J Nepal Health Res Counc 2011 April;9(18):61-6

Fitting Ordinal Regression Analysis to Anthropometric Data

Pradhan A[1]

[1]Department of Community Medicine, KIST Medical College, Lalitpur, Nepal.

ABSTRACT

Background: National level surveys on nutritional status since 1975 to 2006 in Nepal do not indicate the satisfying level of nutrition. Nepal demographic and health survey 2006 uncover that the percent prevalence for underweight and wasted children of under five years of age are 39% and 13% and 49 % of the under five children are stunted. Understanding the factors that affect the nutrition of children is essential. Some studies in other countries show wealth index, size at birth and education as significant contributors. This analysis analyze the factors associated with nutritional status among children of under five years of age.

Methods: This study was cross sectional which used secondary data of the Demographic and Health Survey, 2006 conducted in Nepal. STATA 9, SPSS 13 and SPSS 17 are used for analysis. In this analysis, the outcome variables namely stunting, underweight and wasting are in ordered form. Hence ordinal regression is considered as suitable method.

Results: Ordinal regression well suit the data to model nutritional status through different predictors in case of underweight and wasting however stunting model fails to satisfy the assumption behind ordinal regression. Hence for stunting, model with constraints imposed to certain variables is formed.

Conclusions: Underweight is seen significantly less in households with high wealth index, among children of big and average size at birth and among educated women. Education, wealth index and size at birth are found important factors affecting wasting among children. Wealth index and education of mothers are significantly affecting for stunting among children.

Keywords: ordinal regression, stunting, underweight, wasting.

INTRODUCTION

Nepalese children show evidences of malnutrition indicated by their stunting, wasting or wasting and stunting combined.[1] National surveys on nutrition from 1975 to 1990, indicates the worse picture of nutrition in the country.[2] Nepal Demographic and Health Survey (NDHS) 2001 shows that the prevalence for underweight and wasted children of under five years of age are 48.3% and 9.6% and around 50% are stunted.[3] Similarly, NDHS 2006 reveals that the percent prevalence for underweight and wasted children of under five years of age are 39% and 13%. Forty nine percent of the under five children are stunted.[4] Rapid decline is not observed in stunting and slight decline in underweight can be seen. However percentage of wasted children shows increasing trend.

Understanding the factors that affect the nutrition of children is essential. This analysis tries to analyze the factors associated with nutritional status among children of under five years of age.

Correspondence: Mrs. Amita Pradhan, Department of Community Medicine, KIST Medical College, Lalitpur, Nepal. Email: amiseason@yahoo.com, Phone: 9841302137.

METHODS

This study was cross sectional which used secondary data of the Demographic and Health Survey (DHS), 2006 conducted in Nepal. STATA 9, SPSS 13 and SPSS 17 were used for analysis.

First step of the analysis is to check for the type of outcome variable. In this analysis, the outcome variables namely stunting, underweight and wasting are in ordered form. Hence ordinal regression is considered as suitable method to develop the model. Second step is to determine the predictors. Hence cross table and chi square test is carried out for stunting, underweight and wasting with different predictors such as type of area of residence, sex of household head, sex of child, size at birth, education of mothers, occupation of the mothers, whether mother works at home or away and who decides how to spend money. The checking for empty cells is also determined by the crosstabs of the response variable by each of the categorical predictor variables, and those tables looked fine, hence further analysis is done. The bivariate analysis shows that sex of the child is not significantly associated with response variables. Similarly, works at home or away and who decides how to spend money results with lots of missing cell frequencies. Hence final model is acquired using other variables except these three variables. The reduced model is also tried without some of the explanatory variables like sex of household head and type of area of residence.

RESULTS

Underweight as outcome variable comes with pseudo $R^2 = 0.11$ for full model (Table 1). Model fitting information shows that difference in – 2loglikelihood for final and intercept only model is significant (chi square= 525.7, $p<0.001$). This means the final model is outperforming the model with zero coefficient effect for all the predictors. The additional model fitting statistic, the Pearson's chi-square, ($X^2 = 11828.4$ with d.f. of 12026 and $p = 0.89$) for the complete model with the logit link indicated that the observed data were consistent with the estimated values in the fitted model. Similarly another statistic deviance chisquare($X^2 = 8686.5$ with d.f. of 12026 and $p = 1$) is also consistent with pearson's chi square test. However reduced model comes up with smaller R^2 and violated test of parallel lines assumption. Test of parallel lines come with high insignificant value for full model satisfying the assumption ($X^2=23.34$, df=20, $p=0.27$).

One of the assumptions underlying ordinal logistic regression is that the relationship between each pair of outcome groups is the same. In other words, ordinal logistic regression assumes that the coefficients that describe the relationship between, say, the lowest versus all higher categories of the response variable are the same as those that describe the relationship between the next lowest category and all higher categories, etc. This is called the proportional odds assumption or the parallel regression assumption. Because the relationship between all pairs of groups is the same, there is only one set of coefficients (only one model). If this was not the case, we would need different models to describe the relationship between each pair of outcome groups. So, it is needed to test the proportional odds assumption. The null hypothesis of this chi-square test is that there is no difference in the coefficients between models, so a non-significant result is required.[5] The alternative hypothesis stated that the corresponding regression coefficients were different across all levels of the outcome variable.[6]

Wealth index factor score, sex of household head, size of baby at birth and occupation of mothers (only for formally employed) are found negatively associated with underweight. Statistical significance is seen only in case of wealth index factor score and size at birth. For one unit increase in wealth index factor score, expected ordered log odds decreases by 0.34 as move to more severe category of underweight keeping all other variables in the models constant. Size at birth comes up as another important variable showing that normal and other lower categories of underweight are more likely among normal to big size babies at birth. The odds ratio column explains that there is less odds of higher category of underweight versus normal among big and average sized children at birth as compared to smaller children at birth. Area of residence and educational status of mothers emerges as other main variable with positive association with underweight. This means that higher categories of underweight are more likely among rural area and low educational status, however only education of mothers is observed as statistically significant variable. There is 1.88 and 1.25 times higher odds for higher categories of underweight versus normal among non educated and primary educated mothers as compared to secondary or above level educated mothers. Likewise same higher odds are observed for mild versus higher categories of underweight for lower educational status.

In case of occupation of mothers, mixed results with negative coefficient for formally employed mothers and positive coefficients for other occupational groups, is seen. This shows low likelihood of underweight children among formally employed women and high likelihood for other occupational categories though the effect of occupation of mothers on underweight is not found statistically significant.

Table 1. Parameter estimates for underweight.

Predictor Variables	Estimate	SE	Wald	df	p	95 % CI for estimate		OR	95 % CI for OR	
						Lower	Upper		Lower	Upper
Wealth index	-0.340	0.047	53.392	1	<0.001	-0.431	-0.249	0.71	0.65	0.78
Area of residence										
Rural	0.068	0.077	0.776	1	0.378	-0.083	0.219	1.07	0.92	1.25
Urban*										
Sex of household head										
Female	-0.025	0.067	0.141	1	0.707	-0.157	0.107	0.98	0.85	1.11
Male*										
Size at birth										
Big	-1.161	0.088	175.065	1	<0.001	-1.333	-0.989	0.31	0.26	0.37
Average	-0.690	0.072	91.265	1	<0.001	-0.832	-0.548	0.50	0.44	0.58
Small*										
Educational status of mothers										
No education	0.633	0.081	61.240	1	<0.001	0.474	0.791	1.88	1.61	2.21
Primary	0.224	0.094	5.644	1	0.018	0.039	0.408	1.25	1.04	1.50
Secondary and above*										
Occupational status of mothers										
Formally employed	-0.120	0.123	0.958	1	0.328	-0.361	0.120	0.89	0.70	1.13
Agriculture based work	0.024	0.082	0.089	1	0.765	-0.136	0.185	1.02	0.87	1.20
Domestic and unskilled work	0.264	0.334	0.626	1	0.429	-0.391	0.920	1.30	0.68	2.51
No paid job*										

*Reference category

Wasting is observed that the lower categories are in higher frequency, we tried negative log log function while conducting ordinal regression. However pseudo R^2 decreases to 0.03 for the full model and also for reduced model (without some of the explanatory variables like place of residence and sex of household head). The details are not included. Hence we used the logit link function for wasted as outcome variable.

Final model, that includes the specified predictor variables and has been arrived at through an iterative process that maximizes the log likelihood of the outcomes seen in the outcome variable, does better with significant chi square for difference in -2loglikelihood from reduced model with zero effects of all parameters. The additional model fitting statistic, the Pearson's chi-square, (X^2 = 12072.8 with d.f. of 12755 and p = 1) for the complete model with the logit link indicated that the observed data were consistent with the estimated values in the fitted model. Moreover deviance chi square shows similar results. Likewise comparable results are obtained with negative log log link. Hence no difference observed using different links. Test of parallel lines come with high insignificant value for logit link(p=0.85) as compared to negative log log link(p=0.52). Hence detail results of model using logit link is shown here. Test of parallel lines come with high insignificant value for full model satisfying the assumption (X^2=13.57, df=20, p=0.85).

The negative association of wealth index, sex of household head, size at birth while positive association of area of residence and mixed relation of education and occupation of mothers with wasting among children (Table 2). Moreover, wealth index and size at birth are found as significant contributors in the model while education and occupation of mothers show varied results.

For one unit increase in wealth index expected ordered log odds decreases by 0.17 while moving to more severe category of wasting while all other variables in the model are held constant. Similarly, big to normal size at birth shows less proportional odds for different categories of wasting versus normal or mild versus higher categories of wasting. This means either normal or lower categories of wasting are more likely among big to normal size at birth. Non educated mothers tend to have children with higher categories of wasting, though found insignificant. But, primary educated women are more likely to have normal or other lower categories of wasted children and the proportional odds is found to be lower for higher categories of wasting versus normal and same proportional odds for mild versus other higher categories. Both formally employed and agriculture based working mothers are more likely for having normal or lower categories of wasted children and found statistically significant. However, domestic and unskilled workers shows higher likelihood for higher categories of wasted children but observed insignificant.

Table 2. Parameter estimates for wasting.

Predictor Variables	Estimate	SE	Wald	df	p	95 % CI for estimate		OR	95 CI for OR	
						Lower	Upper		Lower	Upper
Wealth index	-0.171	0.048	12.902	1	<0.001	-0.265	-0.078	0.84	0.77	0.93
Area of residence										
Rural	0.133	0.077	2.933	1	0.087	-0.019	0.285	1.14	0.98	1.33
Urban*										
Sex of household head										
Female	-0.094	0.068	1.9	1	0.168	-0.227	0.04	0.91	0.8	1.04
Male*										
Size at birth										
Big	-0.898	0.087	105.654	1	<0.001	-1.069	-0.726	0.41	0.34	0.48
Average	-0.436	0.07	39.018	1	<0.001	-0.573	-0.299	0.65	0.56	0.74
Small*										
Educational status of mothers										
No education	0.093	0.08	1.342	1	0.247	-0.064	0.25	1.1	0.94	1.28
Primary	-0.199	0.095	4.337	1	0.037	-0.386	-0.012	0.82	0.68	0.99
Secondary and above*										
Occupational status of mothers										
Formally employed	-0.267	0.126	4.451	1	0.035	-0.514	-0.019	0.77	0.6	0.98
Agriculture based work	-0.164	0.082	4.016	1	0.045	-0.324	-0.004	0.85	0.72	1
Domestic and unskilled work	0.148	0.328	0.203	1	0.652	-0.495	0.791	1.16	0.61	2.2
No paid job*										

*Reference category

Stunting as an outcome variable, neither the logit nor the negative log log link satisfies test of parallel lines both for full and reduced model (Table 3, 4). A key problem with the parallel-lines model is that its assumptions are often violated; it is common for one or more *B*'s to differ across values of *j*; i.e., the parallel-lines model is overly restrictive. The common solution can be using multinomial models however the model will be less parsimonus estimating for more parameters than necessary. A special case of the gologit (gologit2) model overcomes these limitations. In the partial proportional odds model, some of the *B* coefficients can be the same for all values of *j*, while others can differ.[7] Firstly 'ologit' command is applied using STATA9 with stunting as dependent variable and wealth index, place of residence, sex of household head, size of the child at birth, education and occupation of mothers as independent variables. Likelihood ratio chi square is 422.67 with $p< 0.0001$ and pseudo $R^2 = 0.04$. 'Ologit' command gives results for simple ordinal model without any constraints. But carrying out brant test shows that the variables 'sex of the household head' and 'size of child at birth' do not satisfy parallel regression assumption. Then model with constraints for parallel lines imposed to variables 'place of residence', 'wealth index', 'occupation of mothers' and ' sex of the household head' is formed, however this model also come up with significant wald test($chi^2(8) = 17.01$, p=0.03). The same model has been tried for 0.01 level of significance and yet the parallel regression assumption is not met. So, finally those two variables, namely 'sex of the household head' and 'size of the child at birth' which are violating parallel assumption as per the brant test, are removed from the model. The final model with imposed constraints come up with insignificant wald test ($chi^2(6) = 11.22$, p=0.08) indicating that final model does not violate the parallel line assumption. The final model is highly significant (LR $chi^2 = 388.64$, $p<0.0001$) with pseudo $R^2= 0.04$.

Table 3. Constraints imposed to different independent variables.

Step 1: Constraints for parallel lines imposed for place of residence (P = 0.3573)
Step 2: Constraints for parallel lines imposed for wealth index (P = 0.2012)
Step 3: Constraints for parallel lines imposed for occupation of mothers (P = 0.0528)
Step 4: Constraints for parallel lines are not imposed for education of mothers (P = 0.01560)

Table 4. Parameter estimates for stunting.

Stunting	Coefficient	Std. Err.	z	P	95% CI	
Normal						
Wealth index	-0.3177	0.043114	-7.37	0	-0.4022	-0.2332
Place of residence	0.022627	0.073824	0.31	0.759	-0.12207	0.167319
Education of mothers	-0.39167	0.039603	-9.89	0	-0.46929	-0.31405
Occupation of mothers	-0.06068	0.033374	-1.82	0.069	-0.12609	0.004733
Constant	0.984278	0.091763	10.73	0	0.804425	1.164131
Mild						
Wealth index	-0.3177	0.043114	-7.37	0	-0.4022	-0.2332
Place of residence	0.022627	0.073824	0.31	0.759	-0.12207	0.167319
Education of mothers	-0.57756	0.068274	-8.46	0	-0.71137	-0.44374
Occupation of mothers	-0.06068	0.033374	-1.82	0.069	-0.12609	0.004733
Constant	-1.61463	0.094886	-17.02	0	-1.8006	-1.42866
Moderate						
Wealth index	-0.3177	0.043114	-7.37	0	-0.4022	-0.2332
Place of residence	0.022627	0.073824	0.31	0.759	-0.12207	0.167319
Education of mothers	-0.33932	0.246825	-1.37	0.169	-0.82309	0.144448
Occupation of mothers	-0.06068	0.033374	-1.82	0.069	-0.12609	0.004733
Constant	-4.69599	0.198861	-23.61	0	-5.08575	-4.30623

Increment in wealth index is associated with more likelihood of normal children than stunted children and relation is significant. Constraints are not imposed for education and it is found with negative and significant coefficients indicating less likelihood of stunted children among educated mothers.

DISCUSSION

Wealth index of household is found related with lesser prevalence of all forms of malnutrition. This notion is supported by different regional level comparisons which show more episodes of malnutrition in developing countries.[8] One study in Bangladesh concludes that household wealth inequality is strongly associated with childhood adverse growth rate stunting.[9] Another study carried out in Nepal shows that the risk of being underweight in the children from the poor socioeconomic status is almost four times as much as in the children from the rich socioeconomic status.[10] Area of residence is not found as significant contributor while explaining the malnutrition while all other variables are held constant in the model for all the three measures of malnutrition. However, in a study carried out in Indonesia, more percentage of rural children are malnourished against urban.[11] Female headed households are observed with less likelihood for underweight and wasting. A study from Botswana conducted by Gobotswang shows that female headed households are 1.5 times (p=.016) more likely to be underweight.[12] Small size at birth is observed as significant predictor of underweight and wasting. The association of size at birth with nutritional status is also supported by the study carried out in Malawi by Madise and Mpoma.[13] Mother's education is another variable significantly affecting for underweight but mixed results are seen for stunting and wasting. Educational effects are shown in studies carried out in Indonesia and in Malawi.[11,13] In case of occupation of mothers as a predictor, the results support other research findings that the mother's working status has both good and bad effect on children.[14,15,16,17,18] The effect depends upon the condition that whether these women are able to utilize their knowledge for choosing appropriate food for their children and if they are empowered enough to have control over their income.[19] In this analysis, for underweight and wasting, positive effects are seen for formally employed women who are believed to be more probable to exercise their rights and negative effects for agriculture based, domestic based and unskilled working women who may not have control over their earning, however the results are not similar in case of stunting.

Ordinal regression well suit the data to model nutritional status through different predictors in case of underweight and wasting however stunting model fails to satisfy the assumption behind ordinal regression. Hence for stunting model with constraints is applied.

CONCLUSION

Underweight is seen significantly less in households with high wealth index, among children of big and average size at birth and among educated women. Education, wealth index and size at birth are also contributing to wasting among children. For stunting, wealth index and education of mothers show significant effects. Hence for all three malnutrition measures have significant association with wealth index and education of mothers. So, the policies should be formed in such a way that it uses both direct and indirect ways to combat malnutrition, the direct being decreasing the poverty and indirect being educating the mothers.

REFERENCE

1. Adhikari RK, Krantz ME. Child Nutrition and Health. III ed. HLMC, Nepal. 2001.
2. World Health Organization. Nutrition in South East Asia. WHO Regional Office for South East Asia. New Delhi. 2000.
3. Ministry of Health and Population, New Era, ORC Macro. Nepal Demographic and Health Survey. 2001.
4. Ministry of Health and Population, New Era, ORC Macro. Nepal Demographic and Health Survey. 2006.
5. UCLA academic technology services. Annotated SPSS Output Ordered Logistic Regression. [Online]. [Cited 2010 Aug 9]; Available from: URL:www.ats.ucla.edu/stat/Spss/output/ologit.htm
6. Chau-Kuang C., John H. Using Ordinal Regression Model to Analyze Student Satisfaction Questionnaires. Association for Institutional Research 2004: (1).
7. Chau-Kuang Chen CK, Hughes Jr J, Using Ordinal Regression Model to Analyze Student Satisfaction Questionnaires. [Online]. 2004 May 26;1. Available from: URL:http://airweb.org/page.asp?page=554
8. William R. Generalized Ordered Logit/partial Proportional Odds Models for Ordinal Dependent Variables. The Stata Journal. 2006: 6(1):58–82.
9. World Bank. Repositioning Nutrition as Central to Development A Strategy for Large-Scale Action. 2006.
10. Hong R, Banta JE, Betancourt JA. Relationship between household wealth inequality and chronic childhood under-nutrition in Bangladesh. International Journal for Equity in Health 2006 [Cited 2010 Aug 9] 5:15. Available from: URL:http://www.equityhealthj.com/content/5/1/15
11. Sapkota VP, Gurung CK. Prevalence and Predictors of Underweight, Stunting and Wasting in Under-Five Children. J Nepal Health Res Counc. 2009 Oct;7(15):120-126.
12. Saadah F, Waters H, Heywood P. Indonesia: Under nutrition in Young Children. East Asia and the Pacific Region Watching Brief 1999;(1).
13. Gobotswang K. Determinants of the Nutritional Status of Children in a Rural African Setting: The Case of Chobe District, Botswana. [Cited 2009 Jun 12]. Available from: URL:www.Unu.edu/unupress/food/v191e/cho9.htm
14. Madise NJ, Mpoma M. Child Malnutrition and Feeding Practices in Malawi. [Cited 2009 Jun 12]. Available from: URL:www.Unu.edu/unupress/food/v182e/begin.htm
15. Popkin BM, Solon E. Income, time, the working mother and child nutriture. Environ Child Health. 1976:22(4):156-66.
16. Blau DM. Investment in child nutrition and women's allocation of time in developing countries. Discussion Paper no. 371. New Haven. Conn. USA: Economic Growth Center, Yale University, 1980.
17. Hart G. Women's participation in the labor force: implications for employment and health nutrition programs. Ithaca. NY, USA: Cornell University Press; 1975.
18. Popkin BM, Bisgrove EZ. Urbanization and nutrition in low-income countries. Food Nutr Bull. 1988;10(1):3-23.
19. Gopaldas T, Patel P, Bakshi M. Selected socioeconomic, environmental, maternal, and child factors associated with the nutritional status of infants and toddlers. Food Nutr Bull. 1988; 10(4):29-34.
20. Abbi R, Christian P, Gujral S, Gopaldas T. The Impact of Maternal Working Status on the Nutrition and Health Status of Children. Food Nutr Bull. 1991; 13(1):20-5.