triprobit and the GHK simulator: a short note
Antoine Terracol∗

1 The trivariate probit

Consider three binary variables $y_1$, $y_2$, and $y_3$, the trivariate probit model supposes that:

\[
\begin{align*}
y_1 &= \begin{cases} 1 & \text{if } X\beta + \varepsilon_1 > 0 \\ 0 & \text{otherwise} \end{cases} \\
y_2 &= \begin{cases} 1 & \text{if } Z\gamma + \varepsilon_2 > 0 \\ 0 & \text{otherwise} \end{cases} \\
y_3 &= \begin{cases} 1 & \text{if } W\theta + \varepsilon_3 > 0 \\ 0 & \text{otherwise} \end{cases}
\end{align*}
\]

with

\[
\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix} \rightarrow N(0, \Sigma)
\]

For identification reasons, the variances of the epsilons must equal 1.

Evaluation of the likelihood function requires the computation of trivariate normal integrals. For example, the probability of observing $(y_1 = 0, y_2 = 0, y_3 = 0)$ is:

\[
\Pr[y_1 = 0, y_2 = 0, y_3 = 0] = \int_{-\infty}^{-X\beta} \int_{-\infty}^{-Z\gamma} \int_{-\infty}^{-W\theta} \phi_3(\varepsilon_1, \varepsilon_2, \varepsilon_3, \rho_{12}, \rho_{13}, \rho_{23}) \, d\varepsilon_3 d\varepsilon_2 d\varepsilon_1
\]

where $\phi_3(\cdot)$ is the trivariate normal p.d.f., and $\rho_{ij}$ is the correlation coefficient between $\varepsilon_i$ and $\varepsilon_j$.

While Stata provides commands to compute univariate and bivariate normal CDF (norm() and binorm()), no command is available for the trivariate case (as a matter of fact, numerical approximations perform poorly in computing high order integrals).

The triprobit command uses the GHK (Geweke-Hajivassiliou-Keane) smooth recursive simulator to approximate these integrals

2 The GHK simulator

Let us illustrate the GHK simulator in the trivariate case (generalization to higher orders is straightforward)

We wish to evaluate

\[
\Pr(\varepsilon_1 < b_1, \varepsilon_2 < b_2, \varepsilon_3 < b_3)
\]

where $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ are normal random variables with covariance structure given in (2)

Equation (4) can be rewritten as a product of conditional probabilities:

\[
\Pr(\varepsilon_1 < b_1) \Pr(\varepsilon_2 < b_2|\varepsilon_1 < b_1) \Pr(\varepsilon_3 < b_3|\varepsilon_1 < b_1, \varepsilon_2 < b_2)
\]

Let $L$ be the lower triangular Cholesky decomposition of $\Sigma$, such that: $LL' = \Sigma$.

∗TEAM, Université de Paris 1 et CNRS, 106-112 boulevard de l’Hôpital, 75647 Paris Cedex 13, France.
Email: terracol@univ-paris1.fr
I am very much indebted to Wolfgang Schwerdt who has provided numerous advices
\[
L = \begin{pmatrix}
  l_{11} & 0 & 0 \\
  l_{21} & l_{22} & 0 \\
  l_{31} & l_{32} & l_{33}
\end{pmatrix}
\]

We get:
\[
\begin{pmatrix}
  \varepsilon_1 \\
  \varepsilon_2 \\
  \varepsilon_3
\end{pmatrix} = \begin{pmatrix}
  l_{11} & 0 & 0 \\
  l_{21} & l_{22} & 0 \\
  l_{31} & l_{32} & l_{33}
\end{pmatrix} \begin{pmatrix}
  \nu_1 \\
  \nu_2 \\
  \nu_3
\end{pmatrix}
\]

where the \( \nu \) are independent standard normal random variables.

By (6), we get:
\[
\varepsilon_1 = l_{11} \nu_1 \\
\varepsilon_2 = l_{21} \nu_1 + l_{22} \nu_2 \\
\varepsilon_3 = l_{31} \nu_1 + l_{32} \nu_2 + l_{33} \nu_3
\]

Thus:
\[
\Pr(\varepsilon_1 < b_1) = \Pr(\nu_1 < b_1 / l_{11})
\]

and
\[
\Pr(\varepsilon_2 < b_2 | \varepsilon_1 < b_1) = \Pr(\nu_2 < (b_2 - l_{21} \nu_1) / l_{22} | \nu_1 < b_1 / l_{11})
\]

and
\[
\Pr(\varepsilon_3 < b_3 | \varepsilon_1 < b_1, \varepsilon_2 < b_2) = \\
\Pr(\nu_3 < (b_3 - l_{31} \nu_1 - l_{32} \nu_2) / l_{33} | \nu_1 < b_1 / l_{11}, \nu_2 < (b_2 - l_{21} \nu_1) / l_{22})
\]

Since \( \nu_1, \nu_2, \nu_3 \) are independent random variables, equation (4) can be expressed as a product of univariate CDF, but conditional on unobservables (the \( \nu \)).

Suppose now that we draw a random variable \( \nu^*_1 \) from a truncated standard normal density with upper truncation point of \( b_1 / l_{11} \), and another one, \( \nu^*_2 \), from a standard normal density with upper truncation point of \( (b_2 - l_{21} \nu^*_1) / l_{22} \). These two random variables respect the conditioning events of equations (8) and (9).

Equation (5) is then rewritten as:
\[
\Pr(\nu_1 < b_1 / l_{11}) \Pr(\nu_2 < (b_2 - l_{21} \nu_1^*) / l_{22}) \Pr(\nu_3 < (b_3 - l_{31} \nu_1^* - l_{32} \nu_2^*) / l_{33})
\]

The \( \text{GHK} \) simulator of (4) is the arithmetic mean of the probabilities given by (10) for \( D \) random draws of \( \nu^*_1 \) and \( \nu^*_2 \):
\[
\widetilde{\Pr}_{\text{GHK}} = \frac{1}{D} \sum_{d=1}^{D} \left\{ \Phi \left[ b_1 / l_{11} \right] \Phi \left[ (b_2 - l_{21} \nu_1^{*d}) / l_{22} \right] \Phi \left[ (b_3 - l_{31} \nu_1^{*d} - l_{32} \nu_2^{*d}) / l_{33} \right] \right\}
\]

where \( \nu_1^{*d} \) and \( \nu_2^{*d} \) are the \( d \)-th draw of \( \nu_1^* \) and \( \nu_2^* \), and where \( \Phi(\cdot) \) is the univariate normal CDF.

The simulated probability (11) is then plugged into the likelihood function, and standard maximisation techniques are used.

### 3 An example on artificial data

```
set obs 5000
local rho12=0.3
local rho13=-0.3
local rho23=0.3
drawnorm eps1 eps2 eps3 corr(1, 'rho12', 'rho13' \ /*
  */ 'rho12', 1, 'rho23' \ /*
  */ 'rho13', 'rho23', 1 )
```
drawnorm x1 x2 x3 x4 x5 x6 x7 x8 x9
gen y3=(1+x6+x7+x8+x9+eps3>0)
gen y2=(1+x4+x5+x6+eps2>0)
gen y1=(1+y2+y3+x1+x2+x3+eps1>0) /*note that y2 and y3 are endogenous*/
triprobit ( y1= y2 y3 x1 x2 x3)(y2= x4 x5 x6)(y3 = x6 x7 x8 x9)

trivariate probit, GHK simulator, 25 draws

Comparison log likelihood = -3876.3152

initial: log likelihood = -3876.3152
<output omitted>
Iteration 5: log likelihood = -3838.0791

Number of obs = 5000
Wald chi2(12) = 3576.34
Log likelihood = -3838.0791 Prob > chi2 = 0.0000

| Coef.  Std. Err.  z    P>|z|     [95% Conf. Interval] |
|-------|------------|------|--------|----------------------|
| y1    |            |      |        |                      |
| y2    |  .9232884  | .0927705 | 9.95  | 0.000  | .7414615  1.105115 |
| y3    |  .9222976  | .0765911 | 12.04 | 0.000  | .7721818  1.072413 |
| x1    |  1.065994  | .0470546 | 22.65 | 0.000  | .9737688  1.158219 |
| x2    |  .991229   | .0449885 | 22.03 | 0.000  | .9030532  1.079405 |
| x3    |  1.037427  | .0449885 | 22.88 | 0.000  | .9485477  1.126307 |
| _cons |  1.085532  | .0735326 | 14.76 | 0.000  | .9414105  1.229653 |
|       |            |      |        |                      |
| y2    |            |      |        |                      |
| x4    |  1.000869  | .0338369 | 29.58 | 0.000  | .9345499  1.067188 |
| x5    |  .963295   | .0342638 | 28.31 | 0.000  | .8966047  1.029985 |
| x6    |  1.066905  | .0352756 | 29.85 | 0.000  | .9977661  1.136044 |
| _cons |  1.065994  | .0470546 | 22.65 | 0.000  | .9737688  1.158219 |
|       |            |      |        |                      |
| y3    |            |      |        |                      |
| x6    |  1.023065  | .0353343 | 29.85 | 0.000  | .9538105  1.092319 |
| x7    |  1.023166  | .0351069 | 29.14 | 0.000  | .9543577  1.091974 |
| x8    |  1.03172   | .0347611 | 29.68 | 0.000  | .9635901  1.099851 |
| x9    |  1.017668  | .0348807 | 29.18 | 0.000  | .9493033  1.086033 |
| _cons |  1.015376  | .0326298 | 31.12 | 0.000  | .951423   1.079329 |

athrho12 |
| _cons   |  .1457736  | .0471507 | 3.09  | 0.002  | .05336    .2381872 |

athrho13 |
| _cons   |  -.278662  | .0546056 | -5.10 | 0.000  | -.38567   -.1716371 |

athrho23 |
| _cons   |  .2598698  | .0348018 | 7.47  | 0.000  | .191659   .32808 |

rho12= .14474975 Std. Err.= .04616273 z= 3.1356413 Pr>|z|= .00171479
rho13= -.27166631 Std. Err.= .05057554 z= -5.3714955 Pr>|z|= 7.80e-08
rho23= .25417374 Std. Err.= .03255343 z= 7.8078952 Pr>|z|= 5.773e-15

LR test of rho12=rho13=rho23=0: chi2(3) = 76.472099 Prob > chi2 = 1.752e-16