HW#7, Nan Deng

1

(a)

```
icu_data <- read.csv("/Users/CandiceDeng\ 1/Desktop/STATS500/HW#7/icu.csv")
icu_data$race <- factor(icu_data$race)
head(icu_data,5)

##   sta age can cpr inf race
## 1   0  27   0   0   1    1
## 2   0  59   0   0   0    1
## 3   0  77   0   0   0    1
## 4   0  54   0   0   1    1
## 5   0  87   0   0   1    1

icu_fit <- glm(sta~.,family=binomial,data=icu_data)
icu_fit

##
## Call:  glm(formula = sta ~ ., family = binomial, data = icu_data)
##
## Coefficients:
## (Intercept)          age          can          cpr          inf
##     -4.1075       0.0284       0.2607       1.5394       0.8800
##       race2        race3
##      0.9163       0.4227
##
## Degrees of Freedom: 199 Total (i.e. Null);  193 Residual
## Null Deviance:       200.2
## Residual Deviance: 176.3     AIC: 190.3

icu_reduce1 <- glm(sta~age+can+race,family=binomial,data=icu_data)
anova(icu_reduce1,icu_fit)

## Analysis of Deviance Table
##
## Model 1: sta ~ age + can + race
## Model 2: sta ~ age + can + cpr + inf + race
##   Resid. Df Resid. Dev Df Deviance
## 1       195     189.13
## 2       193     176.29  2   12.845

1-pchisq(12.845,2)

## [1] 0.00162459

beta <- coef(icu_fit)[4:5]
varb <- vcov(icu_fit)[4:5,4:5]
varbi <- solve(varb)
w <- t(beta) %*% varbi %*% beta
paste('w =',w)

## [1] "w = 12.3727100895835"

paste('p-value = ',1-pchisq(w,2))
```

```
## [1] "p-value =  0.00205731191322589"
```

For the likelihood ratio test, the chi-square test statistics is 12.845 with 2 degree of freedom. The null hypothesis should be rejected because p-value is 0.00162459 which is less than 0.05. For wald test, on the other hand, the test statistic is 12.3727100895835 with 2 degree of freedom. The null hypothesis should also be rejected because p-value is 0.00205731191322589 which is less than 0.05.

(b)

```
library(aod)
x <- model.matrix(icu_fit)[,-1]
new_icu <- data.frame(sta=icu_fit$y,x)
icu_reduce2 <- glm(sta~age+can+cpr+inf+I(race2+race3), family=binomial,data=new_icu)
anova(icu_reduce2,icu_fit)
```

```
## Analysis of Deviance Table
##
## Model 1: sta ~ age + can + cpr + inf + I(race2 + race3)
## Model 2: sta ~ age + can + cpr + inf + race
##   Resid. Df Resid. Dev Df Deviance
## 1       194     177.25
## 2       193     176.29  1  0.96205
```

```
1-pchisq(0.96205, 1)
```

```
## [1] 0.3266709
```

```
wald.test(b=coef(icu_fit),Sigma=vcov(icu_fit),L=cbind(0,0,0,0,0,1,-1))
```

```
## Wald test:
## ----------
##
## Chi-squared test:
## X2 = 0.95, df = 1, P(> X2) = 0.33
```

For the likelihood ratio test, the chi-square test statistics is 0.96205 with 1 degree of freedom. The null hypothesis should not be rejected because p-value is 0.3266709 which is greater than 0.05. For wald test, on the other hand, the test statistic is 0.95 with 2 degree of freedom. The null hypothesis should also not be rejected because p-value is 0.33 which is greater than 0.05.

(c)

```
beta <- summary(icu_fit)$coef
mean <- as.numeric(beta[6] - beta[7])
amatrix <- vcov(icu_fit)[6:7, 6:7]
var_beta <- amatrix[1, 1]+amatrix[2, 2]-2*amatrix[1, 2]
std <- sqrt(var_beta)
paste('95% CI = ', paste(mean-qnorm(0.975,0,1)*std,mean+qnorm(0.975,0,1)*std))
```

```
## [1] "95% CI =  -0.501131458565487 1.48831148719703"
```

The 95% confidence interval for β_{race2}-β_{race3} is to detect how the response sta varies between predictors race2 and race3 while holding other predictors constant. The interval is (-0.501131458565487, 1.48831148719703).

(d)

For the binary response model, unlike the usual form $2[l(\beta_S)-l(\beta_\Omega)]$, the likelihood under a saturated model is not feaible and the Residual Deviance is computed as $-2l(\beta_\Omega)$. Under this circumstance, the Residual Deviance does not follow chi-squared distribution. Therefore, this Residual Deviance cannot be appropriately considered as the sum of the squared deviance residuals, which means it should not be used as a Goodness-of-Fit statistic for this model.

2

(a)

Define $f(p) = \frac{p}{1-p} \Rightarrow \frac{\partial f(p)}{\partial p} = \frac{1}{(1-p)^2}$

$$
\begin{aligned}
Var(f(\hat{p}) - f(p)) &= f'(p)\cdot(\hat{p}-p) \\
&= [f'(p)]^2 \cdot Var(\hat{p}-p) \\
&= \left[\frac{1}{(1-p)^2}\right]^2 \cdot p(1-p) \\
&= \frac{p}{(1-p)^3}
\end{aligned}
$$

$$\sqrt{n}\left(\frac{\hat{p}}{1-\hat{p}} - \frac{p}{1-p}\right) \xrightarrow{d} N\left(0, \frac{p}{(1-p)^3}\right)$$

(b)

$$
\begin{aligned}
Var\left(\frac{\hat{p}}{1-\hat{p}}\right) &= Var\left(\frac{\hat{p}}{1-\hat{p}} - \frac{p}{1-p}\right) = \frac{p}{(1-p)^3} \cdot \frac{1}{(\sqrt{n})^2} \\
&= \frac{p}{n(1-p)^3}
\end{aligned}
$$