Week 3

ZYF

```
library(epitools)
```

3.8 b

```
# Score interval
table <- matrix(c(60, 44, 2, 61), nrow=2)
fisher.test(table, conf.level = 0.95)$conf.int
```

```
[1]   9.89119 362.22970
attr(,"conf.level")
[1] 0.95
```

3.12

The table grand total $n = 422 + 381 + 273 + 299 + 365 + 232 = 1972$ observations. To find the marginal proportion for Republican is $\frac{273+232}{1972} = 0.256$ or 25.6. To find the marginal proportion for female is $\frac{422+381+273}{1972} = 0.545$ or 54.5. Then to find the expected frequency, we do (0.256)(0.545)(1972)275.13 with some rounding difference.

Both the chi squared and likelihood ratio chi squared tests are testing for H_0 : independence between Party ID and gender. Both test have a chi squared test statistic with $df = (I-1)(J-1) = 1 * 2 = 2$. At a 0.05 significance level, we reject the null as both p-values are smaller. This means, we have strong evidence that Gender and Party ID are associated.

Residual tells us how much the actual observation differs from the expected counts. The standardized residual gives us a more systematic way to compare. In multiple cases, the standardized residual have absolute value greater than 2, which tells us abnormality if the two variables are truely independent.

#3.25 Recall the SE of log OR is

$$\hat{\sigma}(\log \hat{\theta}) = \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}$$

Alternatively, with two binomial samples with obs y_1, y_2 and sample sizes n_1, n_2 with corresponding probabilities π_2, π_2, we can rewrite

$$\hat{\sigma}^2(\log \hat{\theta}) = \frac{1}{y_1} + \frac{1}{n - y_1} + \frac{1}{y_2} + \frac{1}{n - y_2}$$

Furthermore,

$$\hat{\sigma}^2(\log \hat{\theta}) = \frac{1}{n_1\pi_1} + \frac{1}{n_1 - n_1\pi_1} + \frac{1}{n_2\pi_2} + \frac{1}{n_2 - n_2\pi_2}$$

Taking partial wrt π_1 and π_2,

$$\frac{\partial}{\partial \pi_1} se(\log \hat{\theta}) = -\frac{1}{n_1\pi_1^2} + \frac{1}{n_1(1-\pi_1)^2}$$

$$\frac{\partial}{\partial \pi_2} se(\log \hat{\theta}) = -\frac{1}{n_2\pi_2^2} + \frac{1}{n_2(1-\pi_2)^2}$$

Solving for either by setting either to 0 gives us $\pi_1 = 0.5$ for $i = 1, 2$. Hence, the asymptotic variance is minimized when each binomial probability is exactly 0.5.