PH242C/STAT247C

Assignment 1

Logistic Regression Review and Simulation

Run the associated dofile from STATA to help you complete the simulation and answer the questions below.

1. Based on the code used to simulate the data, describe the data-generating distribution, including the model of the regression, e.g., $logit\{E(Y|X_1,X_2)\}= b_0+b_1X_1+b_2X_2$, $b_0=-2.0$, $b_1=2.0$, $b_2=-2.25$.
2. Calculate the predicted value at $X_1=0, X_2=1$.
3. What is the true odds ratio when X_1 changes by 0.5, keeping X_2 fixed?
4. Repeat 1-3 for the estimated model but also give a 95% CI for the odds ratio. Do this for both sample sizes.
5. Interpret to the best of your ability every number in the output on the row of results starting with X_2 (for sample size $n=100$).
6. Calculate and describes what happens to the estimated standard deviation (that is, the SE) of the estimate of b_1 when the sample size increases to $n=500$ from $n=100$.