Developing Simulation Scenarios with the Bivariate Probit Model

Okochab Igohn Chike & Isaac Didi Essi
Department of Mathematics/Computer Science
Rivers State University, Port Harcourt.
Rivers State, Nigeria

Abstract
This study focuses on developing simulation scenarios using the bivariate probit model. In this research, we formulate a bivariate probit model by generating two random binary variables W and W_2 and a non- binary variable Z to enable us study the effect of each independent variable on the dependent variable using different scenarios. The artificial dataset is generated using the econometric software, LIMDEP 9.

Key Word: *Developing, Simulation, Scenarios Model*

1.1 Introduction

People and corporate bodies settle on choices that influence their exercises on the long run. One way or the other, most choices made by people and corporate associations depend on decision. These decisions are made to appraise the esteem ascribed to things by uncovering their inclinations and utilities. One tool that helps to make specific decision is the binary choice model. Binary choice models are models that are utilized to relate and choose a choice from among different options (Greene, 2000).
In finance, binary choice models conform to Lancaster's theory which expresses that an individual buys a thing dependent on the highlights that are joined to the things. So also, McFadden (1974) contends that a customer's conduct is best portrayed by the characteristics of the item he wishes to buy. Consider, for example, the decision of whether to make an essential purchase or the decision of which course is best for transportation, or demonstrating work interest.

An advertising organization might need to foresee buyer's devotion to various brands of items. For example, clients may need to pick where to shop, when to purchase an item and which brand to buy. Promotion professionals had throughout the years, contemplated the components that drive purchaser's decisions.
In social insurance, a patient may need to choose whether or not to buy medical coverage. One of such investigations was done via Cardon and Hendell (2001), when they fundamentally assessed a model for medical coverage and medicinal services decisions, utilizing information on single and wedded patients. They tried for perceptible connections between protection status and social insurance utilization.

Furthermore, binary choice models play an important role in our everyday life. In modeling binary choice, we consider a Bernoulli model, where the chances of an event occurring is given as p and the chances of it not occurring is given as (1-p). In most cases, further estimation is usually done by maximizing the likelihood of an event occurring. For example, an individual’s employment status could be full time, part-time or none while, transportation might involve a choice between travelling by air, land or sea. We refer to such a model as a multinomial choice model.

However, situations might arise where there are two binary dependent variables as well as two latent variables and there is need to model them jointly as a function of some explanatory variable i.e. whether to work or not. In such a situation, we refer to such a model as a bivariate probit model (Green (2003)). It goes further to provide a specification for modeling cases where there is an endogenous binary variable in one of the equations.

Kanyi (2018) thinks about the impact of grandparental care on the employment opportunities of women in the UK. Data from the UK's millennium cohort study was analyzed using the bivariate probit model. In the analysis, labor force participation was modeled using a bivariate probit and instrumental variable approach (Grandparent childcare and participation equation). Thereafter, an instrumental variable was used to measure the association between grandparental care and a mother's employment status. The result from the analysis suggests that grandparental care had a positive impact on the employment opportunities of women in the UK.

The aim of our study is to develop different simulation scenarios using the bivariate probit model. In other to achieve this, we shall consider the following objectives:

- To specify and evaluate binary choice models
- To simulate with binary choice models using generated data

1. Methodology

In line with the aim of this research, we shall explain the methods used in carrying out this work and the rationale behind using the specified models. The data set used in this work, were generated artificially generated using Econometric software, LIMDEP 9.

Probit and Logit Models

Specification

According to Wang (2006), the logit model is based on the odds that an event taking place.

$$\text{Logit (p)} = \ln\left(\frac{P}{1-P}\right) \quad (2.1)$$

If P=Pr(Y=1|Xβ) is the probability of an event happening,

Thenln$\left(\frac{P}{(1-P)}\right)$ is the corresponding log odds. The logit model states the log odds of an event happening are a linear function of a given set of explanatory variables, i.e.

$$\ln\left(\frac{P}{1-P}\right) = X\beta \quad (2.2)$$

The probability P=Pr(Y=1|Xβ) is solved by using

$$P=P(Y=1|X\beta)=\frac{\exp(X\beta)}{1+exp(X\beta)} \quad (2.3)$$

The probability of Y being 0 is given as

$$P(Y=0|X\beta) = 1- P(Y=1|X\beta)= 1- \frac{\exp(X\beta)}{1+(X\beta)}$$

Therefore, the likelihood of the logit model is

$$L(B) = \prod_i^n \left[\frac{\exp(X\beta)}{1+\exp(X\beta)}\right]^{Y_i} \left[1-\frac{\exp(X\beta)}{1+\exp(X\beta)}\right]^{1-Y_i} \quad (2.4)$$

The log likelihood LL(B) is

$$LL(B) = \sum_{i=1}^n Y_i \ln \frac{\exp(X\beta)}{1+exp(X\beta)} + (1- Y_i)\ln\left[1-\frac{\exp(X\beta)}{1+\exp(X\beta)}\right] \quad (2.5)$$

Coefficient estimates are obtained by maximizing LL(B)

The Probit model is given as

$$P(Y=1|XB)=\Phi\,(XB) = \int_{-\infty}^{XB} \emptyset(z)dz \quad (2.6)$$

Where Y is a discrete variable that takes the value 1 or 0; X is a vector of explanatory variables, B is a vector of coefficients.$\Phi(z)$is the cumulative normal distribution.
The likelihood function L(B) is given as

$$L(B)= \prod_i^n [\Phi(XB)]^{Y_i} [1-\Phi(XB]^{1-Y_i} \quad (2.7)$$

And the log likelihoodLL(B) for the probit model is given as

$$LL(B) = \sum_{i=1}^n Y_i \ln[\Phi(XB) + (1-Y_i)\ln[1-\Phi(XB] \quad (2.8)$$

By maximizing LL(B), we obtain estimates of the coefficients.
In this study, we shall be using the probit model.

2.9 Simulating With Generated Variables

In this section, we generated two binary variables W, and W_2 and non-binary variable Z, so that we can find the probit regression of W against W_2 and Z. Also, we find the probit regression of W_2 against Z. After the models have been estimated, we go on to simulate with the models using some.

3. Data Analysis

In this chapter, we present results of the statistical analysis, carried out in the work. In Tables 3.1 to 3.5, generated data are used.
The generated data variables are W, Z and W_2.Where W and $W_2$are binary variables while Z is a non-binary variable.

3.1 Probit Regression with Simulated Data

The probit results with generated data are given in Tables 3.2, 3.3, and 3.4.
In Table 3.2, W is regressed against Z and W_2. In Table 3.3, W_2 is regressed against Z.

Table 3.2: Probit Regression for W against Z and W_2
Dependent Variable: W

Variable	Coefficient	P-value
Constant	-27.1225	0.0004
Z	5.3315	0.0003
W_2	-4.2697	0.0006
AIC =0.2892	Prob[ChiSqd> value]	= 0.0000

Table 3.3: Probit Regression of W_2 against Z
Dependent Variable: W_2

Variable	Coefficient	P-value
Constant	-3.8397	0.0000
Z	0.7712	0.0000
AIC = 1.1431	Prob[ChiSqd> value]	= 0.0000

The results of bivariate regression are in Table 3.5. In the first equation of the bivariate regression, W is regressed against Z and W_2 and in the second equation, W_2 is regressed against Z only.

Table 3.4: Bivariate Probit Regression (Simulated Data)
Dependent Variable for 1st Equation: W

Variable	Coefficient	P-value
Constant	-21.5250	0.9999
Z	4.4766	0.9066
W_2	-4.5554	0.9999

Dependent Variable for 2nd Equation: W_2

Variable	Coefficient	P-value
Constant	-4.1144	0.0006
Z	0.7551	0.0006
Rho (1,2)	0.1684	1.0000

3.5 Simulation Analysis with Generated Data

We create scenarios by subjecting Z to increments of 1, 3, and 5 units and the results are shown in Table 3.6.

Table 3.6: Effect of increasing Z by 1, 3 and 5 units.
Scenario 1: Increasing Z by 1 unit

Outcome	Base case	Under Scenario	Change
	0 55=55.00%	18=18.00%	-37
	1 45=45.00%	82=82.00%	37
Total	100.00%	100.00%	0

Scenario 2: Increasing Z by 3 units

Outcome	Base case	Under Scenario	Change
	055=55.00%	0=0.00%	-55
	145=45.00%	100=100.00%	55
Total	100.00%	100.00%	0

Scenario 3: Increasing Z by 5 units

Outcome	Base case	Under Scenario	Change
	055=55.00%	0=0.00%	-55
	1445=45.00%	100=100.00	55
Total	100.00%	100.00%	0

4 Discussion

Probit Regression of binary variables (W and W2) against a non binary variable (Z)

We observe from Table 3.1, that p-values are all significant ($p<0.05$). Among the selection criteria, the AIC appears to be the criterion with the least error; we therefore choose the AIC as our selection criterion. Since Prob[ChiSqd> value] = 0.0000 ($P< 0.05$), it implies that the regression equation is significant. We therefore, model both equations jointly.

Bivariate Probit Regression (Simulated Data)

From Table 3.4, the positive value of the coefficient of Z, suggests that the arbitrary index ($Z = 4.4766$) has a positive effect on the occurrence in the event. A negative value of the estimate of the coefficient ($W2=-4.5554$) implies that there is less chance of the event occurring. Furthermore, from the second equation, the positive value of the arbitrary index ($Z= 0.7551$) also suggests a positive boost in the probability of an event occurring.

Effect of Increasing Z by 1, 3 and 5 units

From Table 3.6, Scenario 1, we observe that when Z is increased by 1 unit, the probability of occurrence increases from 0.45 to 0.82. This signifies an increase in the chances of the event occurring when there is an increase by one unit. Similarly for scenario 2, when Z is increased by 3 units, the probability of an event occurring increases from 0.45 to 1. Finally, for scenario 3, when Z is increased by 5units, the probability of occurrence shifts from 0.45 to 1.We observe here, that as Z increases in units, the probability of occurrence tends to 1.

5 Conclusion

We shall be looking at the summary and findings of our research work and recommendations based on what we have learnt as we consider our aims and objectives. Furthermore, we shall also evaluate our study and recommendations as they may be obvious opportunities for further research other than the work we have done in the course of this study.

From our analysis, we have achieved the following:

1. Generated artificial dataset in which we have used in specifying and estimating binary choice models.
2. We have created scenarios of change of the explanatory variables with corresponding response of dependent variable and these can be applied by researchers.

References

Cardon, J. H., &Hendel, I. (2001). Asymmetric information in health insurance: Evidence from the national medical expenditure survey. *RAND Journal of Economics*, *32*(3), 408-427. DOI: 10.2307/2696362

Castellano R, Punzo. G., Rocca, A. (2013) Women’s Job Search Propensity and Selection Effect in European Labour Markets. *MetoloskiZvezki*, Vol. 10, pp 83-98.

Chun, H., &Mun, S. (2012) *Determinations of Cooperation in Small and Medium Enterprises, Small Business Economics*, 39, 419-436.

Clementine, S. (2016). Working and Women Empowerment in the Egyptian housedhold: The type of work and location matter: <halsh-01525220>.

Coast, J., Flynn, T. N., &Natarafan, L. (2008). Valuing the ICECAP capability index for older people. Social science and medincine, 67,874-882

Colon, M. S. Moles M.C. (2013) Housing and Labor Decisions of Households. *Review of Economics Household* Vol. pp 55-82.

Compton D, Polak K.A. (2013) Family Proximity, Childcare, and Women’s Labor Force Attachment .*Journal of Urban Economics*, http://dx.doi.org/10.1016j.june.2013.03.007

Compton, J,Polak.R. (2014) Family proximity childcare and women labor force attachment. *Journal of Urban Economics*, vol 79,pp 72-90

Connelly, R. Kimmel J. (2001) Participation and Welfare Recipients of Single Mothers: Implications for Welfare Reform: Upjohn Institute Working Paper 07-69.

Corman H. Noonan, K. Reichman N. (2005) Mothers Labor Supply in Tragile Families: The Role of Child Health. *Eastern Economic Journal*, Vol 31, pp 1-616.

Crespo, L. (2006). "Caring for parents and employment status of European mid-wife women." Paper presented at the Second Workshop on Economics of the Family, held at the University of Zaragoza.

Dobis, E.(2011) A Socio Demographic Analysis of Health Care Utilization in the United States of America.

Ekert O. Stier. H. (2009) Normative or Economic Behaviour? Fertility Women's Employment in Isreal. *Social Science Research* Vol. 38, pp 644-655.

Ettner, S. L. (1996). "The opportunity costs of elder care." *Journal of Human Resources, 31* (1), 189-205

Fechner, G. T. (1860). Elemente der Psychophusik. Leipzig: Breitkpof and Hartel.

Finney, D. (1947). Probit Analysis (1st edition) Cambridge University Press, Cambridge, UK.

Finney, D. (1971). *Probit Analysis* (Thirded). Cambridge: Cambridge University Press. First edition in 1947.

Frontiers in Econometrics. New York: Academic Press.

Garganta S. Gaspararini L. Marchison M. Tappata M. (2017) The Effect of Cash Transfers on Fertility: Evidence from Argentina. *Population Residents Policy Review* Vol 36, pp 1-24.

Gensler, H. Walls D. (2006) Labor Force and Welfare Program Participation: The Effects of Welfare. *American Journal of Economics and Sociology*. Vol. 56, pp 229-241.

Green,W., (2003) Econometric Analysis, Prentice Hall, Upper Saddle River. New

Kanyi,S.(2018). Grand Parent Care: A key factor in mothers labor force participation in the UK. *Journal of Sociological Policy, 49*,3, 523–542