Journal of Innovative Construction and Petrochemical Technologies, Vol.3, No.1 (June, 2026), 106-119

P-ISSN: xxxx- xxxx | E-ISSN: 3007-9861 | DOI: https://doi.org/10.71285/icpt.v3i1

Journal of Innovative Construction and Petrochemical Technologies
https://icpt.ouk.kz/index.php/icpt/index



Research Article

Check for updates

A Comparison Between Tweedie Shrinkage And Zero-Inflated Negative Binomial Models

Wakkaa Alil Hadba[1*]
* Correspondence: wakaa2017@uokirkuk.edu.iq

Abstract: The negative-inflated binomial (ZINB) regression is a model of count data, which are overdispersed and contain excessive zero counts. The Tweedie Shrinkage regression (Ridg, Lasso, Elastic) and regression of negative inflated binomial models (ZINB) were compared in this research and it was revealed in the study that the Tweedies Shrinkage regression methods provided better comparison criteria than the regression of negative inflated binomial models (ZINB). Better than Tweedie regression,ols, and Regularization procedure,In this work, the authors tried to compare the performance of zero-inflated negative binomial (ZINB) models with regression models with Shrinkage Tweedy regression (Ridg, Lasso, Elastic). The study revealed that the systematic Tweedie regression provided the highest values of comparison criteria according to the means of prediction, and it is an indication of the possibility of using data. Which is afflicted by distraction as a means of resolving this issue. According to our result of this research, we highly suggest that a Tweedie Shrinkage (Ridg, Lasso, Elastic) regression model should be used to analyze a zero-inflated data to obtain statistically valid results.

Keywords: zero-inflated negative binomial , Tweedie Shrinkage Regression, Ridge ,Lasso ,and Elastic Tweedie Regression

Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).



Introduction

The Poisson distribution has the property of equality of mean and variance, although the counting data that are observed will tend to exhibit over-dispersion (larger variance than mean), or underdispersion (smaller variance than mean), Which is popular in applied data analysis.

Regression model is relevant in a variety of data.It contains assumptions that no significant (multilinear) correlation exists between the predictor variables, the variables utilized in the modeling belong to continuous or normally distributed data.That is, between the response variable and the predictor variable. The regression analysis, least squares method of parameter estimation, is the method applied generally. In actual data a lot of the time assumptions are not fulfilled. Since many predictor variables Cause multilinearity, it is also seen that the response variables of many data are discrete or categorical and they are not normally distributed. Generalized linear modeling can be used to do this in case the response variables are independent and identically distributed of the exponential family. The Poisson model can be used to obtain variables A response with count data that is distributed by a Poisson distribution in most cases. In the calculation of data, however, one is often left with trailing zeros, which means that there is no amount of that observed data. Therefore, trailing zeros are impossible to obtain with the help of the Poisson model since it requires the following assumption to be satisfied within the model: the variance and the mean of the response variable should become equal. In the instance of trailing zeros in which the variable of the response has a larger variance than the mean then this is referred to as excessive dispersion which brings in the linear model. This always inflated Poisson zero (ZIP) model proposed by Lambert (1992), inflated binomial zero (ZIB), inflated Poisson zero (ZIGP) by Famoye (2003), amd zero inflated binomial Negative (ZINB) The zero-inflated negative binomial (ZINB) regression model can be used in solving many types of data. The multilinearity of the modeling may influence the increase of the value of the variance to larger values and the value obtained with the help of the least squares method to become unstable. The Multicollinear could be addressed by removing one of the major correlating variables, or through including new predictors variables. Multicollinear can also be solved by making a modulus model/ regression to zero. Hoerl and Kennard coined the term ridge regression in 1970 which led to the development of the term penalty estimation. The effect of penalty estimators is now felt in the analysis of data through the use of regression modelsThe approach that can be applied in the selection of the coefficient up to the value of zero is the Least Absolute Selection and Contraction Factor (LASSO) by Tishibrani (1996). LASSO method is under development, LASSO successfully identifies subsets in regression model coefficients and, therefore, estimation LASSO was formulated by Zou (2006) by placing an adaptive weight in the LASSO equation thus, it is referred to as adaptive LASSO. Additionally, in 2005. The elastic net was proposed by Zou and Hastie as a regularization technique and variable selection method. In order to compute LASSO and ridge regression estimates, it uses a combination of L1 and L2 penalty functions [1].

Banerjee et al. (2018)[2] used adaptive LASSO to perform Zero Inflated Poisson regression. The reason behind this is that in the instances of ZIP regression, cases of polylinear occurrence are identified in high-dimensional variants. Thus, the model that ZIP has received is volatile. Banerjee employed the simulation data and real data to establish the most appropriate method of quantifying the coefficient of the variable in ZIP regression. In accordance with the study by Sulistyaningsih (2019), it was discovered that ZIP results were superior to ZIGP in over dispersion issue on maternal mortality rate in Bali Province data [3]

To show how the given methods are proposed, we must give a rough concept of the Tweedie regression. In this article, the authors rely on these approaches to introduce the answer to numerous issues connected with the proper organization of the Tweedy model and the application of simulations to get its variable choice and the efficiency of model fitting. Especially, (R T R), (LTR) and (E T R) are selected. Bent Jorgensen called the Tweedies distributions after Maurice Tweedie [4] (from the University of Liverpool, UK). Therefore, in 1984. Jørgensen is the first scholar who presented detailed research about the Tweedie

distributions [5].

In the context of the generalized linear model, the authors in [6] (Wei Qian, Yi Yang & Hui Zou (2016)) investigated the aggregate elastic network method of the Tweedie model. J. Kenneth Ta et al. (2021)[7] examined the elastic modulated lattice regression span of all families of the generalized linear model. Wigena et al. (2021) applied Tweedie compound poisson gamma (TCPG distribution, which does not address multicollinearity), with Lasso regularization [8].

The article is structured in the following fashion; Section 1 offers a background on the study, problem statement and our overall objectives. Section 2 centers on the methodology used in the research such as linear regression methodology of punishment, and statistical nature of the punishment approaches like the Lasso, Rige and Elastic. In section 3, the Tweedy distribution, types, and the estimation of the parameter are presented. Moreover, Tweedie regression model and Tweedie compound Poisson model characterize Tweedie, Negative Binomial GLMs, Zero-Inflated Negative Binomial, and offer punitive (RTR), (LTR) and (ETR) models. Section 4 provides differentiation between models. Section 5 tests the proposal with the help of simulation experiments and actual data. Lastly, Section 6 brings this paper to an end and provides an overview of some future directions.

Shrinkage Methods)

The penalized regression model can be formed with the help of Shrinkage (or Regularization) methods (so it is also referred to as penalized regression). The most common penalized models are Ridge regression, LASSO and Elastic net.

Ridge Regression

As a regression method, the ridge estimator (RE) was used as an alternative to the ordinary least squared (OLSE) regression estimator. It is used to solve the multicollinearity problem in the linear regression model. Consider the data $\{(xi;\ yi)/i\ =\ 1;\ldots.;\ n\}$ with the model, xi refers to the ith row of the X matrix. It is formulated as follows [1].

$$y_i = x_i^{\mathrm{T}}\beta + \epsilon_i; \beta = \left(\beta_1, \ldots, \beta_p\right)^{\mathrm{T}}, \epsilon_i \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2), 1 < i < n \qquad (2.1)$$

where, for β , The least-square estimator is $\tilde{\beta}_n = \left(X^{\mathrm{T}}X\right)^{-1}X^{\mathrm{T}}Y$, where $\mathrm{E}\left(\tilde{\beta}_n\right) = \beta$ and rank(X)=p<n .

From least squares estimation, we recall that β is estimated using the argument that minimizes Ridge Regression Recall that the least square estimates minimize RSS:

$$RSS = \Sigma_{i=1}^{n}\left(y_i - \beta_0 - \Sigma_{j=1}^{p}\beta_j x_{ij}\right)^2 \qquad (2.2)$$

Ridge regression [9] is similar; however, it finds β^{R}that optimizes a slightly different function:

$$\Sigma_{i=1}^{n}\left(y_i - \beta_0 - \Sigma_{j=1}^{p}\beta_j x_{ij}\right)^2 + \lambda\Sigma_{j=1}^{p}\beta_j^2 = RSS + \lambda\Sigma_{j=1}^{p}\beta_j^2 \qquad (2.3)$$

If λ= 0, the ridge regression is the least squares estimate.

As $\lambda \longrightarrow \infty$, ridge regression imposes more emphasis on the penalty, and coefficients will approach 0. Then, we expect that ridge regression will output a different set of coefficients with different λ. Note that β_0 is not included in the constraint because it measures the average response [10]. Additionally, λ > 0 is a tuning parameter.

Ridge regression considers minimizing RSS similar to the least squares. Nevertheless. it inserts a shrinkage penalty $\lambda\Sigma_{j=1}^{p}\beta_j^2$ that considers the number of parameters in the model. It is identical to least squares when λ equals 0. As λ gets larger, the coefficients start shrinking toward zero. When λ approaches ∞, the rest of the coefficients $\beta 1, \ldots, \beta p$ are close to zero. Therefore, the penalty is not applied to $\beta 0$. Every value of λ corresponds to a set of parameter estimates (L2 regularization). Furthermore, the tuning parameter λ is used to adjust the impact of the two parts in equation (2.3).

RE is superior to the normal least square method in the situations when multivariate data is large (p>n) i.e. when there are more predictors than observations. However, RE is afflicted with all the predictors in the final model. Although all the coefficients approach 0 as RE increases, none of them approaches zero. Hence, to address this challenge, LASSO will be the most appropriate.[11]

LASSO

As mentioned earlier, the regression coefficients can be shrunk to 0 by LASSO. With a penalty term called L1-norm, LASSO penalizes the regression model (L1-norm is the sum of the absolute coefficients). The penalty makes some of the coefficient estimates equal to 0 with a little addition to the model in the case of lasso regression. As such to decrease the complexity of the model, the subset selection methods to choose variables can be substituted with the lasso. Like RE, a good λ is important in the lasso. Lasso does better than RE since it is able to create simpler and more interpretable models with a smaller number of predictors.

However, in most cases, none of them out-shines the other. To be more precise, When there are large coefficients in some predictors (i.e., the rest have small coefficients) the lasso is improved. In comparison, the all predictors with approximately equal coefficient give a better performance of RE compared to lasso [12]. On a particular data set, to identify which is better, Cross-validation methods are utilized [11].

In order to fit a regression model, the lasso would be the best choice more so when there is multicollinearity in the data. In short, least squares regression can be used to estimate coefficients that reduce the value of the total squared residuals (RSS).

Similar to RE, lasso adds a penalty. The lasso coefficients $\hat{\beta}_L^{\lambda}$ can be minimized as follows.

$$\Sigma_{i=1}^{n}\left(y_i - \beta_0 - \Sigma_{j=1}^{p}\beta_j x_{ij}\right)^2 + \lambda\Sigma_{j=1}^{p}|\beta_j| = RSS + \lambda\Sigma_{j=1}^{p}|\beta_j| \qquad (2.4)$$

The penalty is the main difference between RE and lasso.

In the statistical method, While $L2$ penalty (β_j^2) is used by ridge, $L1$ penalty ($|\beta j$ |) is utilized by lasso. When the λ is large, the estimates can be reduced to 0. Lasso can be taken as a method of selection of variables. It is a huge strength since it results in a more explicable model. [13]

The elastic net

In this method, L1-norm and L2-norm are penalized in the produced regression model. Consequently, Similar to RE, shrink the coefficients effectively. Additionally, like the lasso, some coefficients are set to 0 [11]. In lasso, specifically, the variables are defined with 0 values and there is an assignment of only one variable in a bunch of correlated variables. RE (shrinkage estimator), on the other hand, brings the variable to the minimum which is not 0. Being a restriction in the lasso aspect, most n variables are selected when p>n. Otherwise, the RE is predominant in LASSO [1].

Consequently, these challenges are raised with the assistance of the elastic net. In practice, the elastic net is superior to the lasso using empirical data and a simulation. Also the elastic net focuses on the grouping effect wherein strong correlated predictors are apt to be out or in the model. The value of the elastic net is much greater with p that is larger than n. However, in the p case there is no need of the lasso [14]. Minimizing the RSS (i.e., eq. 2.2) leads to estimating the parameters.

$$\left(\hat{\beta}_0, \hat{\beta}\right) = \underset{(\beta_0,\beta)\in\mathbb{R}^{p+1}}{\operatorname{argmin}} \frac{1}{2n}\Sigma_{i=1}^{n}\left(y_i - \beta_0 - x_i^T\beta\right)^2 \qquad (2.5)$$

In the last two decades, regularization methods have attracted a lot of researchers. The elastic net [14] is utilized since it minimizes the sum of the RSS and a regularization term. It is a mixture of ℓ_1 and ℓ_2 penalties.

$$\left(\hat{\beta}_0, \hat{\beta}\right) = \underset{(\beta_0,\beta)\in\mathbb{R}^{p+1}}{\operatorname{argmin}} \left[\frac{1}{2n}\Sigma_{i=1}^{n}\left(y_i - \beta_0 - x_i^T\beta\right)^2 + \lambda\left(\frac{1-\alpha}{2}\| \beta \|_2^2 + \alpha \| \beta \|_1\right)\right] \qquad (2.6)$$

where α is a higher-level hyperparameter and $\alpha \in [0,1]$. $\lambda \geq 0$ is a tuning parameter where the path models fit in the λ. However, depending on the prediction model type (we want), the value of α is set [7]. Thus, for lasso (RE), we make α = 1(0). A is set it near to one, where we need a sparse model; but we are concerned with the correlations of features. In general, the final value of λ is determined through cross-validation. The coefficients of the last model are the same as 1/ 0, to obtain the minimum error cross-validated [7].

The Tweedie model

In 1972, Nelder and Wedderburn introduced the generalized linear model (GLM) [15]. The Tweedie model is regarded in exponential dispersion model (EDM) as a special member. Exponential dispersion family as a set of probability distribution are used as response

distribution of GLMs. The probability function is developed in the form of the following equation of EDM family distributions. When y is (continuous) discrete then it is a probability (density) mass function.

$$\mathcal{P}(y;\theta,\phi) = a(y,\phi)\exp\left\{\frac{y\theta-\kappa(\theta)}{\phi}\right\} \quad (3.1) \qquad (3.1)$$

The mean (μ) refers to the canonical parameter (θ) function.

$\phi > 0$ is the dispersion Parameter $(0,+\infty), k(\theta)$ refers to the cumulant function $a(y,\phi)$, which is the normalized quantity that has a size base that is independent of the parameter θ [5,6].

EDM has variance $\mathrm{Var}(y) = \phi k''(\theta) = \phi \mathrm{Var}(\mu)$ and mean feature $\mu = \mathrm{E}[\mathrm{Y}] = k'(\theta) = dk(\theta)/d\theta$. $\mathrm{Var}(y)$ is determined from the 1st and 2nd derivatives of $k(\theta)$ because of the one-to-one mapping between μ and θ. The function of the mean μ is denoted as $k''(\theta)$. The variance function is defined as $\mathrm{V}(\mu) = k''(\theta)$.

Tweedie family densities are characterized by power variance functions of the form $V[\mu] = \varphi\mu^{\xi}$, where $\xi \in (-\infty, 0] \cup [1, \infty)$ is the index determining the distribution [4]. Continuous and discrete densities are existed in the family [16].

The Tweedie Family Tweedie models are exponential dispersion models closed under scale transformations and translations. The Tweedie model is denoted as $\mathrm{Tw}_{\xi}(\mu, \sigma^2)$. Notice how this notation includes a subscript ξ. Throughout this article, we shall refer to ξ as either the power or indexing parameter.

For modeling continuous data, the Tweedie distributions are suitable in case $\xi \leq 0$, where $-\infty < y < \infty$. The normal distribution ($\xi = 0$) is a particular case. By contrast, for modeling discrete data, the Tweedie distributions are suitable in case $\xi \leq 0$, where y = 0, φ, 2φ, 3φ. For example, a positive probability exists for y = 0, 2, 4,..., when φ = 2. When φ = 1, the Poisson distribution is a special case.

For modeling positive continuous data with exact 0s, the Tweedie distributions are suitable for 1 < ξ < 2. Let us take rainfall modeling as an example. When rain does fall, the amount is a continuous measurement; however, an exact zero is recorded when no rain falls. For modeling positive continuous data, the Tweedie distributions are suitable for ξ ≥ 2. Here, the distributions of inverse Gaussian (ξ = 3) and gamma (ξ = 2) are special cases. When p falls between 0 and 1, the Tweedie distribution is not defined. The most interesting interval is the one between 1 and 2 of p. In this range the Tweedie distribution becomes without any mass at the centre of the distribution as it tends to become a gamma distribution instead of a Poisson distribution. A compound Poisson distribution can be used to produce Tweedie random variable Y [17].

Table 3.1: The Tweedie family of distributions for varying values: respective supports (S), parameter spaces (Ω), and index parameter (ξ) [18].

Tweedie EDMs	Ξ	S	Ω	Examples
Positive stable	ξ > 2	R^+	R^+	inverse Gaussian
Gamma	ξ = 2	R^+	R^+	
Poisson-gamma	1 < ξ < 2	$R^+ \cup \{0\}$	R^+	--
Discrete	ξ = 1	$\mathbb{N} \cup \{0\}$	R^+	Poisson
No EDMs exist	0 < ξ < 1	---		----
Normal	ξ = 0	R	R	---
Extreme stable	ξ < 0	R	R^+	--

Table 3.2: Popular members of the Tweedie family of distributions the following parameters: natural parameter (Θ), respective parameter spaces for the mean (Ω), support (S), deviance (d(y, μ), normalizing constant a(y, φ), canonical parameter (θ), dispersion (φ), cumulate function (κ(θ)), variance function (V (μ)) and their index parameter (ξ) [18].

Tweedie EDMs	Ξ	V (μ)	κ(θ)	θ	φ	d(y, μ)	a(y, ø)	S	Ω	Θ
Normal	ξ =0	μ	$\frac{\theta^2}{2}$	μ	σ_2	$(y-\mu)^2$	$\exp\{-(y^2/\sigma^2 + \log 2\pi)/2\}$	R	R^2	R
Poisson	ξ =1	1	$\exp(\theta)$	log(μ)	1	$2\left\{y\log\frac{y}{\mu} - (y-\mu)\right\}$	$\frac{1}{y!}$	$\mathbb{N} \cup \{0$	R^+	R
Poisson-gamma	(1, 2)	μ^{ξ}	$\frac{\{(1-\xi)\theta\}^{(2-}}{2-\xi}$	$\frac{\mu^{1-\xi}}{1-\xi}$	φ	$2\left\{\frac{\max(y,0)^{2-}}{(1-\xi)(2-} - \frac{y\mu^{1-\xi}}{1-\xi} + \frac{\mu^{2-\xi}}{2-\xi}\right\}$	− − −	$R^+ \cup \{0$	R^+	R^-
Gamma	ξ =2	μ^2	$-\log(-\theta)$	$-\frac{1}{\mu}$	φ	$2\left\{-\log\frac{y}{\mu} + \frac{y-\mu}{\mu}\right\}$	$\frac{\phi^{-1/\phi_y/\phi-}}{\Gamma(1/\phi)}$	R^+	R^+	R

Tweedie regression models

Consider, $(y_i, \boldsymbol{x}_i), i = 1, \ldots, n$, where y_i 's are i.i.d. realizations of Y_i according to $Y_i \sim \mathrm{Tw_p}(\mu_i, \phi)$ and $g(\mu_i) = \eta_i = \boldsymbol{x}_i^{\top}\boldsymbol{\beta}$, where $\boldsymbol{x}_i$ and $\boldsymbol{\beta}$ are $(Q \times 1)$ vectors of known covariates and unknown regression parameters, respectively. It is straightforward to see that $\mathrm{E}(Y_i) = \mu_i = g^{-1}(\boldsymbol{x}_i^{\top}\boldsymbol{\beta})$ and the $\mathrm{Var}(Y_i) = C_i = \phi\mu_i^p$. Therefore, by its joint distribution and first two moments, the model is equivalently specified. The Tweedie regression model is parametrized by $\boldsymbol{\theta} = (\boldsymbol{\beta}^{\top}, \boldsymbol{\lambda}^{\top} = (\phi = \exp(\delta), p)^{\top})^{\top}$. It is worth mentioning that, for computational convenience, we adopt the reparametrization $\phi = \exp(\delta)$ [19]. Here, $\hat{\boldsymbol{\theta}}_M$ is the maximum likelihood estimator (MLE) for the parameter vector $\boldsymbol{\theta}$**.** It can be calculated by maximizing the following log-likelihood function.

$$\mathcal{L}(\boldsymbol{\theta}) = \sum_{i=1}^{n} \log\{a(y_i; \boldsymbol{\lambda})\} + \frac{1}{\exp(\delta)}\left(y_i\psi_i - k(\psi_i)\right) \quad (3.2)$$

$$\hat{\theta}_M \sim \mathrm{N}\left(\boldsymbol{\theta}, \mathcal{F}_{\theta}^{-1}\right) \quad (3.3)$$

where the inverse of the Fisher information matrix is denoted as $\mathcal{F}_{\theta}^{-1}$. The approximation $\tilde{\mathcal{F}}_{\lambda}$ is replaced instead of The entry $\mathcal{F}_{\lambda}$ [19]

2.2 The Tweedie compound Poisson distribution

The compound Poisson distribution of Tweedie is a superposition of a degenerate distribution with a point on the origin, and a continuous distribution, which is supported on the positive real line.

$$T \sim Pois(\lambda);\ Xi \sim Gamma(\alpha;\ \gamma)$$

Given a given α parameter coefficient alpha, the distribution can be stated as an exponential-dispersion model with a power variance function $V(\mu) = \mu^p$, where the power index $p = (\alpha + 2) = (\alpha + 1)\,2\,(1;\ 2)$.p

The Tweedie compound-Poisson model is a distribution, used in a given index parameter, with

the representation of exponential dispersion model, with $E(Y)=\mu, \mathrm{Var}(Y)=\phi\cdot V(\mu)=\phi\cdot\mu^{\mathrm{p}}$, where $\phi>0$ is the dispersion parameter. Using this representation, we denote the compound Poisson distribution as $Y\sim \mathrm{CPois}(\mu,\phi,\mathrm{p})$.

we assume the observed response vector is $\boldsymbol{Y}=(Y_1,\cdots,Y_n)'$, where $Y_i\sim$ CPois (μ_i,ϕ,p). In the compound Poisson GLM, the mean $\boldsymbol{\mu}=E(\boldsymbol{Y})$ is stipulated as a function of some linear predictors through a link function η as

$$\eta(\boldsymbol{\mu})=\boldsymbol{X\beta},$$

where X is the design matrix and β is the vector of fixed effects

2. Zero-inflated Tweedie model:

To find out the zero inflated Tweedie model we introduce the Poisson composite Tweedie model, Let N be a Poisson random variable denoted by $\mathrm{Pois}(\lambda)$ with mean λ, and let $\tilde{Z}_d$'s $(d=0,1,\dots,N)$ be i.i.d Gamma random variables denoted by $\mathrm{Gamma}(\alpha,\gamma)$ with mean $\alpha\gamma$ and variance $\alpha\gamma^2$. Assume N is independent of $\tilde{Z}_d$'s. Define a compound Poisson random variable Z by

$$Z=\begin{cases}0 & \text{if } N=0\\ \tilde{Z}_1+\tilde{Z}_2+\cdots+\tilde{Z}_N & \text{if } N=1,2,\dots\end{cases}$$

Then Z is a Poisson sum of independent Gamma random variables.[20]

The power mean-variance relationship is $\mathrm{Var}(Z)=\phi\mu^{\mathrm{p}}$ for some index parameter $\mathrm{p}\in(1,2)$, which gives $\theta=\mu^{1-\mathrm{p}}/(1-\mathrm{p})$, $\kappa(\theta)=\mu^{2-\mathrm{p}}/(2-\mathrm{p})$ and $\ddot{\kappa}(\theta)=\mu^{\mathrm{p}}$. The parameterize of the compound Poisson model by

$$\lambda=\frac{1}{\phi}\frac{\mu^{2-\mathrm{p}}}{2-\rho},\ \alpha=\frac{2-\mathrm{p}}{\rho-1},\ \gamma=\phi(\rho-1)\mu^{\rho-1},$$

then it will have the form of a Tweedie model $\mathrm{Tw}(\mu,\phi,\rho)$ with the probability density function

$$f_{\mathrm{Tw}}(z\mid\mu,\phi,\rho):=a(z,\phi,\rho)\exp\left(\frac{1}{\phi}\left(z\frac{\mu^{1-\rho}}{1-\rho}-\frac{\mu^{2-\rho}}{2-\rho}\right)\right)$$

where

$$a(z,\phi,\rho)=\begin{cases}1, & \text{if } z=0\\ \frac{1}{z}\sum_{t=1}^{\infty}W_t(z,\phi,\rho) & \\ =\frac{1}{z}\sum_{t=1}^{\infty}\frac{z^{t\alpha}}{(\rho-1)^{t\alpha}(2-\rho)^{t}\Gamma(t\alpha)\phi^{t(1+\alpha)t!}}, & \text{if } z>0\end{cases}$$

with $\alpha=(2-\rho)/(\rho-1)$. When $z>0$, the sum of infinite series $\sum_{t=1}^{\infty}W_t$ is an example of Weight's generalized Bessel function.

With the above formula, the Tweedie model has a positive probability mass at zero with P(Z = 0) = P(N = 0) = exp(−λ). A Tweedie distribution with weighted by a probability, which is nondifferent, and an even point mass at zero weighted by the complementary probability is assimilated by the zero-inflated mixture model [20]

1 − q :

$$Y=\begin{cases}Z, & \text{with probability } q, \text{ where } Z\sim \mathrm{Tw}(\mu,\phi,\rho)\\ 0, & \text{with probability } 1-q.\end{cases}$$

We refer to this Tweedie zero-inflate by $Y\sim ZIF-Tw(\mu,\varphi,\rho,q)$. The probability density function of Y can be written as

$$f_{\text{ZIF-Tw}}(y\mid\mu,\phi,\rho,q):=qf_{\mathrm{Tw}}(y\mid\mu,\phi,\rho)+(1-q)I\{y=0\},$$

so that $\mathbb{P}(Y=0)=q\exp\left(-\frac{1}{\phi}\frac{\mu^{2-\rho}}{2-\rho}\right)+(1-q)$ and $\mathbb{E}(Y)=q\mu$.

3. Negative Binomial GLMs

In practice, real-life data can be said to be overdispersed i.e. the variance is greater than the mean. The negative binomial regression model is an extension of the Poisson regression model which permits overdispersion by introducing an unknown heterogeneity term that is denoted as observation i. It is assumed that, the observations are different at random, not explained

completely by the observed covariates. Within the negative binomial model[21], Another approach to modeling the overdispersion using a hierarchical model is used. Rather than assuming that the distribution of a value, say, yi ~ Pois(μi) with mean, mu i, it is possible to add an extra layer of stochasticity and assume that the value, mu i, itself is a random variable. Suppose, consequently, that

$$y_i \mid \lambda_i \sim \text{Pois}(\lambda_i) \text{ and } \lambda_i \sim \text{G}(\mu_i, \psi)$$

where G(μi,ψ) denotes a distribution with mean μi and coefficient of variation ψ.[22]
It is straightforward to show, under the hierarchical model, that

$E[yi] = \mu i \; and \; var[yi] = \mu i + \psi\mu_i^2$, so the variance contains an overdisperion term $\psi\mu_i^2$,. The greater the parameter ψ is, the higher is the overdispersion. The most popular alternative that is adopted is to assume that the mixing distribution G is gamma distributed. Since the dispersion of any gamma distribution is equal to its coefficient of variation, the second level of the hierarchical model presupposes a negative-binomial distribution with the follow on specification: λi ~ Gam(μi,ψ). On this assumption can it be proved that y i follows.

$$\mathcal{P}(y_i; \mu_i, k) = \frac{\Gamma(y_i + k)}{\Gamma(y_i + 1)\Gamma(k)} \left(\frac{\mu_i}{\mu_i + k}\right)^{y_i} \left(1 - \frac{\mu_i}{\mu_i + k}\right)^{k},$$

where $k = 1/\psi$ and $\Gamma()$ is the gamma function, so that $\text{var}\,[y_i] = \mu_i + \mu_i^2/k$.
However, conditional on only $\mathbf{x}_i, Y_i$ is distributed as a negative binomial:[21]

$$P(Y_i = y_i \mid \mathbf{x}_i) = \frac{\theta^\theta \mu_i^{y_i} \Gamma(\theta + y_i)}{\Gamma(y_i + 1)\Gamma(\theta)(\mu_i + \theta)^{\theta + y_i}}$$

The distribution is conditional mean μ i and conditional variance $\mu_i(1 + (1/\theta)\mu_i)$. It can be estimated more easily $\alpha = 1/\theta$ instead of θ. The conditional variance is with this substitution $\mu_i(1 + \alpha\mu_i)$. The relationship between negative binomial and Poisson models has the property that as 0 is approached the negative binomial distribution approaches Poisson. Cameron and Trivedi assume a broad type of negative binomial models having mean. $\mu_i + \alpha\mu_i^p$, where in general $-\infty < p < \infty$ [23]. PROC COUNTREG estimates two negative binomial models, corresponding to $p = 2$ (with variance $\mu_i + \alpha\mu_i^2$) and $p = 1$ (with variance $\mu_i + \alpha\mu_i$). The first is estimated with the option DIST=NEGBIN($\text{p} = 2$), and the second is estimated using DIST = NEGBIN $(p = 1)$. The following statements show how to estimate the first:

3.1 Zero-inflated negative binomial regression: ZINB

Zero-inflated count models also give a mechanism of modeling the excess zeros as well as allowing them overdispersion. and there are two possible data generation mechanisms; a Bernoulli trial decides which mechanism is used. Process 1 is selected with probabilityforsourcei φ_i. and Process 2 with probability $1 - \varphi_i$. Process 1 generates only zero counts, whereas Process $2, g(y_i \mid \mathbf{x}_i)$, generates counts from either a Poisson or a negative binomial model. In general:[21]

$$y_i \sim \begin{cases} 0 & \text{with probability } \varphi_i \\ g(y_i \mid \mathbf{x}_i) & \text{with probability } 1 - \varphi_i \end{cases}$$

One method of explaining trailing zeros in the data is inflated zero-count models which model the data to be a mixture of two distributions: a distribution that is one at zero and a Poisson-type count or negative binomial distribution.
A factor that makes a Poisson model have a count data assumption is that a ZIP model has a mass of p on extra zeroes and a mass of (1-p) on a Poisson distribution, when there are so many extra zeroes in the response variable that overdispersion occurs.This means that with covariates a ZIP regression is proposed instead of a Poisson regression. The convenience of ZIP regression model is that the effect of the covariates can be evaluated at the same time in the additional zeroes and the Poisson part of the model. Nevertheless, a ZIP model cannot be used when non-zero of the distribution is overdispersed relative to a Poisson distribution. The negative binomial component supposes that the Poisson mean is a variable with a gamma distribution. We are going to consider ZINB regression model in this Section.[24] Poisson model and binomial model are the two most widespread count data. Poisson The model can be sued and charged unlimited and The negative binomial form is employed when numbers are constrained. There has been only one parameter of the Poisson distribution, and the negative

binomial distribution He has a negative binomial with two teachers. Due to this property, negative The binomial regression model is more flexible compared to Poisson regression model. Moreover, Poisson

The distribution should be the same in terms of the mean and variance Value This is not what actually occurs or simulation data. Therefore when the data being considered is too distracting, the negative binomial regression The model may be applied instead of the Poisson regression model.[25] A longer form of the negative binomial model of excess count of zeros, the negative binomial of zero count with an extension called zero-inflated negative binomial (ZINB), was proposed by Greene (1994). Then, in such instances, the ZINB regression model usually fits better than the ZIP when the non-homogeneity of data as well as the abundance of zeros lead to the overdispersion. In the case of the ZINB regression model, the probability mass function of the response variable yi must be a random non-negative integer value variable and that yi = 0 must occur at a very high frequency than can be modelled using the standard model. Thus, the regression model is defined as

$$P(Y = y_i \mid x_i, z_i) = \begin{cases} \pi_{i_1} + (1 - \pi_{i_i}) f(0; \theta_i), & y_i = 0, (1) \\ (1 - \pi_{i_i}) f(y_i; \theta_i), & y_i > 0, \end{cases}$$

Where $f(y_i; \theta_i), y_i = 0,1,2, \ldots$ is the pdf of Y_i and $0 < \pi_{i_i} < 1$. Furthermore, the function $\pi_{i_i} = \pi_{i_i}(z_i)$ satisfies $\text{logit}\left(\pi_{i_i}\right) = \log\left(\pi_i \left[1 - \pi_{i_i}\right]^{-1}\right) = \sum_{j=1}^{m} z_{ij}\delta_j$ where $z_i = (z_{i1} = 1, z_{i2}, \ldots, z_{im})$ is the i-th row of covariate matrix Z and $\delta = (\delta_1, \delta_2, \ldots, \delta_m)$ are unknown Column vector m-dimensions of parameters. In this setup, the non-negative function φ_i is modeled Via login link function. This function is linear and other appropriate link functions that allow it π_{i_i} to be negative can be used. [26][25]

With parameters $\mu \geq 0$ for the mean and $\alpha > 0$ over-dispersion ZINB regression model can be expressed in vector form as:

$$\log\left(\mu_{ij}\right) = \beta_o + \sum_{l=1}^{k} \beta_l x_{lj} + U_{oj} + \sum_{l=1}^{k} U_{lj} x_{lj}$$

Here, the covariates X_{ij} and Z_{ij} the β and 7 are the regression vectors that correspond to the 1.5 appearing in the respective negative binomial and logistic components will not necessarily be the same.

4.1Tweedie 1 Shrinkage Regression Model

There are other applications of Tweedie model in general areas like actuarial science. Tweedie 1 Shrinkage regression model is suggested as an alternative to Tweedie regression and regular regression models that could be used to solve the multivariate problem arising out of the linear relationship between the independent variables. The models (RTR), (LTR) and (ETR) are included in the regular regression model provided by Tweedie. It solves this issue primarily when the data is the dependent variable, particularly, non-negative right-skewed data and quasi-continuous data (continuous data with a probability mass at 0). Method of the Tweedie model in the case of the generalized linear model, we analysed the models (RTR), (LTR) and (ETR),

Together with the strict rule, the suggested statistical model is deployed using the R package HDtweedie. Additionally, for the Tweedie model in the context of the generalized linear model, we studied the (RTR), (LTR) and (ETR) methods.

In order to compute the estimation coefficients effectively, we have performed some simulation experiments to examine the variables and model fitting performance of different lasso methods of the Tweedie model. The primary motivation is to offer an effective system of models of various regular regression problems that are Tweedie models, the (RTR), (LTR) and (ETR) and the reliance on simulations in the choice and execution of model fitting by discriminant criteria.

The fact that the application rate is overdispersed and more zero inflated than Poisson distribution was incorporated by the use of the Tweedie compound Poisson model. Poisson model is formulated by Tweedie compound $Y = \sum_{i=1}^{n} Xi$ where Y is the responds vector, Xi is i.i.d. Gamma distributed with parameters α, and γ and n is a Poisson random variable. Equation (3) formulates the penalized negative log-likelihood [27][6]

$$\text{PLL } L(\lambda,\boldsymbol{\beta},\alpha)=\sum_{i=1}^{n}\Big[\ \frac{y_i\exp\left[-(\xi-1)\left(\mathbf{x}_i^T\beta\right)\right]}{\xi-1}+\frac{\exp\left[(2-\xi)\left(\mathbf{x}_i^T\beta\right)\right]}{2-\xi}\Big]$$
$$+\lambda\left[(1-\alpha)\frac{1}{2}\parallel\boldsymbol{\beta}\parallel^2+\alpha\parallel\boldsymbol{\beta}\parallel_1\right]\ldots\ldots(3.4)$$

If ɑ=0 $\text{PLL}\,L(\lambda,\boldsymbol{\beta},\alpha)$ is (RTR)

$$\text{PLL } L(\lambda,\boldsymbol{\beta},\alpha=0)=\sum_{i=1}^{n}\Big[\ \frac{y_i\exp\left[-(\xi-1)\left(\mathbf{x}_i^T\beta\right)\right]}{\xi-1}+\frac{\exp\left[(2-\xi)\left(\mathbf{x}_i^T\beta\right)\right]}{2-\xi}\Big]$$
$$+\lambda\left[\frac{1}{2}\parallel\boldsymbol{\beta}\parallel^2\right]\ldots\ldots(3.5)$$

If ɑ=1 $\text{PLL}\,L(\lambda,\boldsymbol{\beta},\alpha)$ is (LTR)

$$\text{PLL } L(\lambda,\boldsymbol{\beta},\alpha=1)=\sum_{i=1}^{n}\Big[\ \frac{y_i\exp\left[-(\xi-1)\left(\mathbf{x}_i^T\beta\right)\right]}{\xi-1}+\frac{\exp\left[(2-\xi)\left(\mathbf{x}_i^T\beta\right)\right]}{2-\xi}\Big]$$
$$+\lambda[\parallel\boldsymbol{\beta}\parallel_1]\ldots\ldots(3.6)$$

The HDTweedie R package has been utilized to minimize (3). Combined with a blockwise majorization descent (BMD), this method implements an iterative reweighted least squares (IRLS) algorithm [6]. The standard procedure from HDTweedie has been utilized to determine proper values for λ. In the beginning, λmax is computed, such that β = 0. After that, λmin = 0.001λmax is set. for λ, the other m − 2 values are determined by projecting them uniformly on a log scale on the range [λmin, λmax]. With an increase of 0.2, the values of ɑ from 0.1 to 0.9 were used. Finally, we set p as 1, 2, and 3.

5.Model Selection Criteria

5.1Validation of the model

To select the model, determine its accuracy by evaluating the precision of the estimates. In 1970, Japanese Statistician Hirotugu Akaike developed the metric Akaike's Information Criteria (AIC), which is $AIC=2k-2\log\left(\ell(y;\hat{\mu})\right).\ell$ is the likelihood of the fitted model defined above, and k is the number of parameters included in the model [28]. AIC's fundamental concept is to penalize adding new variables to a model. When including additional terms, AIC adds a penalty that increases the error. For a better model, AIC should be lower [28].

$AIC_{Tweedie}$ is used as a performance for a Tweedie regression model. We have used it as the criterion for choosing a Tweedie model in the R package (Tweedie) uses the evaluation of Tweedie Exponential Family Models [29].

5.2Regression Accuracy Metrics [27]

In general, variable selection and prediction accuracy evaluation criteria are used to evaluate the performance of the suggested regularization regression methods and their capacity to compete with other current penalized methods. The quality of the fitted model can be obtained by computing the goodness of fit between the predicted values it generates and the given input data. In regression analysis, the MAE, RMSE, and R-Squared are the most common metrics used to measure prediction error rates and model performance. MAE is the difference between the original and anticipated values calculated by the average absolute difference throughout the data set.

$$MAE=\frac{1}{N}\sum_{i=1}^{N}|y_i-\hat{y}| \qquad (4.1)$$

- **RMSE** (Root Mean Squared Error) is the error rate by the square root of

$$RMSE=\sqrt{MSE}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i-\hat{y})^2} \qquad (4.2)$$

- **R-squared** (Coefficient of determination) is the coefficient that indicates how closely the results fit the original values. The value from 0 to 1 are interpreted as percentages. The model is superior the higher the value. R-squared is written as follows:

$$R^2=1-\frac{\sum(y_i-\hat{y})^2}{\sum(y_i-\bar{y})^2} \qquad (4.3)$$

Where, $\hat{y}$ is the predicted value of y $\bar{y}$ – mean value of y

6Experiment Results and Discussion

6.1Real Application:

The state wildlife biologists desired to simulate the numbers of fish being taken by fishermen in a state park. Tourists receive questions about the duration of their visits, the number of people in the party, whether there were children in the party and the quantity of fish they caught. Other visitors do not fish, but they have no information about whether an individual fished or not. There were some visitors that did not fish and therefore, there are numerous excess zeros in the data.

There were 250 groups that attended a park now we have data. Along with the number of fish caught, there are also questions of the number of children in the group (child), the number of people in the group (persons), whether or not a group had a camper to the park (camper).The occurrence of excess zeros is of interest, i.e. the likelihood that a group caught no fish. Our model will be expressed in variables which we will be using as child, persons, and camper.

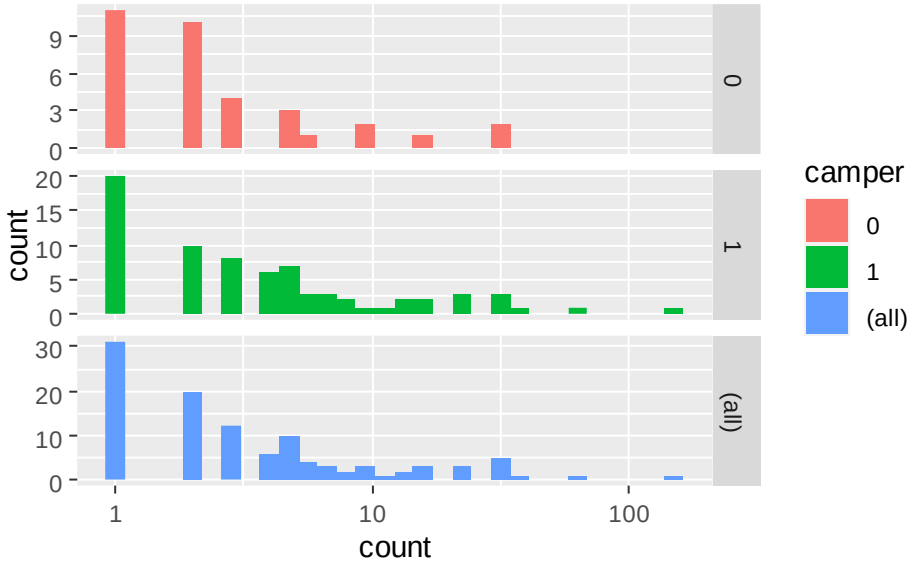


Figure6.1: Histogram of values of the dependent variable

Table 6.1: Criteria for comparing methods

data = zinb	P,alph	AIC	R^2	RMSE	MAE
nb		559.6804	0.0403132	11.37527	3.952754
zinb		561.6805	0.03	17.98655	2.194915
ols		1852.318	0.3327819	9.484862	
Tweedie	1.2, 0.8132008	629.0867	0.03552763	11.4036	1.082508
ridge			0.3327785	9.484886	4.55087
ridge_tweedie			0.7346038	5.981973	1.370801
Lasso			0.3252601	9.538175	4.470756
Lasso tweedie			0.9224108	3.234431	0.999399
elastic			0.3252594	9.53818	4.467894
elastic_tweedie			0.9228969	3.224283	0.9938082

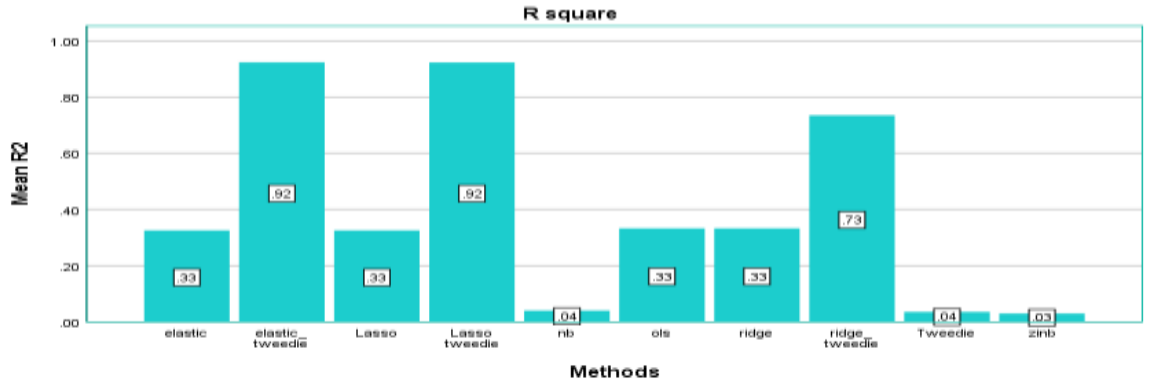


Figure 6.2: Coefficient of determination values(R^2) for all methods

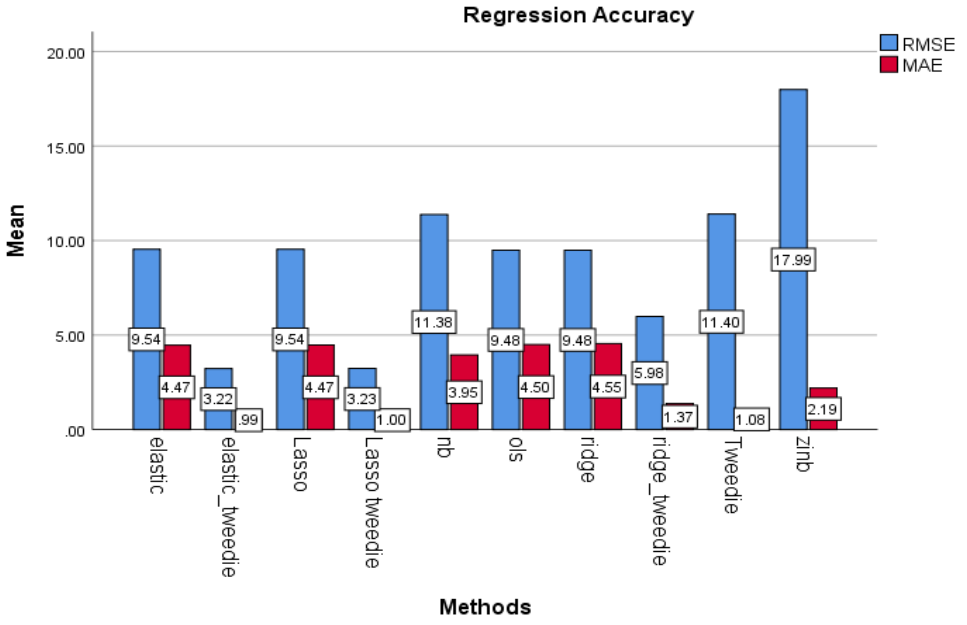


Figure 6.:3 Root Mean Squared Error (RMSE)and Mean Absolute Error (MAE) for all methods

Discussion

which is the evaluation of the comparison of the methods, we observe that NbRegression provided the lowest value of the criterion (AIC) and equal to (559.6804) compared to all other regression methods, . Also, the ridge_tweedie ,Tweedie lasso method and the Tweedie elastic net method gave the highest value for Coefficient of determination values(R^2) (0.7346038,0.9224108, 0.9228969) respectively, and gave the lowest value to Root Mean Squared Error (RMSE) and it is (5.981973,3.234431, 3.224283) respectively and gave the least value to Mean Absolute Error (MAE) and it is (1.370801,0.999399, 0.9938082) respectively and it is less than all the methods and as in the Figure 6.2, Figure 6.3 respectively.

Conclusions and Recommendations:

This research was also done to compare the performance of the zero-inflated negative binomial models (ZINB)and the regression models with the Tweedie Shrinkage regression (Ridg,Lasso, Elastic) It was found that the systematic regression of Toddy gives the best values in the performance based on the criteria of prediction and this is a piece of evidence that data that are affected by dispersion can be used as an alternative to counter the same. It is also possible to make other comparisons. We have utilized most suitable statistical model to research these data and our results will be more precise and readable. As the results of this research indicate, we

strongly suggest that analysis of zero-inflated data should be performed using Tweedie Shrinkage regression (Ridg,Lasso, Elastic) to be able to deliver statistically valid results.

It would be an issue of future inquiry to model Tweedie Shrinkage regression (Ridge Lasso, Elastic) and then to explore the effect of using other statistical techniques instead of the model performance in question.

Reference:

[1] M. Al Dabal and M. Al Dabal, "A Comparative Study of Ridge, LASSO and Elastic net Estimators," pp. 1–49, 2020, [Online]. Available: http://dx.doi.org/10.22215/etd/2021-14573 LB - hgCD.

[2] P. Banerjee, B. Garai, H. Mallick, S. Chowdhury, and S. Chatterjee, "A Note on the Adaptive LASSO for Zero-Inflated Poisson Regression," *J. Probab. Stat.*, vol. 2018, 2018, doi: 10.1155/2018/2834183.

[3] Ismah, K. A. Notodiputro, and B. Sartono, "AMAZonn (A Multicollinearity-adjusted Adaptive LASSO for Zero-infated Count Regression) with Weight of Expectation Maximization Standard Error Adaptive LASSO (SEAL AL) for Zero Inflated Poisson Data," *J. Phys. Conf. Ser.*, vol. 1776, no. 1, 2021, doi: 10.1088/1742-6596/1776/1/012050.

[4] Jørgensen B, "Exponential dispersion models," *London: Chapman and Hall;*, 1997.

[5] Tweedie MCK, "An index which distinguishes between some important exponential families," *Ghosh JK, Roy J, Ed. Stat. Appl. new Dir. Proc. Indian Stat. Inst. Golden Jubil. Int. Conf. Calcutta Indian Stat. Institute;*, 1984.

[6] W. Qian, Y. Yang, and H. Zou, "Tweedie's Compound Poisson Model With Grouped Elastic Net," *J. Comput. Graph. Stat.*, vol. 25, no. 2, pp. 606–625, 2016, doi: 10.1080/10618600.2015.1005213.

[7] J. K. Tay, B. Narasimhan, and T. Hastie, "Elastic Net Regularization Paths for All Generalized Linear Models," 2021, [Online]. Available: http://arxiv.org/abs/2103.03475.

[8] A. J. I. H. Wigena, A. Djuraidah, and A. Kurnia, "A NEW APPROACH TO STATISTICAL DOWNSCALING USING TWEEDIE COMPOUND POISSON GAMMA RESPONSE AND LASSO REGULARIZATION MA'RUFAH," pp. 1–16, 2021.

[9] R. Hoerl, A. and Kennard, "Ridge regression: Biased estimation for nonorthogonal problems.," *Technometrics*, vol. 12, no. 1, pp. 55–67, 1970.

[10] S. Methods, "Regularization Shrinkage Methods," pp. 1–45, 2021.

[11] A. Kassambara, "Machine Learning Essentials. Practical Guide in R.," *Sthda*. p. 210, 2017, [Online]. Available: https://www.m-culture.go.th/mculture_th/download/king9/Glossary_about_HM_King_Bhumibol_Adulyadej's_Funeral.pdf.

[12] and R. T. James, Gareth, Daniela Witten, Trevor Hastie, *An Introduction to Statistical Learning: With Applications in R. Springer Publishing Company, Incorporated.* 2014.

[13] F. Corea, "Introduction to Data," pp. 1–5, 2019, doi: 10.1007/978-3-030-04468-8_1.

[14] H. Zou and T. Hastie, "Erratum: Regularization and variable selection via the elastic net," *J. R. Stat. Soc. Ser. B Stat. Methodol.*, vol. 67, no. 5, p. 768, 2005, doi: 10.1111/j.1467-9868.2005.00527.x.

[15] H. Madsen and P. Thyregod, "Introduction to general and generalized linear models," *Introd. to Gen. Gen. Linear Model.*, no. October, pp. 1–291, 2010, doi: 10.1201/9781439891148.

[16] P. Dunn and G. Smyth, "Tweedie family densities: methods of evaluation," *Proc. 16th Int. Work. …*, no. December 2014, pp. 1–8, 2001.

[17] G. K. Smyth, "Regression Analysis of Quantity Data with Exact Zeros," *Proc. Second Aust. Work. Stoch. Model. Eng. Technol. Manag. Ed. by R. J. Wilson, S. Osaki, D. N. P. Murthy, 572–580. Queensland, Aust. Technol. Manag. Centre, Univ. Queen*, 1996.

[18] A. Halder, S. Mohammed, K. Chen, and D. Dey, "Spatial risk estimation in Tweedie compound Poisson double generalized linear models," 2019, [Online]. Available: http://arxiv.org/abs/1912.12356.

[19] W. H. Bonat and C. C. Kokonendji, "Flexible Tweedie regression models for continuous data," *J. Stat. Comput. Simul.*, vol. 87, no. 11, pp. 2138–2152, 2017, doi: 10.1080/00949655.2017.1318876.

[20] H. Zhou, W. Qian, and Y. Yang, "Tweedie gradient boosting for extremely unbalanced zero-inflated data," *Commun. Stat. Simul. Comput.*, pp. 1–23, 2020, doi: 10.1080/03610918.2020.1772302.

[21] D. Erdman, L. Jackson, and A. Sinko, "SAS Global Forum 2008 SAS Presents Zero-Inflated Poisson and Zero-Inflated Negative Binomial Models Using the COUNTREG Procedure SAS Global Forum 2008 SAS Presents," *Distribution*, vol. 9, no. 1, pp. 1–11, 2008, [Online]. Available: http://ajph.aphapublications.org/cgi/content/abstract/83/8/1163?ijkey=dfd46af33da3cbce90e7e183c4a37b586ed492f6&keytype2=tf_ipsecsha.

[22] P. K. Dunn and G. K. Smyth, *Generalized Linear Models With Examples in R, https://link.springer.com/book/10.1007/978-1-4419-0118-7*. 2018.

[23] P. K. Cameron, A. C. and Trivedi, "Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators," *J. ofApplied Econom. Greene,* vol. 1, pp. 29–53, 1986.

[24] S. M. Mwalili, E. Lesaffre, and D. Declerck, "The zero-inflated negative binomial regression model with correction for misclassification: An example in caries research," *Stat. Methods Med. Res.*, vol. 17, no. 2, pp. 123–139, 2008, doi: 10.1177/0962280206071840.

[25] Seyed Ehsan Saffari and Robiah Adnan, "Zero-Inflated Negative Binomial Regression Model with Right Censoring Count Data," *J. Mater. Sci. Eng. B*, vol. 1, no. 9, 2011, doi: 10.17265/2161-6221/2011.09.020.

[26] D. Kokotsaki, V. Menzies, and A. Wiggins, "Durham Research Online Woodlands," *Crit. Stud. Secur.*, vol. 2, no. 2, pp. 210–222, 2014, doi: 10.1002/bimj.200100000.

[27] S. Bhulai, H. Rusman, and L. Willemsens, "Predicting Candidate Uptake For Online Vacancies," no. c, pp. 63–68, 2016.

[28] J. Lanford, T. Nykodym, A. Rao, and A. Wang, "Generalized Linear Modeling with H2O's R Package," no. February, 2021.

[29] P. K. Dunn and G. K. Smyth, "Evaluation of Tweedie exponential dispersion model densities by Fourier inversion," *Stat. Comput.*, vol. 18, no. 1, pp. 73–86, 2008, doi: 10.1007/s11222-007-9039-6.