

## HW 7 – Math 58B

your name here

not due, 2022

### Assignment Summary (Goals)

- working with relative risk
- working with odds ratios

**Q1. PodQ** Describe one thing you learned from someone in your pod this week (it could be: content, logistical help, background material, R information, etc.) 1-3 sentences.

**Q2. 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. 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. "Success", in the study, means that the guest did agree to reuse the towel.

- Identify the observational units, explanatory variable, and response variable in the study.
- Consider the null hypothesis that states  $H_0 : p_{sn} = p_{ep}$  (each probability represents the true proportion of all people who will choose to re-use their towel when given the particular prompt). Rewrite the hypothesis in three different ways as functions of the two values:  $p_{sn}$  and  $p_{ep}$ . That is, each time, the null hypothesis will be written with a different parameter value. Each null hypothesis should be a function of  $p_{sn}$  and  $p_{ep}$  (i.e., a parameter) on the left of the equal sign and a single number on the right of the equal sign.
  - $H_0$  : the difference in proportions is \_\_\_\_\_
  - $H_0$  : the relative risk is \_\_\_\_\_
  - $H_0$  : the odds ratio is \_\_\_\_\_
- Recall that a natural log transformation creates a sampling distribution of  $\widehat{RR}$  and  $\widehat{OR}$  which are approximately normal. Therefore, we find the SE of the natural log transformed statistics in those cases.

Write down the general formula for the standard error (which is used when the mathematical (instead of computational) approximation is used) for each of the respective statistics:

- $SE(\hat{p}_{sn} - \hat{p}_{ep}) =$
- $SE\left(\ln(\hat{p}_{sn}/\hat{p}_{ep})\right) =$
- $SE\left(\ln\left(\frac{\hat{p}_{sn}/(1-\hat{p}_{sn})}{\hat{p}_{ep}/(1-\hat{p}_{ep})}\right)\right) =$

- (d) Calculate the statistic (and natural log of the statistic), the SE, and the Z score for each of the three measures.

The following table displays the observed data in the 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

- (e) Using any one of the calculated Z-scores (they should all three be *very* close to one another), find the associated two-sided p-value (use `xpnorm()`) and conclude the problem in the words of the study.

**Q3. 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.

- Calculate and interpret the relative risk of HIV positive baby comparing the placebo group (numerator) to the AZT group (denominator).
- Suggest why the researchers might prefer to look at the relative risk in the study rather than the difference in conditional proportions.
- [Part (c) won't be asked on the exam, we might get to it after spring break.] Use the mathematical approximation (i.e., the normal distribution) to 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. (Include/show your work.)
- Briefly explain, in your own words, why we are working with the log of the relative risk in the calculations.
- Use the bootstrap process to 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. (Include/show your work.)
- Based on your confidence interval, do the data provide convincing evidence that placebo takers are more likely to transmit HIV to their babies than AZT takers? Explain how you are deciding.
- To what population are you willing to generalize the results?
- Does the design of the study allow you to conclude that use of AZT causes a lower risk of HIV-transmission? Justify your answer.

```
praise()
```

```
## [1] "You are epic!"
```