PalArch's Journal of Archaeology of Egypt / Egyptology

SOCIOECONOMIC, DEMOGRAPHIC, AND ENVIRONMENTAL ASSOCIATED WITH STUNTING AMONG CHILDREN UNDER FIVE YEARS OLD IN ETHIOPIA: ETHIOPIAN DEMOGRAPHIC AND HEALTH SURVEY, 2016

Getnet Bogale Begashaw [1], Wudneh Ketema Moges [2]*

[1]Wollo University, Dessie, Ethiopia
[2]Debre Berhan University, Debre Berhan, Ethiopia

Getnet Bogale Begashaw [1*], Wudneh Ketema Moges [2]socioeconomic, demographic, and environmental associated with stunting among children under five years old in ethiopia: ethiopian demographic and health survey, 2016 -- Palarch's Journal Of Archaeology Of Egypt/Egyptology 17(9). ISSN 1567-214x

Abstract

Background: Stunting is one of the most serious but least addressed health problems in the world. Adequate nutrition is essential for children's health and development. Globally it is estimated that, directly or indirectly, for at least 35% of deaths in children less than five years of age. Under nutrition is also a major cause of disability preventing children who survive from reaching their full development potential.

Methods: Statistical models that can treat the categorical response variable like binary logistic regression model will be employed. Beside this study will include Socio –economic and demographic factors; Sex and age of child, age of mother, Educational status, occupation, health status, religion, sex of household head, number of children under five years, Household income, family size, land ownership and time of cultivation, income source of household, wealth index as independent variables. Empty model, random intercept and fixed slope with random coefficient are the method of analyzing the dataset.

Result: The prevalence of stunting among children ages under five years old were about 49.3%. Months of breastfeeding, educational level, and wealth index, currently pregnant and child food nutrient are significantly associated with stunting presence. The odds of stunting status of child from women who are pregnant is more likely to be stunted 4.157 compared to non-pregnant women controlling for other variables in the model and random effects at level two. Women who feed nutrient food to their child are 1.239 more likely to be stunted (OR=1.239) than women who didn't feed nutrient food controlling for other variables in the model and random effects at level two.

Conclusions: Age of child, breast feeding, sex, pregnant status, and food nutrient were found to be significantly associated with stunting in multilevel modeling of random coefficient model. Finally random coefficient model best fit the EDHS 2016 dataset. Therefore, interventions that focus on breast feeding, period of next pregnancy, food nutrient taken by children are required for improving child stunting in Ethiopia.

Keywords and Phrases: Multi-Level Logistic Regression Model; Malnutrition; Stunting; Underweight; Wasting

1. Background

Globally, an estimated 171 million children are stunted, including 167 million children in low- and middle-income countries (Black et al. 2013). Globally, the percentage of children under age 5 who are stunted has decreased, from 40% in 1990 to 28% in 2010, with an anticipated further drop to 22% (142 million) by 2020 [1]. In Africa, however, prevalence of stunting among children under age 5 was 36% compared with 27% in Asia, estimated in 2011. It is projected that by 2020, Asia and Africa will have almost similar numbers of stunted children (68 million and 60 million, respectively). These levels are much higher than the number of children stunted in Latin America, at 7 million in 2010 [2].

The higher prevalence of child stunting in Africa and Asia is a public health problem that has often gone unrecognized. Child stunting reflects a failure to receive adequate nutrition over a long period of time and may be affected by intrauterine growth retardation, poor feeding practices, and frequent exposure to infections [3]. When stunting spans generations, it results in grave consequences that include poor quality of life, morbidity, and mortality [4,5]. The 2014 Demographic and Health Surveys (DHS surveys) for Kenya and Cambodia showed that the prevalence of stunting among children under age 5 was 22% and 25% respectively. The prevalence of stunting in children under age 5 in Kenya and Cambodia was higher, at 32% and 26% respectively [6]. Researchers have found that poverty, poor health and nutrition, and social factors are associated with risks to child growth. These factors have prevented over 200 million children in developing countries from attaining their full potential [7]. In developing countries, where mostly women are denied a voice in household decisions, they are most likely to be undernourished themselves and less likely to have access to resources that can be directed toward children’s nutrition [8].

In Ethiopia 40% of children under age five were stunted and 18% of children were severely stunted with regional variation such as in South Nation Nationality Peoples 44.3%, Afar 49.2%, Tigray 44.4%, Amhara National Region State 42.4% children under five were stunted [9]. Stunting is affected by many factors such as: poverty, low parental education, lack of sanitation, low food intake, poor feeding practices, inadequate breastfeeding, repeated infections, family size and birth interval [5].

Stunting remains one of the most common causes of morbidity and mortality among children throughout the world. It has been responsible, directly or indirectly, for 60% of the 10.9 million deaths annually among children under five. Over two-thirds of these deaths, which are often associated with inappropriate feeding practices, occur during the first year of life. Malnutrition is one of the leading causes of morbidity and mortality in children under the age of five in developing countries. Ethiopia being one of these countries malnutrition is an important public health problem. There is no information available on the stated problem. This study is, therefore, aimed at assessing associated factors of stunting children under five years old.

The general objective of this study is to empirically investigate the major factors that are associated with stunting among children below five years old in Ethiopia. The specific objectives of the study is to determine the prevalence of stunting among the children aged below five years, to determine the socio-demographic and economic characteristics of households of children aged below five years and to estimate the within-regional and between-regional level of difference for the incidence of stunting among under five-children in Ethiopia.

2. Methods

2.1 Study Population

This study analyzed the secondary data from the Ethiopian Demographic Health Survey (EDHS), 2016, accessed from the Measure Evaluation Demography, Health Survey 2016 Ethiopia [10] which is freely available online [11] and contains information on a wide range of socioeconomic and demographic factors of the population nationwide. The country has nine regions and two administrative cities. The Ethiopian DHS 2016 utilized a two-stage sample design to select respondents for the study. In the first stage 645 enumeration areas (202 in urban areas and 443 in rural areas) were selected with probability proportional to size. Second stage involved selection of 28 households per cluster with an equal probability systematic selection from the newly formed household list. The EDHS 2016 has three parts: the household questionnaire, the woman's questionnaire, and the man's questionnaire. The data for child mortality and associated factors were taken from a woman's questionnaire. Data were collected by conducting face-to-face interviews with women who met the eligibility criteria (women aged 15–49 years).

Dependent variable: Status of stunting under five years old

Often in many public health studies, binary outcome is preferred as a response of interest for the sake of interpretation. Hence, our two responses were also studied as binary responses:
Stunted verses not stunted.

Y_{1i}= 1 if stunted (Z − score <−2)
0 if not stunted (Z − score >=-2) i.e. normal height for age

Independent variables: Months of breast feeding, sex of child, place of residence, education level, toilet facility, currently pregnant and child food nutrient. The description and coding of the independent variables are listed in **Table 1** below.

Table 1: Variables in the Study

No.	Variable Description	Code (If any)
1.	Breast feeding status	**0**=never breastfed; **1**=inconsistent
2.	Sex of child	**0**=female; **1**=male
3.	Age of child	**0-**>29 months; **1-**<30 months

4.	Residence of child	**0**=urban; **1**=rural
5.	Level of education of Mother	**0=**no education; **1=**primary; **2=**secondary **3=**higher
6.	Use of toilet	**0**=unsafe; **1**=safe
7.	Pregnant status	**0**=no; **1**=yes
8.	Food nutrient status	**0**=no; **1**=yes
9.	Region	Addis Ababa = 0(ref), Tigray = 1, Afar = 2, Amhara = 3, Oromya = 4, Somali = 5, Benishangul-Gumuz = 6, SNNP = 7, Gambella = 8, Harari = 9, Dire Dawa = 10

2.2 Multilevel Logistic Regression Model

Two-Level Model

Multilevel models are statistical models which allow not only independent variable at any level of hierarchical structure but also at least one random effect above level one group [12]. A multilevel logistic regression model can account for lack of independence across levels of nested data (i.e., individuals nested within regions). Conventional logistic regression assumes that all experimental units are independent in the sense that any variable which affects occurrence of stunting has the same effect in all regions, but multilevel models are used to assess whether the effect of predictors vary from region to region.

In this study the basic data structure of the two-level logistic regression is a collection of N groups (regions) and within-group $j(j = 1,2,\dots,N)$, a random sample n_j of level-one units (children). The response variables, i.e., we let $Y_{ij} = 1$ if the i^{th} under five children in j^{th} region has stunting , and $Y_{ij} = 0$ otherwise; with probabilities, $P_{ij} = P(y_{ij} = 1|X_{ij}, u_j)$, is the probability of having stunting for child$i$ in region$j$ and $1 - P_{ij} = P(y_{ij} = 0|X_{ij}, u_j)$ is the probability of having no stunting for child$i$ in region$j$; where u_j is a random cluster effect and often assumed to be $N(0, \sigma_u^2)$. The standard assumption is that Y_{ij} has a Bernoulli distribution. Let P_{ij} be modeled using a logit link function. The two-level model is given by:

$$\text{logit}(p_{ij}) = \log\left(\frac{p_{ij}}{1 - p_{ij}}\right) = \beta_{oj} + \sum_{l=1}^{k} \beta_{lj} x_{lij}\,; \quad l = 1,2,\dots,k \qquad 1$$

Where $\beta_{0j} = \beta_0 + U_{0j}, \beta_{1j} = \beta_1 + U_{1j}, \dots, \beta_{kj} = \beta_k + U_{kj}$

The level-two model (1) can be rewritten as:

$$\text{logit}(p_{ij}) = \log\left(\frac{p_{ij}}{1 - p_{ij}}\right) = \beta_o + \sum_{l=1}^{k} \beta_l x_{lij} + U_{oj} + \sum_{l=1}^{k} U_{lj} x_{lij} \qquad 2$$

where $X_{ij} = (X_{1ij}, X_{2ij}, \dots, X_{kij})$ represent the covariates, $\beta = (\beta_0, \beta_1, \dots, \beta_k)$ are regression coefficients, $U_{0j}, U_{1j}, \dots, U_{kj}$are the random effects of model parameter at level two. It is assumed that the $U_{0j}, U_{1j}, \dots, U_{kj}$ follow a normal distribution with mean zero and variance δ_u^2. Without $U_{0j}, U_{1j}, \dots, U_{kj}$, equation (2) can be considered as a single level logistic regression model. Therefore, conditional on $U_{0j}, U_{1j}, \dots, U_{kj}$, the y_{ij} can be assumed to be independently distributed as Bernoulli random variables [14].

Estimations of Between and Within Group Variance

The true variance between the group dependent probabilities, i.e. the population values of $Var(\mathrm{P}_j)$, is given by:

$$\hat{\tau}^2 = S^2{}_{between} - \frac{S^2{}_{within}}{\tilde{n}} \quad 7$$

where $\tilde{n}$ is defined as: $\tilde{n} = \frac{1}{N-1}\left\{M - \frac{\sum_{j=1}^{N} n_j{}^2}{M}\right\}$

For dichotomous outcome variables, the observed between group variance is closely related to the chi-square test statistic given in equation 5.

$$S^2{}_{between} = \frac{\hat{p}(1-\hat{p})}{\tilde{n}\,(N-1)} X^2$$ Where X^2 is given in equations (5).

The within group variance in case of a dichotomous outcome variable is a function of group averages which is given by:

$$S^2{}_{within} = \frac{1}{M-N}\sum_{j=1}^{N} n_j\, p_j(1-p_j)$$

Multilevel logistic regression can be employed in the simplest case without explanatory variables (usually called empty model) and also with explanatory variables by allowing only the intercept term or both the intercept and the slopes (regression coefficients) to vary randomly. It mainly assumed that the varying coefficients have multivariate normal distribution [14].

2.2.1 The Empty Multilevel Logistic Regression Model

The empty two-level model for a dichotomous outcome variable refers to a population of groups (level-two units) and specifies the probability distribution for group-dependent probabilities $\mathrm{p_j}$ in $\mathrm{Y_{ij}} = \mathrm{p_j} + \varepsilon_{ij}$ without taking further explanatory variables into account. We focus on the model that specifies the transformed probabilities $f(\mathrm{p}_j)$ to have a normal distribution. This is expressed, for a general link function *f* (p), by the formula

$$f(\mathrm{p_j}) = \beta_o + U_{oj} \qquad 8$$

where β_o is the population average of the transformed probabilities and U_{oj} it is the random deviation from this average for group *j*. If *f* (*p*) is the logit function, then $f(\mathrm{p}_j)$ isjust the log-odds for group *j*. Thus, for the logit link function, the log-odds have a normal distribution in the population of groups, which is expressed by:

$$logit(\mathrm{p_j}) = \beta_o + U_{oj} \qquad 9$$

For the deviations U_{oj}it is assumed that they are independent random variables with a normal distribution with mean zero and varianceσ_0^2.This model does not include a separate parameter for the level-one variance [14]. This is because the level-one residual variance of the dichotomous outcome variable follows directly from the success probability which is given by: $Var(\varepsilon_i) = P_j(1 - P_j)$

Denote by $\pi_0$the probability corresponding to the average value β_0, as defined by

$$f(\pi_0) = \beta_0$$

For the logit function, the so-called logistic transformation of β_0, is defined by

$$\pi_0 = logistic(\beta_0) = \frac{\exp(\beta_0)}{1 + \exp(\beta_0)} \qquad 10$$

Note that due to the non-linear nature of the logit link function, there is no a simple relation between the variance of probabilities and the variance of the deviationsU_{0j} [14]. An approximate variance of the probability given by:

$$var(P_j) \approx (\pi_0(1 - \pi_0))^2 \sigma_0^2 \qquad 11$$

Note that an estimate of population variance $var(P_j)$ can be obtained by replacing sample estimates of π_0 and σ_0^2. The resulting approximation can be compared with the nonparametric estimate, $\hat{\tau}^2$ which was given in equation (7).

2.2.2 The Random Intercept Model

In the random intercept model, the intercept is the only random effect meaning that the groups differ with respect to the average value of the response variable, but the relation between explanatory and response variables cannot differ between groups. We assume that there are variables which potentially explain the observed success and failure. These variables are denoted by $X_h, (h = 1,2, \dots, k)$ with their values indicated by X_{hij}. Since some or all of those variables could be level one variables, the success probability is not necessarily the same for all individual in a given group [14]. Therefore, the success probability depends on the individual as well as the group, and is denoted by P_{ij}. the outcome variable is split into an expected value and residual as: $Y_{ij} = P_{ij} + R_{ij}$

The random intercept model expresses the log-odds, i.e. the logit of P_{ij}, as a sum of a linear function of the explanatory variables. That is,

$$\mathrm{logit}(P_{ij}) = \log\left(\frac{p_{ij}}{1 - p_{ij}}\right) = \beta_{oj} + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \cdots + \beta_k x_{kij}$$

$$= \beta_{oj} + \sum_{h=1}^{k} \beta_h x_{hij} \qquad 12$$

Where the intercept term $\beta_{oj}$is assumed to vary randomly and is given by the sum of an average intercept $\beta_o$and group-dependent deviations U_{oj}, that is $\beta_{oj} = \beta_o + U_{oj}$

As a result we have: $\mathrm{logit}(P_{ij}) = \beta_o + \sum_{h=1}^{k} \beta_h x_{hij} + U_{oj}$ 13

Solving for P_{ij} we have: $$P_{ij} = \frac{e^{\beta_o+\sum_{h=1}^{k}\beta_h x_{hij}+U_{oj}}}{1+e^{\beta_o+\sum_{h=1}^{k}\beta_h x_{hij}+U_{oj}}} \quad 14$$

Thus, a unit difference between the X_h values of two individuals in the same group is associated with a difference of β_h in their log-odds, or equivalently, a ratio of $\exp(\beta_h)$ in their odds. Equation (12) does not include a level-one residual because it is an equation for the probability P_{ij}rather than for the outcome Y_{ij}. Note that in the above equation $\beta_o + \sum_{h=1}^{k}\beta_h x_{hij}$ is the fixed part of the model. The remainingU_{oj} is called the random part of the model. It is assumed that the residual U_{oj} are mutually independent and normally distributed with mean zero and varianceσ_o^2.

2.2.3 The Random Coefficient Multilevel Logistic Regression Model

In logistic regression analysis, linear models are constructed for the log-odds. The multilevel analogue, random coefficient logistic regression, is based on linear models for the log-odds that include random effects for the groups or other higher level units. Consider explanatory variables which are potential explanations for the observed outcomes. Denote these variables by $X_1, X_2, \dots, X_k$. The values of $X_h (h = 1, 2, \dots, k)$ are indicated in the usual way by X_{hij}. Since some or all of these variables could be level-one variables, the success probability is not necessarily the same for all individuals in a given group. Therefore, the success probability depends on the individual as well the group, and is denoted byP_{ij}.
Now consider a model with group-specific regressions of logit of the success probability, $\text{logit}(P_{ij})$, on a single level one explanatory variable X,

$$\text{logit}(P_{ij}) = \log\left(\frac{p_{ij}}{1-p_{ij}}\right) = \beta_{oj} + \beta_{1j}x_{1ij} \quad 15$$

The intercepts β_{oj}as well as the regression coefficients or slopes, $\beta_{1j}$are groupdependent. These groupdependent coefficients can be split into an average coefficient and the group dependent deviation:

$$\beta_{oj} = \beta_o + U_{oj}$$
$$\beta_{1j} = \beta_1 + U_{1j}$$

Substitution into (15) leads to the model

$$\text{logit}(P_{ij}) = \log\left(\frac{p_{ij}}{1-p_{ij}}\right) = (\beta_o + U_{oj}) + (\beta_1 + U_{1j})x_{1ij}$$
$$= \beta_o + \beta_1 x_{1ij} + U_{oj} + U_{1j}x_{1ij} \quad 16$$

There are two random group effects, the random intercept U_{oj} and the random slope U_{1j}. It is assumed that the level two residuals $U_{oj}$and U_{1j}have both zero mean given the value of the explanatory variableX. Thus, $\beta_1$is the average regression coefficient like β_o is the average intercept. The first part of equation 16 $\beta_o + \beta_1 x_{1ij}$ is called the fixed part of the model whereas the second part $U_{oj} + U_{1j}x_{1ij}$ is called the random part of the model.

The term $U_{oj} + U_{1j}x_{1ij}$ can be regarded as a random interaction between group and predictors (X). This model implies that the groups are characterized by two random effects: their intercept and their slope. These two groups effects $U_{oj}$and U_{1j} will not be independent. Further, it is assumed that, for different groups, the pairs of random effects (U_{oj}, U_{1j}) are independent and identically distributed. Thus, the variances and covariance of the level-two random effects (U_{oj}, U_{1j}) are denoted by:

$$Var(U_{oj}) = \sigma_{00} = {\sigma_0}^2$$
$$Var(U_{1j}) = \sigma_{11} = {\sigma_1}^2$$
$$Cov(U_{oj}, U_{1j}) = \sigma_{01}$$

The model for a single explanatory variable discussed above can be extended by including more variables that have random effects. Suppose that there are klevel-one explanatory variables $X_1, X_2, \dots, X_k$, and consider the model where all predictor variables have varying slopes and random intercept. That is

$$\mathrm{logit}(P_{ij}) = \log\left(\frac{p_{ij}}{1-p_{ij}}\right) = \beta_{oj} + \beta_{1j}x_{1ij} + \beta_{2j}x_{2ij} + \dots + \beta_{kj}x_{kij} \quad 17$$

Letting $\beta_{oj} = \beta_o + U_{oj}$ and $\beta_{hj} = \beta_h + U_{hj}$ where $h = 1, 2, \dots, k$, we have:

$$\mathrm{logit}(P_{ij}) = \log\left(\frac{p_{ij}}{1-p_{ij}}\right) = \beta_o + \sum_{h=1}^{k} \beta_h x_{ij} + U_{oj} + \sum_{h=1}^{k} U_{hj} x_{ij} \quad 18$$

The first part$\beta_o + \sum_{h=1}^{k} \beta_h x_{ij}$is called the fixed part of the model, and the second part, $U_{oj} + \sum_{h=1}^{k} U_{hj} x_{ij}$ is called the random part of the model.The random variables or effects,$U_{0j}, U_{1j}, \dots, U_{kj}$are assumed to be independent between groups but may be correlated within groups. So the components of the vector $(U_{0j}, U_{1j}, \dots, U_{kj})$ are independently distributed as a multivariate normal distribution with zero mean vector and variances and co-variances matrix Ω given by:

$$\Omega = \begin{pmatrix} \sigma_0^2 & . & \cdots & . \\ \sigma_{01} & \sigma_1^2 & \cdots & . \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{0k} & \sigma_{1k} & \cdots & \sigma_k^2 \end{pmatrix}$$

2.2.4 Intra-class Correlation Coefficient (ICC)

The other fundamental reason for applying multilevel analysis is the existence of intra-class (intra-regional) correlation arising from similarity of incidence of stunting in the same region compared to those of different regions. The intra-class correlation coefficient (ICC) measures the proportion of variance in the outcome explained by the grouping structure. ICC can be calculated using an intercept-only model. This model can be derived from “Eq. (19)” by excluding all explanatory variables, which results in the following equation: (logit (p_j) = $\beta_{o+}U_{oj}$).The ICC is then calculated based on the following formula:

$$ICC = \frac{\delta_{uo}^2}{\delta_{uo}^2 + \delta_e^2} \quad 19$$

whereδ_e^2 variance of individual (lower) level units

In multilevel logit model level one residual variance δ_e^2 =π2/3 ≈ 18 [14] this formula can be reformulated as:

$$ICC = \frac{\delta_{uo}^2}{\delta_{uo}^2 + 3.29} \quad 20$$

For the purpose of model comparison study attempts the concept of maximum likelihood estimation via quadrature, AIC and BIC.

3. Results

3.1 Descriptive Statistics

This research utilized the national wide Ethiopia Demographic and Health Survey (EDHS) 2016 collected data on the stunting of children. The analysis presented in the study is based on 11654 under-five children with complete weight-for-age anthropometric index as indicator of a children's stunting and health status among other indices, since it is an excellent overall indicator of a population's stunting and health status. Table 2: below, shows that the percentage of the severity status of child's stunting

Table 2: Descriptive Statistics of Variables

Characteristics	Category	Not Stunted		Stunted		Total	
		Count	%	Count	%	Count	%
Sex of child	Female	2973	51.26	2827	48.74	5800	49.77
	Male	2949	50.38	2905	49.62	5854	50.23
Residence	Urban	2933	50.18	2912	49.82	5845	50.15
	Rural	2989	51.45	2820	48.54	5809	49.85
Educational level	No education	4324	53.10	3818	46.89	8142	69.86
	Primary	1416	48.33	1514	51.67	2930	25.14
	Secondary	134	34.72	252	65.28	386	3.31
	Higher	48	24.49	148	75.51	196	1.68
Toilet	Not Safe	5610	51.14	5358	48.85	10968	5.87
	Safe	310	45.32	374	54.68	684	5.87
Currently pregnant	No	5509	53.22	4842	46.78	10351	88.82
	Yes	413	31.70	890	68.30	1303	11.18
Child age in months	<29 months	648	32.63	1338	67.37	1986	17.04
	> 30 months	5922	61.32	4394	45.46	9658	82.87
Months of breast feeding	Ever breastfed, not currently breastfed	754	62.72	448	37.27	1202	10.31
	Never breastfed	591	69.45	260	30.55	851	7.30
	Inconsistent	288	43.70	371	56.30	659	5.65
Food nutrient status	No	5711	50.72	5548	49.28	11259	96.61
	Yes	205	66.56	175	46.05	380	3.26

As presented in **Table 2**, the prevalence of stunting found was at 48.74% were female, where males were 49.62%. It shows that 49.82% of urban children were stunted, 48.54% rural of children were stunted and 46.89% of children were stunted their family no education.51.67% of children were stunted their family were primary education, 65.28% of children were stunted their family education were secondary and 75.51% of children were stunted where their family were higher education.

As presented in **Table 2**, the prevalence of stunting found was 48.85% of children were stunted where their toilet was unsafe and 54.68% of children were stunted where their toilet was safe.40.17% of children were stunted where poorest and 69.90% of children were stunted where richest.

Table 2 shows 46.78% of children were stunted whether currently pregnant or not 68.30% of children were stunted where they were currently pregnant. And 52% of children were stunted where they were not used a soup. 47.26% of children were stunted that they used a soup and 49.28% of children were stunted that they were not found nutrients and 46.05% of children were stunted that they were found nutrients.

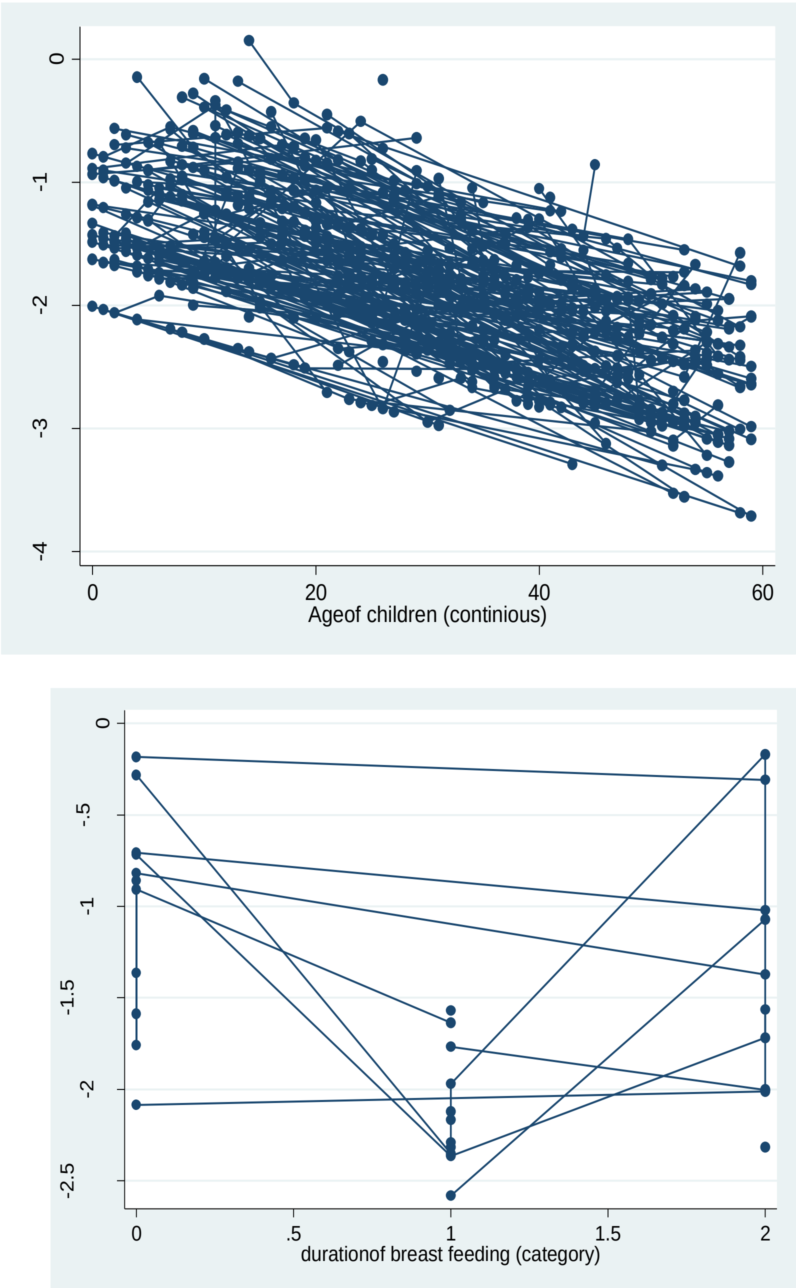

Figure 1: Predicted Probability of child stunting by child age and duration of breastfed vs Region

Figure1: shows that the Predicted Probability of under five children stunting by predictors vs Region. The Maximum predicted log-odds range is considered as regionally varied variables thus variable duration of breast feeding is regionally varied variables and have high random effects on under-five children stunting compare to the other variables. This variable is used in the random slope model.

3.2 Results of Multilevel Logistic Regression Analysis

In the multilevel analysis, a two-level structure is used with regions as the second-level units and under five children as the first level units. This is basically the analysis of region wise variation of stunting among under-five children. Children were nested in regions with a total of 5732 children included in this study.

3.2.1 Multilevel Logistic Regression Model Comparison

The Maximum predicted log-odds range is considered as regionally varied variables thus variable age and stunting are regionally varied variables and have high random effects on under-five children stunting morbidity compare to the other variables. These variables are used in the random slope model.

Table 3: Multilevel Logistic Regression Model for Stunting and their Deviance Based Chi-square Test Statistics.

	Empty model	**Random intercept model**	**Random coefficient model**
-2*log likelihood	7254.4834	6802.2541	6802.544
Deviance based chi-square test	84.1252	245.213	2.1569
P-value	0.0000*	0.0000*	0.6099
Model Fit Diagnostics			
AIC	7295.012	6950.034	6972.0152
BIC	7312.187	7015.312	7096.182

*significant at 5% level

The deviance-based chi-square value for the empty model shown in the above Table 3: is the difference in log likelihoods between an empty model of single level logistic regression and empty model of multilevel logistic regression, which is to be compared with the critical value from the chi-squared distribution with 1 degree of freedom. The significance of this test implies that an empty model with random intercept is better than an empty model without random intercept. The significant deviance-based chi-square value and smallest AIC for random intercept model indicates that the random intercept and fixed slope model is a better fit as compared to the empty model. The deviance-based chi-square test of random effects for random coefficient model is not statistically significant and has larger AIC. This implies that as compared to the model with random intercept and fixed slope model, the random coefficients model is not a better fit. Thus, in the above Table 3: shows that among multilevel logistic regression models, the random intercept and fixed slope model fits significantly better than the other multilevel logistic regression models.

3.2.2 Results of Empty Multilevel Logistic Regression Model

The variance of the random factor is significant which indicates that there is regional variation in experiencing stunting among under-five children (Table 4).The intercept β_O = -1.18546 is interpreted as the odds of stunting in an average region. That is the intercept informs us that the average probability of incidence of stunting everywhere in Ethiopia is exp (-1.18546) / [1+exp (-1.18546)] = 0.234. The intra-region correlation in intercept only model is 0.049 which is significant at 5% level of significance. This result implied that 4.9% of the variation in the incidence of stunting can be explained by grouping the under-five children in regions (higher level units). The remaining (100-4.9%=95.1% of the variation of incidence of stunting is explained within region-lower level units.

Table 4: Results for Multilevel Logistic Regression Model without Explanatory Variables

Fixed part	**Coefficients**	**S.E.**	**t-value**	**P-value**
$\boldsymbol{\beta_{oj} - intercept}$	-1.18546	0.019	-62.4	0.000**
Random part	**Estimate**	**S.E.**	**Z-value**	**P-value**
Randomintercept: $\boldsymbol{\sigma_0}^2 = \boldsymbol{var(U_{oj})}$	0.140688	0.0235	5.987	0.011*
Rho (ρ)	0.048546	0.0204	2.38	0.025*
Deviance= 15,243.52, AIC = 15,247.52, BIC = 15,248.31, Deviance-based Chi-square = 521.19				

**significant at 1% level, *significant at 5% level

The variance of the regional level residuals errors, symbolized by $\boldsymbol{\sigma_0}^2$ is estimated to be 0. 140688. This parameter estimate is larger than the corresponding standard errors and calculation of the Z-test shows that it is significant at p<0. 025. The significance of this residual term indicates that there are regional differences in the women unemployment status in Ethiopia. The deviance-based Chi-square (deviance = 521.19) indicated in table below is the difference in deviance between an empty model without random effect (deviance = 15,764.71) and an empty model with random effect (deviance =15,243.52). This value is compared to chi-square distribution with 1 degree of freedom. The significant of it (X^2 = 521.19, P-value < 0.0001) implies that an empty model with random intercept is better than an empty model without random intercept. The deviance reported in the above Table is a measure of model misfit; when we add explanatory variables to the model, the deviance is expected to go down.

3.2.3 Results of Random Intercept and Fixed Slope Logistic Regression Model

The random intercept and fixed slope logistic regression model is a multilevel model which has random intercept and fixed coefficient of predictors. As can be seen from Table 5: the analysis of multilevel logistic regression revealed that incidence of stunting in under-five children varied among regions. The value of $Var(U_{0j})$ is the estimated variance of the intercept in random intercept and fixed coefficients model. The result displayed that the region-wise difference in the incidence of childhood stunting was statistically significant. In addition, age of child, maternal working status, duration of breastfeeding, stunting, wasting, and underweight were also found to be significant determinants of incidence of stunting among the regions.

Table 5: Results of Random Intercept and Fixed Coefficient Logistic Regression Model

Fixed part				
Fixed effect	$\hat{\beta}$	**S.E.**	**Z-Value**	**p-value**
Breast feeding status (Never breasted=ref.cat)				
Ever breasted, not currently	-.6010689	.2131114	-2.82	**0.005***
Still breast feeding	-.4555433	.2165306	-2.10	**0.035***
Sex of child (female =ref.cat)				
Male	0.7458415	0.14482	5.15	0.002*
Age of child (less than or equals to 29 months =ref.cat)				
Greater than 29 months	0.2452757	0.0639085	3.84	0.000*
Residence of child (urban =ref.cat)				
Rural	0.2021277	0.0778864	2.60	0.009*
Level of education of Mother (no education =ref.cat)				
primary	0.4378549	0.080129	5.46	0.000*
secondary	0.5244647	0.045434	11.54	0.004*
Higher	0.6000154	0.064524	9.30	0.001*
Use of toilet (unsafe =ref.cat)				
Safe	0.400427	0.0910844	4.40	0.000*
Pregnant status (no =ref.cat)				
Yes	0.345115	0.0542164	6.37	0.005*
Food nutrient status (no =ref.cat)				
Yes	0.822101	0.061005	13.48	0.003*
Constant	-.8856897	.2500704	-3.54	0.000*
Random part	**Estimate**	**S.E.**	**Z-value**	**P-value**
Random intercept: $\sigma_0^2 = var(U_{oj})$	0.13432	0.06651	2.02	**0.0217***
Intra-region correlation (rho)	.0392259	.0178844	2.1933	**0.0141***
Deviance based chi-square		928.61		**0.000***
Deviance =14,314.91, AIC = 14,751.85, BIC = 14,762.01				

*significant at 5% level, (ref) = reference category, ICC: Intra-region correlation The deviance-based Chi-square (deviance = 928.61) taken from single logistic regression analysis is the difference in deviance between the empty model with random intercept (deviance = 15,243.52) and fixed slope model with random intercept (deviance = 14,314.91). The significant of it (X^2 =928.61, DF = 15, P-value < 0.0001) implies that fixed slope model with random intercept model is better than empty model with random intercept. Therefore, this model is a better fit as compared to the empty model with random intercept.   Moreover, the AIC and BIC value for fixed slope model with random intercept (AIC=14,751.85, and BIC=14,762.01) are less than those for the empty model with random intercept (AIC = 15,247.52 and BIC = 15,248.31). This indicates that fixed slope model with random intercept is a better fit as compared to the empty model with random intercept model.

3.2.4 Results of Random Coefficient Multilevel Logistic Regression Model

Table 6: Results of Random Coefficient Multilevel Logistic Regression Model

Solutions for Fixed Effects							Odds Ratio Estimates		
Effect	**Level**	**Estimate**	**S.E**	**DF**	t-value	Pr>|**t**|	**Estimate**	95% Confidence Limits	
								LCL	UCL
Intercept		-1.4765	0.3124	58	-4.73	<.0001*	.	.	.
Age of child	>29 months	0	.	.	.	.	.	.	.
	<30 months	0.3513	0.1396	2797	2.52	0.0119*	**0.219**	0.164	0.293
Breast feeding status	Never breasted	0	.	.	.	.	.	.	.
	Ever breasted not currently	-0.0124	0.1282	2797	-0.10	0.9227	0.222	0.164	0.300
	Still breast feeding	0.5095	0.1053	2797	4.84	<.0001*	**0.601**	0.489	0.739
Sex	Female	0	.	.	.	.	.	.	.
	Male	-0.4514	0.1722	2797	-2.62	0.0088*	**1.388**	0.283	0.533
Residence of child	Urban	0	.	.	.	.	.	.	.
	Rural	0.4356	0.1277	2797	3.41	0.0007*	**0.647**	0.504	0.831
Level of education of Mother	no education	0	.	.	.	.	.	.	.
	Primary	0.01740	0.00668	2797	2.61	0.7451	1.033	1.004	1.031
	Secondary	0.4950	0.1769	2797	2.80	0.2721	1.117	0.431	0.862
	Higher	0.7325	0.1449	2797	5.06	0.0862	1.223	0.984	1.634
Use of toilet	Unsafe	0	.	.	.	.	.	.	.
	Safe	-2.3357	1.6853	2797	-1.39	0.1659	0.097	0.004	2.635
Pregnant status	No	0	.	.	.	.	.	.	.
	Yes	1.4249	0.5159	2797	2.76	0.0058*	**4.157**	1.512	11.433
Food nutrient status	Yes	0	.	.	.	.	.	.	.
	No	-0.45	0.17	28	4.06	0.0006*	**1.239**	0.28	0.53
Random effect		**B**	**S.E**	**Z-value**		**P- Value**			
$Var(u_{0j}) = \sigma^2_0$		0.024	0.070	2.12		0.0255			
$Var(u_{2j}) = \sigma^2_2$		0.004	0.002	2.42		0.0041			

Cov(u_{0j}, u_{2j})	-0.001	0.004	-2.34	0.0039
Deviance = 14281.85, AIC= 14720.79, BIC = 14731.35, Deviance based chi-square=25.11				

Table 6: reveals the effect of the intercept on region j is estimated to be -1.4765 (0.3124) + U_{0j} and their variance 0.024 (Standard error 0.070). The intercept variance of 0.024 (Standard error 0.070) is interpreted as the between-region variance when all other variables are held constant (i.e. equal to zero). Their mean is -1.4765 (standard error 0. 3124) and their variance is 0.024 (standard error 0.070). The *between-region* variance of slope of Breast feeding status is estimated to be 0.004 (standard error 0.002). Likewise individual *region* slopes of Breast feeding status vary about with a variance 0.004 (standard error 0.002). The negative covariance estimate of -0.001 (standard error 0.004) between intercept and slopes of Breast feeding status, suggest that regions with a high intercept (above-average) tends to have a flatter-than-average slope.

The quantities AIC and BIC can be used to make an overall comparison of this more complicated model with the random intercept model with fixed slope model. We see that from **Table 6**: the value of fit statistics for random coefficient model (AIC = 14720.79 and BIC = 14731.35) is less than random intercept model **(**AIC=14,751.85 and BIC=14,762.01. This indicates that the random coefficient model is a better fit as compared to the random intercept and fixed effect model.

The odds of stunting of child's from mothers who have still breast feeding were 0.601 (OR=0.601) times higher than the odds of stunting of child's from mothers who never breasted controlling other variables in the model and random effects at level two. Women who live in rural households are 0.647 more likely to be stunted (OR=0.647) than women who reside in urban households controlling for other variables in the model and random effects at level two. The odds of stunting status of child from women who are pregnant is more likely to be stunted 4.157 compared to non-pregnant women controlling for other variables in the model and random effects at level two. Women who feed nutrient food to their child are 1.239 more likely to be stunted (OR=1.239) than women who didn't feed nutrient food controlling for other variables in the model and random effects at level two.

3.3 Discussion

This study analysed the Ethiopian Demographic and Health Survey 2016 dataset, exploring the effect of underlying socioeconomic, demographic, and cultural factors on under-five mortalities in Ethiopia. Under-five children whose mothers had work were 27.8% more likely to experience stunting than under-five children whose mothers had not work. These findings contradict those found in Egypt where stunting was significantly higher among children having mothers not working. This might have the implication that mothers working status affect length of breastfeeding (yilak M. 2014).

The study revealed that incidence of stunting was significantly associated with durations of breastfeeding. Under-five children who had ever been breast fed but not currently were 44.8% less likely to experience stunting as compared to under-five children who were never breasted. Under-five children who are still breastfeeding were 36.4% less likely to experience stunting as compared to under-five children who never breast fed. This present findings is in agreement with a study done in Ghana which found that infants that were either fully breastfed or mixed-fed (fed both breast milk and other foods or liquids) had a lower incidence of stunting than non-breastfed infants [5]. This finding also had confirmed with a study done in Bangladesh which showed than infants who were partially or not breastfed had a high risk of stunting death than exclusively breastfed infants [7]. Not breastfeeding resulted in high exposure of stunting morbidity in comparison to exclusive breastfeeding among infants 0-5 months of age (RR: 10.52) [17] which is also consistent with our study. This might be due to the fact that breast feeding provides vitamins and nutrients that help children develop important antibodies that reduce stunting disease.

This study found that incidence of stunting was significantly associated with nutritional status of under-five children. The prevalence of stunting was higher in stunting under-five children. The odds of having stunting in chronic malnutrition under-five children were 22.6% higher as compared to under-five children who had no chronic malnutrition. This finding is supported by a study done in Zimbabwe and Bangladesh that showed severely stunted children were more likely to have stunting than children of normal height and which had not severe malnutrition [18].

Under-five children who were wasting (acute malnutrition) were 49.2% more likely to experience stunting than under-five children who were not wasted. This present findings is in agreement with a study done in Uganda, which showed that being wasted increases the probability of occurrence of stunting by 14% compared to well-nourished counterparts. The study revealed that incidence of stunting was significantly associated with underweight. Under-five children who were underweight (have low weight-for age) were 54.8% more likely to experience stunting than children who were not underweight. This is consistent with a study in Ghana which showed that stunting was significantly higher for those children who were underweight (yilak M. 2014).

The finding in this study is the identification of variable at the regional level that explains the variation in stunting between the regions of Ethiopia. There are no studies involving multilevel modeling of stunting in Ethiopia that included variables at higher levels. The present study also identified socio-economic indicators of the region as predictors of unemployment. This is the exposure of stunting in different regions of Ethiopia. According to the final model, this level-two variable explains all of the regional-level variation in stunting found in the data.

4. Conclusions

The purpose of this study has been to identify demographic, socio-economic, environmental and nutrition related determinants and to assess regional variation of incidence of childhood stunting in Ethiopia. The descriptive results showed that 15.6% of under-five children have experienced stunting in the two weeks prior to the time of survey (EDHS 2016).

In multilevel logistic regression analysis, under-five children are considered as nested within the various regions in Ethiopia. As a first step in the multilevel approach, non-parametric statistical methods were applied to see if there were differences in the prevalence of stunting in under-five children among the regions. The non-parametric approach based on the chi-square test suggests that prevalence of stunting in under-five children varies among regions. Among the three multilevel logistic regressions models, the random intercept and fixed coefficients model provided the best fit for the data under consideration. It showed that the prevalence of childhood stunting was varying among regions. The significant determinants of prevalence of stunting among regions were age of child, maternal working status, duration of breast fed, stunting, wasting, and underweight.

The main objective of this study is to empirically investigate the major factors that are associated with stunting among children below five years old in Ethiopia. Using the EDHS 2016 data and examines the change in risk factors associated with stunting across the different EDHS years. Age of child, breast feeding, sex, pregnant status, and food nutrient were found to be significantly associated with stunting.

The odds of stunting of child's from mothers who have still breast feeding were 0.601 (OR=0.601) times higher than the odds of stunting of child's from mothers who never breasted controlling other variables in the model and random effects at level two. Women who live in rural households are 0.647 more likely to be stunted (OR=0.647) than women who reside in urban households controlling for other variables in the model and random effects at level two. The odds of stunting status of child from women who are pregnant is more likely to be stunted 4.157 compared to non-pregnant women controlling for other variables in the model and random effects at level two. Women who feed nutrient food to their child are 1.239 more likely to be stunted (OR=1.239) than women who didn't feed nutrient food controlling for other variables in the model and random effects at level two.

Abbreviations

CSA: Central Statistical Agency; *DHS*: Demographic Health Survey; *SE*: Standard Error; *SNNPR*: South Nations Nationalities of Peoples Region; *UNICEF*: United Nations International Children's Emergency Fund; *WHO*: World Health Organization

Declarations

- **Ethics approval and consent to participate:**

Ethics approval for this study was not required since the data is secondary and is available in the public domain.

- **Consent for publication:**

Not applicable

- **Availability of data and materials:**

The datasets used and/ or analysed during the current study are available from the corresponding author on reasonable request.

- **Competing interests:**
  The authors declare that they have no competing interests.

- **Funding:**
  The authors have no support or funding to report.

- **Authors' contributions:**

GB involved in the inception to design, analysis and interpretation and revises critically the manuscript and edit the manuscript for the final submission, WK involved from the inception to design, acquisition of data, analysis and interpretation, drafting the manuscript. Both authors read and approved the final manuscript.

- **Author details**

[1] Department of Statistics, College of Natural Sciences, Wollo University, Dessie, Ethiopia

[2] Department of Statistics, College of Natural and Computational Sciences, Salale University, Fiche, Ethiopia

- **Acknowledgments**

The authors are grateful to CSA (Central Statistical Agency) for providing the 2016 DHS data of Ethiopia.

Reference

[1]. de Onis M., Onyango A.W., Borghi E., Siyam A., Blössner M., Lutter C.K. *et al.* (2012)Worldwide implementation of the WHO Child Growth Standards. Public Health Nutrition 15, 1603–1610. [PubMed] [Google Scholar]

[2]. Mercede s de Onis*, Monika Blo¨ssner and Elaine Borghi, Prevalence and trends of stunting among pre-school children, 1990–2020 , Public Health Nutrition: page 1 of 7

[3]. Long, L.A., and T. Mulatu. 2013. "Getting community-level training right: Important lessons from Ethiopia." Africa Health 35 (3):22-23.

[4]. Black RE, Allen LH, Bhutta ZA et al. for the Maternal and Child Undernutrition Group Maternal and child undernutrition: global and regional exposures and health consequences. Lancet. 2008; 371: 243-260

[5]. Ruksha na Shroff & Sarah Lede rmann, Long-term outcome of chronic dialysis in children, Pediatr Nephrol (2009) 24:463–474 DOI 10.1007/s00467-007-0700-2

[6]. KDHS. 2014. "Kenya Demographic and Health Survey." In, ed Kenya National Bureau of Statistics, Kenya Nairobi, National AIDS Control Council, Kenya Nairobi, National AIDS/STD Control Programme, Kenya Nairobi, Ministry of Public Health and Sanitation, Kenya Nairobi, Kenya Medical Research Institute, Kenya Nairobi, National Coordinating Agency for Population and Development, Kenya Nairobi, ICF Macro MEASURE DHS, Maryland Calverton, U.S.A.,, U.S. Agency for International Development (USAID), Kenya Nairobi and United Nations Population. http://www.dhsprogram.com/publications/publication-fr308-dhs-final-reports.cfm.

[7]. Arifeen S. Exclusive Breastfeeding Reduces Acute Respiratory Infection and Diarrhea Deaths Among Infants in Dhaka Slums, PEDIATRICS is the official journal of the American Academy of Pediatrics. 2001

[8]. Walker et al. 2007

[9]. UNICEF Global Nutrition Database, 2012, MICS, DHS and other national surveys, with additional analysis by UNICEF, 2006–2011.

[10]. 2014, CSA (Central Statistical Agency), Ethiopia Mini Demographic and Health Survey, Addis Ababa: CSA.

[11]. Central Statistical Agency (CSA) [Ethiopia] and ICF, Ethiopia Demographic and Health Survey 2016, CSA and ICF, MD, USA, 2016.

[12]. Measure DHS, Te DHS program: Demographic and Health Surveys. Rockville: ICF Macro, 2016, https://dhsprogram.com/data/availabledatasets.cfm.

[13]. Guang Guo and Hongxin Zhao, Multilevel Modeling for Binary Data, https://www.jstor.org/stable/223452

[14]. Li, M. Zhou, G. Sapiro, and L. Carin. On the integration of topic modeling and dictionary learning. In ICML, 2011.

[15]. Math J. J. M. Candel and Bjorn Winkens, Performance of Empirical Bayes Estimators of Level-2 Random Parameters in Multilevel Analysis: A Monte Carlo Study for Longitudinal Designs, Journal of Educational and Behavioral Statistics, Vol. 28, No. 2 (Summer, 2003), pp.169-194, American Educational Research Association and American Statistical Association, https://www.jstor.org/stable/3701260

[16]. Maxwell, D., C. Ahiadeke, C. Levin, M. Armar-Klemesu, S. Zakariah, and G. M. Lamptey. 1999. Alternative food security indicators: Revisiting the frequency and severity of 'coping strategies'. Food Policy 24 (4): 411–424.

[17]. Laura M Lamberti , Christa L Fischer Walker, Adi Noiman , Cesar Victora , Robert E Black, Breastfeeding and the risk for diarrhea morbidity and mortality, . BMC Public Health 2011, 11(Suppl 3):S15

[18]. Edem M. A. Tette1, Eric K. Sifah and Edmund T. Nartey, Factors affecting malnutrition in children and the uptake of interventions to prevent the condition, Tette et al. BMC Pediatrics (2015) 15:189 DOI 10.1186/s12887-015-0496-3

[19]. Deepesh Bhati, Phazamile Kgosi, Ranganath Narayanacharya Rattihalli, Distribution of Geometrically Weighted Sum of Bernoulli Random Variables, Applied Mathematics, 2011, 2, 1382-1386,

[20]. Fauveau, C., Siddiqui, M., Briend, A., Silimperi, D. R., Begum, N., & Fauveau, V. 1992. Limited impact of a targeted food supplementation programme in Bangladeshi urban slum children. Ann Trop Paediatric, 105(3)