Homework Assignments
Applied Logistic Regression

Week 2

Exercise 2:

Use the ICU study (icu.dta)

Using the results of the output from the logistic regression package used for exercise 2 part (D) of week one, assess the significance of the slope coefficient for AGE using the likelihood ratio test and the Wald test. What assumptions are needed for the p-values computed for each of these tests to be valid? Are the results of these tests consistent with one another? What is the value of the deviance for the fitted model?

```
. logit STA AGE

Iteration 0:   log likelihood = -100.08048     Likelihood of Constants only Model
Iteration 1:   log likelihood = -96.288372
Iteration 2:   log likelihood = -96.153701
Iteration 3:   log likelihood = -96.15319      Likelihood of Full  Model

Logit estimates                                 Number of obs   =        200
                                                LR chi2(1)      =       7.85
                                                Prob > chi2     =     0.0051
Log likelihood = -96.15319                      Pseudo R2       =     0.0392


--------------------------------------------------------------------------------
     STA |      Coef.   Std. Err.       z     P>|z|      [95% Conf. Interval]
---------+----------------------------------------------------------------------
     AGE |   .0275426   .0105645     2.607    0.009      .0068366    .0482487
   _cons |  -3.058513   .6961091    -4.394    0.000     -4.422862   -1.694165
--------------------------------------------------------------------------------
```

        The Likelihood Ratio Test in STATA

```
. estimates store A This saves the test results for the 'full' model in A

. quietly logit STA This Estimates a simpler model without AGE

. lrtest A .        LR test of the last model fit versus model A

likelihood-ratio test                           LR chi2(1)  =      7.85
(Assumption: . nested in A)                      Prob > chi2 =    0.0051

. logistic STA AGE

Logit estimates                                 Number of obs   =        200
                                                LR chi2(1)      =       7.85
                                                Prob > chi2     =     0.0051
Log likelihood = -96.15319                      Pseudo R2       =     0.0392


--------------------------------------------------------------------------------
     STA | Odds Ratio   Std. Err.       z     P>|z|      [95% Conf. Interval]
---------+----------------------------------------------------------------------
     AGE |   1.027925   .0108595     2.607    0.009       1.00686    1.049432
--------------------------------------------------------------------------------
```

Deviance:

$$D = -2\ln\left[\frac{(\text{likelihood of the current model})}{(\text{likelihood of the saturated model})}\right]$$

$$D = -2(-96.15319)$$

$$D = 192.30638$$

Likelihood Ratio Test:

$$H_0 : \beta_1 = 0$$

$$H_A : \beta_1 \neq 0$$

$$G = D(\text{model without variable}) - D(\text{model with variable})$$

$$G = 200.16 - 192.31$$

$$G = 7.85 \qquad\qquad G \sim \chi^2(1) \qquad\qquad p = 0.0051$$

$\therefore$ reject H_0, it is not consistent with the data that $\beta_1 = 0$; we conclude that AGE is a significant predictor of STA.

Assumption: the statistic G will follow a χ^2 distribution with 1 degree of freedom under the null hypothesis.

Wald Test:

$$H_0 : \beta_1 = 0$$

$$H_A : \beta_1 \neq 0$$

$$W = \frac{\hat{\beta}_1}{S\hat{E}(\hat{\beta}_1)} = 2.607 \qquad W \sim N(0,1) \qquad\qquad p = 0.009$$

$\therefore$ reject H_0, it is not consistent with the data that $\beta_1 = 0$; we conclude that AGE is a significant predictor of STA.

Assumption:

The Wald statistic will follow a normal distribution with mean 0 and variance 1.

The results of the likelihood ratio test and the Wald test are consistent. Each test indicates that the model is significant.

The value of the deviance for the fitted model is $D = 192.31$.