**POLS 6481, Spring 2021**

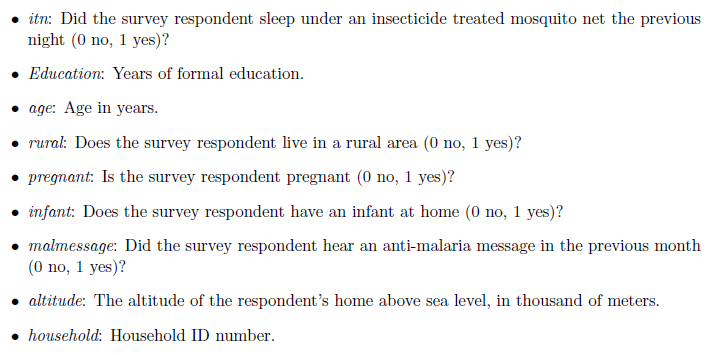
Instructor: Prof. Scott Basinger

Homework Assignment 4

Posted Friday, April 2, 2021

Due Thursday, April 8, 2021

Download the dataset ***Uganda\_ITN.txt*** which describes the behavior of nearly 2000 adult (ages 15-48) women in Uganda. Here is the codebook for this dataset:



Using this dataset, perform the following tasks:

1. Estimate a probit model that uses the three continuous explanatory variables (*education*, *age*, and *altitude*) to predict whether women use insecticide treated mosquito nets. Show the results of your analysis. Identify which variables are statistically significant and explain their effects on *itn*.
2. Plot the predicted probabilities that an individual used a mosquito net across the range of values of *age*; fix *education* and *altitude* equal to their mean or median values (be sure to state which).
3. Re-estimate the probit model in 1., but drop the variable that had statistically insignificant effects and add three binary variables – *rural*, *pregnant*, and *infant*. Explain the effects of each variable.
4. Plot the predicted probabilities that an individual used a mosquito net across the range of values of *age*; but now plot two lines – one with *rural* = 0 (18%) and the other with *rural* = 1 (82%).

Note that 10.4% of women in the sample were pregnant, and 18.4% of women in the sample had an infant at home; set *pregnant* and *infant* equal to their modes.

1. Create a new variable, *atrisk*, that combines women who are pregnant and women who have an infant at home. (You may use this line of code: data$atrisk = 1 - (1-data$pregnant)\*(1-data$infant) )

Run one more probit model that uses *age*, *altitude*, *rural*, and *atrisk* to predict whether women use insecticide treated mosquito nets and show your results.

1. Using this last model, calculate the probabilities of using a mosquito net for a woman who (a) is not pregnant and has no infant, or (b) either is pregnant or has an infant at home. Fix the values of *age* and *altitude* at their means, set *rural* at its mode, and vary *atrisk* between = 0 and = 1. How does being at risk change the probability of using a mosquito net?

Download the dataset ***FERTIL2.DTA*** which describes the behavior of more than 4,000 adult (ages 15-49) women in Botswana (data were collected by a Michigan State University undergraduate student, David Heakins, from a 1988 Demographic and Health Survey). Here is a codebook for key variables:

* *children*: number of children
* *educ*: years of education
* *urban*: whether the respondent lives in an urban ( = 1) or rural ( = 0) community
* *electric*: whether the home has electricity
* *catholic*: whether the respondent belongs to the Roman Catholic church
* *protest*: whether the respondent belongs to a Protestant church
* *spirit*: whether the respondent identifies as a Christian but has assimilated indigenous medical practices (*bongaka*) and religious practices (*badimo*), such as ancestral veneration [especially patriarchal ancestors], into their Christian beliefs

1. For the third and fourth variables shown above, calculate the average numbers of children in:
   1. Residents of rural areas versus residents of urban areas; also answer what percent of homes are in rural areas and urban areas
   2. Homes with without electricity versus homes with electricity; also answer what percent of homes do and do not have electricity

According to *Wikipedia*, Botswana is a Christian majority nation (over 70% currently) however according to one article, “ancestral spirits form a very important part of the religious thought of most Africans…. The present tenacity and resilience of the African traditional ritual and spirituality posed a challenge.”

1. Estimate a linear regression model that uses the one continuous explanatory variables (*educ*) and five binary explanatory variables (*urban*, *electric*, *catholic*, *protest*, and *spirit*) to predict the number of children a woman has. Show the results of your analysis. Identify which variables are statistically significant and explain their effects on *children*.
2. Assuming a linear relationship, plot the predicted number of children that a woman has across the range of values of *education* for two different types of areas: rural (*urban* = 0) and urban ( = 1).

Set the other binary explanatory variables (*electric*, *catholic*, *protest*, and *sprirt*)equal to zero. Beyond what value for *education* is a woman predicted to have zero or fewer children?

1. Re-estimate the regression model in 8. as a Poisson regression, keeping the same variables. Identify which variables are statistically significant and explain their effects on the number of *children*. Use techniques described in lecture to explain how a unit increase in an explanatory variable increases the mean number of children in *percentage* terms.
2. Plot the nonlinear relationship between education and the number of children that a woman has for two different types of areas: rural (*urban* = 0) and urban ( = 1). Set the other binary explanatory variables (*electric*, *catholic*, *protest*, and *sprirt*)equal to zero. How does the effect shown here differ from the linear relationship assumed in 8. and displayed in 9.?
3. Calculate predicted probabilities of the numbers of children for a woman with the median education level (*educ* = 7). Set three binary explanatory variables (*catholic*, *protest*, and *sprirt*)equal to 0, set the *urban* variable equal to 0.5, and vary the *electric* variable between 0 and 1.

In other words, perform the calculations of *mu* and *pr*(*y* = *k* | *mu*) twice to see the impact of electricity on fertility. Let *k* = 0, 1, 2, 3, 4, 5.