Math 150, Spring 2021
Jo Hardin
WU # 6
in-class: Wednesday, 2/10/2021
due: Thursday, 2/11/2021

Your Name: ________________________________

Names of people you worked with: ________________________________

Instructions: Work on this problem in class with your group (if you are attending class synchronously) or out of class (hopefully with a person or two! if you are attending class asynchronously). The problem should be done on a piece of paper with a pencil or on some kind of tablet. The problem should **not** by typed up or done in LaTeX.

Work for a *maximum* of 15 minutes on the problem (regardless of what time you are working). *Do not* come back to the problem to "fix it up" or "finish it." Be sure to write down the names of the people you worked with during class (or outside of class).

Take a picture of your work and use a scanning app to create a pdf (or create a pdf directly from your tablet). Upload your work to Gradescope (via Sakai) within 24 hours of class.

Task: You remember from somewhere (maybe just earlier today) that we can approximate the distribution of the sample ln odds ratio using:

$$\ln(\widehat{OR}) \sim N\left(\ln(OR), \sqrt{\frac{1}{n_1\hat{p}_1(1-\hat{p}_1)} + \frac{1}{n_2\hat{p}_2(1-\hat{p}_2)}}\right)$$

Which leads you to the following 95% CI for the true $\ln(OR)$:

$$\left(\ln(\widehat{OR}) - 1.96 \cdot \sqrt{\frac{1}{n_1\hat{p}_1(1-\hat{p}_1)} + \frac{1}{n_2\hat{p}_2(1-\hat{p}_2)}}, \ln(\widehat{OR}) + 1.96 \cdot \sqrt{\frac{1}{n_1\hat{p}_1(1-\hat{p}_1)} + \frac{1}{n_2\hat{p}_2(1-\hat{p}_2)}}\right)$$

Filling in the pieces, the above 95% CI for the true value of $\ln(OR)$ (the parameter) is (2.71, 4.02).

Argue that $(e^{2.71}, e^{4.02})$ is a 95% confidence interval for the true value of OR by making the following argument (show your work):

1. If you happen to have gotten one of the 95% of samples that lead to a CI which captures the true value, then the first CI capturing the true $\ln(OR)$ implies that the second interval **must** capture the true OR.
2. If you were unlucky and happened to have gotten one of the 5% of samples which do not lead to a capture of the true value, then the first CI will not capture $\ln(OR)$ which implies that the second interval will **not** capture the true OR.
3. Thus, if the first interval (the one for the $\ln(OR)$) has a capture rate of 95%, then the second interval (where the endpoints are exponentiated) **must** also have a capture rate for OR of 95%.