doi.org/10.26398/IJAS.0030-017

APPLICATION OF ZERO ADJUSTED MODELS FOR MODELLING NUMBER OF ANTENATAL CARE VISITS OF PREGNANT WOMEN IN BANGLADESH

Ayan Chowdhury[1], Soma Chowdhury Biswas

Department of Statistics, University of Chittagong, Chittagong, Bangladesh

Abstract: *Antenatal care (ANC) visit is the key to monitor the status of a pregnancy, identify the complications associated with the pregnancy, and prevent adverse pregnancy outcomes. Lack of utilization of antenatal care is a major cause of maternal and neonatal mortality in developing countries including Bangladesh. This study aimed at modelling the number of antenatal care visits using Bangladesh Demography and Health Survey (2014) data. Since several count models such as ZIP, ZINB, HP and HNB have been implemented, this study have selected the best model based on AIC, BIC and Vuong test, and identified the reasonable factors that influence the antenatal care visits of Bangladeshi women. Among these models HNB model is found as the best model for modelling the number of antenatal care visits of women.*

Keywords: *Hurdle models, Vuong test, Zero inflated models.*

1. INTRODUCTION

Antenatal care (ANC) is the medical supervision given to a pregnant woman and her baby starting from the time of conception up to the delivery of the baby by a physician, midwife or obstetrician or a combination of these professionals. It includes regular monitoring of the mother and her baby throughout pregnancy by a variety of routine regular examinations and tests. Its main objective is to ensure a normal pregnancy with delivery of a healthy baby from a healthy mother. Antenatal care contributes to good pregnancy outcomes and oftentimes benefits of antenatal care are dependent on the timing and quality of the care provided. More than half a million women die each year as a result of complications arising from pregnancy and child birth (WHO, 2003). Lack of antenatal care has been identified as one of the risk factors for maternal mortality and other adverse pregnancy outcomes in developing countries (Anandalakshmy et al., 1993; Fawcus et al.,

[1] Corresponding author: Ayan Chowdhury, email: ayan.statcu09@yahoo.com

1996). The use of ANC in developing countries is low compared to developed countries. In developing countries including Bangladesh complications of pregnancy and child birth are the leading causes of deaths among women of reproductive age (WHO, 2012). The World Health Organization (WHO) defines women of reproductive age as those who are aged 15-49 years. They constitute more than one-fifth of the world's population and are repeatedly exposed to the risk of pregnancy and childbearing (WHO, 2003). The WHO recommends that pregnant women make a first visit between 8-12 weeks after conception and make further three visits between 24 and 38 weeks of gestation (WHO, 2002). According to the WHO recommendation, every pregnant woman should receive at least four ANC visits during pregnancy. Four visits for antenatal care are sufficient for uncomplicated pregnancies and more are necessary only in case of complications. Moreover, in developing countries, women often encounter serious health risks during pregnancy either for themselves or for their children (WHO, 2007). It is found that, birth weight is positively correlated with the number of antenatal care visits in Bangladesh (Ahmed and Das, 1992). Antenatal care is a key of reducing maternal and neonatal mortality in developing countries. But the women of developing countries have low interest to take antenatal care due to lack of their awareness about antenatal care during pregnancy.

According to Bangladesh Demographic and Health Survey (BDHS) 2014, 78 percent (68 percent in 2011) of women take antenatal care at least once during pregnancy. But only 31 percent of women have four or more antenatal care visits during the course of pregnancy. Urban women are more likely than rural women to have made four or more antenatal visits (46 percent compared with 26 percent). For urban women this percentage has hardly changed between 2011 and 2014 (from 45 to 46 percent), while in rural areas the percentage of women who made four or more antenatal care visits increased from 20 to 26 percent. Only 42 percent of births in the past three years was assisted by a medically trained provider and 37 percent of births in Bangladesh takes place in a healthy facility. This study is aimed to statistically analyze the determinants of the barriers in number of antenatal care visits of pregnant women in Bangladesh. Furthermore, this study will provide valuable information about several models for modelling count data. This study will also present the best model for analyzing the number of antenatal care visits in the specific case study which is selected by comparing several models based on AIC, BIC and Vuong test.

2. MATERIALS AND METHODS

2.1 SOURCE OF DATA

In this study, the data stems from the response of women’s questionnaire of the Bangladesh Demography and Health Survey (BDHS), 2014. This survey is a part of the worldwide Demography and Health Surveys program, which is designed to collect data on various socio-economic and demographic characteristics of a territory. A total of 18,245 ever-married women aged 15-49 were identified in these households and 17,863 were interviewed, for a response rate of 98 percent. In this study, we have 4396 respondents who gave a live birth preceding 3 years of the survey.

2.2 VARIABLES

In this study the dependent variable is the number of antenatal care visits of women (respondent) during their pregnancy period. From literature reviews, respondent’s age, residence, source of drinking water, respondent’s education, wealth index, husband’s education, decision maker on respondent’s health care and access to mass media are assumed as the determinants of the barriers in number of antenatal care visits.

2.3 STATISTICAL MODELS

The number of times an event occurs within a fixed time interval is known as count event and measured as non-negative and discrete. The modelling of count data is of a primary interest in many fields such as insurance, public health, epidemiology, psychology, and many other research areas. In case of count data modelling, Poisson regression model is commonly used under two principal assumptions: one is that events occur independently over given time and the other is that the conditional mean and variance are equal. However, the equality of the mean and variance rarely occurs; the variance may be either greater than the mean (over-dispersion) or less than the mean (under-dispersion). In that case, the assumption of equality of the mean and variance is violated. In case of over-dispersion problem, Negative Binomial (NB) regression model may be used instead of Poisson regression model. But when count data display not only over-dispersion but also excess of zero, then over-dispersed count data can not be modelled accurately with Negative Binomial model. In this situation, zero-inflated models (i.e. zero-inflated Poisson and zero-inflated Negative Binomial) and hurdle models (i.e. hurdle Poisson and hurdle Negative Binomial) are more appropriate for modelling this kind of over-dispersed count data.

2.3.1 ZERO-INFLATED POISSON REGRESSION MODEL

Introduced by Lambert (1992), zero inflated count models provide a way of modelling the excess zeros in addition to allowing for over-dispersion. Zero inflated Poisson (ZIP) distribution is a mixture distribution assigning a mass of ω_i to the excess zeros and a mass of (1-ω_i) to a Poisson distribution (Chipeta et al., 2014). The probability mass function of zero inflated Poisson (ZIP) distribution is given by:

$$\Pr(Y_i = y_i) = \begin{cases} \omega_i + (1-\omega_i)e^{-\mu_i};\ \mathrm{y_i} = 0 \\ (1-\omega_i)\dfrac{e^{-\mu_i}\mu_i^{y_i}}{y_i!};\ y_i = 1,2,3......, \end{cases} \quad \textbf{(1)}$$

where, $0 \le \omega_i \le 1$ and $\mu_\mathrm{i} \ge 0$. The mean and variance of this distribution are $(1-\omega_i)\mu_\mathrm{i}$ and $\mu_i(1-\omega_i)(1+\mu_i\omega_i)$ respectively. The conditional mean μ_i of the Poisson distribution is expressed as: $\mu_i = \exp(x_i'\beta)$, where x_i is (p+1)×1 a (p+1)×1 vector ofcovariates β is a (p+1)×1 vector of parameters to be estimated and p is the number of covariates in the model.

The parameter ω_i presents the probability of zero inflation. One can observe clearly that the ZIP reduces to the classical Poisson model when $\omega_i = 0$. According to Lambert, we can model ω_i using a Logit model given by: $\mathrm{logit}(\omega_i) = z_i'\gamma$, where z_i is a vector of covariates γ is a (q+1)×1 vector of parameters to be estimated and q is the number of the covariates in the model. In the terminology of generalized linear models (GLMs), log(μ_i) and logit (ω_i) are the log link for the Poisson mean and logit link for Bernoulli probability of success respectively (Lambert 1992).

Zero inflated Poisson (ZIP) regression model is expressed as:

$$\begin{aligned} \log(\mu_i) &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + ... + \beta_p x_{ip} \\ \mathrm{logit}(\omega_i) &= \gamma_0 + \gamma_1 z_{i1} + \gamma_2 z_{i2} + \gamma_3 z_{i3} + ... + \gamma_q z_{iq}. \end{aligned} \quad \textbf{(2)}$$

The log-likelihood function is:

$$\log L = \sum_{y_i=0} \log\left[\omega_i + (1-\omega_i)e^{-\mu_i}\right] + \sum_{y_i \ne 0}\left[\log(1-\omega_i) - \mu_i + y_i \log\mu_i - \log\left(\mathrm{y_i!}\right)\right]. \quad \textbf{(3)}$$

The parameters of this model can be estimated using maximum likelihood estimation.

2.3.2 ZERO-INFLATED NEGATIVE BINOMIAL REGRESSION MODEL

Zero Inflated Negative Binomial (ZINB) model is a mixture of two statistical processes, one always generating zero counts and the other generating both zero and non-zero counts. In case of Zero Inflated Negative Binomial (ZINB) model, a Logit model with Binomial assumption is used to determine if an individual count outcome is from the always zero or the not always zero group and then Negative Binomial model is used to model outcomes in the not always zero group (Liu and Cela, 2008). For respondent *i*, with probability ω_i the only possible response of the first process is zero counts, and with probability of $(1-\omega_i)$ the response of the second process is governed by Negative Binomial distribution with mean μ_i. The probability mass function (Yau et al., 2003) of Zero Inflated Negative Binomial (ZINB) distribution is given by:

$$\Pr(Y_i = y_i) = \begin{cases} \omega_i + (1-\omega_i)(1+\alpha\mu_i)^{-1/\alpha} & ; y_i = 0 \\ (1-\omega_i)\dfrac{\Gamma(y_i + 1/\alpha)(\alpha\mu_i)^{y_i}}{\Gamma(y_i+1)\Gamma(1/\alpha)(1+\alpha\mu_i)^{y_i+1/\alpha}} & ; y_i = 1,2,3..... \end{cases} \tag{4}$$

with mean $(1-\omega_i)\mu_i$ and variance $(1-\omega_i)\mu_i\,(1+\mu_i\,(\omega_i+a))$ where, α is the dispersion parameter. The parameter μ_i is expressed as: $\mu_i = \exp(x_i'\beta)$ where, β is the $(p+1)\times 1$ vector of unknown parameters associated with the known covariates vector x_i and p is the number of covariates. The parameter ω_i is often referred as the zero-inflation factor, which is the probability of zero counts from the binary process. For common choice and simplicity, ω_i is characterized in terms of a logistic regression model by writing as: logit $(\omega_i) = \text{logit}(\omega_i) = z_i'\gamma$, where, γ is the $(q+1)\times 1$ vector of zero-inflated coefficients to be estimated which is associated with the known zero-inflation covariates vector z_i and q is the number of the covariates. In the terminology of generalized linear models (GLMs) $\log(\mu_i)$ and $\text{logit}(\omega_i)$ are the log links for the Negative Binomial mean and logit link for Bernoulli probability of success (Lambert 1992).

Zero Inflated Negative Binomial (ZINB) regression model is expressed as:

$$\begin{aligned} \log(\mu_i) &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + ... + \beta_p x_{ip} \\ \text{logit}(\omega_i) &= \gamma_0 + \gamma_1 z_{i1} + \gamma_2 z_{i2} + \gamma_3 z_{i3} + ... + \gamma_q z_{iq}. \end{aligned} \tag{5}$$

The log-likelihood function is:

$$\log L = \sum_{y_i=0} \log[\omega_i + (1-\omega_i)(1+\alpha\mu_i)^{-1/\alpha}] +$$

$$\sum_{y_i \neq 0} \left[\log(1-\omega_i) + \log\frac{\Gamma(y_i+1/\alpha)}{\Gamma(y_i+1)\Gamma(1/\alpha)} + y_i \log(\alpha\mu_i) - (y_i + 1/\alpha)\log(1+\alpha\mu_i) \right]. \tag{6}$$

The parameters of this model can be estimated using maximum likelihood estimation.

2.3.4 HURDLE NEGATIVE BINOMIAL REGRESSION MODEL

In the hurdle Negative Binomial regression model, the first part is governed by binary process and the second part is governed by truncated Negative Binomial distribution. The unconditional probability mass function for Y_i is:

$$\Pr(Y_i = y_i) = \begin{cases} \omega_i & ; y_i = 0 \\ (1-\omega_i)\dfrac{e^{-\mu_i}\mu_i^{y_i}}{\left(1-e^{-\mu_i}\right)y_i!} & ; y_i = 1,2,3.........., \end{cases} \tag{7}$$

where, $0 \le \omega_i \le 1$ and $\mu_i \ge 0$ The mean and variance of this distribution are $(1-\omega_i)\mu_i$ and $\mu_i(1-\omega_i)(1+\mu_i\omega_i)$ respectively. The conditional mean μ_i of the Poisson distribution is expressed as, $\mu_i = \exp(x_i'\beta)$, where x_i is a (p+1)×1 vector of covariates, β is a (p+1)×1 vector of parameters to be estimated and p is the number of covariates in the model.

The parameter ω_i is the probability of observing a zero count and (1-ω_i) is the probability of observing a positive count. For the hurdle model, the zero hurdle component describes the probability of observing a positive count whereas, for the zero-inflated model, the zero-inflation component predicts the probability of observing a zero count from the pointmass component (Zeileis et al., 2008). We can model (1-ω_i) using a Logit model given by: $\text{logit}(1-\omega_i) = z_i'\gamma$, where z_i is a (p+1)×1 vector of covariates, γ is a (p+1)×1 vector of parameters to be estimated and q is the number of the covariates in the model.

Hurdle Poisson regression model is expressed as:

$$\log(\mu_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + ... + \beta_p x_{ip}$$
$$\text{logit}(1-\omega_i) = \gamma_0 + \gamma_1 z_{i1} + \gamma_2 z_{i2} + \gamma_3 z_{i3} + ... + \gamma_q z_{iq}. \tag{8}$$

The log-likelihood function is:

$$\log L = \sum_{y_i=0} \log \omega_i + \sum_{y_i \ne 0} \left[\log(1-\omega_i) - \mu_i + y_i \log \mu_i - \log\left(1-\mathrm{e}^{-\mu_i}\right) - \log\left(\mathrm{y}_i!\right) \right]. \tag{9}$$

The parameters of this model can be estimated using maximum likelihood estimation.

2.3.4 HURDLE NEGATIVE BINOMIAL REGRESSION MODEL

In the hurdle Negative Binomial regression model, the first part is governed by binary process and the second part is governed by truncated Negative Binomial distribution. The unconditional probability mass function for Y_i is:

$$\Pr(Y_i = y_i) = \begin{cases} \omega_i & ; y_i = 0 \\ (1-\omega_i)\dfrac{\Gamma(y_i+\alpha^{-1})}{\Gamma(y_i+1)\Gamma(\alpha^{-1})}\dfrac{(1+\alpha\mu_i)^{-\alpha^{-1}-y_i}\alpha^{y_i}\mu_i^{y_i}}{1-(1+\alpha\mu_i)^{-\alpha^{-1}}} & ; y_i = 1,2,3...., \end{cases} \quad \textbf{(10)}$$

where, α is the dispersion parameter. The mean and variance of this model are $(1-\omega_i)\mu_i\left[1-(1+\alpha\mu_i)^{-\alpha^{-1}}\right]^{-1}$ and $(1-\omega_i)\mu_i\{1+\mu_i(\omega_i+\alpha)\}\left[1-(1+\alpha\mu_i)^{-\alpha^{-1}}\right]^{-2}$ respectively. The parameter μ_i is expressed as: $\mu_i = \exp(x_i'\beta)$, .where, β is a (p+1)×1 vector of unknown parameters associated with the known covariates vector x_i and p is the number of covariates.

The parameter ω_i is the probability of observing a zero count and (1-ω_i) is the probability of observing a positive count. For common choice and simplicity, (1-ω_i) is characterized in terms of a logistic regression model by writing as: $\text{logit}(1-\omega_i) = z_i'\gamma$, where, γ is a (q+1)×1 vector of hurdle coefficients to be estimated which is associated with the known hurdle covariates vector z_i and q is the number of covariates in the model.

Hurdle Negative Binomial regression model is expressed as:

$$\begin{aligned} \log(\mu_i) &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + ... + \beta_p x_{ip} \\ \text{logit}(1-\omega_i) &= \gamma_0 + \gamma_1 z_{i1} + \gamma_2 z_{i2} + \gamma_3 z_{i3} + ... + \gamma_q z_{iq}. \end{aligned} \quad \textbf{(11)}$$

The log-likelihood function is:

$$\log L = \sum_{y_i=0} \log \omega_i + \sum_{y_i \neq 0}\left[\log(1-\omega_i) + \log h - \log\left\{1-(1+\alpha\mu_i)^{-\alpha^{-1}}\right\}\right] \quad \textbf{(12)}$$

$$\text{where, } h = \frac{\Gamma(y_i+\alpha^{-1})}{\Gamma(y_i+1)\Gamma(\alpha^{-1})}(1+\alpha\mu_i)^{-\alpha^{-1}-y_i}\left(\alpha\mu_i\right)^{y_i}. \quad \textbf{(13)}$$

The parameters of this model can be estimated using maximum likelihood estimation.

2.4 MODEL SELECTION

Among several statistical models to select the best model Akaike information criteria (AIC), Bayesian information criteria (BIC) and Vuong test have been used in this study.

2.4.1 INFORMATION CRITERIA

Akaike information criteria (AIC) is a measure of the relative quality of statistical models for a given set of data. Given a set of candidate models for the data, the preferred model is the one with the minimum AIC value. Akaike information criteria (AIC) is defined as:

$$\text{AIC} = -2\log(\text{L}) + 2\text{k}. \quad \textbf{(14)}$$

Bayesian information criteria (BIC) is also used to select the best model among a set of models. Bayesian information criteria (BIC) is defined as:

$$\text{BIC} = -2\log(\text{L}) + \text{k}\log(\text{n}) \quad \textbf{(15)}$$

where, L is maximum value of the likelihood function for the model, n is sample size and k is the number of parameters to be estimated.

When comparing the Akaike information criteria (AIC) and the Bayesian information criteria (BIC), penalty for additional parameter is more in BIC than AIC. AIC is good for making asymptotically equivalent to cross-validation. On the contrary, BIC is good for consistent estimation.

2.4.2 VUONG TEST

The Vuong test was introduced by Vuong (1989), as a non-nested test for comparing two models. Vuong test is used to select the best model for count data (Liu and Cela, 2008; Moineddin et al., 2011). It is a test that is based on a comparison of the predicted probabilities of two models.

Let's define,

$$m_i = \log\left\{\frac{f_1(Y_i|X_i)}{f_2(Y_i|X_i)}\right\} \quad \textbf{(16)}$$

where, $f_1(Y_i|X_i)$ is the predicted probability of observed count for case *i* from model 1 and $f_2(Y_i|X_i)$ is the predicted probability of observed count for case *i* from model 2.

Null hypothesis underlying Vuong's approach is given by:

$$H_0 : E(m) = 0.$$

Alternative hypotheses are:

$$H_{f_1} : E(m) > 0; \qquad H_{f_1} \text{ is better}$$

$$H_{f_2} : E(m) < 0; \qquad H_{f_2} \text{ is better.}$$

If $E(m) > 0$, then the alternative hypothesis is considered as:

H_A: model 1 > model 2 (i.e. model 1 is better than model 2),

but if $E(m) < 0$, then the alternative hypothesis is considered as:

H_A: model 2 > model 1 (i.e. model 2 is better than model 1).

Under the null hypothesis, the Vuong test statistic is given by

$$V = \frac{\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} m_i\right)}{\sqrt{\operatorname{var}(m)}} \sim N(0,1) \tag{17}$$

where, $Var(m) = \frac{1}{n}\sum_{i=1}^{n} m_i^2 - \left[\frac{1}{n}\sum_{i=1}^{n} m_i\right]^2$ and n is sample size.

Mathematically, if we let confidence interval is $P\left[-Z_\alpha < V < Z_\alpha\right] = 1-\alpha$, then V is greater than Z_α then evidence exists which favors the model 1 relative to the model 2 at α level of significance. Conversely, if V is less than $-Z_\alpha$ evidence exists which favors the model 2 relative to the model 1 at α level of significance. Finally, if V is less than or equal to Z_α and greater than or equal to $-Z_\alpha$ the weak evidence exists, and we cannot decisively determine which model is favored over the other at α level of significance. More specifically, if V > 1.64 then model 1 is better than model 2, if V < -1.64 then model 2 is better than model 1 and if $-1.64 \le V \le 1.64$ then both models are equally appropriate at 5% level of significance.

2.5 SOFTWARE

In this study, **SPSS version 16.0** has been used in descriptive analysis, **R statistical software version 3.2.3 (packages: MASS, pscl)** has been used for fitting several statistical models and models comparison.

3. DATA ANALYSIS AND RESULT

3.1 DESCRIPTIVE ANALYSIS

The number and percentage of the women corresponding to different number of ANC visits during pregnancy period (n = 4396) are given in **Table 1**.

Tab. 1: Frequency table of number of antenatal care visits

No. of ANC visits	0	1	2	3	4	5	6	7	8	9	10
No. of respondents	944	717	731	604	484	299	184	172	157	74	30
Percentage (%)	21.5	16.3	16.6	13.7	11.0	6.8	4.2	3.9	3.6	1.7	0.7

It can be seen that the number of women who have not visited for ANC is maximum. It is also found that the variance of the number of ANC visits is 5.89 which is greater than its mean 2.72.

The number and percentage of the women (n = 4396) of several levels of covariates are given in **Table 2**. In this table, the number and percentage of women who have visited as well as who have not visited for antenatal care during pregnancy for each level of several covariates are also presented.

Tab. 2: Frequency table of women in different levels of several covariates

Independent variables	Levels	Number of respondent (%)	ANC visit (percentage)	
			Yes	No
Respondent's age	Under 19	610 (13.9)	481 (78.9)	129 (21.1)
	19-29	2932 (66.7)	2332 (79.5)	600 (20.5)
	Above 29	854 (19.4)	639 (74.8)	215 (25.2)
Place of residence	Urban	1405 (32.0)	1239 (88.2)	166 (11.8)
	Rural	2991 (68.0)	2213 (74.0)	778 (26.0)
Source of drinking water	Other	579 (13.2)	447 (77.2)	132 (22.8)
	Tap water	321 (7.3)	295 (91.9)	26 (8.1)
	Tube-well water	3496 (79.5)	2710 (77.5)	786 (22.5)
Respondent's education	Illiterate	587 (13.4)	325 (55.4)	262 (44.6)
	Primary	1268 (28.8)	884 (69.7)	384 (30.3)
	Secondary	2137 (48.6)	1851 (86.6)	286 (13.4)
	Above secondary	404 (9.2)	392 (97.0)	12 (3.0)
Wealth index	Poor	1757 (40.0)	1117 (63.6)	640 (36.4)
	Middle	841 (19.1)	673 (80.0)	168 (20.0)
	Rich	1798 (40.9)	1662 (92.4)	136 (7.6)
Husband's education	Illiterate	1023 (23.3)	635 (62.1)	388 (37.9)
	Primary	1342 (30.5)	982 (73.2)	360 (26.8)
	Secondary	1429 (32.5)	1262 (88.3)	167 (11.7)
	Above secondary	602 (13.7)	573 (95.2)	29 (4.8)
Decision maker on respondent's health care	Respondent alone	520 (11.8)	430 (82.7)	90 (17.3)
	Respondent & husband	2122 (48.3)	1690 (79.6)	432 (20.4)
	Husband alone	1417 (32.2)	1053 (74.3)	364 (25.7)
	Someone else	337 (7.7)	279 (82.8)	58 (17.2)
Access to mass media	No	1670 (38.0)	1053 (63.1)	617 (36.9)
	Yes	2726 (62.0)	2399 (88.0)	327 (12.0)

From **Table 2**, it is seen that, the women aged 19-29 are most likely to take antenatal care (79.5 percent), while the women aged above 29 are least likely (74.8 percent). It is observed that, urban women (88.2 percent) are more likely to take antenatal care compared with rural women (74.0 percent). The women who drink tap water are most likely to visit for antenatal care (91.9 percent), while the women who drink water of other sources are least likely (77.2 percent). It is also found that, the uneducated women are least likely to take antenatal care (55.4 percent), while the above secondary educated women are most likely (97.0 percent). Rich women are most mindful about antenatal care (92.4 percent), while poor women are least mindful (63.6 percent). It is observed that, the women whose husbands are uneducated are least attentive (62.1 percent), while the women whose husbands are above secondary educated are most attentive (95.2 percent) about antenatal care during pregnancy. Among the women whose health care decision maker is someone else are most likely (82.8 percent) to visit medically trained provider for antenatal care. It is also found that the women who have access to mass media are more attentive about antenatal care during pregnancy (88.0 percent) compared with the women who have no access to mass media (63.1 percent).

3.2 MODELLING THE NUMBER OF ANC VISITS

From descriptive study, it is found that the variance of number of ANC visits is 5.89 which is greater than its mean 2.72, which indicates the number of ANC visits presents over-dispersed that violates the assumption of equality of the mean and variance of Poisson regression model. It means that the Poisson regression model is not appropriate for modelling the number of ANC visits. In this case, Negative Binomial (NB) regression model may be applied as over-dispersed model. But the number of ANC visits is over-dispersed with excess of zero (i.e. frequency of zero time visits is maximum). Therefore, Negative Binomial regression model is not appropriate for modelling the number of ANC visits accurately. In such situation, zero-inflated models (i.e. ZIP and ZINB) and hurdle models (i.e. HP and HNB) are more appropriate for modelling this kind of over-dispersed count data. Finally, ZIP, ZINB, HP and HNB models have been applied for modelling the number of ANC visits of women in Bangladesh.

3.3 MODELS COMPARISON

To compare ZIP, ZINB, HP and HNB models AIC, BIC and Vuong test have been used.

Tab. 3: Comparison of different models in count data with AIC and BIC

Model	AIC	BIC
Zero-Inflated Poisson Regression Model (ZIP)	17904.11	18134.09
Zero-Inflated Negative Binomial Regression Model (ZINB)	17643.95	17880.32
Hurdle Poisson Regression Model (HP)	17902.34	18132.32
Hurdle Negative Binomial Regression Model (HNB)	17629.45	17865.82

Tab. 4: Comparison of different models in count data with Vuong test

Model	ZIP	ZINB	HP	HNB
ZIP	-------------			
ZINB	V = -8.130 P = 0.000 ZINB is better	-------------		
HP	V = -1.024 P = 0.153 ZIP = HP	V = 8.050 P = 0.000 ZINB is better	-------------	
HNB	V = -8.179 P = 0.000 HNB is better	V = -2.478 P = 0.007 HNB is better	V = -8.158 P = 0.000 HNB is better	-------------

V=Vuong statistic, P = P-value; V > 1.64 indicates that column model had significantly better fit than the row model and V < -1.64 indicates that row model had significantly better fit than the column model at 5% level of significance.

From **Table 3**, it is clear that, Hurdle Negative Binomial (HNB) regression model provides lowest values of AIC and BIC, which indicates Hurdle Negative Binomial (HNB) regression model as the best model. It is also found from **Table 4**, Hurdle Negative Binomial (HNB) regression model is best model based on Vuong test. So this study suggests that, Hurdle Negative Binomial (HNB) regression model is the best model for modelling number of antenatal care visits of women in Bangladesh.

Estimated parameters of count part and zero part of ZIP, ZINB, HP and HNB models are given in the **Table 5(a)** and **Table 5(b)** respectively.

Tab. 5(a): Estimated parameters of ZIP, ZINB, HP, HNB models (count part)

Independent variables (Ref.)	Levels	Estimated parameters (Standard error)			
		ZIP	ZINB	HP	HNB
Intercept		0.85952 (0.067)***	0.75321 (0.084)***	0.86994 (0.067)***	0.79278 (0.086)***
Residence (Urban)	Rural	-0.14185 (0.022)***	-0.14195 (0.028)***	-0.14291 (0.023)***	-0.15014 (0.030)***
Source of drinking water (Other)	Tap water	0.09936 (0.042)*	0.11556 (0.054)*	0.09715 (0.042)*	0.10748 (0.056) .
Respondent's education (Illiterate)	Primary	0.14884 (0.047)**	0.17597 (0.058) **	0.14002 (0.047) **	0.15071 (0.058)**
	Secondary	0.25003 (0.047)***	0.28252 (0.058)***	0.24151 (0.046)***	0.25756 (0.058)***
	Above secondary	0.40194 (0.056)***	0.45436 (0.071)***	0.39184 (0.055)***	0.41803 (0.071)***
Wealth index (Poor)	Rich	0.11664 (0.031)***	0.12150 (0.040)**	0.11372 (0.031)***	0.11649 (0.040)**
Husband's education (Illiterate)	Above secondary	0.13129 (0.042)**	0.13084 (0.054) *	0.13685 (0.042) **	0.14309 (0.055) **
Decision maker on respondent's health care (Respondent alone)	Husband alone	-0.06108 (0.033) .	-0.06337 (0.042)	-0.06097 (0.033) .	-0.06576 (0.043)
	Someone else	-0.09374 (0.044)*	-0.09039 (0.056)	-0.09630 (0.045)*	-0.10265 (0.058) .
Access to mass media (No)	Yes	0.11619 (0.028)***	0.13461 (0.036)***	0.11390 (0.028)***	0.12415 (0.036)***

Ref. = Reference category; Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

From the **Table 5(a)** it is seen that, all independent variables except respondent's age are statistically associated with the ANC visits in ZIP, ZINB, HP and HNB models. Estimated parameters for count part of these models are slightly different to each other.

From the **Table 5(b)** it is seen that, all independent variables without respondent's age and decision maker on respondent's health care are statistically associated with the ANC visits in ZIP, ZINB, HP and HNB models. The absolute value of estimated parameters for zero part of zero-inflated models (ZIP and ZINB) and hurdle models (HP and HNB) are rather different which is not surprising as they pertain to slightly different ways of modelling zero counts. But the signs of the estimated parameters of zero-inflated models

Tab. 5(b): Estimated parameters of ZIP, ZINB, HP, HNB models (zero part)

Independent variables (Ref.)	Levels	Estimated parameters (Standard error)			
		ZIP	**ZINB**	**HP**	**HNB**
Residence (Urban)	Rural	0.18191 (0.137)	0.23859 (0.190)	-0.26018 (0.109)*	-0.26018 (0.109)*
Source of drinking water (Other)	Tube-well water	-0.21646 (0.149)	-0.20964 (0.194)	0.19507 (0.119) .	0.19507 (0.119) .
Respondent's education (Illiterate)	Primary	-0.35360 (0.143)*	-0.30638 (0.180) .	0.41149 (0.115)***	0.41149 (0.115)***
	Secondary	-0.89649 (0.164)***	-0.98204 (0.223)***	0.92143 (0.128)***	0.92143 (0.128)***
	Above secondary	-1.59642 (0.524)**	-1.76916 (1.514)	1.62248 (0.344)***	1.62248 (0.344)***
Wealth index (Poor)	Middle	-0.35651 (0.148)*	-0.47964 (0.211)*	0.26058 (0.110)*	0.26058 (0.110)*
	Rich	-0.95536 (0.181)***	-1.32819 (0.335)***	0.85597 (0.131)***	0.85597 (0.131)***
Husband's education (Illiterate)	Secondary	-0.68794 (0.166)***	-0.79973 (0.250)**	0.56167 (0.123)***	0.56167 (0.123)***
	Above secondary	-1.08243 (0.350)**	-3.32086 (4.557)	0.90175 (0.235)***	0.90175 (0.235)***
Access to mass media (No)	Yes	-0.59457 (0.124)***	-0.66953 (0.173)***	0.56881 (0.095)***	0.56881 (0.095)***

Ref. = Reference category; Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(ZIP and ZINB) and hurdle models (HP and HNB) are inversed. For the hurdle model, the zero hurdle component describes the probability of observing a positive count, whereas for the zero-inflated model, the zero-inflation component predicts the probability of observing a zero count from the point mass component (Zeileis et al., 2008). Overall, both zero-inflated models (ZIP and ZINB) and hurdle models (HP and HNB) lead to the same qualitative results. It is also seen that estimated parameters for zero part of ZIP and ZINB models are slightly different, whereas estimated parameters for zero part of HP and HNB models are same which is not surprising as the probability mass function for zero part of ZIP and ZINB models is different, whereas the probability mass function for zero part of HP and HNB models is same.

This study has been focused on Hurdle Negative Binomial (HNB) regression model. So, it is suitable to interpret estimated parameters of Hurdle

Negative Binomial (HNB) regression model. The average number of antenatal care visits of illiterate and poor women aged under 19, who live in urban area, drink water of other sources without tap and tube-well, take the decision on their health care alone, have no access to mass media (radio, television and newspapers) and whose husbands are uneducated is exp (0.79278) = 2.2095 in the group of women who take antenatal care. It is found that, the average number of antenatal care visits of rural women is exp (-0.15014) = 0.8606 times less than urban women. The average number of antenatal care visits of women during pregnancy who drink tap water is exp (0.10748) = 1.1135 times more as compared to the women who drink water of other sources without tap or tube-well. It is also found that, the average number of antenatal care visits of primary educated women is exp (0.15071) = 1.1627 times more, secondary educated women is exp (0.25756) = 1.2938 times more and above secondary educated women is exp (0.41803) = 1.5190 times more than illiterate women. This means that the average number of antenatal care visits is more in the class of higher educated women. In the group of rich women the average number of antenatal care visits is exp (0.11649) = 1.1235 times more as compared to the group of poor women. The average number of antenatal care visits of the women whose husbands are above secondary educated is exp (0.14309) = 1.1538 times more than the women whose husbands are uneducated. In the group of women whose health care decision is taken by someone else the average number of antenatal care visits is exp (-0.10265) = 0.9024 times less than the group of women who take decision alone about their health care. It also found that, the average number of antenatal care visits of the women who have access to mass media is exp (0.12415) = 1.1322 times more as compared to the women who have no access to mass media.

From the zero part of HNB regression model it is found that rural women visit for antenatal care exp (-0.26018) = 0.7709 times less than urban women. The women who drink tube-well water take antenatal care exp (0.19507) = 1.2154 times more than the women who drink water of other sources without tap and tube-well. The primary educated women exp (0.41149) = 1.5091 times more, secondary educated women exp (0.92143) = 2.5129 times more and above secondary educated women exp (1.62248) = 5.0656 times more visit for antenatal care than illiterate women. The women of middle class family take antenatal care exp (0.26058) = 1.2977 times more and the women of rich family take antenatal care exp (0.85597) = 2.3537 times more as compared to the women of poor family. The women whose husbands are secondary educated visit for antenatal care exp (0.56167) = 1.7536 times more and the women whose husbands are above secondary educated visit for antenatal care exp (0.90175) = 2.4639 times more than the women whose husbands are uneducated. It is also found that, the women who have access to mass media visit for antenatal care exp (0.56881) = 1.7662 times more as compared to the women who have no access to mass media.

4. CONCLUSION

According to the findings of this study, rural women visit less times for antenatal care than urban women as well as the average number of antenatal care visits of rural women is less than urban women who visit for antenatal care during pregnancy. Residence of women has an imperative impact on antenatal care utilization (Islam and Odland, 2011; Ochako et al., 2011; Stephenson et al., 2006). The strong and positive relationship between women's education and maternal health care utilization is well established (Islam and Odland, 2011; Ochako et al., 2011; Zere et al., 2011). Educated women take antenatal care more times than uneducated women as well as the average number of antenatal care visits of educated women is more as compared to uneducated women in the group of women who visit for antenatal care. Knowledge and access to information about risks of pregnancy as well as about available services are increased by education. Husband's education has a significant effect on antenatal care visit (Gabrysch and Campbell, 2009). Husband's education plays an important role on the women visit or not for antenatal care as well as on the number of antenatal care visits of women who take antenatal care during pregnancy. Economic status is a reasonable factor that manipulates the antenatal care visit (Zeine et al., 2010). It is found that the use of antenatal care is related to economic status; women with middle and rich economic level are more likely to attend ANC than poor women as well as the average number of antenatal care visits of women with rich economic level is more than poor women who visit for antenatal care during pregnancy. It is also found that the source of drinking water has significant effect on the women visit or not for antenatal care as well as on the number of antenatal care visits of women who take antenatal care during pregnancy. Women's freedom of taking decision on their health care is also related to the number of antenatal care visits of women who visit for antenatal care. Mass media is the major source of information on risks of pregnancy as well as ANC services (Sanda, 2014). The average number of antenatal care visits of women who have access to mass media is more than the women who have no access to mass media.

The findings of this study suggest that to take suitable public policy by arranging awareness programs on antenatal care among the people specially the women belonging to poor wealth index, to ensure minimum secondary education for both male and female. Better development related to maternal health care in rural area, women's freedom to take decision about their health care and women's involvement to mass media are obligatory to ensure the utilization of antenatal care during pregnancy.

REFERENCES

Ahmed, F.U. and Das, A.M. (1992). Beneficial effects. Three ANC visits might be the divergent point in lowering low birth weight babies, Bangladesh. *Integration,* 33: 50–53.

Anandalakshmy, P.N., Talwar, P.P. and Buckshee, K. (1993). Demographic, socio-economic and medical factors affecting maternal mortality- An Indian experience. *The Journal of Family Welfare*, 1-4.

Chipeta, M.G., Ngwira, B.M., Simoonga, C. and Kazembe L. N. (2014). Zero adjusted models with applications to analyzing helminths count data. *BMC Research Notes 2014*, 7: 856.

Fawcus, S., Mbizvo, M. and Lindmark, G.(1996). A community-based investigation of avoidable factors for maternal mortality in Zimbabwe. *Studies in Family Planning*, Vol. 27, No. 6, 319-327.

Gabrysch, S. and Campbell, O. (2009). Still too far to walk: Literature review of the determinants of delivery service use. *BMC Pregnancy Childbirth.* 9(1): 34.

Islam, M.R. and Odland, J.O. (2011). Determinants of antenatal and postnatal care visits among indigenous people in Bangladesh:a study of Mru Community. *Rural and Remote Health (internet)*, 11: 1672.

Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*, Vol.-34, 1-14.

Liu, W. S. and Cela, J. (2008). Count data models in SAS®. *SAS Global Forum, Statistics and Data Analysis*, paper 371.

Moineddin, R., Meaney, C., Agha, M., Zagorski, B. and Glazier, R.H. (2011). Modeling factors influencing the demand for emergency department services in Ontario: A comparison of methods. *BMC Emergency Medicine*, 11: 13.

Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics*, Vol.-33, 341-365.

Ochako, R., Fosto, J.C., Ikamari, L. and Khasakhala, A. (2011). Utilization of maternal health services among young women in Kenya: Insights from the Kenya Demographic and Health Survey, 2003. *BMC Pregnancy Childbirth*, 11: 1.

Sanda, H.U. (2014). Media awareness and utilization of antenatal care services by pregnant women in Kano State- Nigeria. *Journal of Social Science Studies, ISSN 2329-9150*, Vol. 1, No.2.

Stephenson, R, Baschieri, A, Clements, S, Hennink, M, Madise, N (2006). Contextual influences on the use of health facilities for childbirth in Africa. *America Journal of Public Health.* 2006, 96(1): 84-93.

Vuong, Q.H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, Vol.-57(2): 307–333.

World Health Organization (WHO) (2002). *Reducing Risks, Promoting Health Life.* Geneva: WHO.

World Health Organization (WHO) (2003). *Antenatal Care in Developing Countries:* Promises, Achievements and Missed Opportunities: An Analysis of Trends, *Levels and Differentials, 1990-2001.* Geneva: WHO.

World Health Organization (WHO) (2007). *Standards for Maternal and Neonatal Care*. Geneva, Switzerland: WHO.

World Health Organization (WHO) (2012). *Handbook for Guideline Development*. Geneva: WHO.

Yau, K., Wang, K. and Lee, A. (2003). Zero-inflated negative binomial mixed regression modeling of over dispersed count data with extra zeros. *Biometrical Journal*, Vol.- 45, 437-452.

Zeileis, A., Kleiber, C. and Jackman, S. (2008). Regression models for count data in R. *Journal of Statistical Software,* Volume 27, Issue 8.

Zeine, A., Mirkuzie, W. and Shimeles, O. (2010). Factors influencing antenatal care service utilization in Hadiya Zone. *Ethiopia Journal of Health Science*, Vol.- 20(2): 75-82.

Zere, E., Oluwole, D., Kirigia, J.M., Mwikisa, C.N. and Mbeeli, T. (2011). Inequities in skilled attendence at birth in Namibia: A decomposition analysis. *BMC Pregnancy Childbirth*, 11: 34.