

Homework Assignment 2

CSSS 536: Logistic Regression and Log-Linear Modeling for the Social Sciences

Professor: Kevin Quinn, Political Science and CSSS
Autumn Quarter 2002

Due before class on October 25

Problem 1

Using the Hosmer and Lemeshow low birth weight data (available on the course website) do the following:

Fit a logistic regression model in which the response variable is the low birth weight indicator `low` and the only explanatory variable is the mother's last pre-pregnancy weight `lwt`.

Now fit a probit model in which in which the response variable is `low` and the only explanatory variable is `lwt`.

Plot `low` on `lwt`. Jitter the low birth weight indicator so that the number of positive and negative responses at each weight level can be more easily seen.

In the same plot, graph the fitted values from the logistic regression analysis against `lwt`. In the same plot, but using a different color or line thickness, plot the fitted values from the probit analysis on `lwt`.

Question a) The estimated coefficient on `lwt` from the logistic regression analysis is fairly different from the estimated coefficient on `lwt` from the probit analysis, yet the fitted values are extremely similar from the two models. What's going on here? Explain why this is the case.

Question b) Compute the probability that a 120 pound woman gives birth to a low birth weight child based on both the logit estimates and on the probit estimates. Now compute the probability that a 350 pound woman gives birth to a low birth weight child based on both the logit estimates and the probit estimates. For which level of `lwt` are the logit and probit predicted probabilities more closely in agreement? Why do you think this is the case?