Analysis of Zero-Adjusted Discrete Models

Pushpa L. Gupta
University Of Maine, Department Of Mathematics/Statistics
Orono, Maine 04469-5752, U.S.A.
pgupta@maine.maine.edu

ABSTRACT. In certain applications involving count data, it is sometimes found that zeros are observed with a frequency significantly higher (lower) than predicted by the assumed model. Examples of such applications are cited in the literature from engineering, manufacturing, economics, public health, epidemiology, psychology, sociology, political science, agriculture, road safety, species abundance, use of recreational facilities, horticulture, and criminology .In this paper, a zero-adjusted discrete model is proposed and the effect of such an adjustment is studied. In particular, an adjusted generalized Poisson distribution is studied in detail and all its parameters are estimated. Also, a score test is developed to determine whether such an adjustment is necessary. Examples, with or without the covariates, are provided to illustrate the procedure.

Key Words: Zero-adjusted model; Generalized Poisson distribution; Score test; Covariates

INTRODUCTION: Poisson model has been extensively used for the analysis of the count data. However, in Poisson model, mean-variance relationship is quite restrictive in the presence of extra zeros. It often underestimates the observed dispersion, which may be caused by extra zeros in the data. In the literature, several authors have analyzed such data sets by the use of zero-inflated Poisson, negative binomial or binomial models; see Gupta et al (1996), Lambert (1992), Bohning et al (1999), Dietz and Bohning (2000), Ridout et al (2001), Van den Broek (1995) and the references therin.In this paper, we propose a zero-adjusted discrete model and study the effect of such an adjustment in general. The failure rates and the survival function of the adjusted and the unadjusted models are compared. The relative error incurred by ignoring the adjustment is studied and is shown to be decreasing function of the count. In particular, an adjusted generalized Poisson distribution is studied in detail. And all the parameters are estimated. Also a score test is developed to see if such an adjustment is necessary. Examples, with or without the covariates, will be provided to illustrate the procedure.

Because of the page limitations, we shall only give the development of the model and the basic form of the generalized Poisson distribution.

DEVELOPMENT OF THE MODEL: Let X be a discrete random variable with mass concentrated on the non negative integers. Suppose $X = 0$ is observed with frequency significantly higher (lower) than that predicted by the assumed model. Then the adjusted random variable Y can be described as

$$P(Y=0) = f + (1-f)P(X=0)$$
$$P(Y=j) = (1-f)P(X=j),\ j = 1,2,3,...$$

In case $0 < f < 1$, the new model incorporates extra zeros than given by the original model. Such a distribution can be regarded as a mixture of two distributions, one of which is a degenerate at zero. In case $f < 0$, the new model incorporates fewer zeros than given by the original model, and the inequality $f \geq -P(X=0)/(1-P(X=0))$ should be satisfied.

GENERALIZED POISSON DISTRIBUTION: The probability mass function of the generalized Poisson distribution is given by

$$P(X=j) = \frac{(1+aj)^{j-1}}{j!}\frac{(qe^{-aq})^{j}}{e^{q}},\ j = 1,2,3,...$$
$$= 0,\ j > m \quad \text{when} \quad a < 0$$

and zero otherwise, where $q > 0, \max(-1/q, -1/m) \leq a \leq 1/q$ and m is the largest positive integer for which $q + maq > 0$ when $a < 0$; see Consul (1989).

REFERENCES

[1] Bohning, D., Dietz, E., Schlattman. P., Mendonca,L. and Kirchner,U.(1999). The zero-inflated Poisson model and the decayed, missing and filled teeth index in dental epidemiology. J.R.Statist.Soc.A,162, 195-209.

[2] Consul, P.C. (1989). Generalized Poisson Distributions. Marcel Dekker, New York.

[3] Dietz, E. and Bohning, D. (2000). On estimation of the Poisson parameter in zero-modified Poisson models. Computational Statistics and Data Analysis, 34, 441-459.

[4] Gupta, P.L. ,Gupta, R.C. and Tripathi, R.C. (1996). Analysis of zero adjusted count data. Computational Statistics and Data Analysis, 23, 207-218.

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

[6] Ridout, M., Hinde, J. and Demetrio, C.G.B. (2001). A score test for testing a zero-inflated Poisson regression model against zero-inflated negative binomial alternatives. Biometrics, 57, 219-223.

[7]Van den Broek, .J. (1995).A score test for zero inflation in a Poisson distribution. Biometrics, 51, 731-743.