Epi Assignment #2

```
wcgs$chd69 <- as.integer(wcgs$chd69 == "Yes")
wcgs <- wcgs[wcgs$chol < 645,]
```

1.

```
smoker <- glm(chd69 ~ age + chol + sbp + bmi, data = wcgs)
summary(smoker)
```

```
##
## Call:
## glm(formula = chd69 ~ age + chol + sbp + bmi, data = wcgs)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -0.29597  -0.11055  -0.06736  -0.02485   1.01913
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.6029892  0.0664342  -9.076  < 2e-16 ***
## age          0.0045413  0.0008804   5.158 2.65e-07 ***
## chol         0.0008569  0.0001128   7.599 3.91e-14 ***
## sbp          0.0017235  0.0003375   5.107 3.47e-07 ***
## bmi          0.0024005  0.0019464   1.233    0.218
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.07154479)
##
##     Null deviance: 235.14  on 3140  degrees of freedom
## Residual deviance: 224.36  on 3136  degrees of freedom
##   (12 observations deleted due to missingness)
## AIC: 636.6
##
## Number of Fisher Scoring iterations: 2
```

y = B0 + B1(age) + B2(chol)+ B3(sbp) + B4(bmi) y = -0.603 + 0.005(60) + 0.001(200) + 0.002(150) + 0.002(20)

```
-0.603 + 0.005*60 + 0.001*250 + 0.002*150 + 0.002*20
```

```
## [1] 0.287
```

The log odds is 0.287 for a 60-year old smoker with the characteristics given.

2. y = B0 + B1(age) + B2(chol)+ B3(sbp) + B4(bmi) y = -0.603 + 0.005(60) + 0.001(200) + 0.002(150) + 0.002(20)

```
-0.603 + 0.005*60 + 0.001*200 + 0.002*150 + 0.002*20
```

```
## [1] 0.237
```

The log odds is 0.237 for a 60-year old smoker with the characteristics given a 50 mg/dL decrease.

3.

```
exp(0.237) - exp(0.287)
```

```
## [1] -0.0649831
```

The odds ratio associated with a 50 mg/dL increase in cholesterol is -0.065. This means that the odds of having coronary heart disease is 0.065 times higher with a 50 mg/dL increase in cholesterol if everything else is kept fixed.

4. y = B0 + B1(age) + B2(chol)+ B3(sbp) + B4(bmi) y = -0.603 + 0.005(70) + 0.001(200) + 0.002(150) + 0.002(20)

```
-0.603 + 0.005*70 + 0.001*250 + 0.002*150 + 0.002*20
```

```
## [1] 0.337
```

The log odds is 0.337 for a 70-year old smoker with the characteristics given.

y = B0 + B1(age) + B2(chol)+ B3(sbp) + B4(bmi) y = -0.603 + 0.005(60) + 0.001(200) + 0.002(150) + 0.002(20)

```
-0.603 + 0.005*70 + 0.001*200 + 0.002*150 + 0.002*20
```

```
## [1] 0.287
```

The log odds is 0.287 for a 70-year old smoker with the characteristics given.

```
exp(0.287) - exp(0.337)
```

```
## [1] -0.06831485
```

The odds ratio associated with a 50 mg/dL increase in cholesterol is -0.068. This means that the odds of having coronary heart disease is 0.068 times higher with a 50 mg/dL increase in cholesterol if everything else is kept fixed. The odds are higher as the age has been increased by 10 years.

1.2 Log-binomial 1.

```
betas.smoker <- coef(smoker) # beta coefficients from logistic model
se.beta.smoker <- sqrt(diag(vcov(smoker))) # Standard Error of betas

OR.smoker<-exp(coef(smoker)) # take anti-log of betas
ci.OR.smoker <- exp(confint.default(smoker)) # take anti-log of CI on log-odds scale

# Combine above into single matrix:
output.smoker<-cbind(betas.smoker,
                     se.beta.smoker,
                     OR.smoker,
                     ci.OR.smoker)
colnames(output.smoker)<-c("Beta","SE.beta","OR","95% LL","95% UL")
round(output.smoker,digits=3) # Estimated beta coefficients and ORs
```

```
##               Beta SE.beta    OR 95% LL 95% UL
## (Intercept) -0.603   0.066 0.547  0.480  0.623
## age          0.005   0.001 1.005  1.003  1.006
## chol         0.001   0.000 1.001  1.001  1.001
## sbp          0.002   0.000 1.002  1.001  1.002
## bmi          0.002   0.002 1.002  0.999  1.006
```

```
smoker.1 <- glm(chd69 ~ age + chol + sbp + bmi, data = wcgs, family = binomial())
summary(smoker.1)
```

```
##
## Call:
## glm(formula = chd69 ~ age + chol + sbp + bmi, family = binomial(),
##     data = wcgs)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.2215  -0.4430  -0.3365  -0.2563   2.8167
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -11.560863   0.952454 -12.138  < 2e-16 ***
## age           0.061596   0.011812   5.215 1.84e-07 ***
## chol          0.011463   0.001507   7.608 2.79e-14 ***
## sbp           0.019418   0.004039   4.808 1.52e-06 ***
## bmi           0.039054   0.026278   1.486    0.137
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1774.2  on 3140  degrees of freedom
## Residual deviance: 1635.5  on 3136  degrees of freedom
##   (12 observations deleted due to missingness)
## AIC: 1645.5
##
## Number of Fisher Scoring iterations: 5
```

y = B0 + B1(age) + B2(chol)+ B3(sbp) + B4(bmi)

```
-11.561 + 0.062*60 + 0.011*250 + 0.019*150 + 0.0391*20
```

```
## [1] -1.459
```

The log-risk is -1.459 for a 60-year old smoker with the characteristics given.

2. y = B0 + B1(age) + B2(chol)+ B3(sbp) + B4(bmi)

```
-11.561 + 0.062*60 + 0.011*200 + 0.019*150 + 0.0391*20
```

```
## [1] -2.009
```

The log-risk is -2.009 for a 60-year old smoker with the characteristics given with a 50 mg/dL decrease in cholesterol. 3.

```
exp(-2.009) - exp(-1.459)
```

```
## [1] -0.0983459
```

The relative risk associated with a 50 mg/dL increase in cholesterol is -0.098. This means that the risk of having coronary heart disease is 0.098 times higher with a 50 mg/dL increase in cholesterol if everything else is kept fixed.

There is only a slight difference in the relative risk being higher and this could possibly be due to the increase in cholesterol. The log-risk estimate is higher for 50 mg/dL increase and it means that he has a higher risk of chd.

4. 70 year olds y = B0 + B1(age) + B2(chol)+ B3(sbp) + B4(bmi)

```
-11.561 + 0.062*70 + 0.011*250 + 0.019*150 + 0.0391*20
```

```
## [1] -0.839
```

The log-risk is -0.839 for a 70-year old smoker with the characteristics given.

y = B0 + B1(age) + B2(chol)+ B3(sbp) + B4(bmi)

```
-11.561 + 0.062*70 + 0.011*200 + 0.019*150 + 0.0391*20
```

```
## [1] -1.389
```

The log-risk is -1.389 for a 70-year old smoker with the characteristics given.

```
exp(-0.839) - exp(-1.389)
```

```
## [1] 0.1828179
```

The relative risk associated with a 50 mg/dL increase in cholesterol is 0.183. This means that the risk of having coronary heart disease is 0.183 times higher with a 50 mg/dL increase in cholesterol if everything else is kept fixed for the 70 year old.