POISSON-WEIGHTED EXPONENTIAL UNIVARIATE VERSION AND REGRESSION MODEL WITH APPLICATIONS

Hossein Zamani, Noriszura Ismail and Pouya Faroughi

1Hormozgan University, Iran
2,3Universiti Kebangsaan Malaysia, Malaysia

ABSTRACT
This study introduces a new two-parameter mixed Poisson distribution, namely Poisson-Weighted Exponential (P-WE), which is obtained by mixing Poisson distribution with a new class of weighted exponential distribution. The new P-WE distribution provides a more flexible alternative for modelling over dispersed count data compared to Poisson distribution. The estimation procedures of P-WE distribution via method of moments and maximum likelihood are provided. This study also introduces P-WE regression model which can be fitted to over dispersed count data with covariates. The P-WE distribution and P-WE regression model are fitted to two sets of count data.

Keywords: Mixed Poisson, Weighted Exponential, Poisson-Weighted Exponential, Count Data, Regression Model

1. INTRODUCTION

Mixed distribution can be considered as one of important approaches for obtaining a new distribution for count data in statistics and probability studies. In particular, mixed Poisson and mixed negative binomial distributions provide a more flexible alternative for modelling over dispersed count data compared to Poisson distribution. Examples of mixed Poisson and mixed negative binomial distributions are negative binomial which is a mixture of Poisson and gamma (Klugman et al., 2008; Lawless, 1997), negative binomial-Pareto (Klugman et al., 2008; Meng et al., 1999), Poisson-inverse Gaussian (Klugman et al., 2008; Tremblay, 1992), Poisson-Lindley (Sankaran, 1970; Ghitany et al., 2008), negative binomial-inverse Gaussian (Gomez-Deniz et al., 2008), negative binomial-beta (Wang, 2011) and negative binomial-lindley (Hossein and Ismail, 2010; Lord and Geedipally, 2011). Several applications of mixed Poisson distributions for fitting real data are discussed in Karlis and Xekalaki (2005).

This study introduces a new two-parameter mixed Poisson distribution, namely Poisson-Weighted Exponential (P-WE), which can be considered as an alternative for modelling over dispersed count data. The contents of this study are as follows. In section 2, we study the basic properties of the new P-WE distribution. Section 3 illustrates the estimation of parameters via method of moments and maximum likelihood, section 4 introduces the new P-WE regression model which is applicable to over dispersed count data with covariates. The application of P-WE distribution and P-WE regression model on two sets of count data are provided in section 5. Finally, several conclusions are presented in section 6.

2. MATERIALS AND METHODS

2.1. P-WE Distribution

A new class of two-parameter Weighted Exponential (WE) distribution was recently introduced by Gupta and Kundu (2009) and the p.d.f. is:
where, is the shape parameter and the scale parameter new WE distribution, which was obtained from an idea suggested by Azzalini (1985) who introduced a shape parameter for several symmetric distributions, was proposed by Gupta and Kundu (2009) who introduced a shape parameter to exponential distribution which belongs to non-symmetric distributions. In several cases, the new WE distribution provides better fit than weibull, gamma or generalized exponential distributions. The new WE distribution can also be represented as a sum of two independent exponential distributions.

Assume that the conditional random variable $X|\lambda$ is Poisson distributed with p.m.f.,:

$$Pr(X = x | \lambda) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0,1$$

And the random variable $\Lambda$ is distributed as WE with p.d.f.:

$$u(\lambda) = \frac{\alpha + 1}{\alpha} \beta e^{-\beta \lambda} (1 - e^{-\beta \lambda}), \quad \lambda > 0, \alpha, \beta > 0$$

The p.m.f. of P-WE distribution is Equation (1):

$$Pr(X = x) = \frac{(\alpha + 1) \beta}{\alpha} ((\beta + 1)^{-x} - (\alpha \beta + \beta + 1)^{-x})$$

The m.g.f. of P-WE distribution is Equation (2):

$$M_\lambda(t) = \sum_{x=0}^\infty e^{tx} \frac{(\alpha + 1) \beta}{\alpha} ((\beta + 1)^{-x} - (\alpha \beta + \beta + 1)^{-x})$$

Which is equivalent to the sum of two geometric distributions, each with probability $\frac{\beta}{\beta + 1}$ and $\frac{\alpha \beta + \beta + 1}{\alpha \beta + \beta + 1}$. Therefore, P-WE distribution can be represented, $X = Y + Z$, where $Y \sim Ge \left( \frac{\beta}{\beta + 1} \right)$ and $Z \sim Ge \left( \frac{\alpha \beta + \beta + 1}{\alpha \beta + \beta + 1} \right)$.

Since the P-WE distribution is a convolution of two geometric distributions, it is a special case of a general family of distributions examined by Kemp (1979) involving convolutions of binomial and pseudo-binomial variables. Several models from this family of distributions, such as Non-central Negative Binomial (NNB), Generalized Non-central Negative Binomial (GNNB) and Binomial-Binomial (BB) distributions, were discussed in more details in Kemp (1979) and Ong (1995).

The mean and variance are easier to be obtained by using $M_X(t)$. The mean and variance respectively are Equation (3 and 4):

$$\mu = E(X) = \frac{1}{\beta} \left( 1 + \frac{1}{\alpha + 1} \right)$$

And:

$$\sigma^2 = \frac{1}{\beta^2} \left( 1 + \frac{1}{\alpha + 1} \right) + \frac{1}{\beta^3} \left( 1 + \frac{1}{\alpha + 1} \right)^2$$

Since $\alpha > 0$, this condition implies that $\sigma^2 = \mu + \frac{\alpha^2 + 2 \alpha + 1}{\alpha^2 + 4 \alpha + 4}$ or $\sigma^2 = \mu (1 + \alpha \mu)$, where $0.5 < \alpha < 1$ is the dispersion index.

The coefficient of variation is

$$CV = \frac{\sigma}{\mu} = \sqrt{\frac{(\beta + 1)(\alpha + 1)^2 + \beta (\alpha + 1) + 1}{\alpha + 2}},$$

where the range is $1/\sqrt{2} < CV < \infty$. The skewness is given by

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{(\beta^2 + 3 \beta + 2)(\alpha + 1) + \beta (\alpha + 1)^2 + 3 \beta (\alpha + 1) + 2}{((\beta + 1)(\alpha + 1) + \beta (\alpha + 1) + 1)^3}.$$

Figures 1 and 2 show the coefficient of variation and skewness of P-WE distribution as functions of $z = g(x = \alpha, y = \beta)$. Figures 3-6 show several examples of p.m.f. of P-WE distribution, indicating that the distribution can be considered as an alternative for over dispersed count data.

2.2 Parameter Estimation: Method of Moment Estimators

Assume that $x_1, x_2, \ldots, x_n$ are sample data of size $n$ distributed as P-WE with p.m.f. (1). The Method of Moment (MM) estimators of $\alpha$ and $\beta$ can be obtained by equating the sample and theoretical moments

$$m_1 = \frac{1}{n} \sum x_i$$
$$m_2 = \frac{1}{n} \sum (x_i - \bar{x})^2$$

Figure 1 and 2 show the coefficient of variation and skewness of P-WE distribution as functions of $z = g(x = \alpha, y = \beta)$. Figure 3-6 show several examples of p.m.f. of P-WE distribution, indicating that the distribution can be considered as an alternative for over dispersed count data.
The closed forms for the MM estimators can be obtained and they are Equation (5 and 6):
\[ \beta = \frac{-m_i + \sqrt{2m_i - 2m_i - m_i^2}}{m_i - m_i - m_i^2} \]  
(5)

And:

\[ m_i = m_i + am_i^2 \]  
(6)

Condition \( \alpha > 0 \) in (4) requires \( m_i = m_i + am_i^2 \), where \( 0.5 < \alpha < 1 \). Therefore, the MM estimator \( \hat{\beta} \) can be calculated from (5) if and only if \( m_i = m_i + am_i^2 \), where \( 0.5 < \alpha < 1 \).

2.3. Parameter Estimation: Maximum Likelihood Estimators

The log-likelihood of P-WE distribution is:

\[
\ln L(\alpha, \beta) = n \ln (\alpha + 1) - n \ln \alpha + n \ln \beta + \sum_{i=1}^{\alpha} \ln((\beta + 1)^{(x_i + 1)} - (\alpha \beta + \beta + 1)^{(x_i + 1)}) \]

By taking the partial derivatives with respect to \( \alpha \) and \( \beta \) and equating them to zero, we obtain:

\[
\frac{\partial}{\partial \alpha} \ln L(\alpha, \beta) = \frac{n}{\alpha + 1} - \frac{n}{\alpha} + \beta \sum_{i=1}^{\alpha} (x_i + 1)(\alpha \beta + \beta + 1)^{(x_i + 1)} - (\alpha \beta + \beta + 1)^{(x_i + 1)} = 0
\]

And:

\[
\frac{\partial}{\partial \beta} \ln L(\alpha, \beta) = \frac{n}{\beta} + \sum_{i=1}^{\alpha} (x_i + 1)(\beta + 1)^{(x_i + 1)} - (\alpha + 1)(\alpha \beta + \beta + 1)^{(x_i + 1)} - (\alpha \beta + \beta + 1)^{(x_i + 1)} = 0
\]

where, the Maximum Likelihood (ML) estimators can be solved numerically using MM estimators as initial values.

2.4. P-WE Regression Model

The new P-WE regression model can be derived using different parameterization of P-WE distribution.

Let \( \alpha = v^{-1} \) and \( \beta = \frac{1+v}{\mu_i} \). The p.m.f. of P-WE distribution can be reparameterized as Equation (7):

\[
Pr(Y_i = y_i; \mu_i, v) = \frac{1+v}{\mu_i(v-1)} \left( 1 + \frac{1+v}{\mu_i} \right)^{(y_i+1)} - \left( 1 + \frac{1+v}{\mu_i} \right)^{(y_i+1)}
\]

With mean \( E(Y_i) = \mu_i \) and variance \( Var(Y_i) = \mu_i + \frac{1+v^2}{(1+v)^2} \mu_i^2 \), where \( v > 1 \) and \( \mu_i > 0 \).

The covariates can be incorporated via a log link function, \( E(Y_i | x_i) = \exp(x_i \beta_i) \), where \( x_i \) denotes the exposure, \( x_i \) the vector of covariates and \( \beta_i \) the vector of regression parameters. Hence, the log likelihood of P-WE regression model is:

\[
\log L(y_i; \beta, v) = \log(1+v) - \log(v-1) \sum_i \log \left( 1 + \frac{1+v}{\mu_i} \right)^{(y_i+1)} - \left( 1 + \frac{1+v}{\mu_i} \right)^{(y_i+1)}
\]

The ML estimators of \( \beta \) and \( v \) can be obtained by maximizing the log likelihood.

The new P-WE regression model can be compared with other regression models for count data with covariates such as Poisson, Negative Binomial (NB) and Generalized Poisson (GP). In actuarial literature as well as insurance practice, Poisson regression model has been widely used for modeling claim count data. As examples, (Nasr-Esfahani et al., 1990; Renshaw et al., 1994) respectively fitted Poisson regression model to two different sets of U.K. motor claim count data. For handling over dispersion, several regression models such as NB and GP have been suggested. Several parameterizations have been performed for NB regression models and the two well known models, referred asNB-1 and NB-2 in Greene (2008), have been developed and applied (Cameron and Trivedi, 1986; Lawless, 1987; Ismail and Jemain, 2007; Zulkifli et al., 2013). Several parameterizations have also been performed for GP regression models and the two well known models, referred as GP-1 and GP-2 in Yang et al. (2009), have been developed and applied (Consul, 1989; Ismail and Jemain, 2007; Ismail and Zamani, 2013).
3. RESULTS

Example 1

An insurance count data from Belgium in year 1993 is considered (Denuit, 1997) for fitting Poisson, NB and P-WE distributions, using both ML and MM estimation procedures. It should be noted that the MM estimators of P-WE distribution can be calculated using closed formulas (5 and 6) as long as $0.5 < \alpha < 1$. Based on the sample data, $m_1 = 0.1057$ and $m_2 = 0.1149$, so that $\alpha = 0.8211$. The chi-square and log likelihood are considered as comparison criteria. Table 1 provides the observed values, fitted values, estimated parameters, chi-square and log likelihood.

The results show that the PWE-MLE provides the largest log likelihood and the smallest chi-square. Even though the NB distribution is a strong competitor, the P-WE distribution provides better performances because the log likelihood of PWE-MLE is larger than NB-MLE and the chi-square of PWE-MM is smaller than NB-MM.

Example 2

The US National Medical Expenditure Survey 1987/88 (NMES) data from Deb and Trivedi (1997) is considered. The NMES data was used to model the demand for medical care, captured by the number of physician office visit and the number of hospital outpatient visit. For an illustration purpose, we use only the first 2000 data for fitting the regression models. Our response variable is number of physician visit (OFP) and the covariates are the number of Hospital Stays (HOSP), self-perceived health status (POORHLTH and EXCLHLTH), number of chronic conditions (NUMCHRON), gender (MALE), number of years of education (SCHOOL) and private insurance indicator (PRIVINS). Table 2 shows the mean and standard deviation of the selected variables, whereas Table 3 shows the parameter estimates, standard errors and t-ratios for the fitted models.

The results show that the regression parameters for all models have similar estimates. As expected, NB-2, GP-2 and P-WE regression models provide similar inferences for the regression parameters where the absolute value of t-ratios are smaller than Poisson regression model. Comparison between the standard errors of regression parameters of NB-2, GP-2 and P-WE regression models indicate that the standard errors of NB-2 and P-WE models are equal or smaller than GP-2 model, with the exception of regression variable EXCLHLTH. Based on the log likelihood, AIC and BIC, the P-WE regression model is the best model for fitting the US NMES count data.

| Table 1. Observed and fitted values (example 1) |
|-----------------------------------------------|
| No. of accidents | Frequency | Poisson | NB-ML | NB-MM | PWE-MLE | PWE-MM |
| 0 | 57178 | 56949.52 | 57185.53 | 57197.08 | 57180.20 | 57195.95 |
| 1 | 5617 | 6019.80 | 5583.45 | 5563.16 | 5592.08 | 5566.27 |
| 2 | 446 | 318.16 | 485.86 | 492.69 | 482.14 | 490.05 |
| 3 | 50 | 11.51 | 43.87 | 46.37 | 44.27 | 46.37 |
| 4 | 8 | - | - | - | - | - |

| Parameters |
|------------|
| $\lambda = 0.1057$ |
| $r = 1.2780$ |
| $\alpha = 5.7736$ |
| $\alpha = 8.0683$ |

| Chi-square |
|------------|
| 267.011 |
| 8.02 |
| 8.23 |
| 7.08 |
| 7.34 |

| Log likelihood |
|----------------|
| -22150.54 |
| -22064.31 |
| - |
| -22063.76 |
| - |

| Table 2. Descriptive summary (example 2) |
|-----------------------------------------|
| Variable | Measurement | Mean | Stddev |
|----------|-------------|------|--------|
| OFP | Number of physician visit | 6.046 | 7.561 |
| HOSP | Number of hospital stays | 0.297 | 0.716 |
| POORHLTH | Self-perceived health status; poor = 1, else = 0 | 0.130 | 0.336 |
| EXCLHLTH | Self-perceived health status; excellent=1, else = 0 | 0.071 | 0.257 |
| NUMCHRON | Number of chronic conditions | 1.533 | 1.337 |
| MALE | Gender; male = 1, else = 0 | 0.408 | 0.491 |
| SCHOOL | Number of years of education | 10.355 | 3.640 |
| PRIVINS | Private insurance indicator, yes = 1, no = 0 | 0.794 | 0.405 |
### Table 3. Poisson, NB-2, GP-2 and P-WE regression models (example 2)

| Parameter | Poisson | NB-2 | GP-2 | P-WE |
|-----------|---------|------|------|------|
|           | est.    | s.e. | t-ratio | est.    | s.e. | t-ratio | est.    | s.e. | t-ratio | est.    | s.e. | t-ratio |
| Intercept | 0.99    | 0.04 | 27.51   | 0.88    | 0.09 | 10.17   | 0.82    | 0.09 | 09.12   | 0.86    | 0.09 | 9.91   |
| HOSP      | 0.19    | 0.01 | 20.39   | 0.24    | 0.03 | 7.14    | 0.28    | 0.04 | 06.22   | 0.24    | 0.03 | 7.20   |
| POORHLTH  | 0.21    | 0.03 | 8.28    | 0.31    | 0.07 | 4.32    | 0.39    | 0.09 | 04.53   | 0.30    | 0.07 | 4.21   |
| EXCLHLTH  | -0.21   | 0.04 | -4.84   | -0.18   | 0.10 | -1.92   | -0.17   | 0.09 | -1.80   | -0.19   | 0.10 | -1.96  |
| NUMCHRON  | 0.16    | 0.01 | 24.48   | 0.19    | 0.02 | 10.04   | 0.20    | 0.02 | 09.54   | 0.20    | 0.02 | 10.44  |
| MALE      | -0.10   | 0.02 | -5.31   | -0.12   | 0.05 | -2.52   | -0.13   | 0.05 | -2.54   | -0.12   | 0.05 | -2.59  |
| SCHOOL    | 0.03    | 0.00 | 12.19   | 0.04    | 0.01 | 5.29    | 0.04    | 0.01 | 05.29   | 0.04    | 0.01 | 5.37   |
| PRIVINS   | 0.15    | 0.02 | 5.85    | 0.14    | 0.06 | 2.28    | 0.15    | 0.07 | 02.27   | 0.16    | 0.06 | 2.65   |
| Dispersion| -       | -     | -       | a = 0.04| 25.03|         | 0.28    | 16.37|         |
| Log likelihood | -8813.74 | 5607.195 |       | -5614.72 | 5604.688 |       |
| AIC       | -17611.48 | 11196.39 |       | -11211.44 | 11140.97 |       |
| BIC       | -17566.67 | 11145.98 |       | -11191.38 | 11140.97 |       |

### 4. CONCLUSION

This study has introduced a new two-parameter mixed Poisson distribution, namely Poisson-Weighted Exponential (P-WE), which is obtained by mixing Poisson distribution with a new class of weighted exponential distribution. The P-WE distribution is suitable for over dispersed count data with variance $\sigma^2 = \mu + \frac{\mu^\alpha}{\nu}$, $0.5 < \alpha < 1$.

Besides the univariate version, the regression model of P-WE distribution with mean $E(Y_i) = \mu_i$ and variance $\text{Var}(Y_i) = \mu_i + \frac{\mu_i^\alpha}{(1 + \nu)}$, $\nu > 1$, $\mu_i > 0$, has been derived.

For numerical illustrations, P-WE distribution was fitted using MM and ML estimation procedures to an insurance count data and the results were compared to MM and ML estimators of Poisson and NB distributions. Based on chi-square and log likelihood, the P-WE MLE provide the largest log likelihood and the smallest chi-square. Considering the straightforward manner of obtaining the MM estimators using closed formulas, the P-WE distribution can be considered as an alternative for fitting over dispersed count data.

The P-WE regression model was fitted to the US NMES data. The regression model was compared to Poisson, NB-2 and GP-2 regression models and based on the log likelihood, AIC and BIC, the P-WE regression is the best model for fitting this data. Therefore, the P-WE regression model can also be considered as an alternative for fitting over dispersed count data with covariates.

### 5. REFERENCES

Azzalini, A., 1985. A class of distributions which includes the normal ones. Scandinavian J. Stat., 12:171-178.

Cameron, A.C. and P.K Trivedi, 1986. Econometric models based on count data: Comparisons and applications of some estimators and tests. J. Applied Econometr., 1: 29-53. DOI: 10.1002/jae.3950010104

Consul, P.C., 1989. Generalized Poisson Distribution: Properties and Application. 1st Edn., Marcel Dekker, New York. ISBN-10: 0824778634.

Deb, P. and P.K. Trivedi, 1997. Demand for medical care by the elderly: A finite mixture approach. J. Applied Econometr., 12: 313-336.

Denuit, M., 1997. A new distribution of poisson-type for the number of claims. ASTIN Bull., 27: 229-242. DOI: 10.2143/AST.27.2.542049

Ghitany, M.E., D.K. AL-Mutairi, S. Nadarajah, 2008. Zero-truncated poisson-lindley distribution and its application. Math., Comput. Simulat., 79: 279-287. DOI: 10.1016/j.matcom.2007.11.021

Gomez-Deniz, E., J.M. Sarabia and E.C. Ojeda, 2008. Univariate and multivariate versions of the negative binomial-inverse Gaussian distributions with applications. Math. J. Comput. Simulat., 79: 279-287. DOI: 10.1016/j.matcom.2007.11.021

Gomez-Deniz, E., J.M. Sarabia and E.C. Ojeda, 2008. Univariate and multivariate versions of the negative binomial-inverse Gaussian distributions with applications. Insurance: Math. Econom., 42: 39-49. DOI: 10.1016/j.insmatheco.2006.12.001

Greene, W., 2008. Functional forms for the negative binomial model for count data. Econometric Lett., 99: 585-590. DOI: 10.1016/j.econlet.2007.10.015

Gupta, R.D. and D. Kundu, 2009. A New class of weighted exponential distribution. Statist. J. Theoretical Applied Stat., 43: 621-634. DOI: 10.1080/02331880802605346
Hossein, Z. and N. Ismail, 2010. Negative binomial-Lindley distribution and its application. J. Math. Statist., 6: 4-9.

Ismail, N. and A.A. Jemain, 2007. Handling overdispersion with negative binomial and generalized Poisson regression models. Casualty Actuarial Society Forum Winter.

Ismail, N. and H. Zamani, 2013. Estimation of claim count data using negative binomial, generalized poisson, zero-inflated negative binomial and zero-inflated generalized poisson regression models. Casualty Actuarial Society Forum Spr., 1: 1-29.

Karlis, D. and E. Xekalaki, 2005. Mixed poisson distributions. Int. Stat. Rev., 73: 35-58. DOI: 10.1111/j.1751-5823.2005.tb00250.x

Kemp, A.W., 1979. Convolutions involving binomial pseudo-variables. Sankya Ind. J. Stat., 41: 232-243.

Klugman, S.A, H.H. Panjer and G.E. Willmot, 2008. Loss Models: From Data to Decision. 4th Edn., John Wiley and Sons, USA., ISBN-10: 1118315324, pp: 536.

Lawless, J.F., 1987. Negative binomial and mixed Poisson regression. Canad. J. Statist., 15: 209-225. DOI: 10.2307/3314912

Lord, D. and S.R. Geedipally, 2011. The negative binomial-lindley distribution as a tool for analyzing crash data characterized by a large amount of zeros. Accident Anal. Prevent., 43: 1738-1742. DOI: 10.1016/j.aap.2011.04.004

Meng, S., Y. Wei and G.A. Whitmore, 1999. Accounting for individual over-dispersion in a bonus-malusystem. ASTIN Bull., 29: 327-337. DOI: 10.2143/AST.29.2.504619

Nasr-Esfahani, M.H., J.R. Aitken and M.H. Johnson, 1990. Hydrogen peroxide levels in mouse oocytes and early cleavage stage embryos developed in vitro or in vivo. Development, 109: 501-507.

Ong, S.H., 1995. Some stochastic models leading to the convolution of two binomial variables. Statist. Probability Lett., 22: 161-166. DOI: 10.1016/0167-7152(94)00063-E

Renshaw, B.R., W.C. Fanslow 3rd, R.J. Armitage, K.A. Campbell and D. Liggitt et al., 1994. Humoral immune responses in CD40 ligand-deficient mice. JEM, 180: 1889-1900. DOI: 10.1084/jem.180.5.1889

Sankaran, M., 1970. The discrete poisson-lindley distribution. Biometrics, 26: 145-149.

Tremblay, L., 1992. Using the poisson inverse Gaussian in bonus-malusystems. ASTIN Bull., 22: 97-106.

Wang, Z., 2011. One mixed negative binomial distribution with application. J. Stat. Plann. Inference, 141: 1153-1160. DOI: 10.1016/j.jspi.2010.09.020

Yang, Z., J.W. Hardin and C.L. Addy, 2009. A score test for overdispersion in poisson regression based on the generalized Poisson-2 model. J. Stat. Plann. Inference, 139: 1514-1521. DOI: 10.1016/j.jspi.2008.08.018

Zamani, H. and N. Ismail, 2012. Functional form for the generalized poisson regression model. Commun. Statist. Theory Methods, 41: 3666-3675. DOI: 10.1080/03610926.2011.564742

Zulkifli, M., N. Ismail and A.M. Razali, 2013. Analysis of vehicle theft: A case study in Malaysia using functional forms of negative binomial regression models. Applied Math. Informat. Sci., 7: 389-395. DOI: 10.12785/amis/072L02