Inference for the Bivariate Probit Model Using Eviews

David Giles

Department of Economics, University of Victoria

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. The model that we're going to consider is one which 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 It is particularly easy to deal with the Probit case in EViews, as there is a built-in cumulative bivariate Normal function that we can exploit. The set-up for the model is as follows:

$$y_{i1}^* = x_{i1}'\beta_1 + \varepsilon_{i1} \qquad y_{i1} = 1 \text{ if } y_{i1}^* > 0; \quad = 0, \text{ otherwise} \qquad (1)$$
$$y_{i2}^* = x_{i2}'\beta_2 + \varepsilon_{i2} \qquad y_{i2} = 1 \text{ if } y_{i2}^* > 0; \quad = 0, \text{ otherwise} \qquad (2)$$

$$E(\varepsilon_{i1}) = E(\varepsilon_{i2}) = 0 \quad ; \ Var(\varepsilon_{i1}) = Var(\varepsilon_{i2}) = 1$$
$$Cov(\varepsilon_{i1}, \varepsilon_{i2}) = \rho \qquad ; \quad i = 1, 2, 3, \ldots, n$$

If we had just equation (1), then we could set up a standard Probit model:

$$\Pr[y_{i1} = 1] = \Pr.[y_{i1}^* > 0] = \Pr[x_{i1}'\beta_1 + \varepsilon_{i1} > 0] = \Pr[\varepsilon_{i1} > -x_{i1}'\beta_1] = \Pr[\varepsilon_{i1} < x_{i1}'\beta_1] = \Phi(x_{i1}'\beta_1)$$

where $\Phi(.)$ is the cumulative distribution function for the standard Normal, and we have used the symmetry of the Normal distribution to get the penultimate equality above.

To set up the Bivariate Probit model, based on both equations (1) and (2), we need to consider the following four possible cases:

$$P_{11} = \Pr[y_{i1} = 1, y_{i2} = 1] = \int_{-\infty}^{x_{i1}'\beta_1} \int_{-\infty}^{x_{i2}'\beta_2} \phi_2(z_1, z_2, \rho) dz_1 dz_2$$
$$P_{10} = \Pr[y_{i1} = 1, y_{i2} = 0] = \int_{-\infty}^{x_{i1}'\beta_1} \int_{x_{i2}'\beta_2}^{\infty} \phi_2(z_1, z_2, \rho) dz_1 dz_2$$
$$P_{01} = \Pr[y_{i1} = 0, y_{i2} = 1] = \int_{x_{i1}'\beta_1}^{\infty} \int_{-\infty}^{x_{i2}'\beta_2} \phi_2(z_1, z_2, \rho) dz_1 dz_2$$
$$P_{00} = \Pr[y_{i1} = 0, y_{i2} = 0] = \int_{x_{i1}'\beta_1}^{\infty} \int_{x_{i2}'\beta_2}^{\infty} \phi_2(z_1, z_2, \rho) dz_1 dz_2$$

where the *bivariate* normal density function is

$$\phi_2(z_1, z_2, \rho) = \exp[-0.5(z_1^2 + z_2^2 - 2\rho z_1 z_2)/(1-\rho^2)]/[2\pi(1-\rho^2)^{1/2}].$$

Then, the log-likelihood function for the bivariate model can be obtained in the following way:

(i) Define $q_{i1} = 2y_{i1} - 1$ and $q_{i2} = 2y_{i2} - 1$

(ii) Define $z_{ij} = x_{ij}'\beta_j$ and $w_{ij} = q_{ij}z_{ij}$; j = 1, 2

(iii) Define $\rho_i^* = q_{i1}q_{i2}\rho$

(iv) Then, $\Pr[Y_1 = y_{i1}, Y_2 = y_{i2}] = \Phi_2(w_{i1}, w_{i2}, \rho_i^*)$

(v) $\log L = \sum_{i=1}^{n} \log \Phi_2(w_{i1}, w_{i2}, \rho_i^*)$

The EViews workfile, **bivariate probit.wf1**, contains data for the variables y_1, y_2, x_1 and x_2. The LOGL object, LOGL01, allows us to estimate a Bivariate Probit model for y_1 and y_2. See the **READ_ME** text object in the EViews workfile for more details.

We can use a Wald test to test the hypothesis that the errors in the two equations of the model are independent. This amounts to testing H_0: ρ = 0 against a 2-sided alternative. In terms of the EViews code, we need to test if c(5) = 0. The results are in the table titled **Wald_Test_Result_01**. In this case, we cannot reject H_0, so in fact we would be justified in estimating 2 separate Probit models (see EQ01 and EQ02 in the EViews workfile).

We can also use an LR test to test the hypothesis that the errors in the two equations of the model are independent. In this case we need the maximized value of the log-likelihood function from the bivariate probit model (*i.e.*, the "unrestricted" model), and the maximized value of the log-likelihood function from the "unrestricted" model. The latter is just the sum of the 2 maximized log-likelihood values from the two individual probit models. Here, this number is –(32.96773 + 30.44460) = -63.41233. So, the LR test statistic is 2(-63.16293+63.41233 +) = 0.4988. The p-value based on the (asymptotic) $\chi^2_{(1)}$ distribution is 0.4800, so again, we would not reject the null hypothesis.

Finally, note that the model could be extended in several interesting ways. For example, suppose that we hypothesize that the correlation between the errors in the two equations of the model is proportional to a third variable, x_3. In LOGL02 in the EViews workfile, the code for the log-likelihood function has been altered to allow for this refinement. If we now apply the Wald test to see if the errors of the two equations are independent we get the results in the text object titled **Wald_Test_Result_02**. Note that the p-value has dropped to approximately 7%, which may well alter our earlier conclusion. Of course, this is only an asymptotically valid test and our sample size is just n = 50, so we should be cautious.

March, 2010