Lecture 2

1 Contingency Table Association Measures and Inference for 2x2 Tables

The calculations for the aspirin use vs. heart attack example from class are presented here using R. Note that there is some rounding error in the calculations from class.

```
rm(list=ls())
x = matrix(c(189,104,10845,10933), nrow=2)
rownames(x) = c("Placebo","Aspirin")
colnames(x) = c("Heart Attack","No Heart Attack")
x
```

```
##         Heart Attack No Heart Attack
## Placebo          189           10845
## Aspirin          104           10933
```

1.1 Conditional Probabilities of Heart Attack

```
pi1 = x[1,1] / sum(x[1,])
pi1
```

```
## [1] 0.01712887
```

```
pi2 = x[2,1] / sum(x[2,])
pi2
```

```
## [1] 0.00942285
```

1.2 Difference in Proportions

```
diff = pi1 - pi2
diff
```

```
## [1] 0.007706024
```

```
se_diff = sqrt(pi1*(1-pi1)/sum(x[1,]) + pi2*(1-pi2)/sum(x[2,]))
se_diff
```

```
## [1] 0.001539964
```

```
l = diff - qnorm(1-(.05/2))*se_diff
u = diff + qnorm(1-(.05/2))*se_diff
c(l,u)
```

```
## [1] 0.004687751 0.010724297
```

We are 95% confident that the interval [0.0046878, 0.0107243] contains the true difference in proportions. This interval does not contain 0!

1.3 Relative Risk

```
r = pi1/pi2
r
```

```
## [1] 1.817802
```

```
logr = log(r)
logr
```

```
## [1] 0.597628
```

```
se_logr = sqrt((1-pi1)/x[1,1] + (1-pi2)/x[2,1])
se_logr
```

```
## [1] 0.1213473
```

```
l = logr - qnorm(1-(.05/2))*se_logr
u = logr + qnorm(1-(.05/2))*se_logr
c(l,u)
```

```
## [1] 0.3597917 0.8354642
```

The probability of a heart attack for subjects taking a placebo is 1.8178018 times the probability of a heart attack for subjects taking aspirin. We are 95% confident that the interval [0.3597917, 0.8354642] contains the true log relative risk. This interval does not contain 0!

```
expl = exp(l)
expu = exp(u)
c(expl,expu)
```

```
## [1] 1.433031 2.305884
```

We are 95% confident that the interval [1.4330309, 2.3058842] contains the true relative risk. This interval does not contain 1!

1.4 Odds Ratio

```
OR = (x[1,1]*x[2,2]) / (x[1,2]*x[2,1])
OR
```

```
## [1] 1.832054
```

```
logOR = log(OR)
logOR
```

```
## [1] 0.6054377
```

```
se_logOR = sqrt(sum(1/x))
se_logOR
```

```
## [1] 0.1228416
```

```
l = logOR - qnorm(1-(.05/2))*se_logOR
u = logOR + qnorm(1-(.05/2))*se_logOR
c(l,u)
```

```
## [1] 0.3646726 0.8462029
```

The odds of a heart attack for subjects taking a placebo is 1.8320539 times the odds of a heart attack for subjects taking aspirin. The odds ratio and relative risk are similar because the probability of a heart attack is small. We are 95% confident that the interval [0.3646726, 0.8462029] contains the true log odds ratio. This interval does not contain 0!

```
expl = exp(l)
expu = exp(u)
c(expl,expu)
```

```
## [1] 1.440042 2.330780
```

We are 95% confident that the interval [1.4400424, 2.3307798] contains the true odds ratio. This interval does not contain 1!