2. Data Exercise

**a. The estimated effect on the probability of arrest if *pcnv* goes from 0.25 to 0.75?**

Estimate Std. Error t value Pr(>|t|)

pcnv -0.15438020 0.02093356 -7.3748 2.175e-13 \*\*\*

According to the regression result, if *pcnv* goes from 0.25 to 0.75 (a 0.5 increase), the probability of being arrested will decrease by 0.0772, which is 7.72%.

**b. Test the joint significance of *avgsen* and *tottime.***

Nonrobust result

Chi-squared test:

X2 = 0.36, df = 2, P(> X2) = 0.84

Robust result

Chi-squared test:

X2 = 0.37, df = 2, P(> X2) = 0.83

Both results show that we failed to reject the null

**c. Now estimate the model by probit.**

**d. Obtain the percent correctly predicted.**

predictedarr 0 1

0 1919 697

1 51 58

Percent correctly predicted when narr86 = 0 is 1919/(1919+51) = 0.9741 = 97.41%

Percent correctly predicted when narr86 = 1 is 58/(697+58) = 0.0768 = 7.68%

Percent correctly predicted overall is (1919+58)/2725 = 0.7255 = 72.55%

The model did a decent job on predicting the number of men not arrested but it underestimated the number of arrests.

**e. Add the quadratic terms to the model and test for individual/joint significance**

Individual test result

Estimate Std. Error z value Pr(>|z|)

pcnv 7.5977e-02 8.0340e-02 0.9457 0.344307

pcnvsq -2.4569e-01 8.1258e-02 -3.0235 0.002499 \*\*

pt86sq -1.3998e-02 2.0109e-03 -6.9611 3.375e-12 \*\*\*

inc86sq 3.3106e-06 1.0914e-06 3.0332 0.002419 \*\*

The test shows that all three quadratic terms are individually significant

Joint test result

Chi-squared test:

X2 = 71.9, df = 3, P(> X2) = 1.7e-15

The test shows that the three quadratic terms are jointly statistically significant.

The coefficient on *pcnv* is positive and the coefficient on pcnvsq is negative indicating that initially there is a positive relationship between the probability of arrest and *pcnv* and then the relationship goes negative. The turning point is when the first order condition with respect to *pcnv* is equal to 0.