Gazi University Journal of Science
GU J Sci
23(2):131-136 (2010)

GUJS
www.gujs.org

Analysis of Plant Protection Studies with Excess Zeros Using Zero-Inflated and Negative Binomial Hurdle Models

Abdullah YEŞİLOVA[1♠], Yılmaz KAYA[2], Barış KAKİ[1], İsmail KASAP[3]

[1] *Yuzuncu Yil University, Faculty of Agriculture, Deparment of Biometry-Genetic, Van, Turkey*
[2] *Yuzuncu Yil University, Vocational High School, Department of Computer Programming, Van, Turkey*
[3] *Çanakkale Onsekiz Mart University, Faculty of Agriculture, Department of Plant Protection, Çanakkale, Turkey*

Received: 26.06.2009 Revised: 19.09.2009 Accepted: 21.10.2009

ABSTRACT

In this study, the analysis of data with many zeros for plant protection area was carried out by using the models of Poisson Regression (PR), negative binomial (NB) regression, zero-inflated Poisson (ZIP) regression, zero-inflated negative binomial (ZINB) regression, and negative binomial hurdle (NBH) model. As zero-inflated observations are too much in the studies, done in the plant protection area; models considering zero-inflated observations are frequently required. Mites (Acari: Tetranychidae; Stigmaeidae), the basic material of this study, can reach to quite high amounts under convenient temperature (18-32°C temperature). The fact that deviance obtained from PR model together with Pearson Chi-square and deviance goodness of statistics came about quite higher than the value of (1) represented that there was an overdispersion in data set. In the selection of appropriate regression model, Akaiki information criteria and Bayesian information criteria were used. At the end of these information criteria, ZINB regression was chosen as the best model. In ZINB model, the effects of Zetzellia mali, temperature, and periods were significant on the total P. ulmi number ($p<0.01$), while applying insecticide was insignificant ($p>0.05$).

Key Words: *Zero-inflated data; overdispersion; Negative binomial Hurdle model; Zero-inflated models.*

1. INTRODUCTION

Poisson Regression (PR) is frequently used to analyze based on count data. [1,4,8,21,25]. PR model explains the relation between the independent variables and the based on count dependent variable. In PR, link function between the linear structure of independent variables and the expected value of the dependent variable is given with the logarithmic transformation [4,16,17].

As it is known, means and variance are equal to each other in Poisson distribution. However, in application, it is not always possible to provide this equation. If variance is higher than means, it is described as over dispersion, while if variance is lower than means, it is described as underdispersion [6,20,22,25]. In data sets, generally overdispersion, but rarely underdispersion is seen. In such cases, applying PR causes biased parameter estimations [6]. When there is overdispersion in data set, it is better to use negative binomial (NB) regression model [10,14,18]. In NB regression model, parameter estimations are obtained by considering the effect that stems from overdispersion.

♠Corresponding author, e-mail: yesilova@yyu.edu.tr

Base count observations having Poisson distribution might have excessive zero than expected. In such a case, using zero-inflated Poisson (ZIP) regression model is a suitable approach to analyze the dependent variable having too much zero observations [2,5,13,15,18,25]. ZIP assumes that the population consists of two different type observations. One of them is based on count data having Poisson distribution that can have zero values, while the second one is the data having always zero values [3,23,24]. Together with this, the overdispersion mentioned above is also likely in data sets having excess zero observations. In such cases, zero-inflated negative binominal (ZINB) regression model is an alternative method that is used [10,16,19,25]. As in ZIP regression, in ZINB regression, the observations with zero data and those without zero data are modeled in different ways. Moreover, Hurdle model is used to analyze data set having excess zero values. Hurdle model is composed of two stages. The first one is, binary responses indicating positive counts (1) against zero counts (0); the second one is the period when only positive counts take place. Binary responses are modeled by using binary model [7,19]. On the other hand, positive counts are modeled by using zero-truncated models. Binary part uses logit, probit or complimentary loglog, whereas count part uses Poisson, geometric, and negative binomial [9]. Generally, for the positive counts part, Poisson hurdle or negative binomial hurdle (NBH) is used. In the models of PR, NB, ZIP, ZINB, and NBH; parameter estimations are obtained with the maximum likelihood (ML) method by using EM algorithm. In the selection of the suitable model, Akaiki information criteria (AIC) and Bayesian information (BIC) criteria can be used. The model with the lowest selection criteria is accepted as the best model [7,22].

An Example

Densities for the population of *Panonychus ulmi* Koch. (Acari: Tetranychidae) and its predator *Zetzellia mali* (Ewing) (Acari: Stigmaeidae) were monitored weekly from May 30th to 31 October 31st in 2002 and from May 1st to October 8th in 2003. The experiments were settled according to randomly completed design and two apple cultivars, golden delicious and starking delicious, were planted in mosaic in these orchards and the age of trees were ranged between 15-20 years old. For the experiments, seven trees were marked for each apple cultivar in three orchards. Ten leaves were collected randomly from periphery (1.2-2.3 m height) of the marked trees. The leaves were brought to the lab in plastic containers and stored at 4 ºC in the refrigerator. Populations of *P.ulmi* and *Zetzellia mali* on the leaves were examined under a stereo binocular microscope (magnification 40x) within three days after collection. The apple trees in orchards were sprayed two times with fluvalinate on July 2nd and July 27th in 2002. In this experiment it is aimed that zero-inflated methods are analysed in the cases when dependent variable is obtained depending on zero-inflated count (Total mite (*P. ulmi*)).

2. MEDHODS

2.1. Poisson Regression

Suppose that dependent variable y_i is distributed as Poisson. The logarithm of μ which is an mean of Poisson distribution is supposed to be a linear function of independent variables (x_i) and written as,

$$\log(\mu_i) = (x_i'\beta)$$

Poisson Regression Model with log link function can be written as

$$\Pr(y_i / \mu_i, x_i) = \exp(-\mu_i)\mu_i^{y_i} / y_i!, \text{yi=0,1,}\ldots \quad (1)$$

When the independent variables are given, likelihood function for PR model can be written as,

$$LL(\beta / y_i, x_i) = \sum_{i=1}^{n} \left[y_i x_i'\beta - \exp(x_i'\beta) - \ln y_i! \right] \quad (2)$$

In equation 2, β are unknown parameters. ML parameter estimation for β can be obtained by maximizing log likelihood function [12,20].

2.2. Negative Binomial Regression

NB regression model uses log link function between the dependent variable and independent variables. NB regression model is,

$$\Pr(Y = y_i / x_i) = \frac{\Gamma\left(y_i + \frac{1}{\alpha}\right)}{y_i!\Gamma\left(\frac{1}{\alpha}\right)} \frac{(\alpha\mu_i)^{y_i}}{(1+\alpha\mu_i)^{y_i+\frac{1}{\alpha}}} \quad \alpha > 0 \quad (3)$$

In equation 3, α is an arbitrary parameter showing overdispersion level. Log likelihood function regarding NB regression model is [9,24],

$$LL(\beta,\alpha,y) = \sum_{i=1}^{n} \left[\begin{array}{l} \frac{1}{\alpha}\log(1+\alpha\mu_i) - y_i \log\left(1+\frac{1}{\alpha\mu_i}\right) \\ +\log\Gamma\left(y_i+\frac{1}{\alpha}\right) - \log\Gamma\left(\frac{1}{\alpha}\right) - \log y_i! \end{array} \right]$$

2.3. Zero Inflated Poisson Regression

In order to explain the numbers of y_i extra zeros, ZIP regression is [13],

$$\Pr(y_i / x_i) = \begin{cases} \pi_i + (1-\pi_i)\exp(-\mu_i), & y_i = 0 \\ (1-\pi_i)\exp(-\mu_i)\mu_i^{y_i} / y_i!, & y_i > 0 \end{cases} \quad (4)$$

In equation (4), π_i represents the possibility of extra zeros' existence. Log likelihood function for independent y_i observation value can be given as [22],

$$LL=\sum_{i=1}^{n}\begin{pmatrix} I_{y_i=0}\log\left(\pi_i+(1-\pi_i)e^{-\mu_i}\right) \\ +I_{y_i>0}\log\left((1-\pi_i)\dfrac{\mu_i^{y_i}e^{-\mu_i}}{y_i!}\right)\end{pmatrix}$$

$$LL=\sum_{i=1}^{n}\begin{pmatrix} I_{y_i=0}\log\left(\pi_i+(1-\pi_i)e^{-\mu_i}\right) \\ +I_{y_i>0}\begin{pmatrix}\log(1-\pi_i) \\ +y_i\log\mu_i-\mu_i-\log y_i!\end{pmatrix}\end{pmatrix} \quad (5)$$

$I(.)$, given in equation (5) is the indicator function for the specified event. After that μ_i and π_i parameters can be obtained by using link functions,

$$\log(\mu)=B\beta \quad (6)$$

and

$$\log\left(\frac{\pi}{1-\pi}\right)=G\gamma \quad (7)$$

In equations 6 and 7, B(nxp) and G(nxq) are covariate matrixes. β and γ are respectively unknown parameter vectors with px1 and qx1 dimension. Maximum likelihood estimations for β and γ can be obtained by using EM algorithm.

2.4. Zero Inflated Negative Binomial Regression

In the modeling of y_i dependent variable with many zero values, alternative regression method is ZINB. ZINB regression model is [18],

$$\Pr(y_i/x_i)=\begin{cases}\pi_i+(1-\pi_i)\left(1+\alpha\mu_i\right)^{-\alpha^{-1}}, & y_i=0 \\ (1-\pi_i)\dfrac{\Gamma\left(y_i+\frac{1}{\alpha}\right)}{y_i!\Gamma\left(\frac{1}{\alpha}\right)}\dfrac{(\alpha\mu_i)^{y_i}}{(1+\alpha\mu_i)^{y_i+\frac{1}{\alpha}}}, & y_i>0\end{cases} \quad (8)$$

In equation (8), π_i and μ_i parameters depend on covariates. $(\alpha\geq 0)$ is an overdispersion parameter. ZINB Log likelihood function for independent y_i observation value is [24],

$$LL(\mu,\alpha,\pi;y)=\sum_i\begin{pmatrix} I_{y_i=0}\log(\pi_i \\ +(1-\pi_i)\left(1+\alpha\mu_i\right)^{-\alpha^{-1}}\end{pmatrix}$$

$$+I_{y_i>0}\log\left((1-\pi_i)\frac{\Gamma\left(y_i+\frac{1}{\alpha}\right)}{y_i!\Gamma\left(\frac{1}{\alpha}\right)}\frac{(\alpha\mu_i)^{y_i}}{(1+\alpha\mu_i)^{y_i+\frac{1}{\alpha}}}\right)$$

$$=\sum_i\begin{pmatrix} I_{y_i=0}\log(\pi_i+(1-\pi_i)\left(1+\alpha\mu_i\right)^{-\alpha^{-1}} \\ +I_{y_i>0}\left(\log(1-\pi_i)-\frac{1}{\alpha}\log\left(1+\alpha\mu_i\right)\right. \\ \left.-y_i\log\left(1+\frac{1}{\alpha\mu_i}\right)+\log\Gamma\left(y_i+\frac{1}{k}\right)\right) \\ -\log\Gamma\left(\frac{1}{\alpha}\right)-\log y_i!\end{pmatrix} \quad \ldots\ldots \quad (9)$$

$I(.)$, given in equation 9 is the indicator function for the specified event. The model description suggested by Lambert (1992) can be given as,

$$\log(\mu)=X\beta \text{ and } \log\left(\frac{\pi}{1-\pi}\right)=G\gamma$$

Here, X(nxp) and G(nxq) covariate matrixes, β and γ are respectively unknown parameter vectors with px1 and qx1 dimension. Maximum likelihood estimations for β, α and γ can be obtained by using EM algorithm.

2.5. Negative Binomial Hurdle Model

In Negative Binomial Hurdle, binomial probability model determining the zero or non zero results of based on count dependent variable and truncated count model based on positive count truncated count model are conjoined using the following log-likelihood [9],

$$L=\ln\left(f(0)\right)+\left\{\ln\left[1-f(0)\right]+\ln P(j)\right\} \quad (10)$$

In equation (10), $f(0)$ represents the probability of the binary part and $p(j)$ represents the probability of positive count. In the case where logit model is used, the probability of zero is,

$$f(0)=P(y=0;x)=1/\left(1+\exp(xb1)\right)$$

and

1- $f(0)$ is,

$$\exp(xb1)/\left(1+\exp(xb1)\right)$$

The log likelihood function for both parts of negative binomial Hurdle Model can be written as,

$$\text{L=cond}\left\{y==0,\ln\left(1/1-\exp(xb1)\right),\right.$$

$$\ln\left(\exp(xb1)/\left(1+\exp(xb1)\right)\right)$$

$$\left.+y*\ln\left(\exp(xb)/\left(1+\exp(xb)\right)\right)\right)$$

$$-\ln\left(1+\exp\left(xb\right)\right)/\alpha+\ln\Gamma\left(y+1/\alpha\right)$$

$$-\ln\Gamma\left(y+1\right)-\ln\Gamma\left(1/\alpha\right)$$

$$-\ln\left(1-\left(1+\exp\left(xb\right)\right)\left(-1/\alpha\right)\right)\}$$

2.6. Model Selection

Akaiki Information Criteria and Bayesian Information Criteria are goodness of criteria used for model selection. Many Monte-Carlo simulations indicate that BIC and AIC selection criteria need to be used together [7,22]. The model with the lowest selection criteria is accepted as the best model. Generally, it is described as following,

$$AIC=-2LL+2r \quad . \quad . \quad \textbf{(11)}$$

and

$$BIC=-2LL+r\ln\left(n\right) \quad (12)$$

In equations (11) and (12), LL indicates log likelihood value, *r* indicates parameter number and n indicates sample size.

3. RESULTS

Descriptive statistics for the variables of total mite, feeding (*Zetzellia mali*), and temperature used in the present study are given in Table 1. The 2079 observation values belonging to each variable were used in the study. While the smallest value for total mite was 0, the highest value was detected as 1858. The insecticide applied and non-applied were considered as 1 and 0, respectively. In the year 2002, the 17 periods and year 2003, the 13 periods were observed. Here, the periods were denoted as weeks.

Table 1. Descriptive statistics for variables.

Variables	N	Mean	Min	Max
Total mite	2079	31.87	0	1858
Zetzellia mali	2079	2.842	0	81
Temperature	2079	19.06	12.3	23.1

The 1492 (71.62%) of the 2079 observed values used in the study were zero valued. Total mite numbers given as following in Figure 1. The distribution of the data was skewed to right because of excess zeros.

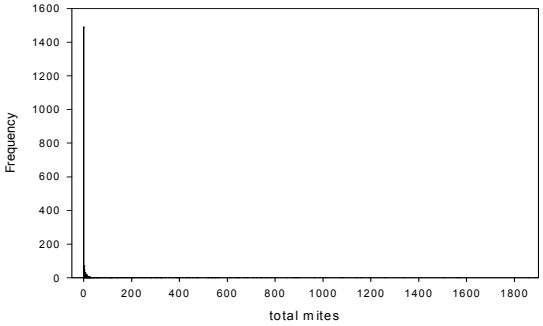

Figure 1. Frequency distribution of the total mite numbers.

In this study, the necessary analyses were done by using the Stata 10 and R statistical software programs. Total mite (*P. ulmi*) was included to the model as dependent variable, while *Zetzellia mali*, applying insecticide, temperature, and periods were integrated as independent variables to the model. In PR analyses, deviance and Pearson Chi-square goodness of statistics indicating over dispersion was obtained as 145.475 and 366.759, respectively. Being higher than (1) of the mentioned goodness of statistics represents that there was an overdispersion in data set. AIC and BIC selection criteria for the models of PR, NB, ZIP, ZINB, and NBH are given in Table 2. The model selection criteria given in Table 2 found extremely different from each other. The excess of zero observations in the data set and the great difference among observations can be indicated as the reason of overdispersion. It was found out that ZIP selection criteria were low as to PR and ZINB selection criteria were low as to NB. The model with the smallest AIC and BIC was ZINB regression. Therefore, ZINB model shown in Table 2 with bold letters was chosen as the best model.

Table 2. Model selection criteria for PR, NB, ZIP, ZINB and NBH.

Models	Log-likelihood	AIC	BIC
PR	-127073.940	254155.880	25178.438
ZIP	-50445.780	100899.560	100922.18
NB	-3928.284	7864.568	7887.126
ZINB	**-2812.363**	**5632.726**	**5655.284**
NBH	-3858.000	7724.000	7746.000

ML parameter estimations and standard errors obtained for the regression models of PR, NB, ZIP, ZINB, and NBH are given in Table 3. According to Table 3, in PR, NB and ZIP models, the effects of *Zetzellia mali*, applying insecticide, temperature, and periods were significant on the total *P. ulmi* number ($p<0.01$). In ZINB and NBH models, the effects of *Zetzellia mali*, temperature, and periods were significant on the total *P. ulmi* number ($p<0.01$), while applying insecticide was insignificant ($p>0.05$).

Table 3. Parameter estimations and standard errors for the models of PR, NB, ZIP, ZINB and NBH.

	PR	NB	ZIP	ZINB	NBH
Parameters	Estimation (stdt. error)	Estimation (stdt. error)	Estimation (stdt. error)	Estimation (stdt. error)	Estimation (stdt. error)
Intercept	3.962[1] (0.233)	3.846[1] (0.062)	7.229[1] (0.024)	7.950[1] (0.632)	7.308[1] (0.745)
Zetzellia mali	0.005[1] (0.0009)	0.091[1] (0.033)	-0.070[1] (0.002)	-0.044[1] (0.011)	-0.041[1] (0.017)
Applying insecticide	-0.603[1] (0.145)	-1.068[1] (0.337)	-0.468[1] (0.147)	0.498 (0.229)	0.033 (0.0.355)
temperature	0.122[1] (0.002)	0.088[1] (0.039)	-0.011[1] (0.002)	-0.122[1] (0.029)	-0.120[1] (0.0.044)
Periods	-0.215[1] (0.0023)	-0.103[1] (0.007)	-0.154[1] (0.001)	-0.077[1] (0.005)	-0.089[1] (0.008)

[1]p<0.01

4. DISCUSSION

Goodness of statistics (deviance and Pearson Chi-square), determining whether regression methods such as Poisson and logistic are applicable are very essential [20]. Generally, it is expected that both goodness of statistics be equal to 1 with regard to the accuracy of applied regression method [20]. In the study, values belonging to both of goodness of statistics were obtained as quite high (145.475 and 366.759, respectively). Besides, as the overdispersion had a big effect, in four different regression models goodness of criteria and parameter estimation values were obtained differently in each model in the study. Thus, as overdispersion had a great effect, the PR goodness of criteria given on Table 2 was obtained higher than the other regression models. Some reasons of the overdispersion can be explained as the use of wrong link function, difference of observations, lack of important terms that need to be in the model, and small sample size [22].

In cases which zero valued observations are not much in dependent variables obtained from based on count data, the use of PR and NB regression models are more applicable. If in such data set, there is overdispersion, NB regression model is preferred to PR model [1,4,20,21]. In Standard Poisson regression, overdispersion has a small effect on parameter estimations, but it causes wrong estimations of standard errors [6].

In general, the regression model which has the smallest AIC and BIC values is regarded as the best model [7,22]. When model selection criteria are considered, ZINB model was determined as the best applicable model. Computed Voung test statistical values demonstrated that ZIP model was more effective in preventing overdispersion as to PR and ZINB model was more effective as to NB. Besides, at the end of likelihood ratio test, it was seen that ZINB model gave better results than ZIP model. In point of obtained results, goodness of criteria, the Voung statistics, and likelihood ratio test were in parallel with each other. The fact that PR model selection criteria are much bigger than other regression models indicates the extent of overdispersion.

Hurdle models zero-inflated data by using a different approach. Hurdle model resembles ZIP regression model. Both models combine binominal likelihood with Poisson distribution. However, Hurdle model differentiates non zero inflated values from zero values by modeling with truncated distribution [7]. In this study, truncated negative binomial was used for non zero y_i.

Since the study was conducted in live materials, evaluating the zero observations in a different way (ZIP, ZINB, and NBH) is important for the credibility of the study. Mites, the basic material of this study, can reach to quite high amounts under convenient temperature (18-32°C temperature). Yet, their reproduction slows down in rates of these conditions, and moreover their reproduction stops and they start dying in temperatures higher than 35°C and lower than 15°C. Zero observations emerge in periods when mite reproduction is the lowest and deaths are the highest.

The 1492 (71.62%) of the 2079 observation values in this study are zero valued. Therefore, evaluating the zero data in a different way is important in terms of population development. Compared to other regression types, in hurdle and ZINB regression types, the effect of applying insecticide alone were found insignificant. Mean mite number in insecticide applied group was 24.95, while the number was 28.82 in the group that no insecticide was applied. In terms of mean mite number,

the values of insecticide applied and non-applied groups were similar. Predator mite *Zetzellia mali* was feeding densely on *P. ulmi* and may suppress their population. *Zetzellia mali*, observed in the test garden is particularly available in pesticided gardens and throughout the weeks it increases its population while feeds itself from *P. ulmi* eggs and thus may suppress their population [11].

REFERENCES

[1] Agresti, A., "Categorical Data Analysis", New Jersey, Canada, ***John and Wiley & Sons Incorporation,*** (1997).

[2] Böhning, D., "Zero- Inflated Poisson Models and C. A. MAN. A Tutorial Collection of Evidence", ***Biometrical Journal***, 40(7): 833-843 (1998).

[3] Böhning, D., Dietz, E., Schlattmann, P., "The Zero-Inflated Poisson Model and the Decayed, Missing and Filled Teeth Index in Dental Epidemiology", ***Journal of Royal Statistical Society A***, 162: 195–209 (1999).

[4] Cameron, A.C., Trivedi, P.K., "Regression Analysis of Count Data", New York, ***Cambridge University Pres,*** (1998).

[5] Cheung, Y.B., "Zero-Inflated Models for Regression Analysis of Count Data", ***A Study of Growth and Development. Statistics in Medicine***, 21: 1461-1469 (2002).

[6] Cox, R., "Some Remarks on Overdispersion", ***Biometrika***, 70: 269-274 (1983).

[7] Dalrymple, M.L., Hudson, I.L., Ford, R.P.K., **"**Finite Mixture, Zero-Inflated Poisson and Hurdle Models with Application to SIDS", ***Computational Statistics & Data Analysis***, 41: 491-504 (2003).

[8] Frome, E.D., Kutner, M.H., Beauchamp, J.J. "Regression Analysis of Poisson- Distributed Data", ***Journal of American Statistical Association***, 68(344): 935-940 (1973).

[9] Hilbe, J.M., "Negative Binomial Regression" ***Cambridge***, UK (2007).

[10] Jansakul, N., "Fitting a Zero-inflated Negative Binomial Model via R. In Proceedings 20th International Workshop on Statistical Modelling", Sidney, Australia, 277-284 (2005).

[11] Kasap, İ., Çobanoğlu, S., "Mite (Acari) Fauna in Apple Orchards of Around The Lake Van Basin of Turkey", ***Türk. Entomol. Journal***. 31 (1): 135-149 (2007).

[12] Khoshgoftaar, T.M., Gao, K., Szabo, R.M., "Comparing Software Fault Predictions of Pure and Zero- inflated Poisson Regression Models", ***International Journal of Systems Science***, 36(11): 705-715 (2005).

[13] Lambert, D., "Zero-Inflated Poisson Regression, with an Application to Defects in Manaufacturin", ***Technometrics***, 34(1): 1-13 (1992).

[14] Lawles, J.F., "Negative Binomial and Mixed Poisson Regression", ***The Canadian Journal of Statistcs***, 15(3): 209-225 (1987).

[15] Lee, A.H., Wang, K., Yau, K.K.W., "Analysis of Zero-Inflated Poisson Data Incorporating Extent of Exposure", ***Biometrical Journal***, 43(8):963-975 (2001).

[16] Long, J.S., Freese, J., "Regression Models for Categorical Depentent Variable Using Stata", ***A Stata Pres Publication***, USA, (2006).

[17] McCullagh, P., Nelder, J.A., **"**Generalized Linear Models"**.** Second Edition, ***Chapmann and Hall***, London, (1989).

[18] Ridout, M., Hinde, J., Demetrio, C.G.B., "A Score Test for a Zero-Inflated Poisson Regression Model Against Zero-Inflated Negative Binomial Alteratves". ***Biometrics***, 57: 219-233 (2001).

[19] Rose, C.E, Martin, S.W., Wannemuehler, K.A., Plikaytis, B.D., "On the of Zero-inflated and Hurdle Models for Medelling Vaccine Adverse event Count Data", ***Journal of Biopharmaceutical Statistics***, 16: 463-481(2006).

[20] SAS. SAS/Stat. Software, Hangen and Enhanced, USA: SAS, Institute. Incorporation (2007).

[21] Stokes, M.E., Davis, C.S., Koch, G.G., "Categorical Data Analysis Using the SAS System", USA; ***John and Wiley & Sons Incorporation*** (2000).

[22] Wang, P., Puterman, M.L., Cockburn, I.M., Le, N., "Mixed Poisson Regression Models with Covariate Dependent Rates", ***Biometrics,*** 52: 381-400 (1996).

[23] Yau, K.K.W., Lee, A.H., "Zero-Inflated Poisson Regression with Random Effects to Evaluate an Occupational Injury Prevention Programme", ***Statistics in Medicine***, 20: 2907-2920 (2001).

[24] Yau, Z., "Score Tests for Generalization and Zore-Inflation in Count Data Modeling", Unpublished Ph. D. Dissertation, ***University of South Caroline***, Columbia (2006).

[25] Yeşilova, A., Kaki, B., Kasap, İ., "Regression methods Used in Modelling of Dependent Variable Obtained Based on Zero-Inflated Count Data"**,*Journal of Statistical Research***, 05:1-9 (2007).