International Journal of Multidisciplinary Research and Development

IJMRD 2015; 2(3): 20-26
www.allsubjectjournal.com
Impact factor: 3.672
Received: 08-02-2015
Accepted: 22-02-2015
E-ISSN: 2349-4182
P-ISSN: 2349-5979

Muhammad Alamgir Islam
Department of Statistics,
University of Chittagong,
Chittagong- 4331, Bangladesh

Md. Shahidul Hoque
Department of Statistics,
University of Chittagong,
Chittagong- 4331, Bangladesh

Correspondence:
Muhammad Alamgir Islam
Department of Statistics,
University of Chittagong,
Chittagong- 4331, Bangladesh

Multinomial logistic model for mode of occupational choices

Muhammad Alamgir Islam, Md. Shahidul Hoque

Abstract
In this paper we have derived an unordered multinomial logistic model for mode of occupational choices for the outgoing graduate students of Bangladesh. The paper finds out the factors which motivate a student after completion of graduation to choose one's profession. For derivation of multinomial logistic model for mode of occupation choices, we have considered the factor that is likely to influence choosing occupation of a student as a response variable and considered five factors such as, sex of the respondents, family economic status, place of permanent residence, discipline of study of the respondents and result obtained in honors program as determinants of the mode of occupational choices. In this study information from a sample of 1000 students from each of four major faculties of six well established non-technical public universities through well defined questionnaire. Stratified random sampling technique with proportional allocation scheme has been used in order to determine sample size under each selected faculty of each university. However, after scrutiny, 690 students with consistent and valid information have been finally used for analysis. The parameters of our multinomial logistic models for mode of occupation choices have been estimated by using maximum likelihood method. From this study it has been found that the likelihood of male respondents would choose business as occupation is 5.930 times more than female respondents. The respondents who do not have good result would choose occupation in a private firm 1.575 times more likely than respondents who have excellent result and so on. The findings of this study may be helpful to the students to build up themselves according to the likelihood of professions to be chosen on the basis of their socio-economic and academic backgrounds. This would also help the government to create job opportunity accordingly.

Keywords: Occupation, Multinomial Logistic Model, Maximum Likelihood Method, Likelihood Ratio Test.

1. Introduction

Bangladesh is one of the least developed countries in the world. In most of the developing countries like Bangladesh the frame of education is in the face of multifarious problems. Education directly or indirectly depends on society and economic situation. There are thirty-four public universities in Bangladesh which are providing higher education to the bulk of students. These universities are funded by the government while managed as self-governed organizations. Occupation selection is an important factor for any student after completion of his/her graduation from University. Although most of the students relish an occupation in mind for livelihood that they think would be better career of life, even then due to different practical constraints, after completion of their graduation from Universities, they stay in a dilemma to take decision regarding the occupation they will choose. Our present study will mainly emphasize on graduate students' choice of occupation. It will also help to find out the factors which motivate a student to choose his/her profession after successful completion of graduation. Multinomial logistic regression models are used to explain the association between a polytomous response variable and a set of regressors. These polytomous response models can be classified into two distinct types, depending on whether the response variable has an ordered or unordered structure. In the ordered logistic models the response variable has more than two ordered or ranked categories. But there are situations where the regressand is unordered. For example, the choice of transportation mode to an employee may be plane, rickshaw, car, bus or train. Several theoretical and empirical works on this multinomial logistic modeling technique are available "Ref. [3, 4, 5, 6, and 7]".The present study attempts to derive an unordered multinomial logistic model for mode of occupational choices for the outgoing graduate students of Bangladesh. The estimation of the parameters of our postulated multinomial logistic model has been carried out by using maximum likelihood procedure "Ref. [2]".

Multinomial logistic model

Multinomial logistic models "Ref. [1, 8]"are used to explain the relationships between a polytomous response variable and a set of qualitative and/or quantitative regressors .Multinomial logistic regression, is a regression model which generalizes logistic regression by allowing more than two discrete outcomes. That is, it is a model that is used to predict the probabilities of the different possible outcomes of a categorically distributed dependent variable, given a set of independent variables (which may be real-valued, binary-valued, categorical-valued, etc.).

Suppose a dependent variable has M categories. One value (typically the first, the last) of the dependent variable is designated as the reference category. The probability of membership in other categories is compared to the probability of membership in the reference category.

For a dependent variable with M categories, this requires the calculation of M-1 equations, one for each category relative to the reference category, to describe the relationship between the dependent variable and the independent variables.

Hence, if the first category is the reference, then, for m= 1, 2 …M

$$\ln \frac{p(Y_i=m)}{p(Y_i=1)} = \alpha_m + \sum_{k=1}^{K} \beta_{mk} X_{ik} = Z_{mi}$$

Hence, for each case, there will be M-1 predicted log odds, one for each category relative to the reference category. (Note that when m = 1 you get ln (1) = 0 = Z_{1i}, and exp (0) = 1.)

When there are more than 2 groups, computing probabilities is a little more complicated than it was in logistic regression. For m = 2… M,

$$P(Y_i = m) = \frac{\exp(Z_{mi})}{1 + \sum_{h=2}^{m} \exp(Z_{hi})}$$

For the reference category,

$$P(Y_i = 1) = \frac{1}{1 + \sum_{h=2}^{m} \exp(Z_{hi})}$$

Note that, when M = 2, the mlogit and logistic regression models (and for that matter the ordered logit model) become one and the same.

Objectives of the study

The foremost objectives of our research work are as follows:

(1) To derive a multinomial logistic model for the mode of occupational choices.
(2) To recognize the factors, which motivate a graduate student to choose his/her occupation
(3) To find the odds of choosing different types of occupation under different types of socio-economic and academic backgrounds.

2. Data and Methodology

In this study the multinomial logistic modeling technique has been employed to find the likelihood of choice of occupation of the students. The area of 6 public non-technical Universities is covered by the survey. The universities are Dhaka, Chittagong, Rajshahi, Jahangirnagar, Khushtia Islamic University and Comilla University. The faculties considered in this study are Science, Commerce, Arts and Social Sciences. Here, sampling units are the outgoing M.S. students of public Universities. We have collected information using a stratified random sampling technique with proportional allocation "Ref. [9]", considering each faculty of each university as a stratum. In this study, we have considered a convenient sample of 1000 students from these universities through a questionnaire method which contained open & structured questions, only 690 consistent respondents have been considered in final analysis. The final data comprise with the students of Science faculty (24.8%), Arts faculty (20.6%), Business Administration faculty (27.20%) and Social Science faculty (27.40%).

3. Derivation of Multinomial Logistic Model for Mode of Occupational Choices

Thus for derivation of a multinomial logistic model for the mode of occupational choices, variables are categorized as below.

3.1. Categories of Response Variable Occupational Choices

Following three categories of occupational choices are considered

1) Government services
2) Private firm
3) Business

3.2. The Possible Factors that Might Influence the Occupational Choices

There are several factors that might influence the occupational choices of students. Here we have considered only five such factors.

1) Sex of the respondents.
2) Family economic status.
3) Place of permanent residence.
4) Discipline (faculty) of study of the respondents
5) Result (in Honors).

3.3. Categories of regressor variables

Sex of the Respondents:
(a) Male
(b) Female
(c) Social science and
(d) Science

Discipline of Study:
(a) Arts
(b) Commerce

Family Economic Status:
(a) Not good (below GPA -3.0)
(b) Just good (GPA -3.0 to 3.5)
(c) Excellent (above GPA 3.5)

Result in Honors:
(a) Lower class
(b) Middle class
(c) Upper class

Place of Permanent Residence:
(a) Rural area
(b) Semi urban area
(c) Urban area

3.4. Operational Definition

Dependent variable

Reference category: Government Services
1= business
0= otherwise
1=private firm
0= otherwise

Independent variables

1) Sex of the respondents: Reference category female
   1= male
   0= otherwise
2) Discipline of study of the respondents: Reference category: Science faculty
   1=Arts
   0= otherwise
   1=Commerce
   0= otherwise
   1=Social science
   0= otherwise
3) Family economic status: Reference category: Upper class family
   1= lower class
   0= otherwise
   1= middle class
   0= otherwise
4) Result in Honors: Reference category: Excellent result
   1= not good
   0= otherwise
   1= Just good
   0= otherwise
5) Place of permanent residence: Reference category: Urban area
   1= rural
   0= otherwise
   1= semi urban
   0= otherwise

3.5. Selected Model

The postulated multinomial logistic model for mode of occupational choices is as follows:

$$\ln(\frac{p_1}{p_0}) = \beta_{10}+\beta_{11}\ SEX_1 +\beta_{12}STATUS_1 + \beta_{13}\ STATUS_2 +\beta_{14}\ RESI_1 + \beta_{15}\ RESI_2 + \beta_{16}\ FACULTY_1 + \beta_{17}\ FACULTY_2 + \beta_{18}\ FACULTY_3 + \beta_{19}\ RESULT_1 + \beta_{1\,(10)}\ RESULT_2 + U_1$$

And

$$\ln(\frac{p_2}{p_0}) = \beta_{20}+ \beta_{21}\ SEX_1 + \beta_{22}\ STATUS_1 + \beta_{23}\ STATUS_2 + \beta_{24}\ RESI_1 + \beta_{25}\ RESI_2 + \beta_{26}\ FACULTY_1 + \beta_{27}\ FACULTY_2 + \beta_{28}\ FACULTY_3 + \beta_{29}\ RESULT_1 + \beta_{2\,(10)}\ RESULT_2 + U_2$$

Here first equation is defined for business vs. government service and second equation is defined for private firm vs. government service where β_{10} and β_{20} denotes the constant terms for all reference categories of respondents, such as female, have upper family status, come from urban area, study at science faculty and have excellent honors result.

Sex_1 is a dummy variable denoting the two categories for sex: male and female.

$STATUS_1$ and $STATUS_2$ are set of two categories representing the dummy variable family status of the respondents, such as: lower class, middle class respectively, where upper class is considered as reference category.

$RESI_1$ and $RESI_2$ are set of two categories representing the dummy variable permanent place of residence of the respondents, such as: rural area, semi urban area respectively, where urban area is considered as reference category.

$FACULTY_1$, $FACULTY_2$ and $FACULTY_3$ are set of three categories representing the dummy variable faculty, such as: Arts, Commerce and Social science respectively, where science faculty is considered as reference category.

$RESULT_1$ and $RESULT_2$ are set of two categories representing the dummy variable honors result of the respondents, such as: not good, just good respectively, where excellent result is considered as reference category.

4. Estimation and Test

Here the estimation of the parameters of the multinomial logistic model has been carried out by using maximum likelihood procedure. Likelihood ratio test and Wald test have been employed for testing the formulated hypothesis.

Summary of Sample Data

We have considered 690 filled-in questionnaires for final study. The summaries of sample data are as follows

Table1: Summary of sample data

Background Characteristics	Category	N	Marginal Percentage
Choices of Occupation	Business	108	15.7%
	Private firm	283	41.0%
	Government Service	299	43.3%
Sex of the respondents	Male	440	63.8%
	Female	250	36.2%
Family economic status of the respondents	Lower class	96	13.9%
	Middle class	431	62.5%
	Rich class	163	23.6%
Type of residence	Rural	205	29.7%
	Semi Urban	272	39.4%
	Urban	213	30.9%
Discipline of study of the respondents	Arts	142	20.6%
	Commerce	188	27.2%
	Social science	189	27.4%
	Science	171	24.8%
Honors result of the respondents	Not good(below GPA-3.0)	115	16.7%
	Just good(GPA3.0 to3.5)	391	56.7%
	Excellent(above GPA- 3.5)	184	26.7%
Valid		690	100.0%
Missing		0	
Total		690	

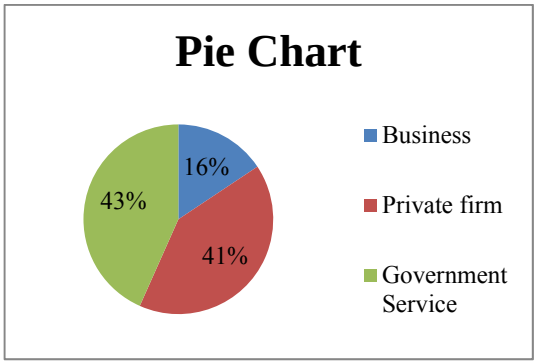

Fig 1: Pie chart for the occupational choices of the respondents.

Comment

From the above table and diagram it is found that, 16% respondents would like to choose their occupation as business, 41% respondents would like to choose their occupation in a private firm, and 43% respondents would like to choose their occupation as government service. Therefore we may conclude that most of the respondents would choose their occupation as government service.

Table2: Case processing summary

		N
Choices of occupation	1	108
	2	283
	3	299
Sex of the respondents	1	440
	2	250
Family status of the respondents	1	96
	2	431
	3	163
Type of residence	1	205
	2	272
	3	213
Discipline of study of the respondents	1	142
	2	188
	3	189
	4	171
Honors result of the respondents	1	115
	2	391
	3	184
Missing		0
Total		690

Model Fitting Information

By using SPSS program we get the following model fitting information

Likelihood Ratio Tests

Hypothesis:

H_0: There is no significant difference among the effects of selected variables.

Table 3: Likelihood ratio test

Model	-2 Log Likelihood	Chi-Square	Df	Sig.
Intercept Only	993.704			
Final	811.160	182.544	20	.000

Likelihood Ratio Test

Effect	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	8.112E2[a]	.000	0	.
sex	842.079	30.919	2	.000
status	864.481	53.320	4	.000
resi	831.072	19.912	4	.001
faculty	866.682	55.522	6	.000
result	834.187	23.027	4	.000

Here the reduced model is formed by omitting an effect from the final model. This reduced model is equivalent to the final model because omitting the effect does not increase the degrees of freedom.

Table 4: Estimation and Wald test

Occupation[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
								Lower Bound	Upper Bound
1	Intercept	-1.415	.498	8.078	1	.004			
	[sex=1]	1.780	.368	23.348	1	.000	5.930	2.881	12.209
	[sex=2]	0[b]	.	.	0	.	.	.	.
	[status=1]	-.975	.406	5.764	1	.016	.377	.170	.836
	[status=2]	-2.005	.305	43.326	1	.000	.135	.074	.245
	[status=3]	0[b]	.	.	0	.	.	.	.
	[resi=1]	-.428	.325	1.736	1	.188	.652	.345	1.232
	[resi=2]	-1.311	.313	17.533	1	.000	.270	.146	.498
	[resi=3]	0[b]	.	.	0	.	.	.	.
	[faculty=1]	.838	.363	5.342	1	.021	2.312	1.136	4.707

	[faculty=2]	.299	.381	.617	1	.432	1.349	.639	2.847
	[faculty=3]	-.419	.371	1.279	1	.258	.658	.318	1.360
	[faculty=4]	0[b]	.	.	0	.	.	.	.
	[result=1]	.297	.482	.380	1	.537	1.346	.524	3.461
	[result=2]	1.061	.329	10.379	1	.001	2.888	1.515	5.505
	[result=3]	0[b]	.	.	0	.	.	.	.
2	Intercept	.268	.330	.659	1	.417			
	[sex=1]	.122	.183	.448	1	.503	1.130	.790	1.617
	[sex=2]	0[b]	.	.	0	.	.	.	.
	[status=1]	-.645	.315	4.198	1	.040	.525	.283	.972
	[status=2]	-.297	.232	1.629	1	.202	.743	.471	1.172
	[status=3]	0[b]	.	.	0	.	.	.	.
	[resi=1]	-.326	.235	1.916	1	.166	.722	.455	1.145
	[resi=2]	-.503	.222	5.164	1	.023	.604	.392	.933
	[resi=3]	0[b]	.	.	0	.	.	.	.
	[faculty=1]	-.165	.272	.366	1	.545	.848	.497	1.446
	[faculty=2]	1.031	.248	17.259	1	.000	2.803	1.724	4.558
	[faculty=3]	-.370	.241	2.342	1	.126	.691	.431	1.109
	[faculty=4]	0[b]	.	.	0	.	.	.	.
	[result=1]	.454	.271	2.811	1	.094	1.575	.926	2.678
	[result=2]	-.058	.203	.082	1	.774	.944	.634	1.404
	[result=3]	0[b]	.	.	0	.	.	.	.

a. The reference category is: 3 (Government service)

b. This parameter is set to zero because it is redundant (Reference category)

Comment:

Based on likelihood ratio test, it has been found that all parameters are statistically significant. However, Wald test indicates that there are some parameters those are not significant, namely: $RESI_1$, $FACULTY_2$, $FACULTY_3$, and $RESULT_1$ for the first equation (business vs. government service); and SEX_1, $STATUS_2$, $RESI_1$, $FACULTY_1$, $FACULTY_3$, $RESULT_1$ and $RESULT_2$ for the second equation (private firm vs. government service).

Estimated Model

Based on the table 4, the estimated multinomial logistic equations:

$$\ln(\frac{p_1}{p_0}) = -1.415+1.780SEX_1 - 0.975\ STAUS_1-2.005STAUS_2-0.428RESI_1-1.311RESI_2+.838FACULTY_1 + 0.299\ FACULTY_2 - 0.419\ FACULTY_3 + 0.297\ RESULT_1+ 1.061\ RESULT_2$$

And

$$\ln(\frac{p_2}{p_0})= 0.268 + 0.122SEX_1 -0.645STAUS_1 - 0.297\ STAUS_2-0.\ 326RESI_1- 0.503RESI_2-0.165FACULTY_1 +1.031\ FACULTY_2- 0\ .370\ FACULTY_3- 0.454\ RESULT_1 - 0.058\ RESULT_2$$

5. Interpretation of Estimated Model

The fitted model can be interpreted as follows:

5.1 Interpretation of the Intercepts

For those respondents who are female, have upper family status, come from urban area, study at science faculty, and have excellent honors result, all the selected regressors will have no effect.

For the first equation we have,

$\ln(p_1/ p_0) = -1.415$

$(p_1/ p_0) = \text{Exp}(-1.415)$

$(p_1/ p_0) = 0.2429$

$P_1= 0.2429\ p_0$

Therefore, the respondents who are female, have upper family status, come from urban area, study at science faculty, and have excellent honors result, have the likelihood of choosing business as occupation is 0.2429 times than the government service. That means, this section of graduates would prefer joining in a business farm to the government service.

For the second equation we have,

$\ln(p_2/ p_0) = 0.268$

$(p_2/ p_0) = \text{Exp}(0.268)$

$(p_2/ p_0) = 1.3073$

$P_2 = 1.3073p_0$

Therefore, the respondents who are female, have upper family status, come from urban area, study at science faculty, and have excellent honors result, have chance of choosing private firm 1.3073 times more than the government service. That means, this section of graduates would like to join in a private firm more than government service.

5.2 Interpretation of the Slope Coefficients

The slope coefficients of our estimated model based on table 4, for parameter estimates and Wald test may be interpreted in terms of odds ratio as follows:

For the first equation

For the coefficient of SEX_1,we have Exp (B) = 5.930, this suggests that respondents who are male, would choose business as occupation likely 5.930 times more than respondents who are female.

For the coefficient of $STATUS_1$, we have Exp(B) = 0.377, this suggests that respondents who come from lower class family, would choose occupation as business likely 0.377

times more than respondents who come from upper class family and so on.

For the coefficient of $RESI_1$, we have Exp(B) = 0.652, this suggests that respondents who come from rural area, would choose occupation as business likely 0.652 times more than respondents who come from urban area and so on.

For the coefficient of $FACULTY_1$, we have Exp(B) = 2.312, this suggests that respondents who study at Arts faculty, would choose business as occupation likely 2.312 times more than respondents who study at Science faculty and so on.

For the coefficient of $RESULT_2$, we have Exp (B) = 2.888, this suggests that respondents who have just good result, would choose occupation as business likely 2.888 times more than respondents who have excellent result and so on.

For the second equation

For the coefficient of SEX_1, we have Exp (B) = 1.130, this suggests that respondents who are male, would choose occupation in a private firm likely 1.130 times more than respondents who are female.

For the coefficient of $STATUS_1$, we have Exp(B) =0.525, This suggests that respondents who come from lower class family, would choose occupation in a private firm likely 0.525 times more than respondents who come from upper class family and so on.

For the coefficient of $RESI_2$, we have Exp(B) = 0.604, this suggests that respondents who come from semi urban area, would choose occupation in a private firm likely 0.604 times than respondents who come from urban area and so on.

For the coefficient of $FACULTY_2$, we have Exp(B) = 2.803,this suggests that respondents who study at commerce faculty, would choose occupation in a private firm likely 2.803 times more than respondents who study at Science faculty and so on.

For the coefficient of $RESULT_1$, we have Exp (B) = 1.575, this suggests that respondents who have not good result, would choose occupation in a private firm likely 1.575 times more than respondents who have excellent result and so on.

6. Summary and Focus of Current Research

From this study it is found that, 16% respondents would like to choose their occupation as business, 41% respondents would like to choose their occupation in a private firm, and 43% respondents would like to choose their occupation as government service. Therefore we may conclude that most of the respondents would choose their occupation as government service.

Based on likelihood ratio test, it has been found that all parameters are statistically significant. However, Wald test indicates that there are some parameters those are not significant, namely: $RESI_1$, $FACULTY_2$, $FACULTY_3$, and $RESULT_1$ for the first equation (business vs. government service); and SEX_1, $STATUS_2$, $RESI_1$, $FACULTY_1$, $FACULTY_3$, $RESULT_1$ and $RESULT_2$ for the second equation (private firm vs. government service).

The likelihood of choosing business as occupation for the respondents who are female, have upper family status, come from urban area, study at science faculty, and have excellent honors result, have the likelihood of choosing business as occupation is 0.2429 times than the government service.

On the other hand the respondents who are female, have upper family status, come from urban area, study at science faculty, and have excellent honors result, have likelihood of choosing private firm 1.3073 times more than that of government service.

The likelihood that the male respondents would choose business as occupation is 5.930 times more than female respondents.

The respondents who come from lower class family would choose occupation as business likely 0.377 times more than respondents who come from upper class family.

The respondents who come from rural area would choose occupation as business likely 0.652 times more than respondents who come from urban area.

The respondents who study at Arts faculty would choose business as occupation 2.312 times more likely than respondents who study at Science faculty.

The respondents who do not have good result would choose occupation as business is 2.888 times more likely than respondents who have excellent result.

Again, the likelihood that the male respondents would choose occupation in a private firm likely 1.130 times more than female respondents.

The respondents who come from middle class family would choose occupation in a private firm likely 0.743 times more than respondents who come from upper class family.

The likelihood of respondents who come from semi urban area would choose occupation in a private firm likely 0.604 times more than respondents who come from urban area**.**

The respondents who study at commerce faculty would choose occupation in a private firm likely 2.803 times more than respondents who study at Science faculty.

The respondents who do not have good result would choose occupation in a private firm 1.575 times more likely than respondents who have excellent result.

Policy implications

The study leads us to the following policies to regard awareness and creation of employment opportunities:

The findings of this study may be helpful the students to build up themselves according to the likelihood professions to be chosen on the basis of their socio-economic and academic background. The outcome of the thesis may also make the government aware of the service sectors students are most likely to seek and can help the government create opportunities accordingly.

The thesis at the same time can help the government take proper reformations for those jobs which are neglected so that students also make a rush to those jobs.

The government should also take necessary steps for creating a business friendly environment for those students who drawn of being successful business entrepreneurs.

References

1. Ying So,Warren F. Kuhfeld: Multinomial Logit Models, MR-2010G.
2. Herman J. Bierens: *The LogitModel:Estimation,Testing and Interpretation*, Lecture notes, Undergraduate econometrics, October 25, 2008.

3. D. McFadden and K. Train. Mixed mnl models of discrete response. Journal of Applied Econometrics, 15(5):447–470, 2000.
4. Leonardo Grilli · Carla Rampichini: "A multilevel multinomial logit model for the analysisof graduates' skills" Stat. Meth. & Appl. (2007) 16:381–393 DOI 10.1007/s10260-006-0039-z.
5. AbolfazlMohammadian, Pavlos S. Kanaroglou: Applications of Spatial Multinomial Logit Model to Transportation Planning, 10th International Conference on Travel BehaviourResearchLucerne, 10-15, August 2003.
6. Ivan Pryanishnikov, Katarina Zigova: Multinomial Logit Models for the Austrian Labor Market, AUSTRIAN JOURNAL OF STATISTICS Volume 32 (2003), Number 4, 267–282.
7. Berkson, J. (1944).Application of the logistic function to bio-assay. Journal of the American Statistical Association 39, 357-365.
8. Multinomial Logit Models – Overview (This is adapted heavily from Menard's Applied Logistic Regression analysis ;)
9. William J. Cochran: Sampling Technique, 3rd Edition, July 1977.