Journal of Statistical Modelling: Theory and Applications
Vol. 2022, No. 01, 2022, pp. 83-101
Yazd University Press 2022

JSMTA

Research Paper

A new alpha power type-1 family of distributions and modelling the overdispersed count outcome

Getachew Tekle [1], Rasool Roozegar*[2], Hamid Bidram [3]
[1,2]Department of Statistics, Yazd University, Yazd, Iran
[3]Department of Statistics, University of Isfahan, Isfahan, Iran

Received: October 28, 2022/ Revised: February 20, 2023/ Accepted: February 23, 2023

Abstract: In this paper, we introduce a novel family of statistical models called a new alpha power type-1 family of distributions. Three sub-cases of the family are discussed. Based on the novel family, a special model, explicitly, a new alpha power type-1-Weibull distribution is studied in depth. The new model has very interesting patterns of failure rates like increasing, decreasing, bathtub, and parabola-down. Hence, it is so flexible. Based on the comparison analysis, among five well-known models, it has an impact on health data analysis. Furthermore, the count data models capable of handling overdispersion and zero-inflation are discussed and applied the real health data. The zero-inflated negative binomial model in the frequentist approach has shown its popularity in handling both overdispersion and zero-inflation simultaneously, while the discrete Weibull model with the logit(q) link in the Bayesian approach outperformed its counterparts.

Keywords: Count models; Discrete Weibull distribution; Family of distributions; New alpha power type-1 family; Overdispersion; Statistical modeling; Zero-inflation.
Mathematics Subject Classification (2010): 60E05, 62E15.

1 Introduction

In the training and implementation of health care and the health sector in general (Malehi et al., 2015), statistical modeling and predicting real-life events are vital issue. It is noted that the classical and modified statistical models have been applied to the data in health applications. However, these models do not provide the best fit when

*Corresponding author: `rroozegar@yazd.ac.ir`

the data show non-monotonic failure rates. This clearly demands the generalized or extended versions of these classical models. This demand motivated many applied researchers to propose a new flexible extensions of distributions by adding one or more additional parameters to the baseline distribution. Therefore, this paper introduces a novel new more flexible family of distributions called a new alpha power type-1 family of distributions by introducing a new parameter α to the exponential type of family of distributions. It is more suitable for skewed data with non-monotonic failure rates.

There are numerous recent developments in the distribution theory. Very recently, Chesneau et al. (2022) introduced a new extended family of distributions called an alternative to the Marshall-Olkin family of distributions. They used five different estimation methods to reveal the alternating capacity of the new family to the existing one and they used the sub-cases for the regression purpose.

Shehata et al. (2021) introduced a flexible family of distributions for the asymmetric left-skewed bimodal real-life data with special attention to the flexibility patterns of the probability density and hazard functions and they called it a novel two-parameter G family of distributions.

It is observed that researchers used transformation, extension, and compounding techniques to introduce a new family of distributions. Recently, Chesneau et al. (2019) proposed a new family of probability distributions, based on a cosine-sine transformation by compounding a baseline distribution with the cosine and sine functions. Other authors, Ahmad et al. (2019), used the same approach by adding an additional parameter to introduce a new class of probability distributions. They named the newly suggested model the extended alpha power transformed family of distributions and the extended alpha power transformed Weibull distribution is studied as a special model. Another recent study on the discrete version is presented by Nekoukhou and Bidram (2020). They introduced a similar class of distributions in the discrete case based on geometric odds ratio.

The first part of this study introduces a new family of distributions called a new alpha power type-1 family of distributions to the health data with the different patterns of the data. The second part discusses the count models for overdispersion and zero-inflation in the frequentist and Bayesian approaches. The rest of the paper is organized as: Section 2 introduces a new alpha power type-1 family of distributions, Section 3 discusses special cases, Section 4 discusses basic statistical properties of the newly proposed family, Section 5 presents an estimation of the model parameters of the new sub-family of distributions, Section 6 illustrates the application to the new dataset. Section 7 presents models for count outcomes, Section 8 gives the results for the count models, and followed by summary points with concluding remarks are presented.

2 A new alpha power type-1 family of distributions

As we stated above numerous methods have been suggested in the literature to introduce a new family of distributions. For instance, Nekoukhou and Bidram (2020) introduced a class of distributions in the discrete case based on geometric odds ratio with the cumulative distribution function (CDF)

$$F\left(z;\mathbf{\Omega},\theta,\alpha\right)=1-\theta^{\left(\frac{R(z;\mathbf{\Omega})}{S(z+1;\mathbf{\Omega})}\right)^{\alpha}},\qquad z\in\mathbb{N}_0,0<\theta<1,\alpha>0,$$

where $\mathbf{\Omega}$ is the vector of parameters for the baseline distribution, $R\,(.;.)$ is an arbitrary discrete CDF, and $S\,(.;.)$ denotes its corresponding survival function. A well-known method to introduce new distributions is given by the CDF and the probability density function (PDF)

$$\begin{aligned} F\,(z;\boldsymbol{\xi}) &= 1 - e^{-W[G(z;\boldsymbol{\xi})]}, \quad z \in \mathbb{R}, \\ f\,(z;\boldsymbol{\xi}) &= \frac{d}{dz}\left\{W\left[G(z;\boldsymbol{\xi})\right]\right\} e^{-W[G(z;\boldsymbol{\xi})]}, \quad z \in \mathbb{R}, \end{aligned} \tag{1}$$

respectively, where $\boldsymbol{\xi}$ is the vector of parameters and $G\,(.;.)$ is the CDF for the baseline distribution. In this paper, we suggest another new method to introduce a new family of distributions called, a new alpha power type-1 family (NAPF1) of distributions. The CDF of the NAPF1 is given by

$$F\,(z;\alpha,\boldsymbol{\xi}) = 1 - \alpha^{-W[G(z;\boldsymbol{\xi})]}, \quad \alpha > 1, z \in \mathbb{R}, \tag{2}$$

with PDF

$$f\,(z;\alpha,\boldsymbol{\xi}) = \frac{d}{dz}\left\{W\left[G(z;\boldsymbol{\xi})\right]\right\} \alpha^{-W[G(z;\boldsymbol{\xi})]} \log(\alpha), \quad \alpha > 1, z \in \mathbb{R}. \tag{3}$$

If $\alpha = e$, the CDF in (2) becomes similar to (1). The function $W\left[G(z;\boldsymbol{\xi})\right]$ fulfils the conditions given in (i) and (ii), and it is straightforward that $0 \le F\,(z;\alpha,\boldsymbol{\xi}) \le 1$. By setting the odds ratio function $W\left[G(z;\boldsymbol{\xi})\right] = \frac{G(z;\boldsymbol{\xi})}{1-G(z;\boldsymbol{\xi})}$ in (2) and (3), we obtain the $F\,(z;\alpha,\boldsymbol{\xi})$ and $f\,(z;\alpha,\boldsymbol{\xi})$ of the NAPF1 as follows

$$\begin{aligned} F\,(z;\alpha,\boldsymbol{\xi}) &= 1 - \alpha^{-\left(\frac{G(z;\boldsymbol{\xi})}{1-G(z;\boldsymbol{\xi})}\right)}, \quad \alpha > 1, z \in \mathbb{R}, \\ f\,(z;\alpha,\boldsymbol{\xi}) &= \frac{\log(\alpha) g(z;\boldsymbol{\xi})}{\left[1-G(z;\boldsymbol{\xi})\right]^2} \alpha^{-\left(\frac{G(z;\boldsymbol{\xi})}{1-G(z;\boldsymbol{\xi})}\right)}, \quad \alpha > 1, z \in \mathbb{R}, \end{aligned} \tag{4}$$

respectively. Next, the survival function $S\,(z;\alpha,\boldsymbol{\xi}) = 1 - F\,(z;\alpha,\boldsymbol{\xi})$, and hazard function $h\,(z;\alpha,\boldsymbol{\xi}) = \frac{f(z;\alpha,\boldsymbol{\xi})}{S(z;\alpha,\boldsymbol{\xi})}$ are given by

$$\begin{aligned} S\,(z;\alpha,\boldsymbol{\xi}) &= \alpha^{-\left(\frac{G(z;\boldsymbol{\xi})}{1-G(z;\boldsymbol{\xi})}\right)}, \quad \alpha > 1, z \in \mathbb{R}, \\ h\,(z;\alpha,\boldsymbol{\xi}) &= \frac{\log(\alpha) g(z;\boldsymbol{\xi})}{\left[1-G(z;\boldsymbol{\xi})\right]^2}, \quad \alpha > 1, z \in \mathbb{R}, \end{aligned}$$

respectively.

The extra parameter which replaces the exponent term in (4) gives the NAPF1 of distributions with greater flexibility in fitting the application data.

3 Special cases

In this section, the special cases of the new family are discussed. Three important members with Lomax distribution (NAPF1-Lomax), Gompertz distribution (NAPF1-Gomp), and with Weibull distribution (NAPF1-Weib) are discussed (see Table 1, which displays the baseline CDFs for the three distributions together with the CDFs and hazard functions for the three special cases). The analytical and graphical results for the first two special cases are presented in the supplementary material. Particularly, as an important case, the special case with two-parameter Weibull distribution is discussed in detail.

Table 1: Some members of the NAPF1 of distributions.

Name	$G(z;\boldsymbol{\xi})$	$F(z;\alpha,\boldsymbol{\xi})$	$h(z;\alpha,\boldsymbol{\xi})$
Lomax	$1-(1+\lambda z)^{-\beta}$	$1-\alpha^{-\left((1+\lambda z)^{\beta}-1\right)}$	$\frac{\log(\alpha)\left(\lambda\beta/(1+\lambda z)^{(1+\beta)}5\right)}{(1+\lambda z)^{-2\beta}}$
Gompertz	$1-e^{-\lambda/\beta\left(e^{\beta z}-1\right)}$	$1-\alpha^{-\left(e^{\lambda\beta^{-1}}(e^{\beta z}-1)-1\right)}$	$\lambda\log(\alpha)e^{\beta z}e^{\lambda/\beta(e^{\beta z}-1)}$
Weibull	$1-e^{-\beta z^{\lambda}}$	$1-\alpha^{-\left(e^{\beta z^{\lambda}}-1\right)}$	$\lambda\beta\log(\alpha)z^{\lambda-1}e^{\beta z^{\lambda}}$

3.1 A new alpha power type-1-Weibull distribution

Let Z be a Weibull random variable with the CDF

$$G(z;\boldsymbol{\xi}) = 1-e^{-\beta z^{\lambda}}, \qquad \lambda,\beta>0, z\geq 0, \tag{5}$$

where λ and β are the shape and scale parameters, respectively.

Based on (5), the $g(z;\boldsymbol{\xi})$, $S(z;\boldsymbol{\xi})$, and $h(z;\boldsymbol{\xi})$ are given by

$$\begin{aligned} g(z;\boldsymbol{\xi}) &= \lambda\beta z^{\lambda-1}e^{-\beta z^{\lambda}}, \\ S(z;\boldsymbol{\xi}) &= e^{-\beta z^{\lambda}}, \\ h(z;\boldsymbol{\xi}) &= \lambda\beta z^{\lambda-1}, \end{aligned}$$

respectively.

By inserting (5) into (4), we obtain the CDF of the NAPF1-Weibull (NAPF1-Weib, for short) distribution as follows:

$$F(z;\alpha,\boldsymbol{\xi}) = 1-\alpha^{-\left(e^{\beta z^{\lambda}}-1\right)}, \qquad \alpha>1, \lambda,\beta>0, z\geq 0.$$

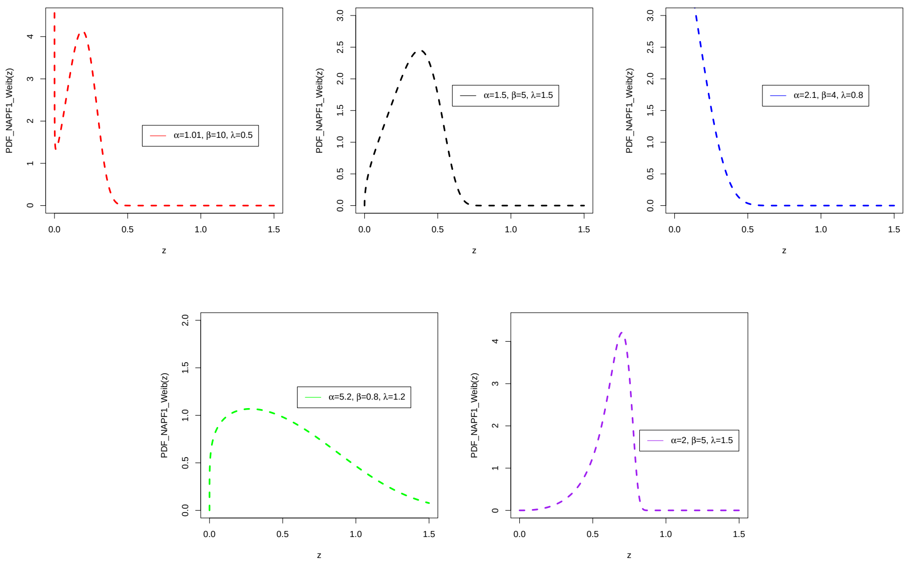
In addition, the $f(z;\alpha,\boldsymbol{\xi})$, $S(z;\alpha,\boldsymbol{\xi})$, and $h(z;\alpha,\boldsymbol{\xi})$ are given by:

$$\begin{aligned} f(z;\alpha,\boldsymbol{\xi}) &= \lambda\beta\log(\alpha)z^{\lambda-1}e^{\beta z^{\lambda}}\alpha^{-\left(e^{\beta z^{\lambda}}-1\right)}, \\ S(z;\alpha,\boldsymbol{\xi}) &= \alpha^{-\left(e^{\beta z^{\lambda}}-1\right)}, \\ h(z;\alpha,\boldsymbol{\xi}) &= \lambda\beta\log(\alpha)z^{\lambda-1}e^{\beta z^{\lambda}}, \end{aligned} \tag{6}$$

respectively. Following this, the graphical expression of the NAPF1-Weib model for different parameter values is displayed as follows:

The $f(z;\alpha,\boldsymbol{\xi})$ of the NAPF1-Weib is elicited in Figure 1 (a) and it has attractive flexible patterns like (i) decreasing-increasing-decreasing-constant or a polynomial type (ii) right-skewed (iii) decreasing (iv) parabola-down (v) left-skewed. Figure 1 (b) illustrates the $h(z;\alpha,\boldsymbol{\xi})$ for different parameter values, to show the different patterns and how flexible the distribution is. It has the patterns like (i) increasing (ii) decreasing (iii) bathtub (iv) parabola-down.

(a) Visual illustration of the $f(z;\alpha,\boldsymbol{\xi})$ for the NAPF1-Weib.



(b) Visual illustration of the $h(z;\alpha,\boldsymbol{\xi})$ for the NAPF1-Weib.

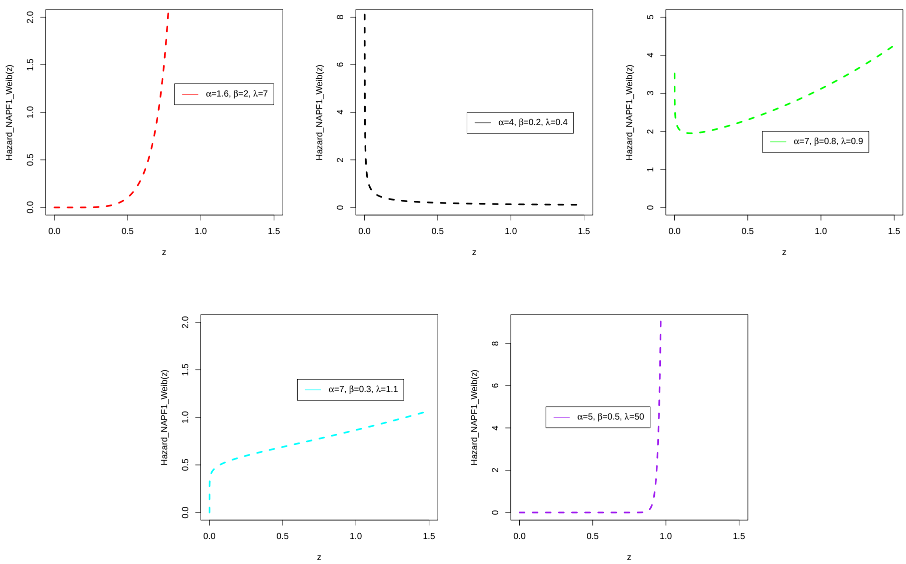


Figure 1: Plots of the $f(z;\alpha,\boldsymbol{\xi})$ and $h(z;\alpha,\boldsymbol{\xi})$ for the NAPF1-Weib distribution for different scenarios.

4 Basic statistical properties of the new alpha power type-1 family of distributions

In this section, we discuss the basic statistical properties of the NAPF1 distributions.

4.1 Quantile function

Quantile function can be used for many purposes in theory and numerical applications in statistics. For example, it can be used to draw simulations. The quantile function of the NAPF1 distribution is obtained by applying the inversion technique. Thus,

$$Qx_n\left(u\right) = F_{NAPF1}(z;\boldsymbol{\xi})^{-1},$$
$$1-\alpha^{-\left(\frac{G(z;\boldsymbol{\xi})}{1-G(z;\boldsymbol{\xi})}\right)} = u, \quad 0<u<1.$$

By solving the non-linear equation, it is derived as:

$$Qx_n\left(u\right) = \frac{\log\left(\alpha^{-(1-u)}\right)}{1+\log\left(\alpha^{-(1-u)}\right)}. \tag{7}$$

It should be noted that u$\sim$ uniform$(0,1)$. The median(Med) is obtained by substituting $u=1/2$ in (7) as

$$Med = \frac{\log\left(\alpha^{-1/2}\right)}{1+\log\left(\alpha^{-1/2}\right)}.$$

Similarly, we can get the lower and upper quartiles by substituting $u=1/4$ and $u=3/4$ in (7), respectively.

4.2 Skewness and Kurtosis

The numerical results for Skewness(Sk) and Kurtosis(K) with the help of quartiles from (7) can be given by

$$Sk = \frac{(Q_3-2Q_2+Q_1)}{(Q_3-Q_1)} = \frac{(q_{(0.75)}-2q_{(0.5)}+q_{(0.25)})}{(q_{(0.75)}-q_{(0.25)})},$$
$$K = \frac{Q_{\frac{7}{8}}-Q_{\frac{5}{8}}+Q_{\frac{3}{8}}-Q_{\frac{1}{8}}}{Q_{\frac{6}{8}}-Q_{\frac{2}{8}}} = \frac{q_{(0.875)}-q_{(0.625)}+q_{(0.375)}-q_{(0.125)}}{q_{(0.75)}-q_{(0.25)}},$$

respectively.

4.3 Order statistics

Order statistics are widely used in applied statistics such as reliability and lifetime, records, etc. Suppose that $Z_1, Z_2, \ldots, Z_n$ is a random sample of size n following the NAPF1 distribution with parameters (λ,β) and $Z_{1:n}, Z_{2:n}, \ldots, Z_{n:n}$ are its corresponding order statistics. Then, the density function of $Z_{i:n}$ for $(i=1,2,\ldots,n)$ is given by

$$f_{i:n}\left(z\right) = \frac{n!}{(i-1)!(n-i)!}\sum_{j=0}^{n-i}(-1)^j\binom{n-i}{j} f_{NAPF1}\left(z;\boldsymbol{\xi}\right)\left[F_{NAPF1}\left(z;\boldsymbol{\xi}\right)\right]^{i+j-1}.$$

By substituting the CDF and PDF of the NAPF1 (see (4) and (5)) into $f_{i:n}\left(z\right)$, we obtain

$$f_{i:n}\left(z\right) = \frac{n!}{(i-1)!(n-i)!}\sum_{j=0}^{n-i}(-1)^j\binom{n-i}{j}\left[\log(\alpha)\right]^{i+1}\left[g\left(z;\boldsymbol{\xi}\right)\right]^{i-1}\left[\overline{G}(z;\boldsymbol{\xi})\right]^{-2i-1}$$

$$\times\alpha^{-\left[\frac{G(z;\boldsymbol{\xi})}{\overline{G}(z;\boldsymbol{\xi})}\right]^{i-1}}\left[1-\alpha^{-\frac{G(z;\boldsymbol{\xi})}{\overline{G}(z;\boldsymbol{\xi})}}\right]^{i+j-1}, z>0,$$

$$f_{i:n}(z) = \sum_{j=0}^{n-i}\Psi_j f^*_{NAPF1}(z;\boldsymbol{\xi}), \tag{8}$$

where $\Psi_j = \frac{n!}{(i-1)!(n-i)!}(-1)^j\binom{n-i}{j}$, and $f^*_{NAPF1}(z;\boldsymbol{\xi})$ is the product of the PDF and the CDF for the i^{th} order statistics. The PDF of the order statistics for NAPF1 can be obtained from (8) and its r^{th} central moments and moments generating function are given in the next section.

4.4 Moments and moment generating functions

Based on the PDF of the NAPF1 distribution given above, its r^{th} central moments, $E(z^r)=\mu_r^{'}$ is obtained as follows

$$\begin{aligned}\mu_r^{'} &= \int_0^{+\infty} z^r\frac{\log(\alpha)g(z;\boldsymbol{\xi})}{[1-G(z;\boldsymbol{\xi})]^2}\alpha^{-\left(\frac{G(z;\boldsymbol{\xi})}{1-G(z;\boldsymbol{\xi})}\right)}dz,\\ &= \int_0^{+\infty}[\log(\alpha)]^{i+1}z^{ri-1}\left[g(z;\boldsymbol{\xi})\right]^{i-1}\left[\overline{G}(z;\boldsymbol{\xi})\right]^{-2i-1}\alpha^{-\left[\frac{G(z;\boldsymbol{\xi})}{\overline{G}(z;\boldsymbol{\xi})}\right]^{i-1}}dz,\\ &= \sum_{i=0}^{+\infty}\left[\frac{\log(\alpha)}{i!}\right]^{i+1}\Phi_{r,i+1},\end{aligned}$$

where

$$\Phi_{r,i+1} = \int_0^{+\infty} z^{ri-1}\left[g(z;\boldsymbol{\xi})\right]^{i-1}\left[\overline{G}(z;\boldsymbol{\xi})\right]^{-2i-1}\alpha^{-\left[\frac{G(z;\boldsymbol{\xi})}{\overline{G}(z;\boldsymbol{\xi})}\right]^{i-1}}dz.$$

The moment generating function of NAPF1 can be obtained by using the last result of $\mu_r^{'}$ in $M_z(t)$ as

$$M_z(t) = \sum_{s=0}^{+\infty}\frac{t^s}{s!}\mu_r^{'}.$$

The next section deals with the maximum likelihood estimation for the model parameters of the NAPF1-Weib model.

5 Estimation

The method of maximum likelihood estimation for the model parameters for the NAPF1-Weib is discussed in this section.

5.1 Maximum likelihood estimation

This sub-section deals with the computation of maximum likelihood estimators (MLEs) for the model parameters of NAPF1-Weib. Let $z_1, \cdots, z_n$ be n observations of a random sample drawn from NAPF1-Weib with parameters α, λ, and β. Using the PDF of NAPF1-Weib (see (6)), the likelihood function is written as

$$L\left(z_1, \cdots, z_n \mid \boldsymbol{\xi}\right) = (\lambda\beta\log(\alpha))^n \prod_{i=1}^{n} z_i^{\lambda-1} e^{\beta\sum_{i=1}^{n} z_i^{\lambda}} \alpha^{-\sum_{i=1}^{n}\left(e^{\beta z_i^{\lambda}}-1\right)},$$

and its log-likelihood function is given by

$$\begin{aligned}\log L\left(z; \alpha, \lambda, \beta\right) &= n\log(\lambda) + n\log(\beta) + n\log(\log(\alpha)) + (\lambda-1)\sum_{i=1}^{n}\log(z_i)\\ &\quad - \log(\alpha)\sum_{i=1}^{n}\left(e^{\beta z_i^{\lambda}} - 1\right) + \beta\sum_{i=1}^{n} z_i^{\lambda}.\end{aligned}$$

The model parameters are estimated by taking the first partial derivatives of the $\log L(z; \alpha, \lambda, \beta)$ with respect to each model parameter and equating them to zero.

Therefore,

$$\begin{aligned}\frac{\partial \log L(z;\alpha,\lambda,\beta)}{\partial\alpha} &= \frac{n}{\log(\alpha)\alpha} - \frac{1}{\alpha}\sum_{i=1}^{n}\left(e^{\beta z_i^{\lambda}} - 1\right),\\ \frac{\partial \log L(z;\alpha,\lambda,\beta)}{\partial\lambda} &= \frac{n}{\lambda} + \sum_{i=1}^{n}\log(z_i) - \log(\alpha)\beta\sum_{i=1}^{n} e^{\beta z_i^{\lambda}}\log(z_i) z_i^{\lambda} + \beta\sum_{i=1}^{n}\log(z_i) z_i^{\lambda},\\ \frac{\partial \log L(z;\alpha,\lambda,\beta)}{\partial\beta} &= \frac{n}{\beta} + \log(\alpha)\sum_{i=1}^{n} e^{\beta z_i^{\lambda}} z_i^{\lambda} + \sum_{i=1}^{n} z_i^{\lambda}.\end{aligned}$$

Subsequently, the MLEs of the parameters α, λ, and β can be obtained by solving the non-linear equation

$$\boldsymbol{U_n} = \left(\frac{\partial \log L(z;\alpha,\lambda,\beta)}{\partial\alpha}, \frac{\partial \log L(z;\alpha,\lambda,\beta)}{\partial\lambda}, \frac{\partial \log L(z;\alpha,\lambda,\beta)}{\partial\beta}\right)^T = 0,$$

using numerical methods such as Newton-Raphson or Broyden's methods.

6 An application to the breast cancer data

The new data on time-to-recovery of 686 breast cancer patients are taken from patient's medical record cards that were enrolled from October 2012 to April 2017 in Nigist Elleni Mohamad memorial referral comprehensive hospital (NEMMRCH), Hossana, south Ethiopia.

We illustrate the fitting capacity of the NAPF1-Weib model to the data by comparing it to three-parameter exponential flexible Weibull extension (EFWE) of El-Desouky et al. (2016), three-parameter Poisson inverse Weibull (PIW) of Joshi and Kumar (2021), two-parameter truncated exponential-exponential (TEE) of Mahdavi and Oliveira (2017), five-parameter exponentiated Weibull-Weibull (EWW) of Hannan and Elgarhy (2016), and a four-parameter exponentiated generalized Frechet (EGF) of Cordeiro et al. (2013).

The information criteria (IC) such as (i) AIC (Akaike, 1974), (ii) CAIC (Bozdogan, 1987), (iii) BIC (Schwarz, 1978), and (iv) HQIC (Hannan and Quinn, 1979) are used to discriminate the best model. In addition to these criteria, the log-likelihood $(-2\log L)$ of the fitted models is also calculated. In all these, the model with the least value is taken to be the best model to fit the data.

Table 2: MLE of the parameters and the corresponding standard errors (SE. in the parentheses) for the fitted models.

Dist.	$\hat{a}$ (SE.)	$\hat{\alpha}$ (SE.)	$\hat{\beta}$ (SE.)	$\hat{\gamma}$ (SE.)	$\hat{\sigma}$ (SE.)	$\hat{\lambda}$ (SE.)
NAPF1-Weib		32.000	0.012			3.566
		(8.747)	(0.002)			(0.112)
EGF		123.438	0.475		11.770	1.045
		(40.335)	(0.100)		(2.175)	(0.114)
EWW	2.402	17.498	0.768	3.045		0.003
	(0.447)	(2.253)	(0.055)	(0.132)		(0.002)
TEE		23.514				1.730
		(1.942)				(0.047)
PIW		1.316	3.673			3.682
		(0.026)	(0.293)			(0.099)
EFWE		0.516	0.538			0.055
		(0.007)	(0.112)			(0.007)

Table 3: Model adequacy measures for the fitted models.

Dist.	AIC	CAIC	BIC	HQIC	$-2\log L$
NAPF1-Weib	1215.46	1215.49	1229.05	1220.72	604.73
EGF	1230.97	1231.03	1249.10	1237.99	611.48
EWW	1290.48	1290.57	1313.14	1299.25	640.24
TEE	1317.45	1317.47	1326.51	1320.96	656.72
PIW	1440.46	1440.49	1454.05	1445.71	717.23
EFWE	1435.40	1435.43	1448.99	1440.66	714.70

Table 2 displays the MLEs and standard errors of the NAPF1-Weib model along with the five competing models (EGF, EWW, TEE, PIW, and EFWE). Table 3 gives the model comparison result (model adequacy measures) for all models considered in this section. The new proposed model NAPF1-Weib, based on the five criteria, is shown to be the best-performing model among the five competing models. This shows the new proposed model outperforms the set of similar competing models and is applicable to health data.

The second part of this study is concerned with the adaptation and implementation of the count models for the overdispersed and zero-inflated count health data in both frequentist and Bayesian settings.

Discrete and count data are collected in many application areas, mainly medicine, health, Biology, sometimes in finance and industry (Haselimashhadi et al., 2016; Cameron and Trivedi, 2013). The are two precise examples in the health sector (medicine). The first one is the length of stay in the hospital, usually used as a pointer of the quality of healthcare and planning capacity within a hospital (Atienza et al., 2008; Carter and Potts, 2014). The second one is the number of visits to a doctor (specialist) (José and Santos, 2005), sometimes considered as a measure of demand in healthcare. The are also some other examples in different areas of application like high-throughput genomic data generated by next-generation sequencing experiments (Bao et al., 2014; Ozsolak and Milos, 2011; Robinson and Smyth, 2008) or lifetime data, such as the number of cycles before a machine breaks down (Haselimashhadi et al., 2016; Nakagawa and Osaki, 1975). This kind of data can also be collected from reproductive health services. A good example is the number of antenatal care (ANC) service visits of pregnant women from early pregnancy to their 9 months of pregnancy period. This kind of study was done by Workie and Lakew (2018) in Ethiopia, in which they used Bayesian count regression models to determine the potential determining factors for this service in health facilities.

Most researchers in the literature use the models Poisson (P), QuasiPoisson, Negative binomial (NB), Zero-inflated-Poisson (ZIP), zero-inflated-negative-binomial (ZINB), Conway-Maxwell-Poisson (COMP), discrete Weibull (DW) and their mixtures with another truncated form to model count data.

As implemented by Adesina et al. (2021), DW distribution gives hopeful results for modelling count data as compared to the models P and NB distributions and their mixtures such as Poisson Tweedie, zero-inflated regression, and COMP distribution. It is also observed that this model can capture over and under-dispersion simultaneously and gives a closed-form analytical expression of the quantiles of the conditional distribution (Sellers and Shmueli, 2010).

The Quasi-maximum likelihood estimation or Poisson maximum likelihood is taken to be the well-known method of estimation of the model parameters in the frequentist approach. This is because (i) it gives convenient or satisfactory (but may not be accurate) results; (ii) it has no computational burden and is easily available in many software packages; (iii) it is recommended when doubt exist about the form of the variance function.

Recently, some studies have shown interest in adopting the Bayesian techniques for estimating and fitting count data, as it is observed that it is efficient in its estimation. This technique also has the capacity to deal with complex models that lacking analytically manageable likelihood functions, and are flexible to be adapted to produce estimates that are excellent and perfect substitutes for maximum likelihood estimates (Adesina et al., 2017; Cameron and Trivedi, 2005).

It is clearly known that analysing discrete count data by using the ordinary regression models will cause problem in the analysis results mainly invalid and or incorrect estimates, confidence intervals (CI), and P-values. This can happen because these kinds of data possess the commonly known feature of overdispersion and sometimes zero-inflation and typically are highly skewed. Hence, it is not an easy task for researchers to do this without evaluation of the nature of the data. In this study, we are interested in exploring the appropriate fitting capacity of those models for count data in the health application area.

7 Models for count outcomes

Assume that the random variable Y follows an exponential family of distributions with the following PDF

$$f(y|\eta, \phi) = \exp\{\frac{[y\eta - \psi(\eta)]}{\phi} + c(y, \phi)\}, \qquad y \in \mathbb{R}^+, \tag{9}$$

for definite fixed unknown parameters η and ϕ (often called natural or canonical and dispersion parameter, respectively) and for known functions $\psi(.)$ and $c(.,.)$. Then, it is easy to write the first two moments by following the function $\psi(.)$ as: $E(Y) = \mu = \psi^{'}(\eta)$ and $Var(Y) = \sigma^2 = \phi\psi^{''}(\eta)$, based on the computation of Molenberghs et al. (2007). From this calculation, the mean and variance can be related as: $\sigma^2 = \phi\psi''[\psi^{'-1}(\mu)] = \phi\nu(\mu)$, where $\nu(.)$ is a variance function defining the mean-variance relationship.

A basic example of exponential families for the normal, binary, count, and time-to-event cases is given by Molenberghs et al. (2010). The normal one is seen as a case of the exponential family for its particular feature that it needs an overdispersion parameter which should exceed 1. Thus, it lacks a mean-variance relationship while others are held to exist (Roozegar et al., 2022).

It is flexible to explain the relationship between the two variables $\boldsymbol{Y}$ and $\boldsymbol{X}$, where $\boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots, \boldsymbol{Y}_N$ are an N number of outcomes and $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_p$ are a p dimensional vectors of covariates. This finds a member of the class called GLM, with an assumption of $\boldsymbol{Y}_i$ having $f(y_i \mid \theta_i, \phi)$ densities belonging to an exponential family. Model specification can be done by modeling the means: $\mu_i = h(\eta_i) = h(\boldsymbol{x}_i^T\boldsymbol{\xi})$, where $h(.)$ is a known function, and $\boldsymbol{\xi}$ is a vector of p fixed and unknown regression parameters (coefficients). Usually, $h^{-1}(.)$ is taken to be a link function, while the natural link function is taken to be $h(.) = \psi^{'}(.)$.

7.1 Models for overdispersion and zero-inflation

Let $\boldsymbol{Y}_i$ be Poisson distributed random variables, then the probability mass function (PMF) is given by

$$f(y_i \mid \lambda_i) = \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!}, \qquad y_i = 0, 1, 2, \ldots, i = 1, 2, \ldots, n, \lambda_i > 0, \tag{10}$$

for which the natural parameter $\eta_i = \log(\lambda_i)$, the mean $\mu_i = \lambda_i$, dispersion parameter $\phi = 1$, and the variance function $\nu(\mu_i) = \mu_i = \lambda_i$. The natural logarithm is for the link function which gives an ordinary Poisson regression with $\log(\lambda_i) = \boldsymbol{x}_i^T\boldsymbol{\xi}$.

From (10), it is observed that it presents the equality of variance-mean relationship. However, in practice (in real data) it fails to hold true for the observed or sample variance is greater than the sample mean and which leads to overdispersion.

This can happen due to the absence of the relevant covariates, heterogeneity of sampling units, hierarchical structures, and excess of zeros (Demétrio et al., 2014; Grunwald et al., 2011). This initiates us to adapt models that take into account this problem and to overcome incorrect and misleading inferences (Oliveira, 2014).

Thus, it is reasonable to extend the Poisson model by adding a dispersion parameter ϕ, by assuming λ_i vary according to some distributions and let's take λ_i as continuous for its positive support, then this leads to λ_i to follow a gamma distribution whose PMF given by

$$f(y_i \mid \lambda_i, \phi) = \frac{\Gamma(y_i + \frac{1}{\phi})}{\Gamma(\frac{1}{\phi})\Gamma(y_i + 1)} \left(\frac{\phi\lambda_i}{1 + \phi\lambda_i}\right)^{y_i} \left(\frac{1}{1 + \phi\lambda_i}\right)^{\frac{1}{\phi}}, \qquad y_i = 0, 1, 2, ..., \tag{11}$$

$$\phi > 0, \lambda_i > 0.$$

The mean and variance, respectively given as: $E(\boldsymbol{Y}_i) = \lambda_i$ and $Var(\boldsymbol{Y}_i) = \lambda_i + \lambda_i^2\phi$. The model given by (11) is said to be NB which relatively is more flexible and can accommodate more additional shapes than the one given in (10).

One of the prominent data issues in this study is zero-inflation and it is carried out by ZIP and ZINB models which approach is verified by other scholars like Abdullah and Ahmad (2013); Güneri and Durmuş (2021); Hu et al. (2011); Famoye and Singh (2006); Altun (2018) and Adesina (2021).

7.2 Discrete Weibull distribution

The DW distribution was first introduced by Nakagawa and Osaki (1975) as a discretized form of a continuous Weibull distribution, which is analogous to the geometric distribution that is the discretized form of the exponential and is termed to be a type I DW. Latter, Chanialidis (2015) introduced other types of DW distributions as type II and III. Type I has unbounded support as compared to type II, also type I has a more straightforward interpretation as compared to type III (Adesina et al., 2017).

Let the random variable Y has a type-I DW distribution, then the CDF of Y is defined as

$$F(y;q,\beta) = \begin{cases} 1 - q^{(1+y)^\beta}, & y = 0,1,2,\ldots \\ 0, & if \quad y < 0, \end{cases}$$

and the PMF of Y is given by

$$f(y;q,\beta) = q^{y^\beta} - q^{(1+y)^\beta}, \quad y = 0,1,2,\ldots.$$

For the DW regression model, let the random variable Y be the vector of the response variables with the possible assumed values $0,1,2,\ldots$ and let $X_1, X_2, \ldots, X_p$ be a p-dimensional vector of covariates. It is further assumed that the conditional distribution of Y given X follows a DW distribution with parameters q and β. There are a number of possible ways to link the parameters q and β to the covariates, the commonly suggested are a logit and log links that follow
(i) q is related to X as follows

$$\log\left(\frac{q}{1-q}\right) = \theta_0 + \theta_1 x_1 + \cdots + \theta_p x_p,$$

or

$$\log(-\log(q)) = \theta_0 + a_1 x_1 + \cdots + \theta_p x_p. \tag{12}$$

(ii) β is related to X as follows $\log(\beta) = A_0 + A_1 x_1 + \cdots + A_p x_p$,or in the matrix form

$$\log(\beta) = \boldsymbol{XA}, \tag{13}$$

where $\boldsymbol{A} = (A_0, A_1, \cdots, A_p)^T$.

In order to handle the more intricate correlations, Bolstad (2007) suggested one extra parametrization for q by a logit link function and a linkage between the second parameter β and the covariates, which further has presented it to be reasonably operational for the statistical conclusion.

7.2.1 Bayesian case for discrete Weibull model

The regression model for the Bayesian DW model can be devised by considering the regression parameters $\theta = (\theta_0, \theta_1, \ldots, \theta_p)^T$ and $A = (A_0, A_1, \ldots, A_p)^T$. This approach has two advantages: (i) there is a likelihood of considering the prior information, (ii) this procedure gives an automatic display of credible intervals (CI) for all the model parameters. To formulate the likelihood function for the approach, let us suppose the following pieces of informations are given, the response variable y_i and the row vector of covariates $x_i = (1, x_{i1}, x_{i2}, \ldots, x_{ip})$, where $i = 1, 2, \ldots, n$ and Y is for the vector of the response variable and X is for the vector of covariates. Hence, with the help of this idea and combining (12) and (13), the likelihood function is given by

$$L(\theta, A \mid X, Y) = \prod_{i=1}^{n} \left[\left(\frac{e^{x_i\theta}}{1 + x_i\theta} \right)^{y_i^{x_i A}} - \left(\frac{e^{x_i\theta}}{1 + e^{x_i\theta}} \right)^{(1+y_i)^{x_i A}} \right]. \tag{14}$$

It is noted from (14) that $q = \frac{e^{x_i\theta}}{1+e^{x_i\theta}}$. The prior distribution on the regression parameters is considered as the DW model has no conjugate prior. Thus, the Laplace prior for θ and A are considered (Sellers and Shmueli, 2010), which take the following form

$$\begin{aligned} p(\theta \mid \mu) &= \frac{\mu}{2} e^{-\mu|\theta|}, \quad \mu > 0, \\ p(A \mid \gamma) &= \frac{\gamma}{2} e^{-\gamma|A|}, \quad \gamma > 0. \end{aligned}$$

The maximization of the posterior density under those priors, for the pre-specified μ and γ, resembles maximising the $\log L$ penalised log-likelihood as

$$\log L(\theta, A \mid X, Y) - \mu \sum_{j=1}^{p} \mid \theta_j \mid - \gamma \sum_{k=1}^{p} \mid Ak \mid .$$

Further consideration of Gamma (α, λ) hyper prior for μ and γ to get the posterior density (Sellers and Shmueli, 2010) can be done as

$$p(\theta, A \mid X, Y) \propto L(\theta, A \mid X, Y) \times p(\theta \mid \mu) \times p(A \mid \gamma) \times p(\mu) \times p(\gamma).$$

8 Results for count models

The data for this part is collected on congestive heart failure (CHF) disease patients from NEMMRCH at Hossana, South Ethiopia. The data is taken from 107 CHF patients following their treatments from September 30, 2018 to June 25, 2020. The descriptive summary measures for the data, the numerical and graphical confirmation of overdispersion and zero-inflation in the data are given in the supplementary material. The variable of interest (response), here, is the number of doctor visits (the number of times the patients have seen their doctor). The important results of the models will be presented in this section.

8.1 Test result for overdispersion and zero-inflation

The presence of overdispersion is tested by using the function “dispersiontest(chf, trafo = 1)” of the R package AER and it is distributed asymptotically normal with mean zero and variance 1. The input “chf” of the function represents the response variable and “trafo=1” represents the linear transformation. The test statistic with the p-value and alpha (z = 1.862, P-value = 0.025, and alpha= 1.120) confirms that the data is overdispersed due to the known feature in the data. This feature as shown next is due to the excess structural zeros in the data. Thus, since the P-value is much smaller than the z-value and the alpha is greater than 1, over-dispersion in the data is confirmed.

The test for excess zeros with the test statistic distributed as Chi-square with 1 df (Chi-square = 19.713, df = 1, and p-value=8.99e-06) shows that there is zero-inflation in the data.

8.2 Model comparison for the count models for the frequentist approach

This sub-section presents the numerical results for the count models in the frequentist settings and the comparison among these models is based on the -2logL, AIC, and BIC measures. The significant covariates are identified by ∗.

Table 4: Comparison of P, NB, ZIP, ZINB, and COMP models. The table summarizes the parameter estimates and standard errors (SE.) together with model comparison criteria results.

		P	NB	ZIP
Effect	Parameter	Estimate (SE.)	Estimate (SE.)	Estimate (SE.)
Intercept	ξ_0	-0.155(0.395)	-0.694(0.274)∗	-0.594(0.310)
sex-male	ξ_1	-0.309(0.141)∗	-0.203(0.101)∗	0.033(0.117)
time	ξ_2	0.570(0.056)∗	0.315(0.031)∗	0.368(0.029)∗
Place-Urban	ξ_3	-0.015(0.141)	-0.019(0.097)	-0.019(0.110)
wieghts	ξ_4	0.006(0.007)	0.005(0.004)	0.004(0.005)
heart-disease-yes	ξ_5	0.240(0.142)	0.155(0.095)	0.027(0.110)
-2logL		530.6	959.4	448.8
AIC		1075.2	973.4	923.5
BIC		1101.8	1000.0	972.9
		ZINB	COMP	
Effect	Parameter	Estimate (SE.)	Estimate (SE.)	
Intercept	ξ_0	-2.911(1.224)∗	-0.785(0.375)∗	
sex-male	ξ_1	1.067(0.467)∗	-0.167(0.142)∗	
time	ξ_2	0.711(0.139)∗	0.302(0.034)∗	
Place-Urban	ξ_3	-0.040(0.428)	0.035(0.127)	
wieghts	ξ_4	-0.010(0.020)	0.004(0.006)	
heart-disease-yes	ξ_5	-0.742(0.475)	0.059(0.128)	
-2logL		448.8	476.3	
AIC		921.5	964.7	
BIC		967.1	987.5	

Based on the results of Table 4, the five models are compared based on the model comparison criteria. Accordingly, the model improvement is done in the order ZINB, ZIP, COMP, NB, and P, respectively. Thus, the first four models are better performing compared to the P model. Furthermore, the ZINB is the best frequentist count model to deal with the count overdispersed data (the number of doctor visits that the patients committed).

Based on the ZINB model, the male patients were longer visited by doctors than the female patients, and the patients with a longer time until they get recovered longer visited their doctors than patients with relatively shorter follow-up times.

8.3 Model comparison for the count models for the Bayesian approach

In this sub-section, Bayesian DW with log and logit link is compared to the Bayesian P without and with random effects, Geometric (Ge), and NB models. The R package MCMCglmm with MCMCglmm() function does not support the mixture models ZIP, ZINB, and COMP. Hence, the potential candidate models in the Bayesian setting are P, NB, Ge, and DW.

Table 5: Comparison of the six models in the Bayesian approach using the ICs, where the table displays posterior mean for the parameter, and the 95% CIs (lower CI (l-95%), upper CI (u-95%)) for the posterior mean of the parameters.

	DW(Log)			DW(Logit)			P-MCMCglmm		
Par.	Post.mean	Lower	Upper	Post.mean	Lower	Upper	Post.mean	Lower	Upper
$\widehat{\theta}_0$	-0.051	-0.560	0.650	-0.310	-0.945	0.374	-0.849	-1.316	-0.442 *
$\widehat{\theta}_1$	0.174	-0.075	0.382	-0.235	-0.452	0.072	-0.193	-0.362	-0.032 *
$\widehat{\theta}_2$	-0.564	-0.664	-0.446*	0.587	0.477	0.709*	0.327	0.279	0.384*
$\widehat{\theta}_3$	0.015	-0.226	0.234	-0.049	-0.261	0.246	-0.003	-0.172	0.146
$\widehat{\theta}_4$	-0.005	-0.016	0.004	0.007	-0.006	0.018	0.007	0.001	0.016 *
$\widehat{\theta}_5$	-0.147	-0.392	0.063	0.186	-0.061	0.450	0.160	-0.079	0.332
$\widehat{\beta}$	2.099	1.884	2.345*	2.081	1.860	2.315 *			
AIC	937.21			935.92			974.43		
BIC	963.81			962.51			1001.02		
DIC	936.84			934.79			969.64		
	P-MCMCglmmR			Ge-MCMCglmm			NB-MCMCglmm		
Par.	Post.mean	Lower	Upper	Post.mean	Lower	Upper	Post.mean	Lower	Upper
$\widehat{\theta}_0$	-0.447	-0.899	-0.157*	0.309	-0.297	1.040	-0.149	-0.897	0.619
$\widehat{\theta}_1$	-0.163	-0.371	-0.059*	0.152	-0.009	0.519	-0.315	-0.572	-0.032*
$\widehat{\theta}_2$	0.304	0.260	0.354*	-0.291	-0.365	-0.224*	0.567	0.461	0.686*
$\widehat{\theta}_3$	-0.074	-0.185	0.094	0.084	-0.087	0.313	-0.018	-0.282	0.238
$\widehat{\theta}_4$	0.002	-0.002	0.010	0.001	-0.011	0.013	0.007	-0.007	0.020
$\widehat{\theta}_5$	0.033	-0.062	0.209	-0.174	-0.289	-0.079*	0.243	-0.029	0.501
AIC	976.12			1071.02			1082.14		
BIC	1002.72			1097.61			1108.73		
DIC	969.40			1061.03			1100.70		

Table 5 displays the results for the Bayesian approach for the models DW (both for log and logit links), P-MCMCglmm without random effect (P-MCMCglmm), P-MCMCglmm with a normal random effect (P-MCMCglmmR), Ge-MCMCglmm and NB-MCMCglmm. The normal random effect is imposed on the P model to refer to the correlation in the data. The models P, Ge, and NB were treated with the help of the R package MCMCglmm. The discrimination of the models is done based on the three ICs, where the model with the least ICs is chosen to be the best model. In the class of MCMCglmm, the P model without random effect outperformed the P model with random effect, Ge, and NB. The DW model (for which the R package BDWreg is implemented) with both link functions is best model to fit the data and it handles the overdispersion in the data.

Furthermore, DW with logit(q) link function or DW(Logit) is a better model as compared to DW(Log). Hence, further Bayesian inference can be done with this model. Time (time until a patient recovers) and β are significant in DW models for both link functions. Most of the covariates including β have a positive relationship with the response variable.

Additionally, as displayed in the supplementary material, for simulated data, the 95% HPD shows the parameters interval where they are significant. Thus, the red lined cross like interval shows the significant parameters which Table 5 displays.

Discussion and conclusions

In this study, a new data set (breast cancer) is considered and the proposed model is compared to the recent models. The NAPF1-Weib, EFWE, EGF, EWW, TEE, and PIW models were applied to the above mentioned public health data. The NAPF1-Weib model has shown its supremacy based on the five adequacy measures.

The results of the frequentist approach are obtained by using the maximum likelihood estimation technique. The count models are very important for the regression modelling of count data as discussed by Adesina et al. (2017); Abdullah and Ahmad (2013); Güneri and Durmuş (2021); Hu et al. (2011); Famoye and Singh (2006); Altun (2018) and Adesina (2021). In this study, the known five-count models (P, NB, ZIP, ZINB, COMP) were compared. The ZINB outperformed the others and it handles both overdispersion and zero-inflation in the present data.

The Bayesian approach outperformed the frequentist counterpart. In the Bayesian approach, the DW model with the logit(q) link outperformed the DW model with the log(q) link and the models in the class of MCMCglmm, which is supported by the results of Adesina et al. (2017) and Hadfield (2010). This makes it the best model for overdispersed count data in public health. In the absence of the Bayesian DW model, the Bayesian P model with normal random effects can treat correlation besides the overdispersion in the count data. The DW model is a discrete model used for both under-dispersed and overdispersed count data.

The newly proposed family of distributions is taken to be an alternative performing family to the exponential type family. The DW model with the logit(q) link is the best model capable of dealing with the overdispersion in the count data.

Acknowledgement

The authors would like to thank Yazd University for supporting this study. We would finally like to thank anonymous reviewers, referees, and also the Editor-in-Chief.

References

Abdullah, S.A. and Ahmad, W.M.A.W (2013). Modelling the incident and risk factors of streptococcus pneumonia transmission among children in Malaysia. *International Journal of Physical, Chemical and Mathematical Sciences*, **2**(2):106–116.

Adesina, O.S. (2021). Models for zero truncated count data in medicine and insurance. *Theory and Practice of Mathematics and Computer Science*, **6**:129–141.

Adesina, O., Agunbiade, D. and Osundina, S. (2017). Bayesian regression model for counts in scholarship. *Journal of Mathematical Theory and Modelling*, **7**(9):46–57.

Adesina, O.S., Agunbiade, D.A. and Oguntunde, P.E. (2021). Flexible Bayesian Dirichlet mixtures of generalized linear mixed models for count data. *Scientific African*, **13**:e00963.

Ahmad, Z., Ampadu, C.B., Hamedani, G.G., Jamal, F. and Nasir, M.A. (2019). The new exponentiated TX class of distributions: properties, characterizations and application. *Pakistan Journal of Statistics and Operation Research*, **XV**(IV):941–962.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**(6):716–723.

Altun, E. (2018). A new zero-inflated regression model with application. *İstatistikçiler Dergisi: İstatistik ve Aktüerya*, **11**(2):73–80.

Atienza, N., Garcia-Heras, J., Munoz-Pichardo, J.M. and Villa, R. (2008). An application of mixture distributions in modelization of length of hospital stay. *Statistics in Medicine*, **27**(9):1403–1420.

Bao, Y., Vinciotti, V., Wit, E. and Hoen, P. (2014). Joint modelling of ChIP-seq data via a Markov random field model. *Biostatistics*, **15**(2):296–310.

Bolstad, W.M. (2007). *Introduction to Bayesian Statistics*. New York: AJohn Willey & Sons. Inc.

Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, **52**(3):345–370.

Cameron, A.C. and Trivedi, P.K. (2013). *Regression Analysis of Count Data*. Vol. 53, Cambridge university press.

Cameron, A.C. and Trivedi, P.K. (2005). *Microeconometrics Methods and Application*. Cambridge University Press.

Carter, E.M. and Potts, H.W. (2014). Predicting length of stay from an electronic patient record system: a primary total knee replacement example. *BMC Medical Informatics and Decision Making*, **14**(1):1–13.

Chanialidis, C. (2015). *Bayesian mixture models for count data*. Doctoral dissertation, University of Glasgow.

Chesneau, C., Bakouch, H.S. and Hussain, T. (2019). A new class of probability distributions via cosine and sine functions with applications. *Communications in Statistics-Simulation and Computation*, **48**(8):2287–2300.

Chesneau, C., Karakaya, K., Bakouch, H.S. and Kuş, C. (2022). An alternative to the Marshall-Olkin family of distributions: bootstrap, regression and applications. *Communications on Applied Mathematics and Computation*, **4**:1–29.

Cordeiro, G.M., Ortega, E.M. and da Cunha, D.C. (2013). The exponentiated generalized class of distributions. *Journal of Data Science*, **11**(1):1–27.

Demétrio, C.G., Hinde, J. and Moral, R.A. (2014). Models for overdispersed data in entomology. *Ecological Modelling Applied to Entomology*, 219–259.

El-Desouky, B.S., Mustafa, A. and Al-Garash, S. (2016). The exponential flexible Weibull extension distribution. *arXiv preprint arXiv:1605.08152*.

Famoye, F. and Singh, K.P. (2006). Zero-inflated generalized Poisson regression model with an application to domestic violence data. *Journal of Data Science*, **4**(1):117–130.

Grunwald, G.K., Bruce, S.L., Jiang, L., Strand, M. and Rabinovitch, N. (2011). A statistical model for under-or overdispersed clustered and longitudinal count data. *Biometrical Journal*, **53**(4):578–594.

Güneri, Ö.İ. and Durmuş, B. (2021). Models for overdispersion count data with generalized distribution: An application to parasites intensity. *Journal of New Theory*, **35**:48–61.

Hadfield, J.D (2010). MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. *Journal of Statistical Software*, **33**(2):1–22.

Haselimashhadi, H., Vinciotti, V. and Yu, K. (2016). A new Bayesian regression model for counts in medicine. *arXiv:1601.02820*.

Hannan, E.J. and Quinn, B.G. (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society: Series B (Methodological)*, **41**(2):190–195.

Hassan, A.S. and Elgarhy, M. (2016). A new family of exponentiated Weibull-generated distributions. *International Journal of Mathematics and its Applications*, **4**(1-D):135–148.

Hu, M.C., Pavlicova, M. and Nunes, E.V. (2011). Zero-inflated and hurdle models of count data with extra zeros: examples from an HIV-risk reduction intervention trial. *The American Journal of Drug and Alcohol Abuse*, **37**(5):367-375.

José, A.F.M. and Santos Silva, J.M.C. (2005). Quantiles for counts. *Journal of the American Statistical Association*, **100**(472):1226–1237.

Joshi, R.K. and Kumar, V. (2021). Poisson inverse Weibull distribution with theory and applications. *International Journal of Statistics and Systems*, **16**(1):1–16.

Mahdavi, A. and Oliveira Silva, G. (2017). A method to expand family of continuous distributions based on truncated distributions. *Journal of Statistical Research of Iran*, **13**(2):231–247.

Malehi, A.S., Pourmotahari, F. and Angali, K.A. (2015). Statistical models for the analysis of skewed healthcare cost data: a simulation study. *Health Economics Review*, **5**(1):1–16.

Molenberghs, G., Verbeke, G., Demétrio, C.G. and Vieira, A.M. (2010). A family of generalized linear models for repeated measures with normal and conjugate random effects. *Statistical Science*, **25**(3):325–347.

Molenberghs, G., Verbeke, G. and Demétrio, C.G. (2007). An extended random-effects approach to modelling repeated, overdispersed count data. *Life time Data Analysis*, **13**(4):513–531.

Nakagawa, T. and Osaki, S. (1975). The discrete Weibull distribution. *IEEE Transactions on Reliability*, **24**(5):300–301.

Nekoukhou, V. and Bidram, H. (2020). A new discrete distribution based on geometric odds ratio. *Journal Statistical Modelling: Theory and Applications*, **1**(2):153-166.

Oliveira, I.R.C.D. (2014). *Modeling strategies for complex hierarchical and overdispersed data in the life sciences.* Doctoral dissertation, Universidade De São Paulo, Uhasselt.

Ozsolak, F. and Milos, P.M. (2011). RNA sequencing: advances, challenges and opportunities. *Nature Reviews Genetics*, **12**(2):87–98.

Remi, J.D., Olumide, S.A., Pelumi E.O. and Olasunmbo O.A. (2019). Adaptive regression model for highly skewed count data. *International Journal of Mechanical Engineering and Technology*, **10**(1):1964–1972.

Robinson, M.D. and Smyth, G.K. (2008). Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics*, **9**(2):321–332.

Roozegar, R., Tekle, G., and Hamedani, G. (2022). A new generalized-X family of distributions: Applications, characterization and a mixture of random effect models. *Pakistan Journal of Statistics and Operation Research*, **18**(2):483–504.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, **6**(2):461-464.

Sellers, K.F. and Shmueli, G. (2010). Predicting censored count data with COM-Poisson regression. *Robert H. Smith School Research Paper No. RHS-06-129.*

Shehata, W.A., Yousof, H. and Aboraya, M. (2021). A Novel generator of continuous probability distributions for the asymmetric left-skewed bimodal real-life data with properties and copulas. *Pakistan Journal of Statistics and Operation Research*, **17**(4):943–961.

Workie, M.S. and Lakew, A.M. (2018). Bayesian count regression analysis for determinants of antenatal care service visits among pregnant women in Amhara regional state, Ethiopia. *Journal of Big Data*, **5**(1):1–23.