

<http://www.pjbs.org>

PJBS

ISSN 1028-8880

Pakistan Journal of Biological Sciences

ANSI*net*

Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Length of Hospital Stay at Arak (Central Iran) Maternity Clinics Using Proposed Zero-Inflated Negative Binomial Modeling

¹M. Rafiei, ²S.M.T. Ayatollahi and ³J. Behboodian

¹Department of Biostatistics and Health, Arak University of Medical Sciences, Arak, Iran

²Department of Biostatistics and Epidemiology, School of Public Health,

Shiraz University of Medical Sciences, Shiraz, Iran

³Department of Statistics, Shiraz University, Shiraz, Iran

The objective of the present article is to suitably model the hospitalization time of the mother and compare the different models. An observational and cross-sectional study was done with a randomized sample of 1600 mothers admitted for delivery at Arak maternity clinics. The hospitalization time was regarded as dependent variable; mother's age and its square, mother's job, having abnormal child, ordinal pregnancy or delivery and its square, number of abortions and its square, number of present children and its square, mother's residency, type of delivery, twice and triplets, all were considered as independent variables. Advanced recent methods of countable data modeling were used. An innovative method was introduced for the data analysis. The modeling of mother's hospitalization time was shown to be the negative binomial model. It was a suitable model due to unequal variance and means of dependent variables for mother's hospitalization time. Having abnormal child, type of delivery (NVD, C and S) and twice delivery were significant variables in this model. More specific models (Zero-truncated Poisson and negative binomial) were shown to be more suitable for the age and its square, having an abnormal child, type of delivery, delivery of twice or triplet which was significant variables in determining mother's hospitalization time. A suitable statistical model for determination and modeling of mothers' hospitalization time was achieved with a simple change of these times. This model included more variables with higher specificity.

Key words: Post delivery mother's hospitalization time, countable data models, Poisson and negative binomial regression

INTRODUCTION

Modeling is an essential tool in the explanation of health and medical phenomena in applied and medical statistics. Modeling the length of hospital stay is useful to find different factors influencing different hospitalization times. Obstetrical delivery is the most frequent cause of hospitalization in many countries. However, the issue of appropriate Length Of Stay (LOS) after delivery is complex and hotly debated (Udom and Betley, 1998). Research has shown that early discharge after delivery not only has positive effects in women's physical and emotional health, but also is a facility in terms of economical considerations (Shorten, 1995). Patient's length of stay is used as a measure of hospital efficiency. It was also considered as a reasonable proxy of resource consumption (Solomon, 1996; Wang *et al.*, 2002). From the health service management perspective, it is

imperative to assess the relative performance of hospitals by modeling the distribution of maternity LOS. Accounting for distributional characteristics could assist in the preparation of prescriptive policies for more efficient utilization of resources, but the heterogeneity of LOS within DRG poses a problem for statistical analysis (Marazzi *et al.*, 1996).

Hospitalization time has positive skewness and the experimental distribution also has the same direction (Silberbach *et al.*, 1996; Wolfe *et al.*, 1995). Hospitalization modeling has been done by logarithmic change in many studies to adjoin to the normal distribution. Multiple regression methods were used to mentioned variables modeling (Silberbach *et al.*, 1996; Wolfe *et al.*, 1995; Lave and Frank, 1998; Melfi *et al.*, 1995). The other studies have pointed to separateness of heterogeneous hospitalization rate and purposed that two or more distributions were factors in the length of stay

distribution (Bender and McGuire, 1995; Palmer and Aisbett, 1996). These methods are not completely valid. It is more advisable to use countable variables like Poisson or negative binomial distribution for hospitalization data (Wang *et al.*, 2002; Cameron and Trivedi, 1998; Zelterman, 2004). Poisson is used when mean and variance are equal; and negative binomial is suitable to clarify countable response variable based on independent variables while variance distribution is bigger than mean (overdispersion) (Xio and Xio, 1996; Xio and Vermura, 1999). When the countable response of variance is smaller than mean, hurdle Poisson or generalized linear model is used (Xio and Vermura, 1999; Christopher and Zorn, 1996). In these studies, the variable in actual countable data is not analyzed and noting else was mentioned. This study has been done to modeling hospitalization (in Arak University of Medical Sciences Hospital, Central Province, Iran); and some demographic factors which affect the hospitalization rate by some new and advanced methods to describe the response countable variables like Zero-Truncated Poisson models (Lee *et al.*, 2003), Zero-Truncated negative binomial models, Zero-inflated Poisson model (Lee *et al.*, 2004), Zero-inflated negative binomial model (Yau *et al.*, 2003), Modified Poisson models, generalized countable model and Tobit models (Min and Agresti 2002). Also, a initiative model, which clarifies the inequality of mean and variance of response variable expensively were applied. In these models different response models were compared to clarify the length of hospitalization and efficacy of the mentioned initiative model.

MATERIALS AND METHODS

This is an observational, cross-sectional study, which was done randomly on a group of mothers after delivery in Taleghani Hospital in Arak Iran. This hospital is the only educational center affiliated to Arak University of Medical Sciences. Women who refer to this hospital are from the people of different social stratum. Sample size was considered to 1600 women. Then by random sampling, 1600 mothers who referred to Taleghani Hospital were studied in 2004. A questioner was completed by mothers which comprises the length of stay as response variable and independent variables include mother's age and its square, mother's job (1 = employed, 0 = housewife), abnormal newborn (1 = yes, 0 = no), number of gravid and para and its square, mother's abortion rate and its square, number of child and its square, place of living (1 = city, 0 = village), delivery type (1 = NVD, 0 = C/S), existing twice (1 = yes, 0 = no) existing triple children (1 = yes, 0 = no). The hospitalization rate comprised the number of the days that mothers stayed in

the hospital from reception to discharge. The square of quantity variables does not appear linear in models since one of the objectives of this study was the comparison between different models to clarify hospitalization after delivery. Second order has been used in all models and there is no any bias or error. SAS and STATA8 were used for data analysis; and the written programs were used for modeling. Deviance statistics was used to study the suitability of models and comparing them with observed values of response variables. The model with less deviance is more suitable to clarify or explain the hospitalization rate based on the other independent variables (Xio and Vermura, 1999). Some incessant link functions (Gaussian, inverse Gaussian, Gamma) were used to present the supremacy-expressed models.

Zero-inflated poisson mixed regression model: Zero-Inflated Poisson (ZIP) regression is a model for counts data with excess zeros. Consider a discrete random variable Y with ZIP distribution (Johnson *et al.*, 1992):

$$\begin{aligned} \Pr(Y = 0) &= \phi + (1 - \phi)e^{-\theta} \\ \Pr(Y = y) &= (1 - \phi) \frac{e^{-\theta}\theta^y}{y!}; y = 1, 2, \dots \end{aligned}$$

Where $0 < \phi < 1$ so that it incorporates more zeros than those allowed by the Poisson. A graphical representation of this distribution is given by Bohning *et al.* (1999).

The ZIP distribution may also be regarded as a mixture of a Poisson (θ) and a degenerate component putting all its mass at zero. A plausible interpretation in the context of LOS is thus in terms of its (unobserved) two-point heterogeneity: A sub-population of patients who are scheduled to treated on a same-day basis ($LOS = 0$) and another sub-population whose members are susceptible and may stay at least one night. It can be shown that:

$$\begin{aligned} E(Y) &= (1 - \phi)\theta \\ \text{Var}(Y) &= E(Y) + E(Y)\{\theta - E(Y)\} \end{aligned}$$

So that the ZIP model incorporates the extra variation unaccounted for by the Poisson.

For independent counts Y_i , $i = 1, \dots, n$, a ZIP regression model (Lambert, 1992) was proposed to study the effects of risk factors or confounders by assuming both $\log(\theta_i)$ and $\text{logit}(\theta) = \log(\theta/(1-\theta))$ to be linear functions of some covariates. Maximum likelihood estimation of the regression coefficients is performed via the EM algorithm.

The truncated Poisson model: Let n denote the number of LOS observed. We assume that n follows a Poisson distribution with parameter q . Because the study is designed to include participants who had at least one LOS, the Poisson process among participants is truncated at zero. Thus

$$\Pr(n = y) = \frac{q^y e^{-q}}{y!(1 - e^{-q})}; y = 1, 2, \dots$$

$$\xi = E(n) = \frac{q}{1 - e^{-q}}$$

$$\text{Var}(n) = \frac{q(1 - (1 + q)e^{-q})}{(1 - e^{-q})^2} \quad (1)$$

Note that the equation (1) is the conditional probability of n given $n > 0$.

The maximum likelihood estimator (MLE) of ξ is

$$\xi = \frac{1}{N} \sum_{i=1}^N n_i$$

whereas the MLE of q is the solution of the equation

$$\frac{q}{1 - e^{-q}} = \frac{1}{N} \sum_{i=1}^N n_i$$

(Cousul and Famoye, 1992).

RESULTS

The following variables were seen among 1600 mothers who referred to Taleghani Hospital; the average of mothers' hospitalization rate was 1.54 days with 0.57 variance. The hospitalization rate is shown in Table 1. According to this table the percentages of hospitalization were 58.1, 33.2 and 8.7% for one day, two days or more than two days, respectively. The variance of response variable (the day number of stay) is less than mean (underdispersion). 1.8% mothers were employed and 93.3% were housewives. First and second groups were 28 and 1572 women, respectively. From total of newborns; 924(57.8%) were boys and 676(42.3%) were girls. Fifteen mothers (0.9%) gave birth to abnormal children. The birth gradation of child in selected cases was 2.1 with 1.31 standard deviation. The abortion mean was 1.01 newborn with 1.21 standard deviation. From total mothers, 834 (52.1%) were rural and 766(47.9%) were urban. Five hundred and eighty eight mothers (36.8%) had

Table 1: Frequency distribution of mother's length of stay (day) in selected samples from Taleghani Hospital in Arak

Length of stay (day)	Frequency	(%)	Cumulative frequency (%)
1	930	58.10	58.1
2	531	33.20	91.3
3	106	6.60	97.9
4	22	0.40	99.3
5	7	0.70	99.8
6	3	0.20	99.9
7	1	0.10	100.0
Total	1600	1.00	

The mean of length hospital stay = 1.54, The variance of length hospital stay = 0.57, Skewness coefficient = 1.87

normal vaginal delivery (NVD) and 1012(63.3%) mothers had cesarean section (C/S). Thirteen women (0.8%) delivered twice and 13 delivered triple. Since response variable is mother's hospitalization rate, Poisson distribution (the number of events which happen in a special place or time have Poisson distribution) must be used. Poisson regression was used to study response variable. In Poisson distribution mean and variance are equal use. In this study the variance of hospitalization rate was (0.57) less than the mean (1.54). This manner is defined as underdispersion of Poisson distribution in countable response variable. Since the mean and the variance of countable response variable were not equal, Poisson regression is not a suitable method to determine hospitalization based on the other independent variables. In this regard, generalized linear models or hurdle Poisson and mixture Poisson distribution were used. Table 2 demonstrates coefficients of generalized linear models and the values of the significance of these coefficients. In order to clarify mother's hospitalization rate according to other independent variables identity link function with different distribution. Deviance (suitability of used model) of negative binomial modeling has the least value among other models (102.6). To clarify mothers' hospitalization in comparison with the other models, it is clear that negative binomial model is a suitable model. Hospitalization rate is affected by delivery type, existing normal newborn and the number of newborns (twins). If this model is used with negative binomial of variance function, the deviance is reduced to 102.2, no change will appear yet no significant change will appear in the coefficient. If generalized linear models are used with logarithmic link function, negative binomial is the best model in purporting the values and the dispersion of hospitalization. Since hospitalization is distributed more than one day and these values are zero (Table 1), it is necessary to use the hurdle models, which does not include zero. The results of using Zero-truncated Poisson (ZTP) and Tobit quasi-continuous, order logit model and order Table 3 explains probit model. This table and deviance zero-truncated negative binomial

Table 2: The estimation of GLM of coefficient and statistical significances of these coefficient based on identity and logarithmic link function with different distribution of mothers' hospitalization by other variable

		Negative binomial negative binomial variance function	Gamma	Negative Binomial	Poisson	Inverse Gaussian	Gaussian
Link function	Link function	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)
Age	Identity	-0.011(0.017)	-0.020 (0.019)	-0.026 (0.066)	-0.037 (0.043)	-0.007 (0.0176)	-0.055 (0.0202)**
	Logarithmic	—	-0.023 (0.013)	-0.027 (0.045)	-0.032 (0.027)	-0.013 (0.013)	-0.037 (0.012)**
Square of age	Identity	0.000 (0.000)	0.000 (0.000)	0.000 (0.001)	0.000 (0.000)	0.000 (0.000)	0.001 (0.000)**
	Logarithmic	—	0.000 (0.000)*	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)***
Job	Identity	-0.005 (0.095)	-0.175 (0.107)	-0.156 (0.369)	-0.106 (0.229)	-0.217 (0.113)	0.027 (0.112)
	Logarithmic	—	0.078 (0.072)	-0.050 (0.249)	0.015 (0.156)	0.134 (0.077)	0.162 (0.083)
Gender of newborn Sex	Identity	-0.004 (0.026)	0.041 (0.026)	0.032 (0.093)	0.016 (0.606)	0.056 (0.025)*	-0.017 (0.030)
	Logarithmic	—	0.011 (0.019)	0.003 (0.067)	-0.109 (0.042)	0.031 (0.020)	-0.083 (0.019)***
Abnormal newborn	Identity	0.153 (0.075)*	1.110 (0.283)***	1.138 (0.884)	1.222 (0.457)**	1.011 (0.371)**	1.337 (0.154)***
	Logarithmic	—	0.520 (0.099)***	0.528 (0.307)	0.552 (0.153)***	0.491 (0.138)***	0.593 (0.049)***
Number of gravid & para	Identity	0.011 (0.120)	0.163 (0.148)	0.161 (0.522)	0.144 (0.331)	0.163 (0.143)	0.096 (0.154)
	Logarithmic	—	0.088 (0.102)	0.074 (0.349)	0.043 (0.203)	0.119 (0.110)	0.000 (0.088)
Square of number of gravid and para	Identity	-0.005 (0.012)	-0.029 (0.014)*	-0.029 (0.047)	-0.029 (0.029)	-0.278 (0.0145)	-0.028 (0.014)
	Logarithmic	—	-0.018 (0.009)	-0.017 (0.031)	-0.015 (0.020)	-0.019 (0.010)	-0.011 (0.008)
Number of abortion	Identity	-0.004 (0.113)	-0.098 (0.160)	-0.076 (0.540)	-0.042 (0.328)	-0.173 (0.169)	0.000 (0.148)
	Logarithmic	—	-0.021 (0.098)	-0.011 (0.327)	-0.006 (0.193)	-0.064 (0.118)	0.025 (0.085)
Square of number of abortion	Identity	0.004 (0.020)	0.064 (0.048)	0.053 (0.138)	0.040 (0.066)	0.109 (0.070)	0.029 (0.025)
	Logarithmic	—	0.021 (0.017)	0.019 (0.055)	0.014 (0.033)	0.038 (0.036)	0.007 (0.017)
Existing child	Identity	-0.006 (0.112)	-0.133 (0.141)	-0.129 (0.496)	-0.110 (0.031)	-0.138 (0.135)	-0.059 (0.146)
	Logarithmic	—	-0.069 (0.096)	-0.055 (0.322)	-0.026 (0.190)	-0.099 (0.104)	0.014 (0.082)
Square of existing child	Identity	0.005 (0.013)	0.036 (0.016)	0.036 (0.055)	0.035 (0.035)	0.036 (0.017)*	0.033 (0.016)*
	Logarithmic	—	0.021 (0.011)*	0.019 (0.036)	0.017 (0.022)	0.024 (0.011)*	0.013 (0.009)
Place of living	Identity	-0.008 (0.026)	-0.028 (0.026)	-0.030 (0.093)	-0.034 (0.061)	-0.024 (0.024)	-0.044 (0.030)
	Logarithmic	—	-0.022 (0.019)	-0.023 (0.067)	-0.025 (0.042)	-0.020 (0.014)	-0.028 (0.019)
Delivery type	Identity	-0.219 (0.026)***	-0.918 (0.037)***	-0.916 (0.123)***	-0.915 (0.072)***	-0.925 (0.044)***	-0.919 (0.032)***
	Logarithmic	—	-0.571 (0.021)***	-0.570 (0.069)***	-0.570 (0.043)***	-0.576 (0.023)***	-0.575 (0.020)***
Twins delivery	Identity	0.137 (0.099)	0.970 (0.242)***	0.948 (0.770)	0.887 (0.419)*	1.034 (0.321)***	0.787 (0.165)***
	Logarithmic	—	0.509 (0.106)***	0.484 (0.340)	0.418 (0.187)*	0.589 (0.142)***	0.326 (0.076)***
Triple delivery	Identity	-0.008 (0.118)	-0.065 (0.213)	-0.063 (0.696)	-0.062 (0.403)	-0.073 (0.255)	-0.065 (0.168)
	Logarithmic	—	-0.023 (0.108)	-0.024 (0.351)	-0.027 (0.203)	-0.023 (0.129)	-0.036 (0.084)
Constant	Identity	-0.260 (0.275)	2.374 (0.309)***	2.430 (1.089)*	2.539 (0.699)***	2.267 (0.297)***	2.69 (0.333)***
	Logarithmic	—	1.029 (0.217)***	1.068 (0.741)	1.100 (0.454)*	0.929 (0.225)***	1.058 (0.207)***
Deviance	Identity	102.2	166.5	102.6	284.8	106.7	543.13
	Logarithmic	—	166.2	102.2	282.3	106.9	534.87

*p-value<0.05, **p-value<0.01, ***p-value<0.001

Table 3: The estimation of truncated countable models coefficient in ZTP and ZTNB, quasi-continuous models (Tobit model, Order Probit model, Order Logit model) and statistical significance of these coefficient to study mothers' Hospitalization on the other variables

	Order probit model	Order logit model	Tobit model	ZTNB	ZTP
Model variables	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)
Age	-0.091 (0.044)*	-0.198 (0.078)*	-0.088 (0.046)	-0.013 (0.087)**	-0.058 (0.022)**
Square of age	0.001 (0.000)*	0.004 (0.001)**	0.002 (0.000)*	0.000(0.000)***	0.001 (0.000)**
Job	-0.098 (0.239)	-0.116 (0.474)	-0.127 (0.000)	0.139(0.118)	0.030 (0.124)
Newborn sex	-0.008 (0.069)	-0.096 (0.122)	0.036 (0.068)	-0.071(0.241)	-0.019 (0.033)
Abnormal newborn	1.296 (0.294)***	2.899 (0.661)***	1.577 (0.306)***	0.991(0.089)***	1.353 (0.162)***
Number of gravid and para	0.221 (0.333)	0.349 (0.574)	0.226 (0.333)	0.181(0.192)	0.101 (0.168)
Square of number of gravid and para	-0.053 (0.031)	-0.090 (0.051)	-0.050 (0.031)	-0.019(0.011)	-0.029 (0.015)
Number of abortion	-0.002 (0.315)	0.030 (0.540)	-0.051 (0.313)	-0.064(0.011)***	0.000 (0.160)
Square of number of abortion	0.054 (0.052)	0.099 (0.086)	0.057 (0.051)	0.418(0.852)	0.030 (0.027)
Existing child	-0.187 (0.312)	-0.361 (0.538)	-0.194 (0.310)	-0.223(0.312)	-0.062 (0.158)
Square of existing child	0.067 (0.035)	0.119 (0.059)*	0.063 (0.035)	0.029(0.024)	0.034 (0.018)
Place of living	-0.074 (0.067)	-0.123 (0.123)	-0.062 (0.068)	-0.020(0.041)	-0.048 (0.033)
Delivery type	-1.874 (0.076)***	-3.405 (0.145)***	-1.920 (0.078)***	-0.576(0.078)***	-0.965 (0.035)***
Twins delivery	1.515 (0.306)***	2.505 (0.648)***	1.395 (0.317)***	0.589(0.124)***	0.818 (0.173)***
Triple delivery	-0.025 (0.325)	-0.169 (0.536)	0.047 (0.316)	0.023(0.011)*	-0.063 (0.176)
Constant	—	—	3.090 (0.748)***	0.929(0.219)***	2.731 (0.364)***
Deviance	157.2	179.7	93.7	89.9	90.4

*p-value<0.05, **p-value<0.01, ***p-value<0.001

Table 4: The estimation of countable models coefficient with ZTP and ZTNB and statistical significance of these coefficient to study mothers' Hospitalization based on the other variables

Model variables	ZIP	ZINB
	Coefficient (SE)	Coefficient (SE)
Age	-0.005 (0.002)*	-0.072 (0.031)**
Square of age	0.000 (0.000)*	0.001 (0.000)**
Job	0.162 (0.521)	0.266 (0.302)
Newborn sex	-0.166 (0.084)*	-0.158 (0.098)
Abnormal newborn	0.934 (0.178)***	0.886 (0.157)***
Number of gravid and para	-0.021 (0.874)	0.071 (0.497)
Square of number of gravid and para	-0.019 (0.452)	-0.029 (0.510)
Number of abortion	0.035 (0.564)	0.026 (0.741)
Square of number of abortion	0.024 (0.810)	0.027 (0.815)
Existing child	0.420 (0.761)	0.005 (0.189)
Square of existing child	0.019 (0.214)	0.024 (0.679)
Place of living	-0.076 (0.687)	-0.062 (0.861)
Delivery type	-1.013 (0.154)	-1.022 (0.142)
Twins delivery	0.671 (0.288)**	0.697 (0.283)***
Triple delivery	0.121 (0.547)	0.059 (0.025)**
Constant	0.090 (0.009)***	0.635 (0.007)***
Deviance	84.4	76.1

*p-value<0.05, **p-value<0.01, ***p-value<0.001

are used as the best method and variables such as age and its square, abnormal child, delivery type, twins and triple newborns as significant variables in mothers hospitalization. Mother hospitalization will have zero observation if number one subtracts from every existing value while a new variable exists ($LOS^* = LOS - 1$). Since by changing this variable, zero day hospitalization has great percentage (58.1%), one parameter response distribution like Poisson or negative binomial are is not a suitable model to clarify mothers' length of stay. It is better to use ZIP and ZINB to describe zero observation with more percentage. Table 4 shows the result of using these models to estimate the hospitalization rate. Considering this table, age square, abnormal existing child, delivery type, twins and triple delivery variables are effective factors in mothers' hospitalization.

DISCUSSION

Countable response variables are used for modeling and clarifying the length of stay, because they are countable response variables in comparison with the traditional models such as multiple linear regressions. We will neglect discreteness the length of stay in hospital; so, estimation is not valid (Cameron and Trivedi, 1998). For analyzing countable variable data and its modeling based on the other independent variable, Poisson and negative binomial distribution regarded as standard models. Mean and variance are equal in Poisson regression. Mixture modeling is used if variance is more than mean. Special modeling is used as explained in this article and if variance is more than mean, mixture modeling should be used. A few studies have been carried out in this condition for the

purpose of comparison; and every study defined one model like Katz or Geck, all the models have goodness-of-fit in special situations (Xio and Vermura, 1996; Xio and Vermura, 1999; Cousul and Famoye, 1992; Cameron and Johansson, 1997). In this study the mean of hospitalization was 1.54 and the variance was 0.57. The range of the length of stay was 1-7 days in this study. In Sweden, it was 1-3 days (Persson and Dykes, 2002), in England 6-48 h (Winterburn and Fraser, 2000), in Australia 1-4 days (Rice *et al.*, 2000). The above distribution has right skewness (skewness coefficient 1.87). Therefore, the common method for modeling and determining effective factors on hospitalization based on the other variables is by generalized linear modeling. In addition to Zero-Truncated Negative Binomial (ZTNB), ZTP, a special model has been studied by changing the variable of hospitalization in this study. The achieved results by GLM models showed that; however, deviance model is less than other models; negative binomial and regression, which emerged from it, are the best models to describe hospitalization rate. In this model, the abnormal child has direct effect on hospitalization; and giving birth to an abnormal child results in hospitalization for 3.1 days after delivery. Delivery type has direct effect on hospitalization; and C/S will increase it to 2.5 days. Twins delivery will increase it to 2 days. If we use truncated countable model in ZTP, ZTNB and Tobit models (Table 3), the truncated model in ZTNB is more suitable than the other models (deviance model is less than the others). Using this model, in addition to negative binomial, age variable and its square and triple delivery are effective factors in hospitalization. Mothers with less age had more hospitalization rate and triple delivery has increased it to 1.02 days. By using goodness-of-fit, this model is a significant model to clarify the length of stay in hospital. The high efficacy of this model has been mentioned in modeling of hospitalization rate in the previous studies (Christopher and Zorn, 1996; Lee *et al.*, 2003; Rice *et al.*, 2000; Dietz and Boehning, 2000; Xio and Aickin, 1997). At the end to sum up, defined model (subtracting 1 from the length of stay) was used to clarify mothers' hospitalization. Whenever this change of variable has been done for response variable, zero day hospitalization shows relative frequency (58.1%); so, it is possible to use zero inflated Poisson and zero inflated negative binomials. ZINB is a better model when models are used for data and it is significant in goodness-of-fit. By using this model (ZINB), age and its square, abnormal newborn, delivery type, twin's delivery and triple delivery variables are effective factors in the length of stay. +, - Signs show direct and reverse effects (Table 4). Appropriateness of ZIP, ZINB in description and clarification of zero inflated

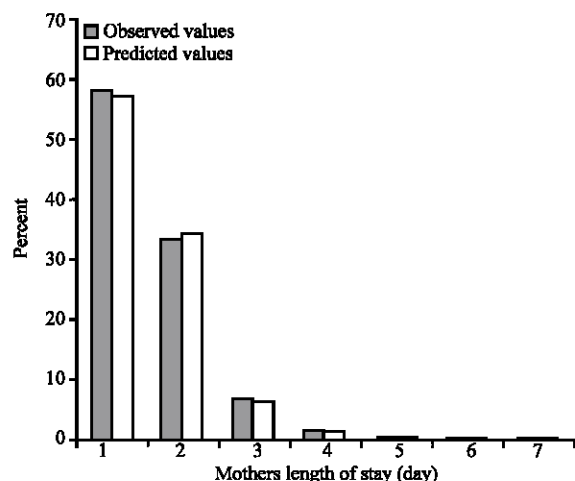


Fig. 1: Relative frequency distribution of patients length of stay based on observed and predicted values of zero-inflated negative binomial modeling to purport its appropriateness

observation in medical and health studies has been expressed (Cameron and Trivedi, 1998; Christopher and Zorn, 1996; Lee *et al.*, 2003; Lee *et al.*, 2004; Yau *et al.*, 2003; Min and Agresti, 2002; Johnson *et al.*, 1992; Bohning *et al.*, 1999; Lambert, 1992; Cousul and Famoye, 1992; Cameron and Johansson, 1997; Persson and Dykes, 2002; Dietz and Bohning, 2000). Therefore, when hurdle Poisson was defined for observation (innovative model), zero-inflated negative binomial was a suitable model based on the other variables. So, when we use a simple variable, Poisson and ZINB models might be used instead of truncated Poisson or negative binomial models since their estimations are complicated. Previous studies have pointed to unknown factors in increasing or decreasing the hospitalization (Persson and Dykes, 2002). Figure 1 shows prediction of the length of stay and observed one by using independent variables in the proposed ZINB model. According to this Fig. 1, suitability of the innovative model to show hospitalization based on compared with the other variables is clear because there is a very little difference between estimated and observed values. At the end, we recommend the applied suggestion models for countable response variables and their modeling based on independent variable in health and medical studies, especially whenever response variable is hurdle Poisson (the variance of countable response variable is less than the mean of this variable, zero-inflated model, lack of zero) in the observation.

REFERENCES

- Bender, J.A. and T.E. McGuire, 1995. A Focussed Look at the L3h3 Exception Policy. Hofdijk, J. (Ed.). Proceedings of Patient Classification System/Europe, 11th Working Conference. Oslo, pp: 266-77.
- Bohning, D., F. Dietz, P. Schlattmann and L. Mendonca, 1999. The zero-inflated Poisson model and the decayed, missing and filled teeth index in dental epidemiology. *J. R. Stat. Soc. A.*, 162: 195-209.
- Cameron, C. and P. Johansson, 1997. Count data regression using series expansions: With applications. *J. Applied Econometrics*, 12: 203-23.
- Cameron, C. and P. Trivedi, 1998. *The Analysis of Count Data*. Cambridge University Press, New York.
- Christopher, J.W. and W. Zorn, 1996. Evaluating zero-inflated and hurdle Poisson specifications. *Midwest Political Sci. Assoc.*, 18: 1-16.
- Cousul, P.C. and F. Famoye, 1992. Generalized Poisson regression model. *Communications in Statistics: Theory Method*, 21: 89-109.
- Dietz, E. and D. Bohning, 2000. On estimation of the Poisson parameter in zero-modified Poisson models. *Computational Statistics and Data Analysis*, 34: 441-459.
- Johnson, N.L., S. Kotz and A.W. Kemp, 1992. *Univariate Discrete Distribution*. 2nd Edn., Wiley, New York.
- Lambert, D., 1992. Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*, 34: 1-14.
- Lave, J.R. and R.G. Frank, 1998. Factors affecting Medicaid patients length of stay in psychiatric units. *Health Care Financing Rev.*, 10: 57-66.
- Lee, A.H., K. Wang and K.K.W. Yau, 2003. Truncated negative binomial mixed regression modeling of ischaemic stroke hospitalizations. *Statistics in Medicine*, 22: 1129-1139.
- Lee, A.H., L. Xiang and W.K. Fung, 2004. Sensitivity of score tests for zero-inflation in count data. *Statistics in Medicine*, 23: 2757-2769.
- Marazzi, A., F. Paccaud, C. Ruffeux and C. Beguin, 1996. Fitting the distributions of length of stay by parametric models. *Medical Care*, 36: 915-927.
- Melfi, C., E. Holleman and D. Arthur, 1995. Selecting a patient characteristics index for the prediction of medical outcomes using administrative claims data. *J. Clin. Epidemiol.*, 48: 917-926.
- Min, Y. and A. Agresti, 2002. Modeling nonnegative data with clumping at zero: A survey. *J. Iranian Stat. Soc.*, 1: 7-35.

- Palmer, G. and C. Aisbett, 1996. Defining and Paying for Outliers: An Evidence-based Clarification of Conceptual Issues. Hofdijk, J. (Ed.). Proceedings of Patient Classification System. 12th Working Conference. Sydney, pp: 12-21.
- Persson, E.K. and A.K. Dykes, 2002. Parents' experience of early discharge from hospital after birth in Sweden. *Midwifery*, 18: 53-60.
- Rice, P.L., C. Naksook and L.E. Watson, 2000. The experiences of postpartum hospital stay and returning home among Thai mothers in Australia. *Midwifery*, 15: 47-57.
- Shorten, A., 1995. Obstetric early discharge versus traditional hospital stay. *Australian Health Rev.*, 18: 19-39.
- Silberbach, M., D. Shumaker and V. Menash, 1996. Predicting hospital charge and length of stay for congenital heart disease Surgery. *Am. J. Cardiol.*, 72: 985-963.
- Solomon, G.L., 1996. Length of the hospital stay for mothers and newborns. *New England J. Med.*, 334: 1134-1139.
- Udom, N.U. and C.L. Betle, 1998. Effects of maternity-stay legislation on drive-through deliveries. *Health Affairs*, 17: 208-215.
- Wang, K., K.W.Y. Kelvin and A.H. Lee, 2002. A zero-inflated Poisson mixed model to analyze diagnosis related groups with majority of same-day hospital stays. *Computer Methods and Programs. Biomedicine*, 68: 195-203.
- Winterburn, S. and R. Fraser, 2000. Does the duration of postnatal stay influence breast-feeding rates at one month in women giving birth for the first time? A randomized control trial. *J. Adv. Nursing*, 32: 1152-1157.
- Wolfe, M.W., G.S. Roubin and M. Schweiger, 1995. Length of hospital stay and complications after percutaneous transluminal coronary angioplasty. *Circulation*, 92: 311-319.
- Xio, W. and M. Xio, 1996. A mixed Poisson Model and its application to attribute testing data. *J. Microelectron and Reliability*, 36: 133-140.
- Xio, T. and M. Aickin, 1997. A truncated Poisson regression model with application to occurrence of adenomatous polyps. *Stat. Med.*, 16: 1845-1857.
- Xio, J., A. Lee and S. Vemura, 1999. Mixture distribution analysis of length of hospital stay for efficient funding. *J. Socio- Economic Planning Sci.*, 33: 39-59.
- Yau, K.K.W., K. Wang, A.H. Lee, 2003. Zero-inflated negative binomial mixed regression modeling of over-dispersed count data with extra zeros. *Biometrical J.*, 4: 437-452.
- Zelterman, D., 2004. Discrete Distributions: Applications in health. John Wiley, New York.