Part 6: Assignment

1. Evaluate the birthweight2 data

```
setwd("/Users/akamau/Work/OneDrive - Kemri Wellcome Trust/Stats forum/Stat training")
bw.data<-read.csv("Data/birthweight2.csv", header=TRUE)

names(bw.data)
class(bw.data$matage)
class(bw.data$ht)
class(bw.data$gestwks)
```

```
library(epitools)
library(epiR)
```

2. Look at the association between LBW and gestational weeks. Divide gestational week in two groups i.e. group1 < 35 weeks and group2 >= 35 weeks. Then calculate the odds ratio between LBW and gestational weeks.

```
mytable1 <- table(bw.data$lbw, bw.data$gestwks)
mytable1
```

```
##
##                  25  26  28  29  30  31  32  33  34  35  36  37  38  39  40
##   Normal 2500+    0   0   0   0   0   0   0   0   5   8  11  30  87 167 146
##   Weight<2500     1   1   3   1   3   5   7   6   7   9   6  11  14   3   3
##
##                  41  42
##   Normal 2500+   87  20
##   Weight<2500     0   0
```

```
bw.data$gestwks1 <- NULL
bw.data$gestwks1[bw.data$gestwks < 35] <- 1
bw.data$gestwks1[bw.data$gestwks > 34] <- 2
bw.data$gestwks1 <- factor(bw.data$gestwks1 , labels = c("< 35 weeks" , ">= 35 weeks"))

mytable2 <- table(bw.data$gestwks1, bw.data$lbw)
mytable2
```

```
##
##                Normal 2500+ Weight<2500
##   < 35 weeks              5          34
##   >= 35 weeks           556          46
```

```
chisq.test(mytable1)
```

```
## Warning in chisq.test(mytable1): Chi-squared approximation may be incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  mytable1
## X-squared = 301.99, df = 16, p-value < 2.2e-16
```

```
chisq.test(mytable2)
```

```
## Warning in chisq.test(mytable2): Chi-squared approximation may be incorrect
```

```
##
```

```
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  mytable2
## X-squared = 204.92, df = 1, p-value < 2.2e-16
```

```
fisher.test(mytable2)
```

```
##
##  Fisher's Exact Test for Count Data
##
## data:  mytable2
## p-value < 2.2e-16
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.003574824 0.033832466
## sample estimates:
## odds ratio
## 0.01236775
```

```
mytable3<-cbind(mytable2[,2],mytable2[,1])
colnames(mytable3)<-c("Weight<2500","Normal 2500+")
mytable3
```

```
##             Weight<2500 Normal 2500+
## < 35 weeks           34            5
## >= 35 weeks          46          556
```

```
mytable2
```

```
##
##               Normal 2500+ Weight<2500
##   < 35 weeks             5          34
##   >= 35 weeks          556          46
```

```
epitab(mytable3, method="oddsratio", rev= "both")
```

```
## Warning in chisq.test(xx, correct = correction): Chi-squared approximation
## may be incorrect
```

```
## $tab
##             Normal 2500+          p0 Weight<2500    p1 oddsratio    lower
## >= 35 weeks          556 0.991087344          46 0.575    1.0000       NA
## < 35 weeks             5 0.008912656          34 0.425   82.1913 30.67057
##                upper     p.value
## >= 35 weeks       NA          NA
## < 35 weeks  220.2571 4.58709e-29
##
## $measure
## [1] "wald"
##
## $conf.level
## [1] 0.95
##
## $pvalue
## [1] "fisher.exact"
```

Using epiR package

```r
epi.2by2(mytable3, method="cohort.count",conf.level=0.95)
```

```
##              Outcome +    Outcome -      Total        Inc risk *
## Exposed +           34            5         39             87.18
## Exposed -           46          556        602              7.64
## Total               80          561        641             12.48
##                   Odds
## Exposed +       6.8000
## Exposed -       0.0827
## Total           0.1426
##
## Point estimates and 95% CIs:
## -------------------------------------------------------------------
## Inc risk ratio                               11.41 (8.43, 15.44)
## Odds ratio                                   82.19 (30.67, 220.26)
## Attrib risk *                                79.54 (68.83, 90.24)
## Attrib risk in population *                  4.84 (1.52, 8.16)
## Attrib fraction in exposed (%)               91.24 (88.14, 93.52)
## Attrib fraction in population (%)            38.77 (27.10, 48.58)
## -------------------------------------------------------------------
##  Test that odds ratio = 1: chi2(1) = 212.138 Pr>chi2 = < 0.001
##  Wald confidence limits
##  CI: confidence interval
##  * Outcomes per 100 population units
```

3. Check the Odds ratio for the association between LBW and gender

```r
mytable4 <- table( bw.data$sex, bw.data$lbw)
mytable4
```

```
##
##           Normal 2500+ Weight<2500
##   Female           270          45
##   Male             291          35
```

```r
mytable5<-cbind(mytable4[,2],mytable4[,1])
colnames(mytable5)<-c("Weight<2500","Normal 2500+")
mytable5
```

```
##        Weight<2500 Normal 2500+
## Female          45          270
## Male            35          291
```

```r
epitab(mytable5, method="oddsratio", rev = "both")
```

```
## $tab
##        Normal 2500+        p0 Weight<2500     p1 oddsratio     lower
## Male            291 0.5187166          35 0.4375  1.000000        NA
## Female          270 0.4812834          45 0.5625  1.385714 0.8645583
##           upper   p.value
## Male         NA        NA
## Female 2.221023 0.1895456
##
## $measure
## [1] "wald"
##
```

```
## $conf.level
## [1] 0.95
##
## $pvalue
## [1] "fisher.exact"
```

Using epiR package

```r
epi.2by2(mytable5, method="cohort.count",conf.level=0.95)
```

```
##              Outcome +    Outcome -      Total        Inc risk *
## Exposed +           45          270        315              14.3
## Exposed -           35          291        326              10.7
## Total               80          561        641              12.5
##                 Odds
## Exposed +      0.167
## Exposed -      0.120
## Total          0.143
##
## Point estimates and 95% CIs:
## -------------------------------------------------------------------
## Inc risk ratio                               1.33 (0.88, 2.01)
## Odds ratio                                   1.39 (0.86, 2.22)
## Attrib risk *                                3.55 (-1.57, 8.67)
## Attrib risk in population *                  1.74 (-2.48, 5.97)
## Attrib fraction in exposed (%)               24.85 (-13.66, 50.31)
## Attrib fraction in population (%)            13.98 (-8.59, 31.85)
## -------------------------------------------------------------------
##  Test that odds ratio = 1: chi2(1) = 1.848 Pr>chi2 = 0.174
##  Wald confidence limits
##  CI: confidence interval
##  * Outcomes per 100 population units
```