Joint Regression and Association Modelling of Under-nutrition Outcomes among Children

Leonard K. Atuhaire*
School of Statistics and Planning, Makerere University, Kampala, Uganda
Email:latuhaire@bams.mak.ac.ug

Abstract

The presence of under-nutrition in children is usually assessed using three anthropometric parameters (weight-for-age, height-for-age and weight-for-height) and particularly by comparing them with internationally accepted reference standards. Undernourished children are then categorized as "stunted", "underweight", or "wasted". Studies to identify determinants of under-nutrition have to decide which of the three measures to use, and usually carry out separate analyses for the three indicators of under-nutrition, or choose only one of them. Such methods make the implicit assumption that the three indicators of under-nutrition are independent variables, which is quite implausible. Such a problem can be avoided if a multivariate regression approach is considered. Multivariate regression models consist of two parts: (i) a regression model for each marginal response, and (ii) an association structure to account for correlation between multiple response variables. In this paper we illustrate the application of such joint regression and association modelling for the three measures of under-nutrition using Uganda Demographic and Health Survey (UDHS) data. In short, we consider under-nutrition as a trivariate binary outcome and carry out a multivariate regression for binary responses.

Keywords: Nutritional anthropometry; General Estimating Equations; Dependence ratio.

1. Introduction

Height and weight are the most commonly used indicators of the nutritional status of a child. According to a WHO Working Group (1986), appropriate height-for-age of a child reflects linear growth and can measure long- term growth faltering or stunting, while appropriate weight-for-height reflects proper body proportion or the harmony of growth. Weight-for-height is particularly sensitive to acute growth disturbances and is useful to detect the presence of wasting. Weight-for-age represents a convenient synthesis of both linear growth and body proportion and thus can be used for the diagnosis of underweight children.

The presence of under-nutrition in children is assessed using these three anthropometric parameters (weight-for-age, height-for-age and weight-for-height) and by comparing them with internationally accepted reference standards, i.e. National Center for Health Statistics (NCHS) /WHO international reference population (WHO, 1983). If a child has a low height-for-age, i.e. a Z-score below two standard deviations of the reference population mean (-2 Z- score), such a child is categorized as "stunted ". Similarly, a low weight-for-age is diagnostic of an "underweight" child, while a low weight- for-height is indicative of "wasting ".

Studies to identify determinants of under-nutrition have to decide which of the three measures to use, and usually carry out separate analyses for the three indicators of under-nutrition, or choose only one of them (Madise et al., 1999; Kandala et al, 2009).

Such methods make the implicit assumption that the three indicators of under-nutrition are independent variables, which is quite implausible.

Such a problem can be avoided if a multivariate regression approach is considered. Multivariate regression models consist of two parts: (i) a regression model for each marginal response *i*,

$$h\{E(Yij)\} = \underline{x}_{ij}^{t}\,\underline{\beta} \quad (1)$$

and (ii) an association structure to account for correlation between multiple response variables.

In equation (1), h() is a suitable 'link' function.

In this paper we illustrate the application of such joint regression and association modeling for the three measures of under-nutrition using DHS data. We therefore consider under-nutrition as a trivariate binary outcome (Y_1, Y_2, Y_3) and carry out a multivariate regression for binary responses.

Because each of the three variables is binary, logistic regression is usually employed for this purpose (Bloss et al., 2004). We shall also consider models based on the logit link.

2. Methods

GEE Models

One approach is to use the generalized estimation equations (GEE) initially introduced for the analysis of longitudinal data (Liang and Zeger,1986). This approach allows one to carry out simultaneous inference on marginal response rates and pair-wise correlations (Prentice, 1988). Under this approach, stunting, underweight, and wasting are taken as 'repeated' measures of under-nutrition.

GEE models can be fitted using modules for panel data in standard statistical packages such as STATA (StataCorp, 2014).

Dependence Ratio

Another approach is to use the dependence ratio, a novel and interesting association structure introduced by Ekholm et al. (1995), Ekholm et al.(2000), and Ekholm et al. (2003).

These dependence ratios provide 'simple and sensible association models for third- and higher- order moments' (Jokinen, 2006).

The dependence ratio is defined as the joint success probability divided by the joint success probability assuming independence. For example, the second order dependence ratio is

$$\tau_{12} \qquad Pr(Y_1{=}1, Y_2{=}1)/\{pr(Y_1{=}1).Pr(Y_2{=}1)\} \quad (2)$$

In words τ_{12} measures how many times more probable it is to observe both (Y_1=1) and (Y_2=1) than would be expected if the two events were independent. This generalizes to dependency ratios of higher orders.

A package 'drm'(Jokinen,2009) in R (R Development Core Team, 2006) is available for fitting such models.

3. Results

Table 1 shows the results of fitting joint regression models using both the GEE and Dependency Ratio approaches to data on 2,004 children with complete anthropometric data in the Uganda DHS of 2011.

Table 1. Results for joint regression and association models

	GEE			Dependence Ratio		
	B	S.E.	Z value	β	S.E.	Z value
Intercept	-0.154	0.372	-0.41	1.960	0.495	3.96
Underweight	-1.136	0.067	-16.95	-1.712	0.115	-14.93
Wasted	-2.230	0.113	-19.73	-2.972	0.143	-20.83
Age	0.002	0.002	0.83	-0.008	0.004	-2.08
Sex(female)	-0.365	0.087	-4.20	-0.188	0.116	-1.62
Preceding birth interval:baseline 1st birth						
< 24 months	0.185	0.109	1.70	0.204	0.142	1.44
24+ months	-0.128	0.120	-1.07	0.015	0.152	0.10
Size at birth(small)	0.745	0.097	7.68	0.857	0.127	6.73
Mother's BMI	-0.048	0.014	-3.47	-0.082	0.018	-4.66
Residence(rural)	0.593	0.134	4.43	0.692	0.174	3.98
Mother's Education (Sec.+)	-0.402	0.124	-3.24	-0.506	0.159	-3.18
Had fever in last 2 weeks	-0.008	0.094	-0.08	0.137	0.126	1.09
Had cough in last 2 weeks	0.048	0.092	0.52	-0.006	0.119	-0.05
Had diarrhea recently	0.179	0.104	1.73	0.184	0.137	1.34

The coefficients for 'Underweight' and 'Wasted' reflect the different prevalence rates for these components of under-nutrition.

Apart from the effects of Age and Sex, which have curiously contradictory results, the two models are in broad agreement that the major determinants of under-nutrition are Size at birth, Mother's BMI, Residence, and Mother's education.

It is also evident that the GEE model has consistently lower standard errors.

The regression parameters have a population-averaged interpretation. This is not a serious issue if the covariates are confounders. If, however, interest is on the effect of covariates on individual components, this can be accommodated by forming interaction terms with the different components.

Table 2(a) shows the results of fitting separate regression for the different components of under-nutrition while Table 2(b) shows the results of a joint regression model with interaction terms that can be used to study the effects of the covariates on the individual components.

Table 2 (a) Results of independent logistic regression models

	Stunted			Underweight			Wasted		
	β	S.E.	Z value	β	S.E.	Z value	β	S.E.	Z value
Age	0.013	0.003	4.46	-0.007	0.004	-1.69	-0.045	0.007	-6.21
Sex(female)	-0.489	0.100	-4.88	-0.240	0.135	-1.77	0.092	0.207	0.44
Preceding birth interval:baseline 1st birth									
< 24 months	0.227	0.126	1.81	-0.287	0.165	1.74	-0.041	0.269	-0.15
24+ months	-0.236	0.138	-1.70	0.045	0.195	0.23	0.132	0.297	0.44
Size at birth(small)	0.629	0.116	5.45	1.037	0.141	7.34	0.888	0.214	4.15
Mother's BMI	-0.026	0.015	-1.69	-0.097	0.023	-4.14	-0.123	0.037	-3.32
Residence(rural)	0.734	0.152	4.82	0.669	0.228	2.93	-0.097	0.295	-0.33
Mother's Education (Sec.+)	-0.422	0.140	-3.01	-0.309	0.200	-1.54	-0.446	0.311	-1.43
Had fever in last 2 weeks	-0.166	0.109	-1.53	0.251	0.147	1.71	0.475	0.226	2.10
Had cough in last 2 weeks	0.164	0.106	1.55	0.078	0.144	0.54	-0.512	0.223	-2.30
Had diarrhea recently	0.186	0.121	1.54	0.250	0.156	1.60	0.066	0.234	0.28

Table 2(b) Results of joint regression model with interaction terms

	Stunted			Underweight			Wasted		
	β	S.E.	Z value	β	S.E.	Z value	β	S.E.	Z value
Age	0.013	0.003	4.40	-0.023	0.004	-5.44	-0.058	0.008	-7.54
Sex(female)	-0.499	0.100	-4.99	0.281	0.136	2.06	0.609	0.227	2.68
Preceding birth interval:baseline 1st birth									
< 24 months	0.226	0.126	1.80	0.033	0.168	0.19	-0.258	0.293	-0.88
24+ months	-0.239	0.138	-1.73	0.300	0.195	1.53	0.355	0.323	1.10
Size at birth(small)	0.626	0.115	5.42	0.407	0.147	2.77	0.251	0.241	1.04
Mother's BMI	-0.025	0.015	-1.62	-0.079	0.023	-3.41	-0.103	0.040	-2.59
Residence(rural)	0.749	0.152	4.92	0.001	0.229	0.01	-0.842	0.328	-2.57
Mother's Education (Sec.+)	-0.436	0.140	-3.12	0.123	0.200	0.62	0.010	0.335	0.03
Had fever in last 2 weeks	-0.166	0.109	-1.52	0.419	0.148	2.83	0.650	0.248	2.62
Had cough in last 2 weeks	0.164	0.106	1.55	-0.115	0.145	-0.79	-0.680	0.244	-2.79
Had diarrhea recently	0.193	0.121	1.60	0.061	0.159	0.39	-0.110	0.260	-0.42

In Table 2(b) the results for 'Underweight' and 'Wasted' are contrasts with stunting. Other types of contrasts are possible.

4. Conclusions

We have demonstrated that it is fairly straightforward to fit a joint regression model for under_nutrition outcomes, and that such a model can be used to identify factors specific to each outcome.

References

Bloss, E., Wainaina, F., and Bailey, R. C. (2004) Prevalence and predictors of underweight, stunting, and wasting among children age 5 and under in western Kenya. *J. Trop. Pediatr., 50(5)*, 260-270.

Ekholm, A., Smith, P.W.F., McDonald, J.W. (1995). Marginal regression analysis of a multivariate binary response. *Biometrika, 82*, 847–854.

Ekholm, A., McDonald, J.W., Smith, P.W.F. (2000). Association models for a multivariate binary response. *Biometrics 56*, 712–718.

Ekholm, A., Jokinen, J., McDonald, J. W., amd Smith, P. W. F. (2003) Joint Regression and Association Modelling of Longitudinal Ordinal Data. *Biometrics, 59*, 795-803

Jelliffe, D.B. (1966). *The assessment of the nutritional status of the community.* WHO Monograph No. 53. Geneva.

Jokinen, J. (2006) Fast estination algorithm for likelihood-based analysis of repeated categorical responses. *Computational Statistics and Data Analysis, 51*, 1509-1522.

Jokinen, J. (2009) *drm:regression and association models for repeated categorical data*. http://www.helsinki.fi/~jtjokine/drm

Kandala. N., Fahrmeir, L., Klasen, S., Priebe, J. (2009) Geo-additive models of childhood undernutrition in three sub-Saharan African . *Population, Space and Place, 15(5)*, 461-473

Liang, K. Y. & Zeger, S. L. (1986) Longitudinal data analysis using generalized linear models. *Biometrika, 73*, 13-22

Madise,N. J., Mathews, Z., and Margretts, B. (1999) Heterogeneity of Child Nutritional Status between Households: A Comparison of six Sub-Saharan Countries. *Population Studies, 53*, 331-343.

Prentice, R.L. (1988), Correlated Binary Regression with Covariate Specific to Each Binary Observation. *Biometrics, 44*, 1033-1048.

R Development Core Team (2006) *R:A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

StataCorp (2014) *Stata Statistical Software: Release 13*. College Station, TX: Stata Corporation

WHO. (1983). *Measuring change in nutritional status. Guidelines for assessing the nutritional impact of supplementary feeding programmes for vulnerable groups.* Geneva.

WHO. (1995). *Physical status: the use and interpretation of anthropometry.* WHO Technical Report No. 854. Geneva.

WHO Working Group. (1986). Use and interpretation of anthropometric indicators of nutritional status. *Bull. World Heath Org.*, 4: 924 -941.