**Bivariate Probit**

The bivariate probit model considers two binary outcomes. The outcomes are potentially related after conditioning on the regressors. How related they are occurs through correlation of the errors that appear in the model formulation of the binary outcome model. These notes describe how to estimate a **Bivariate Probit Model** by Maximum Likelihood Estimation (MLE), and use the Likelihood Ratio (LR) and the Wald tests to test an interesting hypothesis. Bivariate Probit model involves two equations, rather than one – and each equation is a binary choice model. Although we’ll be looking at the Bivariate Probit model, everything that follows could be done for a Bivariate Logit model.

**Interpretation of Results**

The mean (proportion) for excellent health status (**hlthe**) is 0.54 and the mean (proportion) for visiting the doctor (**dmdu**) is 0.67. As shown in Table 1, the correlation is -0.01, so the two outcomes are practically uncorrelated (higher correlation is needed to apply bivariate probit).

**Correlation between excellent health status (hlthe) and visiting the doctor (dmdu)**

|  |  |  |
| --- | --- | --- |
| **Variable** | **hlthe** | **dmdu** |
| **Hlthe** | 1.0000 |  |
| **Dmdu** | -0.0110 | 1.0000 |



Where 𝑟ℎ𝑜 (𝜌) is the correlation coefficient of the error terms; i.e. 𝑟ℎ𝑜 denotes the correlation between the two sets of unobserved factors or error terms in Equation 5 and Equation 6. If 𝑟ℎ𝑜 = 0 the outcomes are independent and therefore the two equations are best modelled separately. When 𝑟ℎ𝑜 ≠ 0 the two outcomes are correlated as the probability of one outcome depends on the probability of the other. As the value of 𝑟ℎ𝑜 rises from 0 to 1, the correlation of the two error terms is increasing. Similarly, values between 0 and -1 indicate a negative correlation.

* Coefficient interpretation: Younger individuals, individuals with higher incomes, and those with lower number of chronic disease are more likely to be in an excellent health status. Individuals with higher incomes and those with higher number of chronic diseases are more likely visit the doctor.

The correlation coefficient between the bivariate outcomes is 0.02 and not significant thus rejecting the hypothesis that the two dependent variables are not jointly determined.

* Therefore, we can proceed by estimating separate probit models instead of a bivariate probit model. The decisions are not interrelated and can be estimated independently.
* Comparing the separate univariate probit models to the Bivariate probit model, it would be observed that the results from the separate probit models are almost identical to those from the bivariate probit model, so in this case there is no need to perform the bivariate probit model.

Marginal Effects

The marginal effects are interpreted similarly to those of binary probit and logit models, but the effect is on the joint probability of the two outcomes.

* The marginal effects sum up to zero across the four joint probabilities.
* When age increases by 1 year, the probability P(y1=0, y2=1) of a person not being in an excellent health and visiting the doctor increases by 0.5%.
* Notice that the other marginal effects are negative.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **P(y1=0, y2=0)** | **P(y1=0, y2=1)** | **P(y1=1, y2=0)** | **P(y1=1, y2=1)** |
| Age | 0.001\* | 0.005\* | -0.003\* | -0.004\* |
| Log income | -0.037\* | -0.015\* | -0.006 | 0.059\* |
| Number of chronic diseases | -0.012\* | 0.015\* | -0.011\* | -0.003\* |

