Math 58B (only) - Introduction to Biostatistics

your name here

due Feb 27, 2020

Homework 5

Assignment Summary (Goals)

- relative risk
- odds ratios

1. **Praising Intelligence or Effort** (HW 3.6 ISCAM)

Psychologists investigated whether praising a child's intelligence, rather than praising his/her effort, tends to negative consequences such as undermining their motivation (Mueller & Dweck, 1998). Children participating in the study were given a set of problems to solve. After the first set of problems, half of the children were randomly assigned to be praised for their intelligence, whereas the other half was praised for their effort. The children were then given another set of problems to solve and later told how many they got right. They were then asked to write a report about the problems for other children to read, including information about how many they got right. Some of the children misrepresented (i.e., lied about) how many they got right, as shown in the following table:

	Praised for intelligence	Praised for effort	Total
Misrepresented their score (lied)	11	4	15
Did not misrepresent (did not lie)	18	26	44
Total	29	30	59

(a) Identify the explanatory and response variables in this study.

(b) For each group, determine the proportion who lied, and identify them with appropriate symbols.

Recall how index cards could be used to conduct a simulation analysis for determining whether the difference between these proportions is statistically significant.

```
i. how many cards you would use?
ii. how many would be marked how?
iii. how many you would deal out?
iv. which kinds of cards you would count?
v. what you would compare the results to, after you conducted a large number of repetitions?
```

59 cards are needed, one for each child. Mark 15 of the cards to represent the ones who lied, and mark the other 44 to represent the ones who did not lie. Shuffle up the cards and deal out 29 for the "intelligence" group and the other 30 for the "effort" group. Count how many of the cards marked as lying children are in the "intelligence" group. After repeating this process a large number of times, see how unusual it is to have 11 or more of the lying children/cards in the "intelligence" group.

(c) Use `infer` to conduct a simulation hypothesis test with 1000 repetitions. Provide the histogram which represents the sampling distribution of the difference in proportions under the null hypothesis, and report the empirical p-value.

(d) Provide a complete, detailed interpretation (in one or two sentences) of what this p-value means in this context (i.e., what is it the probability of, assuming what?)

(e) Based on this p-value, is the observed difference between the groups statistically significant at the $\alpha = 0.05$ level? Explain how you know.

(f) Summarize and justify your conclusion about whether the data provide evidence that praising a child's intelligence leads to more negative consequences than praising his/her effort.

2. **Conserving Hotel Towels?** (HW 3.8 ISCAM)

Many hotels have begun a conservation program that encourages guests to re-use towels rather than have them washed on a daily basis. A recent study examined whether one method of encouragement might work better than another. Different signs explaining the conservation program were placed in the bathrooms of the hotel rooms, with random assignment determining which rooms received which sign. One sign mentioned the importance of environmental protection, whereas another sign claimed that 75% of the hotel's guests choose to participate in the program. The researchers suspected that the latter sign, by appealing to a social norm, would produce a higher proportion of hotel guests who agree to re-use their towels. Researchers used the hotel staff (a mid-sized, mid-priced hotel in the Southwest that was part of a well-known national hotel chain) to record whether guests staying for multiple nights agreed to reuse their towel after the first night.

(a) Identify the observational units, explanatory variable, and response variable in this study.

(b) State the null and alternative hypotheses in symbols, and be sure to define the parameter in the context of this study.

The following table displays the observed data in this study:

towel	Social norm	Environmental protection	Total
Guest opted to re-use towel	98	74	172
Guest did not opt to re-use towel	124	137	261
Total	222	211	433

(c) Calculate the conditional proportions of re-use in each group. Also calculate the difference between them and the ratio of these proportions. (Feel free to use R as a calculator, in R the natural log function is `log()`.)

(d) Interpret what this ratio reveals in this context.

(e) Produce and interpret a 90% confidence interval for the ratio of probabilities of re-using towels (relative risk) between these two signs.

2. **Effectiveness of AZT (cont.)** (HW 3.12 ISCAM)

In 1993, one of the first studies aimed at preventing maternal transmission of AIDS to infants gave the drug AZT to pregnant, HIV-infected women (Connor et al., 1994). Roughly half of the women were randomly assigned to receive the drug AZT, and the others received a placebo (a "fake" treatment, same appearance as the drug but with no active ingredients). The HIV-infection status was then determined for 363 babies, 180 from the AZT group and 183 from the placebo group. Of the 180 babies whose mothers had received AZT, 13 were HIV-infected, compared to 40 of the 183 babies in the placebo group.

(a) Calculate and interpret the relative risk of HIV comparing the placebo group to the AZT group.

(b) Suggest why the researchers might prefer to look at the relative risk in this study rather than the difference in conditional proportions.

(c) Calculate and interpret a 95% confidence interval for the relative risk of HIV transmission between the population of placebo takers and the population of AZT takers. (Make sure you show your work, also, the natural log in R is the function `log()`.)

(d) Briefly explain, in your own words, why we are working with the log of the relative risk in these calculations.

(e) State the null and alternative hypotheses (in symbols and in words) for testing whether the risk of HIV-transmission is higher among placebo users than AZT users).

(f) Based on your confidence interval, do these data provide convincing evidence that placebo takers are more likely to transmit HIV to their babies than AZT takers? Explain how you are deciding.

(g) What population are you willing to generalize these results to?

(h) Does the design of this study allow you to conclude that use of AZT causes a lower risk of HIV-transmission? Justify your answer.