



## Modelling the number of abortions having excess zero values using zero-inflated generalized Poisson regression

\*<sup>1</sup> Abdullah Yeşilova, <sup>2</sup> Numan ÇİM, <sup>3</sup> Yıldız Bora

<sup>1,3</sup> Biometry and Genetics Unit, Faculty of Agriculture, University of Yuzuncu Yil, Van, Turkey

<sup>2</sup> Department of Obstetrics and Gynecology, Faculty of Medicine, University of Yuzuncu Yil, Van, Turkey

### Abstract

The purpose of this study was to plan for zero-inflated generalized Poisson regression (ZIGP) in the modelling of abortion data that include excess values of zero. The data were collected using the questionnaire technique. It was ascertained that 68.67% (206 observations) of the total number of abortions taken as a model-dependent variable had zero values. A fterward, ZIGP was used to model the dataset. The results of  $ZIGP(m_i, j, w_i)$ , as mean regression and zero-inflated regression, were determined in two stages. As independent variables were taken into account, it was obtained that zero-inflated data had an important effect on abortion numbers. Therefore, the zero-inflated level was found to have a statistically significant effect ( $p < 0.01$ ). It was determined that over dispersion did not have an important effect on abortion numbers ( $p > 0.05$ ). In mean regression, the effects of age, number of pregnancies and educational level were found to be statistically significant on abortion numbers ( $p < 0.01$ ).

**Keywords:** overdispersion, zero-inflated data, Poisson regression, abortion

### Introduction

In 1997, the World Health Organization (WHO) introduced a new definition of abortion by taking the weight of the pregnancy product and the pregnancy process as a base. Based on this definition, prior to the 20th pregnancy week, an event during which an embryo weighing less than 500 g, including a foetus and its joints, is partially or completely removed from the uterine cavity is considered abortion (miscarriage) (Cunningham *et al.* 1997) [5]. Abortion numbers are based on count data. Such observations include many zero values (when there is no miscarriage during pregnancy). Poisson regression is commonly used in modelling an acquired dependent variable based on count data. In Poisson distribution, mean and variance are equal. However, it is not always possible to ensure this equality. Overdispersion (variance is greater than mean) is commonly encountered in such cases (Consul and Famoye 1992; Böhning 1998; Chen *et al.* 2013; Zamani and Ismail 2014; Zhao *et al.* 2014) [4, 1, 2, 18, 19]. To eliminate the effects of overdispersion, negative binomial and generalized Poisson regressions are used (Famoye and Singh 2003; Famoye and Karan 2006; Czado *et al.* 2007) [7, 8, 6]. Furthermore, generalized Poisson regression is used in the event of underdispersion (variance is less than mean).

There might also be zero-inflated data other than overdispersion in the data set obtained from counting. In other words, a major portion of the data set might consist of zero values. Zero-inflated Poisson (Lambert 1992; Böhning 1998) [1, 9], zero-inflated negative binomial (Consul and Famoye 1992) [4] and Hurdle models (Rose and Martin 2006) [12] have been commonly used to model the effect of large number of zero values. The zero-inflated generalized Poisson (ZIGP) regression model, which is based on the generalized Poisson distribution, has also been used recently (Consul and Famoye 1992; Famoye and Karan 2006; Czado *et al.* 2007; Zamani and Ismail 2014) [4, 18, 7, 8]. However, separate regression analyses are performed on the mean, overdispersion and zero-inflated levels using ZIGP (Czado *et al.* 2007) [6]. In other words, the ZIGP regression evaluates the dataset in three levels as mean, overdispersion and zero-inflated regressions. The Akaike information criterion (AIC) and Vuong statistics are widely used in determining the proper regression model (Consul and Famoye 1992) [4].

The aim of this study is to apply zero-inflated generalized Poisson regression in analysis of abortion incidence (miscarriage) which are composed of many zero values.

### Materials and Methods

#### Materials

This study was performed on pregnant women who attended Van Yüzüncü Yıl University Training and Research Hospital, The Clinic of Obstetrics and Gynecology. A total of 300 pregnant women were participated in the study between January 1, 2011 and February 1, 2012. The data were collected using a questionnaire containing single-answer questions regarding number of abortions, age, weight, educational level (1: not literate, 2: primary school, 3: high school, 4: higher education) and socio-economic level (1: good, 2: average, 3: poor).

**Methods**

Let  $Y_i$  ( $i=1,2,\dots,n$ ) shows a zero-inflated generalized Poisson distribution with parameters  $(\mu, \phi$  and  $\omega)$ , denoted by  $Y \sim ZIGP(\mu, \phi, \omega)$ . The probability mass function for ZIGP is

$$P(Y = y | m, j, w) = 1_{\{y=0\}} \left[ w + (1-w)e^{-\frac{m}{j}} \right] + 1_{\{y>0\}} \left[ (1-w) \frac{m(m+(j-1)y)^{y-1}}{y!} j^{-y} e^{-\frac{1}{j}(m+(j-1)y)} \right] \tag{1}$$

(Czado *et al.* 2007) [6]. In equation 1,  $\omega, \phi$  and  $\mu$  are the zero-inflated, overdispersion and mean parameters, respectively. In some datasets, a constant overdispersion and/or zero-inflated parameters can be very restrictive. Therefore, the zero-inflated generalized Poisson regression is used for the case in which the overdispersion and zero-inflated parameters are not fixed (Czado *et al.* 2007) [6]. This regression model can be represented as  $Y_i \sim ZIGP(m_i, j_i, w_i)$ . In this model,  $X_i = (1, x_{i1}, x_{i2}, \dots, x_{ip})^t$ ,  $J_i = (1, j_{i1}, j_{i2}, \dots, j_{ir})^t$  and  $Z_i = (1, z_{i1}, z_{i2}, \dots, z_{iq})^t$  are the vectors of independent variables for the mean, overdispersion and zero-inflation parameters, respectively. The stages of the  $ZIGP(\mu_i, \phi_i, \omega_i)$  regression model can be written as follows.

**Random Component**

In this stage,  $\{Y_i, 1 \leq i \leq n\}$  is independent variable (abortion numbers) with ZIGP distribution, denoted by  $Y \sim ZIGP(\mu, \phi, \omega)$ .

**Systematic Component**

The linear component is defined as it is for traditional linear models. The differentiable link function describes how the expected value of  $Y_i$  is related to the linear predictor  $h_i$  (Ser *et al.* 2013; SAS, 2015). In ZIGP regression, three linear predictors  $h_i^m(b) = x_i^t b$ ,  $h_i^\phi(\alpha) = \omega_i^t \alpha$  and  $h_i^w(g) = z_i^t g$  impact the dependent variable. Here,  $b = (b_0, b_1, \dots, b_p)^t$ ,  $a = (a_0, a_1, \dots, a_r)^t$  and  $g = (g_0, g_1, \dots, g_q)^t$  are unknown regression parameters.  $X = (x_1, x_2, \dots, x_n)^t$ ,  $W = (\omega_1, \omega_2, \dots, \omega_n)^t$  and  $Z = (z_1, z_2, \dots, z_n)^t$  are called design matrices (Czado *et al.* 2007) [6].

**Parametric Link Component**

The linear predictors  $h_i^m(b)$ ,  $h_i^\phi(\alpha)$  and  $h_i^w(g)$  are related to the parameters  $\mu_i(\beta)$ ,  $\phi_i(\alpha)$  and  $\omega_i(\gamma)$ , respectively. Log link, shifted log link and logit link are used here to linearize  $h_i^m(b)$ ,  $h_i^\phi(\alpha)$  and  $h_i^w(g)$  for the mean, overdispersion and zero-inflated levels, respectively. In other words, log link, shifted log link and logit link are assumed to be linear functions of the covariates  $X_i, J_i$  and  $\omega_i$ , respectively. In the  $ZIGP(\mu_i, \phi_i, \omega_i)$  regression model, these levels can be written as follows for parametric link component.

**Mean Level**

The log link function for the mean level is

$$E(Y_i | b) = m_i(b) = E_i e^{x_i^t b} = e^{x_i^t b + \log(E_i)} > 0$$

$$\Leftrightarrow h_i^m(b) = \log(m_i(b)) - \log(E_i)$$

Where  $E_i$  indicates exposure variables for individual observation periods ( $E_i > 0 \quad \forall i$ ).

**Over Dispersion Level**

In the case of  $\varphi_i(\alpha) = 1$ , there is no overdispersion in the dependent variable. However, by using the shifted log link function in the case of overdispersion ( $\varphi_i(\alpha) > 1$ ), the overdispersion level is

$$j_i(a) = 1 + e^{j_i a} > 1$$

$$h_i^j(a) = \log(j_i(a) - 1)$$

**Zero-Inflated Level**

Here, the logit link function is used in the case where the dependent variable is equal to zero ( $Y_i = 0$ ) or greater than zero ( $Y_i > k \quad k = 1, 2, \dots, n$ ). Hence, the logit link function for the zero-inflated level is

$$w_i(g) = \frac{e^{z_i g}}{1 + e^{z_i g}} \in (0, 1)$$

$$\Leftrightarrow h_i^w(g) = \log\left(\frac{w_i(g)}{1 - w_i(g)}\right)$$

The ZIGP likelihood function for dependent variables ( $y_i$ ) can be written as follows (Famoye and Singh 2003; Famoye and Karan 2006; Czado *et al.* 2007; Zamani and Ismail 2014) [18, 8, 6].

$$l(d) = \sum_{i=1}^n I_{(y_i=0)} \left[ \log \left( e^{z_i g} + \exp \left( - \frac{E_i \cdot e^{x_i b}}{1 + e^{w_i a}} \right) \right) - \log(1 + e^{z_i g}) \right]$$

$$+ I_{(y_i>0)} \left[ - \log(1 + e^{z_i g}) + \log(E_i) + x_i b - \log(y_i!) - y_i \log(1 + e^{w_i a}) \right]$$

$$+ (y_i - 1) \log \left( E_i e^{x_i b} + e^{w_i a} y_i \right) - \frac{E_i e^{x_i b} + e^{w_i a} y_i}{1 + e^{w_i a}} \quad (2)$$

where  $(b^t, a^t, g^t)$  is the unknown regression parameters vector. The unknown parameters  $(b^t, a^t, g^t)$  are obtained using the maximum likelihood method after maximizing the likelihood function.

**Results and Discussion**

The R 2.11.1 statistical software program was used in this study. The age, weight, number of pregnancy experiences, socio-economic conditions and educational level were taken as independent variables, while abortion numbers was taken as dependent variable in the model. The descriptive statistics for age and weight were given in Table 1.

**Table 1:** Descriptive statistics for age and weight

| Variable | Mean (standard deviation) | Minimum | Maximum |
|----------|---------------------------|---------|---------|
| Age      | 26.694 (6.324)            | 16      | 48      |
| Weight   | 71.385 (8.780)            | 41      | 93      |

It was obtained that 68.67% (206) of the 300 observations regarding the abortion numbers had zero values. In this context, their distribution was positively skewed because there were many zero observations in the abortion numbers.

The AIC values for the Poisson (PR), negative binomial (NBR), zero-inflated Poisson (ZIP), zero-inflated negative binomial (ZINB), Poisson hurdle (PH), negative binomial hurdle (NBH), generalized Poisson (GP) and ZIGP regressions were given in Table 2. The AIC obtained from the  $PR(\mu_i)$  model was determined to be greater than that obtained from the other regression models. The model with the smallest AIC is defined the best model (Böhning 1998) [1]. In this case, the model with the smallest AIC was obtained to be  $ZIGP(\mu_i, \varphi, \omega_i)$ . Following this model, the model observed to have the smallest AIC value was the ZINB model. In terms of the AIC, in cases in which ZIGP regression are not used, ZINB regression usually yields better results than the

other regressions (Famoye and Karan 2006; Czado *et al.* 2007; Zamani and Ismail 2014; Zhao *et al.* 2014) [18, 19, 7, 6]. The AIC of the  $ZIGP(\mu_i, \phi, \omega)$  model was also presented in Table 2. The difference between the  $ZIGP(\mu_i, \phi, \omega)$  and  $ZIGP(\mu_i, \phi, \omega_i)$  models is that the zero-inflated parameter ( $\omega$ ) can vary. In Table 2,  $\omega$  indicates the fixed-value zero-inflated parameter, while  $\omega_i$  indicates the varying zero-inflated parameter.

**Table 2:** AIC for different regression models

| Model   | AIC      |
|---|----------|
| Poisson regression ( $PR(m_i)$ )                                      | 491.213  |
| Negative binomial regressions $NBR(\mu_i)$                            | 488.319  |
| Zero-inflated Poisson regression $ZIP(m_i, \omega)$                   | 476.621  |
| Zero-inflated Poisson regression $ZIP(\mu_i, \omega_i)$               | 471.195  |
| Zero-inflated negative binomial regression $ZINB(\mu_i, \omega)$      | 469.072  |
| Zero-inflated negative binomial regression $ZINB(\mu_i, \omega_i)$    | 466.014  |
| Poisson Hurdle regression $PH(\mu_i, \omega)$                         | 479.142  |
| Negative binomial Hurdle regression $NBH(\mu_i, \omega)$              | 477.056  |
| Generalized Poisson regression $GP(m_i, j)$                           | 489.875  |
| Zero-inflated generalized Poisson regression $ZIGP(m_i, j, \omega)$   | 472.513  |
| Zero-inflated generalized Poisson regression $ZIGP(m_i, j, \omega_i)$ | 465.028* |

\*Lowest AIC value explain the best model.

In Table 2, firstly, the event for which we include the zero-inflated parameter ( $\omega$ ) in the model, we compared the nested regression models  $PR(m_i)$  vs.  $ZIP(m_i, \omega)$ ,  $NBR(m_i)$  vs.  $ZINB(\mu_i, \omega)$  and  $GP$  vs.  $ZIGP(m_i, j, \omega)$ . Comparing  $PR(m_i)$  with  $ZIP(m_i, \omega)$ , the AIC decreased from 491.213 to 476.621 and the Vuong statistics was determined to be  $v = 5.24$ . According to the zero-inflated parameter, both the AIC and the Vuong statistics displayed that the  $ZIP(\mu_i, \omega)$  regression model is to be preferred to the  $PR(\mu_i)$  model ( $p < 0.01$ ). Comparing  $NBR(m_i)$  with  $ZINB(m_i, \omega)$ , the AIC falled from 488.319 to 469.072, and  $v = 6.87$ . With respect to the zero-inflated parameter, the  $ZINB(\mu_i, \omega)$  regression model is to be preferred to the  $NBR(m_i)$  model ( $p < 0.01$ ). Comparing  $GP$  with  $ZIGP(m_i, j, \omega)$ , the AIC decreased from 489.875 to 472.513, and  $v = 4.87$ . Both the AIC and the Vuong statistics indicated that the  $ZIGP(m_i, j, \omega)$  model is to be preferred to the  $GP$  model ( $p < 0.01$ ).

Scondly, we compared the constant zero-inflated ( $\omega$ ) and the varying zero-inflated ( $\omega_i$ ) parameters. We compared the nested models  $ZIP(m_i, \omega)$  vs.  $ZIP(m_i, \omega_i)$ ,  $ZINB(m_i, \omega)$  vs.  $ZINB(m_i, \omega_i)$ , and  $ZIGP(m_i, j, \omega)$  vs.  $ZIGP(m_i, j, \omega_i)$ . Comparing  $ZIP(m_i, \omega)$  with  $ZIP(m_i, \omega_i)$ , the AIC falled from 476.621 to 471.195, and the Vuong statistics was obtained to be  $v = 3.76$ . So, the  $ZIP(m_i, \omega_i)$  regression model is to be preferred to  $ZIP(m_i, \omega)$  model ( $p < 0.05$ ). Moreover, Both the AIC and the Vuong statistics indicated that the  $ZINB(m_i, \omega_i)$  regression model is to be preferred to the  $ZINB(m_i, \omega)$  model ( $p < 0.05$ ). Eventually, we compared the nested models  $ZIGP(m_i, j, \omega)$  and  $ZIGP(m_i, j, \omega_i)$ . The AIC decreased from 472.513 to 465.028 and  $v = 3.84$ . This difference was statistically significant ( $p < 0.05$ ). Thus, the  $ZIGP(\mu_i, \phi, \omega_i)$  regression model is to be preferred to the  $ZIGP(\mu_i, \phi, \omega)$  model.

Third, to determine whether or not the fixed overdispersion parameter ( $\phi$ ) is important in the event that it is included in the model, we compared  $PR$  vs.  $GP$ ,  $PR$  vs.  $NBR$ ,  $PH$  vs.  $NBH$ ,  $ZIP$  vs.  $ZINB$ , and  $ZIP$  vs.  $ZIGP$ . Comparing  $PR$  with  $GP$ , the AIC falled from 491.875 to 489.875, and the Vuong statistics was obtained to be 1.13. It was concluded that there was no difference between the two regression models with respect to the overdispersion parameter. Comparing  $PR$  with  $NBR$ , the AIC

failed from 491.213 to 488.319. Both the AIC and the Vuong statistics displayed no difference between the PR and NBR regression models according to the overdispersion parameter ( $p > 0.05$ ). Comparing PH with NBH, both the AIC and the Vuong statistics showed no difference between PH and NBH. Eventually, we compared ZIP vs. ZINB and ZIGP. There was no difference between these regression models ( $p > 0.05$ ). That is, these outcomes displayed that overdispersion in the abortion numbers was not important effect.

Finally, we compared  $ZIGP(\mu_i, \phi, \omega_i)$  model vs. other regression models. With respect to the zero-inflated parameter, both the AIC and the Vuong statistics displayed that the  $ZIP(\mu_i, \omega)$  regression model is to be preferred to the other zero-inflated regression models ( $p < 0.01$ ). Since the  $ZIGP(\mu_i, \phi, \omega_i)$  model was selected to be the best model, we must interpret the parameter estimations according to this model. The parameter estimations and standard errors with respect to the  $ZIGP(\mu_i, \phi, \omega_i)$  model were given in Table 3.

**Table 3:** Parameter estimation and standard error for  $ZIGP(m_i, j, w_i)$  regression model

| Independent variables                 | Mean regression      |                | Zero-inflated regression |                |
|---------------------------------------|----------------------|----------------|--------------------------|----------------|
|                                       | Parameter estimation | Standard error | Parameter estimation     | Standard error |
| Intercept                             | -0.346               | 0.678          | 0.212                    | 4.619          |
| Age                                   | -0.078               | 0.021**        | 0.023                    | 0.143          |
| Number of pregnancies                 | 0.432                | 0.044***       | -2.712                   | 0.969**        |
| Weight                                | -0.007               | 0.008          | 0.061                    | 0.059          |
| Socio-economic conditions             | -0.143               | 0.163          | -0.441                   | 1.058          |
| Education level                       | 0.658                | 0.123***       | 0.984                    | 0.869          |
| Mean range ( $M$ )                    |                      |                | (0.06, 4.12)             |                |
| Zero-inflated parameter range ( $W$ ) |                      |                | (0.0, 1.0)               |                |

\*\*\* $p < 0.001$  \*\* $p < 0.01$

In the mean regression, the effects of age ( $p < 0.01$ ), the number of pregnancy experiences ( $p < 0.001$ ) and the educational level ( $p < 0.001$ ) on the abortion numbers were statistically significant, while the effects of the socio-economic conditions and weight were not significant ( $p > 0.05$ ). Considering the variables that are important in terms of the number of abortions, it was determined that a one-unit increase in age caused the number of abortions in pregnant women as increase by more than 6.8%. Furthermore, it was obtained that a one-unit increase in the number of pregnancies (the number of previous pregnancy experiences) increased the number of abortions by 55%, whereas an increase in the educational level reduced the number of abortions by 89%.

In the zero-inflated regression, the effect of the number of pregnancy experiences on the number of abortions was found to be statistically significant ( $p < 0.01$ ), while effects of all the other independent variables were obtained not to be significant ( $p > 0.05$ ). In zero-inflated regression, it was detected that a one-unit increase in the number of pregnancies increased the number of miscarriages by 94.3%.

The mean of the number of abortions (dependent variable) was obtained to be 0.71, while the variance was obtained to be 0.82. Additionally, one of the major reasons for overdispersion is the fact that the observations vary substantially (Consul and Famoye 1992; Famoye and Singh 2003; Famoye and Karan 2006)<sup>[4, 7, 8]</sup>. However, in this study, the smallest value for the number of abortions was zero, while the greatest value was five. Both the AIC and the Vuong values indicated that overdispersion was not significant in this case ( $p > 0.05$ ).

One major characteristic of the human reproductive system is that it functions with a high loss rate. 10-12% of pregnancies were recognized as clinical losses. Both retrospective and prospective studies have suggested that ratios between 15% and 40% taking the previous data as the basis were not accurate (Salem *et al.* 1984)<sup>[13]</sup>. Abortion risk increases with the mother's age as well as parity (Warburton and Fraser 1964; Wilson *et al.* 1986)<sup>[16, 17]</sup>. According to the present evidence, the risk of loss gradually increases with the number of losses in pregnancies (Toksöz *et al.* 1991; Clifford *et al.* 1997; Poland *et al.* 1997; Polat *et al.* 2000)<sup>[15, 3, 10, 11]</sup>.

In this paper, as far as the number and the rate of abortion were concerned, the number of those who had one miscarriage (abortion) was determined to be 62 (19.87%), the number who had two miscarriages were determined to be 21 (6.73%) and those who had three and above miscarriages were obtained to be 16 (5.23%). In addition, the incidence of spontaneous abortions was ascertained to be 31.73%. In contrast to the literature (Salem *et al.*, 1984)<sup>[13]</sup>, the high incidence of abortions in this study might have been because our hospital is a third-grade hospital and there are too many high-risk pregnancy follow-ups. Besides, the fact that the increase in the maternal age leads to 6.7% rise in miscarriage frequency rates, and also the increase in the number of pregnancies (the sequence of pregnancy) causes 53% rise in number of miscarriages, and both circumstances are consistent with literature (Warburton and Fraser 1964)<sup>[16]</sup>. In two different studies, the number of surviving children was determined to be associated with the mother's age and educational background (Toksöz *et al.* 1991)<sup>[15]</sup>. In this study, conducted on women with similar fertility and educational status in our country, we ascertained that progress in the educational level reduced the number of

abortions at a rate of 87%. This situation is in accordance with the literature. We also obtained that socio-economic status had no effect on the incidence of abortion. However, we believe that improving the socio-economic status would make a positive contribution to improving nutrition, pregnancy follow-up and adverse conditions.

### Conclusion

In this study, the PR regression model has a greater AIC value than the ZIP, ZINB, PH, NBH, GP and ZIGP regressions because of the large number of zero values in the dependent variable. The regression model with the smallest AIC value was obtained to be  $ZIGP(\mu_i, \phi, \omega_i)$ . Consequently, the regression model parameter estimates and standard errors were interpreted based on this model. 68.67% (206) of abortion numbers were obtained to be zero. The zero-inflated parameter varied between 0% and 100% (table 3). Thus, the zero-inflated parameter range was detected to be quite high. Therefore, both the AIC and the Vuong statistics showed that the presence of exceedingly many zero values in the number of abortions was statistically significant in  $ZIGP(\mu_i, \phi, \omega_i)$  regression. However, the mean and the variance of the abortion number were obtained as 0.71 and 0.82, respectively. Similarity of two values suggests that there is no overdispersion in the number of abortions. Therefore, overdispersion was not considered important effect in  $ZIGP(\mu_i, \phi, \omega_i)$  regression models.

### References

1. Böhning D. Zero- inflated Poisson models and C.A.MAN: A tutorial collection of evidence. *Biometrical Journal*. 1998; 40(7):833-843.
2. Chen XD, Ying-Zi Fu YZ, Wang XR. Local influence measure of zero-inflated generalized Poisson mixture regression models. *Statistics in Medicine*. 2013; 32:1294-1312.
3. Clifford K, Rai R, Regan L. Future pregnancy outcome in unexplained recurrent first trimester miscarriage. *Human Reproduction*. 1997; 12(2):387-389.
4. Consul P, Famoye F. Generalized poisson regression model. *Comm. Statist. Theory Methods*. 1992; 21(1):89-109.
5. Cunningham FG, MacDonald PC, Gant NF, Leveno KJ, Gilstrap LC, Hankins GDV, *et al. Williams obstetrics*. 20 th ed. London: Prentice-Hall International Inc, 1997.
6. Czado C, Erhardt V, Min A, Wagner S. Dispersion and zero-inflation level applied to patent outsourcing rates Zero-inflated generalized Poisson models with regression effects on the mean. *Statistical Modelling*. 2007; 7(2):125-153.
7. Famoye F, Singh KP. On inflated generalized poisson regression models. *Advanced Applied Statistics*. 2003; 3(2):145-158.
8. Famoye F, Karan PS. Zero- inflated generalized poisson regression model with an application to domestic violence data. *Journal of Data Science*. 2006; 5(4):117-130.
9. Lambert D. Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*. 1992; 34(1):1-13.
10. Poland BJ, Miller JR, Jones DC, Trimble BK. Reproductive counseling in patients who have had a spontaneous abortion. *American Journal of Obstetrics & Gynecology*. 1997; 27(7):685-691.
11. Polat SA, Açık Y, Gürateş B. Fertility characteristics and family planning methods used by the women visiting the Elazığ state hospital family planning clinic. *Journal of Inonu University Medical Faculty*. 2000; 7(3):260-265.
12. Rose CE, Martin SW, Wannemuehler KA, Plikaytis BD. On the of Zero-inflated and Hurdle Models for Medelling Vaccine Adverse event Count Data. *Journal of Biopharmaceutical Statistics*. 2006; 16:463-481.
13. Salem HT, Ghaneimah SA, Shaaban MM. Prognostic value of biochemical tests in the assessmant of fetal outcome in threatened abortion. *Br Obstetrics & Gynecology*. 1984; 91(4):382-385.
14. Ser G, Yeşilova A, Yılmaz A. Comparing covariance structures using different optimization techniques in GLMM on some sexual behaviors of male lambs. *The Journal of Animal & Plant Sciences*. 2013; 23(6):1583-1587.
15. Toksöz P, Çelik MY, Aydınalp A, Esen İ. Diyarbakir province in the center of fertility characteristics of women ages 15-49. *Dicle Medical Journal*. 1991; 18(2):21-7.
16. Warburton D, Fraser FC. Spontaneous abortion risks in man: Data from reproductive histories collected in a medical genetics unit. *American Journal of Human Genetics*. 1964; 16(1):1-25.
17. Wilson RD, Kendrick V, Wittmann BK, McGillivray B. Spontaneous abortion and pregnancy outcome after normal first-trimester ultrasound examination. *Obstetrics & Gynecology*. 1986; 67(3):352-5.
18. Zamani H, Ismail N. Functional form for the zero-inflated generalized Poisson regression model. *Communication in Statistics-Theory and Methods*. 2014; 43(3):515-529.
19. Zhao W, Zhang R, Liu J, Lv Y. Semi varying coefficient zero-inflated generalized Poisson regression model. *Communication in Statistics-Theory and Methods*. 2014; 44(1):171-185.