Course Title: Quantitative Techniques for Economics
Course Code:ECON6002
Topic: The logit regression models

Ph.D. Economics (1st Semester)

Dr. Kailash Chandra Pradhan

Mahatma Gandhi Central University,

Department of Economics

Qualitative response regression models

- There are three approaches to developing a probability model for a binary response variable:

  1. The linear probability model (LPM)

  2. The logit model

  3. The probit model

The logit model

- Recall that in explaining home ownership in relation to income, the LPM was

  $P_i = E(Y = 1 | X_i) = \beta_1 + \beta_2 X_i$ **(1)**

  where X is income and Y = 1 means the family owns a house. But now consider the following representation of home ownership:

$$P_i = E(Y = 1 | X_i) = \frac{1}{1 + e^{-(\beta_1 + \beta_2 X_i)}} \qquad (2)$$

The logit model

- For ease of exposition, we can write as

$$P_i = E(Y=1 \mid X_i) = \frac{1}{1+e^{-(Z_i)}} = \frac{e^Z}{1+e^Z} \qquad (3)$$

where $Z_i = \beta_1 + \beta_2 X_i$.

- Equation (3) is known as the (cumulative) **logistic distribution function**
- Here, Z_i ranges from $-\infty$ to $+\infty$, P_i ranges between 0 and 1 and that P_i is nonlinearly related to Z_i (i.e., X_i).
- We cannot use the OLS procedure to estimate the parameters. But this problem is more apparent than real because equation (2) can be linearized, which can be shown as follows.
- If P_i , the probability of owning a house, is given in equation (3), then $(1-P_i)$, the probability of not owning a house is

$$1-P_i = \frac{1}{1+e^{Z_i}} \qquad (4)$$

The logit model

- Therefore, we can write

$$\frac{P_i}{1-P_i} = \frac{1+e^{Z_i}}{1+e^{-Z_i}} = e^{Z_i} \qquad (5)$$

- Now $P_i/(1 - P_i)$ is simply the odds ratio in favor of owning a house—the ratio of the probability that a family will own a house to the probability that it will not own a house. Thus, if $P_i = 0.8$, it means that odds are 4 to 1 in favor of the family owning a house.
- Now if we take the natural log of equation (5), we obtain as follows.

$$L_i = \ln\left(\frac{P_i}{1-P_i}\right) = Z_i = \beta_1 + \beta_2 X_i \qquad (6)$$

- L is the log of the odds ratio, is not only linear in X, but also linear in the parameters. Here, L is called the logit model.
- Logit regression is estimated using the method of maximum likelihood.

Properties of logit model

1. As P goes from 0 to 1 (i.e., as Z varies from $-\infty$ to $+\infty$), the logit L goes from $-\infty$ to $+\infty$.
2. Although L is linear in X, the probabilities themselves are not. This property is in contrast with the LPM model where the probabilities increase linearly with X.
3. Although we have included only a single X variable, or regressor, in the preceding model, one can add as many regressors as may be dictated by the underlying theory.
4. The logit becomes negative and increasingly large in magnitude as the odds ratio decreases from 1 to 0 and becomes increasingly large and positive as the odds ratio increases from 1 to infinity.
5. The interpretation of the logit model is as follows: β_2, the slope, measures the change in L for a unit change in X, that is, it tells how the log-odds in favor of owning a house change as income changes by a unit, say, Rs.1000. The intercept β_1 is the value of the logodds in favor of owning a house if income is zero. Like most interpretations of intercepts, this interpretation may not have any physical meaning.
6. Whereas the LPM assumes that P_i is linearly related to X_i, the logit model assumes that the log of the odds ratio is linearly related to X_i .

Empirical logit model

- Example: Letting Y = 1 if a student's final grade in an intermediate microeconomics course was A and Y = 0 if the final grade was a B or a C, Spector and Mazzeo used grade point average (GPA), TUCE, and Personalized System of Instruction (PSI) as the grade predictors. The logit model here can be written as:

$$L_i = \ln\left(\frac{P_i}{1-P_i}\right) = \beta_1 + \beta_2 GPA_i + \beta_3 TUCE_i + \beta_4 PSI_i + u_i$$

Results

Variable	Coefficient	Std. error	*Z statistic*	*Probability*
C	*-13.0213*	*4.931*	*-2.6405*	*0.0082*
GPA	2.8261	1.2629	2.2377	0.0252
TUCE	0.0951	0.1415	0.6722	0.5014
PSI	2.3786	1.0645	2.2345	0.0255

McFadden R^2 = *0.3740* *LR statistic (3 df) = 15.40419*

Empirical logit model

Some of general observations:

1. Since we are using the method of maximum likelihood, which is generally a large-sample method, the estimated standard errors are asymptotic.
2. As a result, instead of using the t statistic to evaluate the statistical significance of a coefficient, we use the (standard normal) Z statistic. So inferences are based on the normal table. Recall that if the sample size is reasonably large, the t distribution converges to the normal distribution.
3. As noted earlier, the conventional measure of goodness of fit, R^2, is not particularly meaningful in binary regressand models. Measures similar to R^2, called pseudo R^2, are available. Eviews presents one such measure, the McFadden R^2.
4. To test the null hypothesis that all the slope coefficients are simultaneously equal to zero, the equivalent of the F test in the linear regression model is the likelihood ratio (LR) statistic. Given the null hypothesis, the LR statistic follows the $\chi 2$ distribution with df equal to the number of explanatory variables, three in the present example. (Note: Exclude the intercept term in computing the df).

Empirical logit model

Interpretation:

- Each slope coefficient in this equation is a partial slope coefficient and measures the change in the estimated logit for a unit change in the value of the given regressor (holding other regressors constant).
- Thus, the GPA coefficient of 2.8261 means, with other variables held constant, that if GPA increases by a unit, on average the estimated logit increases by about 2.83 units, suggesting a positive relationship between the two.
- As we can see, all the other regressors have a positive effect on the logit, although statistically the effect of TUCE is not significant.
- However, together all the regressors have a significant impact on the final grade, as the LR statistic is 15.40, whose p value is about 0.0015, which is very small.
- More meaningful interpretation is in terms of odds, which are obtained by taking the antilog of the various slope coefficients. Thus, if we take the antilog of the PSI coefficient of 2.3786 we will get 10.7897.
- This suggests that students who are exposed to the new method of teaching are more than 10 times likely to get an A than students who are not exposed to it, other things remaining the same.

Reference:

Reference

Gujarati, D (1995), *Basic Econometrics*, 4th Edition, New York: McGraw Hill