ECOM40006/ECOM90013 Econometrics 3
Department of Economics
University of Melbourne

Week 12 Tutorial Exercise

Semester 1, 2025

The primary function of this week's tutorial is revision and to provide you with the opportunity to ask your tutor questions that you may have from anywhere in the course. Challenge yourselves to challenge your tutor!

In the event that there is time remaining then you should attempt the following exercises.

1. The file `loanapp.dta` contains information on mortgage loan applicants in the US. The variables in the data set are described in the file. The pool of loan applicants is comprised of white, black and Hispanic people. Whether or not the application was approved is recorded in the variable *approve* which is equal to one if the application was successful and is zero otherwise.

   Load the package `haven`, and any dependencies, so that R can read Stata files and then read in `loanapp.dta`.

   (a) To test for discrimination in the mortgage loan market, estimate a probit model with *approve* as the dependent variable and a list of regressors which includes *hrat*, *obrat*, *loanprc*, *unem*, *male*, *married*, *dep*, *sch*, *cosign*, *chist*, *pubrec*, *mortlat1*, *mortlat2*, *vr*, and *white*. Do you find evidence of racial discrimination in the success of loan applications? Justify your answer.

   (b) Suppose that you add a new regressor that is an interaction (the product of the two variables) between *obrat* and *white*. Call the new variable *wobrat*. Does your conclusion about discrimination in the loan market changes? Justify your answer.

   (c) Explain how you might construct prediction intervals for the probability of success of a given applicant.

2. Deb and Trivedi (1997) analyze data on 4406 individuals, aged 66 and over, who are covered by Medicare, a public insurance program. Originally obtained from the US National Medical Expenditure Survey (NMES) for 1987/88, the data are available from the data archive of the Journal of Applied Econometrics at `http://qed.econ.queensu.ca/jae/1997-v12.3/deb-trivedi/`. The objective is to model the demand for medical care — as captured by the number of physician/non-physician office and hospital outpatient visits — by the covariates available for the patients.

Here, we adopt the number of physician office visits *ofp* as the dependent variable and use the health status variables *hosp* (number of hospital stays), *health* (self-perceived health status), *numchron* (number of chronic conditions), as well as the socioeconomic variables *gender*, *school* (number of years of education), and *privins* (private insurance indicator) as regressors.

(a) Create a subset of the data comprised only of columns 1, 6, 7, 8, 13, 15, 18. Call this dataset dt.

(b) Use the following plotting command and describe the features of the resulting plot.

   plot(table(dt$ofp), main="*mtext*", xlab="*xtext*", ylab="Frequency")

   Note: so that the command would fit on a single line of this document I have factored out *main text* and *xtext*. You should replace these bits (between the quotes) by the following text in the command that you provide to R:

   mtext = Frequency distribution for number of physician office visits.
   xtext = Number of physician office visits

(c) Use the following command to fit a Poisson regression model to this data:

   PReg=glm(ofp $\sim$., data = dt, family = poisson)
   summary(PReg)

   Note that the '.' in the glm command simply says: use all the remaining variables in dt as regressors.

(d) Interpret the estimated coefficient on *school*.

(e) Explain how you would construct standard errors on predicted values from this model.

Final Remarks

I should highlight the role that tutorials have played this semester. Although sometimes they have been used to reinforce ideas introduced during the lectures, they have frequently been used to expose you to material and ideas that we weren't going to get to during the regular lectures. As such your exam revision should probably include material covered during the tutorials.

Finally, although this is clearly your most important subject, your other subjects are important too. Good luck with all of your exams.

Cheers, Chris.