STAT511 HW#10

Reading: Read Chapter 10 of Ott & Longnecker
See Canvas Calendar for Due Date.

26 points total, 2 points per problem part unless otherwise noted.

1. Bacillus Calmette-Guerin (BCG) is a vaccine for preventing tuberculosis. For this question, we will examine data from 3 studies (Vandiviere et al 1973, TPT Madras 1980, Coetzee & Berjak 1968). The data is summarized below.

Study	Trt Status	TBneg	TBpos
1	Ctrl	619	10
1	Trt	2537	8
2	Ctrl	87892	499
2	Trt	87886	505
3	Ctrl	7232	45
3	Trt	7470	29

    A. Calculate the odds ratio (corresponding to TBpos for Trt vs Ctrl) for each study separately. (**4 pts**)

    B. Use the Breslow-Day test to test for equality of odds ratios across the 3 studies. State your p-value and conclusion. Can we conclude that the odds ratios are equal across the 3 studies? Based on this test, should be combine information across studies? (**4 pts**)

A note about the BCG vaccine from Wikipedia:
The most controversial aspect of BCG is the variable efficacy found in different clinical trials that appears to depend on geography. Trials conducted in the UK have consistently shown a protective effect of 60 to 80%, but those conducted elsewhere have shown no protective effect, and efficacy appears to fall the closer one gets to the equator.

2. Problem 10.36 involves bomb hits during WWII. Bomb hits were recorded in n = 576 grids in a map of a region of South London.

# Bomb Hits	0	1	2	3	≥4	Total
# Grids	229	211	93	35	8	576

    A. Find the sample mean ($\hat{\mu}$) bomb hits per grid.

    B. Use the GOF test to test whether the number of bomb hits per grid follows the Poisson distribution. Calculate the GOF test statistic, df, p-value and give a conclusion using $\alpha = 0.05$. (**6 pts**)

3. The data "PoissonData.csv" gives observations Y (counts or events) for n = 50 (units) generated from the Poisson distribution (using the rpoiss() function).
   A. Calculate the sample mean and sample standard deviation. Also construct a histogram and qqplot of the data and include them in your assignment. (**4 pts**)
      NOTE: Because the data comes from the Poisson distribution, you should find that the mean and the sample variance (s^2) are close. However, you should also find from the histogram and qqplot that the data looks approximately normal.
   B. Give a standard t-based 95% confidence interval for μ.
   C. Following the example on CH10 Slide 106 (Death by Mule Kick CI), construct a 95% confidence interval for μ based on the normal approximation to the Poisson distribution. (**4 pts**) In order to do this, you will start by constructing a CI on the total number of events, then divide by the number of units.
   NOTE: The CIs from parts B and C should be similar.