Homework Solutions
Applied Logistic Regression

WEEK 5

f. Use the Wald test to assess the significance of the coefficient for **female**.

The test is given in the computer output after fitting the regression **("logit nas135 female")**. Z is 4.39, which yields a p-value<0.0001.

Alternatively, you can hand-calculate the Wald statistic:

```
. di 1.22596/0.279546
4.3855394

. di 1-normal(4.3855)
5.786e-06
```

g. Fit a model with **runtime** as the only independent variable. Assess the significance of the model.

```
. logit nas135 runtime, nolog

Logistic regression                                     Number of obs   =        477
                                                        LR chi2(1)      =      25.35
                                                        Prob > chi2     =     0.0000
Log likelihood = -167.77184                             Pseudo R2       =     0.0702

------------------------------------------------------------------------------
      nas135 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     runtime |   .0155019   .0030909     5.02   0.000     .0094439    .0215599
       _cons |  -5.592594    .771282    -7.25   0.000    -7.104278   -4.080909
------------------------------------------------------------------------------
```

This model is significantly better than the naïve model. The LR test gives a p-value<0.0001.

h. Calculate the probability of hyponatremia of a runner who takes 4 hours (240 minutes) to complete the marathon.

First, calculate the estimate for the logit when x=240 given the coefficients in the fitted model:

```
. di -5.5926+0.0155*240
-1.8726
```

Next. calculate the odds by exponentiating the coefficient.

```
. di exp(-5.5926+0.0155*240)
.15372346
```

Finally, calculate the probability given that

$$Probability = \frac{Odds}{1 + Odds}$$

```
. di 0.1537/(1+0.1537)
.13322354
```

The probability of hyponatremia of a runner who takes 4 hours is 13.3%

i. Fit a model with **female** and **runtime** as independent variables. Assess the significance of the model. Which null hypothesis is tested?

```
. logit nas135 female runtime, nolog

Logistic regression                             Number of obs   =        477
                                                LR chi2(2)      =      36.42
                                                Prob > chi2     =     0.0000
Log likelihood = -162.23985                     Pseudo R2       =     0.1009

------------------------------------------------------------------------------
      nas135 |      Coef.   Std. Err.       z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female |   .9638364    .291049     3.31   0.001     .3933908    1.534282
     runtime |   .0142136   .0032947     4.31   0.000     .0077562     .020671
       _cons |  -5.721056    .823284    -6.95   0.000    -7.334663   -4.107449
------------------------------------------------------------------------------
```

The model is significant as it is shown by the LR test (p<0.0001). The null hypothesis being tested is that the coefficient for **female** and the coefficient for **runtime** are both zero.