last lesson but this time use the hypergeometric distribution.

Is yawning contagious?

An experiment conducted by the *MythBusters*, a science entertainment TV program on the Discovery Channel, tested if a person can be subconsciously influenced into yawning if another person near them yawns. 50 people were randomly assigned to two groups: 34 to a group where a person near them yawned (treatment) and 16 to a group where there wasn't a person yawning near them (control). The following table shows the results of this experiment.

		Group		
		Treatment	Control	Total
Result	Yawn	10	4	14
	Not Yawn	24	12	36
	Total	34	16	50

The data is in the file "yawn.csv".

a. What are the hypotheses?

b. Calculate the observed statistic, pick a cell.
c. Find the p-value using the hypergeometric distribution.
d. Plot the the sampling distribution.
e. Determine the conclusion of the hypothesis test.
f. Compare your results with the randomization test.

2. Repeat the golf ball example using a different test statistic.

Use a level of significance of 0.05.

a. State the null and alternative hypotheses.

b. Compute a test statistic.

c. Determine the p-value.

d. Draw a conclusion.

3. Body Temperature

Shoemaker[1] cites a paper from the American Medical Association[2] that questions conventional wisdom that the average body temperature of a human is 98.6. One of the main points of the original article – the traditional mean of 98.6 is, in essence, 100 years out of date. The authors cite problems with Wunderlich's original methodology, diurnal fluctuations (up to 0.9 degrees F per day), and unreliable thermometers. The authors believe the average temperature is less than 98.6. Test the hypothesis.

a. State the null and alternative hypotheses.

b. State the significance level that will be used.
c. Load the data from the file "temperature.csv" and generate summary statistics and a boxplot of the temperature data. We will not be using gender or heart rate for this problem.
d. Compute a test statistic. We are going to help you with this part. We cannot do a randomization test since we don't have a second variable. It would be nice to use the mean as a test statistic but we don't yet know the sampling distribution of the sample mean.

Let's get clever. If the distribution of the sample is symmetric, this is an assumption but look at the boxplot and summary statistics to determine if you are comfortable with it, then under the null hypothesis the observed values should be equally likely to either be greater or less than 98.6. Thus our test statistic is the number of cases that have a positive difference between the observed value and 98.6. This will be a binomial distribution with a probability of success of 0.5. You must also account for the possibility that there are observations of 98.6 in the data.

e. Determine the p-value.

f. Draw a conclusion.

[1]L. Shoemaker Allen (1996) What's Normal? – Temperature, Gender, and Heart Rate, Journal of Statistics Education, 4:2

[2]Mackowiak, P. A., Wasserman, S. S., and Levine, M. M. (1992), "A Critical Appraisal of 98.6 Degrees F, the Upper Limit of the Normal Body Temperature, and Other Legacies of Carl Reinhold August Wunderlich," Journal of the American Medical Association, 268, 1578-1580.