Modelli Statistici Avanzato

a.a. 2017/2018

Homework n. 6

With reference to: Sections 5.1.1 - 5.1.3 of the textbook (Fahrmeir *et al.*), sections D.6.1-D.6.4 in the Appendix of Fox and class slides on binary regression models (part I, II and III, downloadable from Moodle), briefly answer the following questions.

1. Define the linear probability model for a binary response and discuss the drawbacks of this model.
2. Describe the elements of a GLM (response distribution, linear predictor and link function) and how these elements specialize for the logit, probit and c-log-log models, respectively.
3. What are the main differences between the logit, probit and c-log-log models? Consider both the response functions and the parameters.
4. Show how the probit model can be derived by considering a latent continuous response variable.
5. Given the probability model $P(y_i = 1|\mathbf{x}_i) = h(\mathbf{x}_i'\boldsymbol{\beta})$ and the linear predictor $\eta_i = \mathbf{x}_i'\boldsymbol{\beta}$ derive the (conditional) effect of a continuous covariate x_k on the probability: (i) for the logit model, (ii) for the probit model.

APPLIED EXERCISE
(to be solved 'by hand', i.e. without using Stata or other statistical software)

The following Table reports the results of the estimation of the logit model for labor force participation of married women analyzed during the lessons of the last week.

```
Logistic regression                             Number of obs     =        753
                                                LR chi2(7)        =     226.22
                                                Prob > chi2       =     0.0000
Log likelihood = -401.76515                     Pseudo R2         =     0.2197

------------------------------------------------------------------------------
           inlf |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
       nwifeinc |  -.0213452   .0084214    -2.53   0.011    -.0378509   -.0048394
           educ |   .2211704   .0434396     5.09   0.000     .1360303    .3063105
          exper |   .2058695   .0320569     6.42   0.000     .1430391    .2686999
                |
c.exper#c.exper |  -.0031541   .0010161    -3.10   0.002    -.0051456   -.0011626
                |
            age |  -.0880244    .014573    -6.04   0.000     -.116587   -.0594618
        kidslt6 |  -1.443354   .2035849    -7.09   0.000    -1.842373   -1.044335
        kidsge6 |   .0601122   .0747897     0.80   0.422     -.086473    .2066974
          _cons |   .4254524   .8603697     0.49   0.621    -1.260841    2.111746
------------------------------------------------------------------------------
```

A.1 Write down the expression of the estimated linear predictor and of the estimated probability.
A.2 Compute the odds for a woman aged 40, without children, educ=10, experience=0, nwifeinc=0.
A.3 What is the effect of one more year of education on the odds?
A.4 Compute the predicted probability fixing the covariate values as in A.2.
A.5 Estimate the marginal effect of age on the probability computed in A.4.
A.6 Compute the discrete change in the probability in A.4 for one more year of education. Comment the result.
A.7 What is the maximum marginal effect of education on the probability?
A.8 Suppose we fit the model in the Table using the probit link: what do you expect for the estimates of the regression coefficients?