PhD Empirical Methods
Spring 2021

Problem Set 5 – Maximum Likelihood
Due 4/22/2021

In 1969, the popular magazine *Psychology Today* published a 101-question survey on extramarital affairs. Professor Ray Fair (1978) extracted a sample of 601 observations on men and women who were currently in their first marriage and analyzed their responses to the survey. He used the “Tobit” model as his estimation framework for this study. The dependent variable is a count of the number of affairs, so instead of a Tobit, a standard Poisson model may be a better choice. Download the data set from github (ps5.mat), and estimate the parameters to the model below using the methods that I describe below.

Data Description

- y - count data: number of affairs in the past year.
- $\boldsymbol{x}$ - constant term=1, age, number of years married, religiousness (1-5 scale), occupation (1-7 scale), self-rating of marriage (1-5 scale)

Assingment

The following is the data generating assumptions for the Poisson model, where j is the number of affairs:

$$Pr[y_i = j] = \frac{e^{-\lambda_i}\lambda_i^j}{j!} \tag{1}$$

$$log(\lambda_i) = \boldsymbol{x}_i'\boldsymbol{\beta} \tag{2}$$

$$E(y_i \mid x_i) = e^{\boldsymbol{x}_i'\boldsymbol{\beta}} \tag{3}$$

for some $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5)'$.

The log-likelihood function is:

$$\begin{aligned}
\ln L &= \sum_{i=1}^{n} \ln f\left(y_i \left| x_i \right., \beta\right) \\
&= \sum_{i=1}^{n} \ln \frac{e^{-\lambda_i}\lambda_i^j}{j!} \\
&= \sum_{i=1}^{n} \left[-\lambda_i + y_i \ln \lambda_i - \ln y_i!\right] \\
&= \sum_{i=1}^{n} \left[-e^{\mathbf{x}_i{}'\beta} + y_i x_i'\beta - \ln y_i!\right]
\end{aligned}$$

1. Estimate the parameter vector $\boldsymbol{\beta}$ using maximum likelihood. Use as the starting value a vector of zeros. Use four algorithms:

1. nelder-mead;
2. Quasi-newton with BFGS Hessian without a user-supplied derivative;
3. Quasi-newton with BFGS with a user-supplied derivative;

4. the MLE technique we went over in class that uses the score function to approximate the Hessian (this is also referred to as BHHH).

Report the estimated parameters for each case, the number of iterations, and the number of function evaluations.

2. Report the eigenvalues for the Hessian approximation for the BHHH MLE method from the last question. Report the eigenvalues for the initial Hessian, and the Hessian at the estimated parameters.

3. Now estimate the model using the NLLS method we went over in class, starting from the same initial point. Report the results.

4. Report the standard errors for the BHHH and the NLLS methods.

Turn in your results in a way that I can easily read your code and match it to your results, eg. a Jupyter notebok or a well organized script.