ECON 3161 Fall 2023

Homework 5 Solutions

Vidit D. Pokharna Due: Dec 3, 2023

- **Instruction: There are 60 points in total. This homework needs computer coding. Stata software is recommended, but you can use other software you prefer. You can either type or handwrite your answer. Then upload your answer file on canvas.**

Problem 1

Part 1

Given: $y_i^* = \beta_0 + \beta_1 x_i + u_i$ and $y_i = \begin{cases} 1 & \text{if } y_i^* > 0 \\ 0 & \text{if } y_i^* \leq 0 \end{cases}$. Since $u_i \sim N(0,1)$,

$$\begin{aligned} P(y_i = 1|x_i) &= P(y_i^* > 0|x_i) = P(\beta_0 + \beta_1 x_i + u_i > 0) \\ &= P(u_i > -(\beta_0 + \beta_1 x_i)) \\ &= 1 - P(u_i \leq -(\beta_0 + \beta_1 x_i)) \\ &= 1 - \Phi(-(\beta_0 + \beta_1 x_i)) \\ &= \Phi(\beta_0 + \beta_1 x_i) \end{aligned}$$

Part 2

The model, represented by $P(y_i = 1|x_i) = \Phi(\beta_0 + \beta_1 x_i)$, is known as the Probit Model.

Part 3

$$\begin{aligned} \text{Let } L(\beta_0, \beta_1) &= \prod_{i=1}^{n} P(y_i|x_i; \beta_0, \beta_1) \\ &= \prod_{i=1}^{n} \left[\Phi(\beta_0 + \beta_1 x_i)^{y_i} \cdot (1 - \Phi(\beta_0 + \beta_1 x_i))^{1-y_i}\right] \end{aligned}$$

Then, the log-likelihood function is:

$$\log L(\beta_0, \beta_1) = \sum_{i=1}^{n} \left[y_i \log \Phi(\beta_0 + \beta_1 x_i) + (1 - y_i) \log(1 - \Phi(\beta_0 + \beta_1 x_i))\right]$$

Part 4

The main disadvantage of the linear probability model, represented as $y_i = \beta_0 + \beta_1 x_i + u_i$, is that it can predict probabilities outside the [0, 1] range, which are not meaningful for a probability.

Problem 2

Part 1

If discrimination against minorities is present and controlled for, the sign of β_1 in the model approve $= \beta_0 + \beta_1 \text{white} + u$ would be expected to be negative, indicating a lower likelihood of loan approval for minority applicants.

Part 2

```
. regress approve white
```

Source	SS	df	MS	Number of obs	=	1,989
				F(1, 1987)	=	102.23
Model	10.4743407	1	10.4743407	Prob > F	=	0.0000
Residual	203.59303	1,987	.102462521	R-squared	=	0.0489
				Adj R-squared	=	0.0485
Total	214.067371	1,988	.107679764	Root MSE	=	.3201

approve	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
white	.2005957	.01984	10.11	0.000	.1616864	.239505
_cons	.7077922	.0182393	38.81	0.000	.6720221	.7435623

The regression coefficient for `white` is 0.2009, indicating that the probability of loan approval for white applicants is 20.09 percentage points higher compared to non-white applicants. The p-value is less than 0.01, which suggests that this coefficient is statistically significant. The R-squared value of 0.0489 means that approximately 4.89% of the variability in loan approval is explained by the race of the applicant.

Part 3

```
. probit approve white

Iteration 0:  Log likelihood = -740.34659
Iteration 1:  Log likelihood = -701.33221
Iteration 2:  Log likelihood = -700.87747
Iteration 3:  Log likelihood = -700.87744

Probit regression                                 Number of obs =  1,989
                                                  LR chi2(1)    =  78.94
                                                  Prob > chi2   = 0.0000
Log likelihood = -700.87744                       Pseudo R2     = 0.0533
```

approve	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
white	.7839465	.0867118	9.04	0.000	.6139946	.9538985
_cons	.5469463	.075435	7.25	0.000	.3990964	.6947962

```
. margins, at(white=(0 1))

Adjusted predictions                              Number of obs = 1,989
Model VCE: OIM

Expression: Pr(approve), predict()
1._at: white = 0
2._at: white = 1
```

	Margin	Delta-method std. err.	z	P>|z|	[95% conf.	interval]
_at						
1	.7077922	.0259133	27.31	0.000	.657003	.7585814
2	.9083879	.007036	129.10	0.000	.8945975	.9221782

In the probit model, the coefficient for `white` is 0.7839. This coefficient is not directly interpretable as a change in probability due to the nonlinear nature of the model. However, it indicates a positive association between being white and the latent propensity for loan approval. According to the margins output:

- The predicted probability of approval for non-white applicants is approximately 70.77%.
- The predicted probability of approval for white applicants is approximately 90.88%.

These findings are consistent with the linear model and may suggest racial discrimination in loan approvals.

Part 4

```
. logit approve white

Iteration 0:  Log likelihood = -740.34659
Iteration 1:  Log likelihood =  -709.1878
Iteration 2:  Log likelihood =  -700.9007
Iteration 3:  Log likelihood = -700.87744
Iteration 4:  Log likelihood = -700.87744

Logistic regression                                     Number of obs =  1,989
                                                        LR chi2(1)    =  78.94
                                                        Prob > chi2   = 0.0000
Log likelihood = -700.87744                             Pseudo R2     = 0.0533
```

approve	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
white	1.409422	.1511511	9.32	0.000	1.113172	1.705673
_cons	.8846854	.1252927	7.06	0.000	.6391162	1.130255

```
. margins, at(white=(0 1))

Adjusted predictions                                    Number of obs = 1,989
Model VCE: OIM

Expression: Pr(approve), predict()
1._at: white = 0
2._at: white = 1
```

	Margin	Delta-method std. err.	z	P>|z|	[95% conf.	interval]
_at						
1	.7077922	.0259133	27.31	0.000	.657003	.7585814
2	.9083879	.007036	129.10	0.000	.8945975	.9221782

The logistic regression output yields a coefficient for `white` of 1.4094, indicating that the log odds of loan approval for white applicants are higher compared to non-white applicants. The margins output, which is identical to the probit model, shows:

- The predicted probability of approval for non-white applicants is approximately 70.77%.
- The predicted probability of approval for white applicants is approximately 90.88%.

All three models (linear probability, probit, and logit) indicate that white applicants are more likely to be approved for a loan than non-white applicants, pointing to potential discrimination.