Parameter estimation in the probit model

The probit model assumes that the response of each unit in the population is 0 or 1. The probability of response 1 depends on covariates $X_1, X_2, \ldots, X_m$ through

$$P(Y = 1|X_1 = x_1, X_2 = x_2, \ldots, X_m = x_m) = \Phi(\alpha + \beta_1 x_1 + \cdots + \beta_m x_m)\,,$$

where $\Phi(x)$ is the CDF of the standard normal distribution. The likelihood function is given by

$$\begin{aligned} \ell(\mathbf{y}, \mathbf{x}, \alpha, \beta) &= \sum_{i=1}^{n} y_i \log(\Phi(\alpha + \beta_1 x_{i1} + \cdots + \beta_m x_{im})) \\ &+ \sum_{i=1}^{n} (1 - y_i) \log(1 - \Phi(\alpha + \beta_1 x_{i1} + \cdots + \beta_m x_{im}))\,. \end{aligned}$$

Denote the sample size by n.

Attached you will find the printout from the LINDEP program. On the basis of the printout do the following.

a. Check that the parameters are estimated by maximum likelihood.

b. Compute the expectations

$$\begin{aligned} f_{11} &= E\left(-\frac{\partial^2 \ell(\mathbf{Y}, \mathbf{X}, \alpha, \beta)}{\partial \alpha^2}\right) \\ f_{12} &= E\left(-\frac{\partial^2 \ell(\mathbf{Y}, \mathbf{X}, \alpha, \beta)}{\partial \alpha \partial \beta}\right) \quad \text{and} \\ f_{22} &= E\left(-\frac{\partial^2 \ell(\mathbf{Y}, \mathbf{X}, \alpha, \beta)}{\partial \beta^2}\right)\,. \end{aligned}$$

Here $\ell(\mathbf{Y}, \mathbf{X}, \alpha, \beta)$ is the full likelihood function. Check that the standard errors of parameter estimates are given by

$$\mathbf{F}^{-1} = \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix}^{-1}$$

Explain why the above matrix can be used to produce standard errors. Would you expect the sampling distribution to be approximately normal. Why?

c. The printout contains the $\chi^2(1)$ statistic. Formulate the null-hypothesis that you think this statistic is testing. Check that it is the value of the likelihood ratio statistic for this hypothesis. Why does the statistic have one degree of freedom?

```
MODEL COMMAND: PROBIT;LHS==Y;RHS==ONE,X$


Method==NEWTON; Maximum iterations == 25
Convergence criteria: Gradient  ==    .1000000E-03
Function  ==    .1000000E-05
Parameters==    .1000000E-04
Starting values:    .7212      -.2109

====> NEWTON Iterations

Iteration  1     Function    212.1327
Param    .721     -.211
Gradnt   19.5      87.1

Iteration  2     Function    189.2619
Param    .686     -.694
Gradnt  -4.33      9.75

Iteration  3     Function    188.7442
Param    .728     -.778
Gradnt  -.276      .416

Iteration  4     Function    188.7430
Param    .731     -.782
Gradnt  -.760E-03  .102E-02

Iteration  5     Function    188.7430
Param    .731     -.782
Gradnt  -.510E-08  .644E-08
** Gradient has converged.
```

```
** Function has converged.
** B-vector has converged.

Maximum Likelihood Estimates

Log-Likelihood.............     -188.74
Restricted (Slopes==0) Log-L.     -233.30
Chi-Squared ( 1)...........      89.121
Significance Level..........      .32173E-13

Variable Coefficient Std. Error T-ratio Prob|t|F2x  Mean of X
Std.D.of X
-------------------------------------------------------------
ONE      .730790     .775747E-01   9.420  .00000  1.0000
.00000
X       -.782305     .925799E-01  -8.450  .00000 -.41779E-01
.94192
```