Econometrics HW5

Michael B. Nattinger[*]

April 18, 2021

1 Question 25.1

The estimated probabilities of purchase and not purchase will be the same across the two models, so we can write the probabilities as follows:

$$P(\text{purchase}|X = x) = 1 - P(\text{no purchase}|X = x)$$
$$\Phi(x'\beta_1) = 1 - \Phi(x'\beta_2)$$
$$= \Phi(-x'\beta_2)$$
$$\Rightarrow \beta_1 = -\beta_2.$$

2 Question 25.3

Let $Y = P(X) + e$ and note that $Y = \begin{cases} 1, \text{ w.p. } P(X) \\ 0, \text{ w.p. } 1 - P(X) \end{cases}$.

Solving for e we have $e = \begin{cases} 1 - P(X), \text{ w.p. } P(X) \\ -P(X), \text{ w.p. } 1 - P(X) \end{cases}$. It is trivial to see that $E[e|X] = 0$.

$$E[e^2|X] = (1 - P(X))^2 P(X) + (-P(X))^2(1 - P(X))$$
$$= (1 - 2P(X) + P(X)^2)P(X) + P(X)^2 - P(X)^3$$
$$= P(X) - P(X)^2$$
$$= P(X)(1 - P(X)),$$
$$\Rightarrow Var(e|X) = P(X)(1 - P(X)).$$

3 Question 25.9

From (25.5), the conditional probability mass function is

$$\pi(Y|X) = \Lambda(Z'\beta)$$
$$= (1 + exp(-Z'\beta))^{-1},$$

where $Z = \begin{cases} X, & \text{if } Y = 1 \\ -X, & \text{if } Y = 0 \end{cases}$. Given data we can write the log likelihood function as the following:

$$L(\beta) = \sum_{i=1}^{n} log((1 + exp(-Z_i'\beta))^{-1})$$

$$= -\sum_{i=1}^{n} log(1 + exp(-Z_i'\beta))$$

Taking FOCs of the log likelihood function,

$$\sum_{i=1}^{n}(1 + exp(-Z_i'\beta))^{-1} exp(-Z_i'\beta)Z_i = \vec{0}$$

$$\sum_{i=1}^{n}(1 - \Lambda(Z_i\beta))Z_i = \vec{0}$$

Then, $\hat{\beta}_{MLE}$ is the solution to this set of equations. Note that this solution has no closed form.

4 Question 25.12

Given data, the NLLS estimator would minimize the sum of square errors:

$$\hat{\beta}_{NLLS} = \arg\min_{\beta} \sum_{i=1}^{n}(Y_i - \Phi(X_i\beta))^2$$

5 Question 25.14

5.1 Part A

$$P(Y > 0) = P(Y^* > 0)$$
$$= P(m(X) + e > 0)$$
$$= P(e > -m(X))$$
$$= 1 - \Phi_{\sigma^2(X)}(-m(X))$$

where $\Phi_{\sigma^2(X)}$ is the CDF of $N(0, \sigma^2(X))$.

5.2 Part B

No. Assume for the purpose of contradiction that $m(X)$ and $\sigma^2(X)$ were uniquely identified. Then, note that $m'(X) = cm(X), \sigma^{2'}(X) = c^2\sigma^2(X)$ yield the same response probability so the identification is not unique.

5.3 Part C

Normalize $\sigma^2(X) = 1$. Then we can identify $\tilde{m}(X)$, which is the m function transformed appropriately s.t. the normalization of e.

5.4 Part D

No, $\sigma^2(X) = 1 \forall X$ so there is no heteroskedasticity in the model.

# 6   Question 25.15

Below are the estimated coefficients.

	Coeff	SE	T
intercept	1.95	0.107	18.2
age	-0.00793	0.00142	-5.58
education	0.0255	0.00622	4.1
hispanic	0.298	0.0569	5.24
black	0.0541	0.06	0.901

   From the coefficients and standard errors which we were asked to provide, we can see that age, education, and the indicator for hispanic are significant coefficients in the logit model we were asked to run. Note that the marginal effect is not directly seen in the information we were asked to provide.

```matlab
mkdir('pings')
clear; close all; clc
read = 0;
if read
    cd('..\PS4')
    [x,xt] = xlsread('cps09mar.xlsx','Sheet1');
    save 'data'
    cd('..\PS5')
else
    cd('..\PS4')
    load 'data'
    cd('..\PS5')
end

men = logical(1-x(:,2)); % first matlab question
Y = categorical(x(men,8));
X = x(men,[1 4 3]);
race = x(men,11);
black = race==2; % is this right?
X = [X black];
[B,~,stats] = mnrfit(X,Y);
tab = table(B,stats.se,stats.t,'VariableNames',{'Coeff','SE','T'}, ...
    'RowNames',{'intercept','age','education','hispanic','black'});
table2latex(tab,'tab1.tex')
```

7   Question 25.17

estimated marriage probabilities for women with degrees
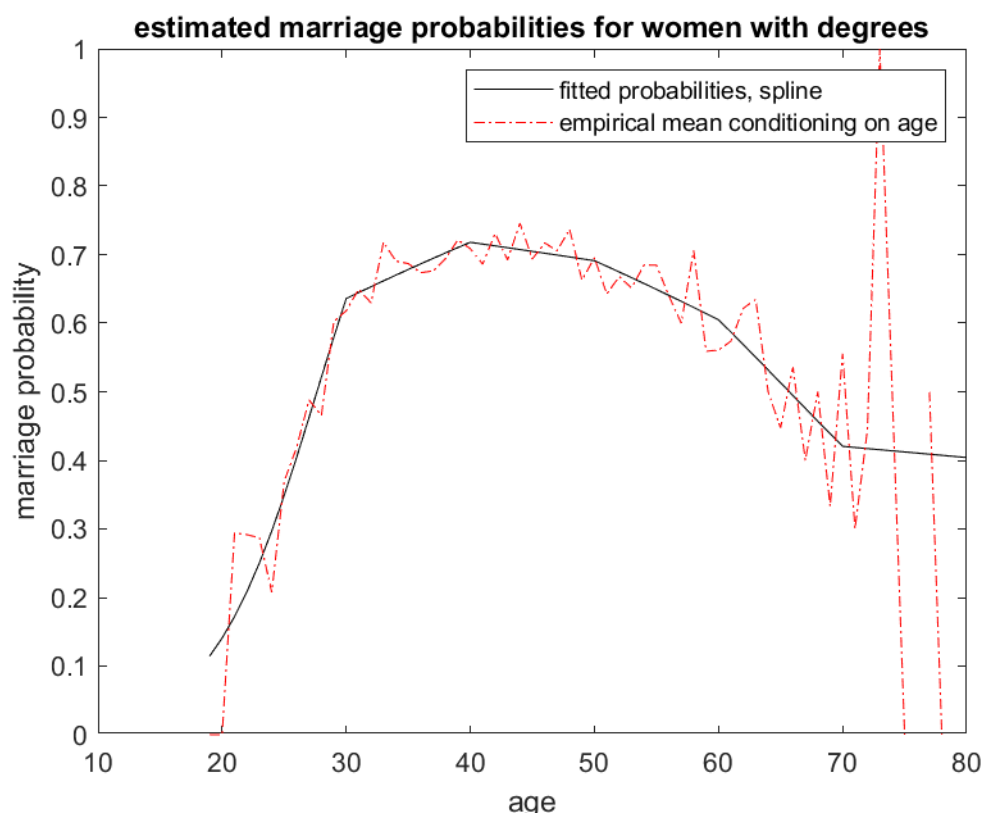
Above we see the results for the model that I decided on. I included linear splines for age above 30, 40, 50, 60, and 70. To visually check the model results, I estimated the mean of the marriage indicator at each age. This is plotted on the figure in the red dashed line. From comparing the fitted logit model to the nonparametric conditional estimate we can see that the model appears to fit the data reasonably well.

My reasoning for the linear splines at decade intervals was an intuition that each decade of life may, for a variety of reasons, lead to different relationships between age and marriage. I used a linear spline rather than a quadratic spline as the number of knots on the spline was 5, which already seemed fairly large and so I did not want to include the extra quadratic terms. Additionally, the splines seemed to fit the data well from just the linear terms.

Comparing to figure 25.1, it is clear that the marriage probabilities are lower for educated women than men. Moreover, for educated women there appears to be a significant drop-off in marriage probabilities after around the age of 60. This does not appear in the distribution of men. If anything, the distribution of men appears to begin to rise after 60.

```matlab
women = logical(x(:,2)); % second matlab question
college = logical(x(:,4)>13);
woco = logical(women.*college);
Y = categorical(x(woco,12)<4);
age = x(woco,1);
ov30 = 0*age;
ov30(age>30,:) = age(age>30,:)-30;
ov40 = 0*age;
ov40(age>40,:) = age(age>40,:)-40;
ov50 = 0*age;
ov50(age>50,:) = age(age>50,:)-50;
ov60 = 0*age;
ov60(age>60,:) = age(age>60,:)-60;
ov70 = 0*age;
ov70(age>70,:) = age(age>70,:)-70;
X = [age ov30 ov40 ov50 ov60 ov70];
%X = [age ov30 ov30.^2 ov50 ov50.^2 ov70 ov70.^2];
B = mnrfit(X,Y);
aG = (min(age):max(age))';
aG30 = 0*aG;
aG30(aG>30,:) = aG(aG>30,:)-30;
aG40 = 0*aG;
aG40(aG>40,:) = aG(aG>40,:)-40;
aG50 = 0*aG;
aG50(aG>50,:) = aG(aG>50,:)-50;
aG60 = 0*aG;
aG60(aG>60,:) = aG(aG>60,:)-60;
aG70 = 0*aG;
aG70(aG>70,:) = aG(aG>70,:)-70;
G = [ones(size(aG)) aG aG30 aG40 aG50 aG60 aG70];
fit = (1+exp(G*B)).^(-1);
emp = 0*fit;
for i=1:length(emp)
    emp(i) = mean(double(Y(age==aG(i)))-1);
end
figure
plot(aG,fit,'k')
hold on
plot(aG,emp,'r-.')
hold off
title('estimated marriage probabilities for women with degrees')
xlabel('age')
ylabel('marriage probability')
legend('fitted probabilities, spline','empirical mean conditioning on age')
set(gcf,'Color',[1 1 1])
cd('pings')
saveas(gcf,'fig1.png')
cd('..')
```

8 Question 26.1

$$P_j(x) = \frac{exp(x'\beta_j)}{\sum_{l=1}^{J} exp(x'\beta_l)}.$$

$$exp(a) > 0$$

$$\Rightarrow P_j(x) > 0.$$

$$\frac{exp(x'\beta_j)}{\sum_{l=1}^{J} exp(x'\beta_l)} < \frac{exp(x'\beta_j)}{exp(x'\beta_j)}$$

$$= 1.$$

$$\Rightarrow 0 \le P_j(x) \le 1$$

$$. \sum_{j=1}^{J} P_j(x) = \sum_{j=1}^{J} \frac{exp(x'\beta_j)}{\sum_{l=1}^{J} exp(x'\beta_l)}$$

$$= \frac{\sum_{j=1}^{J} exp(x'\beta_j)}{\sum_{l=1}^{J} exp(x'\beta_l)}$$

$$= 1.$$

9 Question 26.3

$$\frac{\partial P_j(x)}{\partial x} = \frac{\partial}{\partial x} \frac{exp(x'\beta_j)}{\sum_{l=1}^{J} exp(x'\beta_l)}$$

$$= \frac{exp(x'\beta_j)}{\sum_{l=1}^{J} exp(x'\beta_l)} \beta_j - \frac{exp(x'\beta_j)}{\sum_{l=1}^{J} exp(x'\beta_l)} \frac{\sum_{i=1}^{J} (exp(x'\beta_i)\beta_i)}{\sum_{l=1}^{J} exp(x'\beta_l)}$$

$$= P_j(x) \left(\beta_j - \sum_{i=1}^{J} P_i \beta_i \right)$$

10 Question 26.7

The average marginal effects can be written in the following form:

$$AME_{jj} = E[\delta_{jj}(W, X)]$$

$$= E[\gamma P_j(W, X)(1 - P_j(W, X))]$$

Replacing expectations with sample averages and estimated coefficients and probabilities, we get the following:

$$\hat{AME}_{jj} = \hat{\gamma} \frac{1}{n} \sum_{i=1}^{n} \hat{P}_j(W_i, X_i)(1 - \hat{P}_j(W_i, X_i)).$$

11 Question 26.8

For a multinomial logit model we have the following:

$$\frac{P_j(W, X|\theta)}{P_l(W, X|\theta)} = \frac{\frac{exp(W'\beta_j + X_j'\gamma)}{\sum_{i=1}^{J} exp(W'\beta_i + X_i'\gamma)}}{\frac{exp(W'\beta_l + X_l'\gamma)}{\sum_{i=1}^{J} exp(W'\beta_i + X_i'\gamma)}}$$

$$= \frac{exp(W'\beta_j + X_j'\gamma)}{exp(W'\beta_l + X_l'\gamma)}.$$