Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 53 (5) (2024), 1419 – 1437
DOI : 10.15672/hujms.1261283

Research Article

Iterative stochastic restricted $r - d$ class estimator in generalized linear models: application to binomial, Poisson and negative binomial distributions

Atif Abbasi*[1], M. Revan Özkale [2]

[1] *The University of Azad Jammu and Kashmir Muzaffarabad, Department of Statistics, 13100, AJK, Pakistan*

[2] *Çukurova University, Faculty of Science and Letters, Department of Statistics, 01330, Adana, Turkey*

Abstract

In this paper, we provide an iterative stochastic restricted $r - d$ (SR-rd) class estimator that incorporates prior and sample information to address the multicollinearity problem. The newly proposed estimator is a manifold estimator that contains various estimators under specific conditions. The new estimator is compared to the maximum likelihood, principal components regression, and $r - d$ class estimators. To assess the performance, two numerical examples and two simulation studies are performed where the scalar mean square error and expected mean square error are the performance evaluation criteria. The analysis results show that the value of d affects the performance of the estimators. The farther the d value is from zero, the better the SR-rd estimator is compared to other estimators, and the SR-rd estimator is a good estimator at the optimal d value.

Mathematics Subject Classification (2020). 62J07, 62J12, 62F10

Keywords. Stochastic restrictions; shrinkage parameter, mean square error, binomial distribution, Poisson distribution, negative binomial distribution, r-d class estimator

1. Introduction

Generalized linear model (GLM) is an advanced statistical modeling technique formulated by [18]. It is an umbrella term that encompasses many other models, which allows the response variable Y to have a distribution other than a normal distribution. The GLMs include linear regression, logistic regression, Poisson regression, gamma regression, negative binomial regression, and Log-linear models etc. The purpose of the GLM is to describe the relationship between the response variable and the set of covariates. This is done through the linear predictor and mean, and a nonlinear relationship can be reconstructed as a linear relationship.

*Corresponding Author.
Email addresses: atifbbs3@gmail.com (A.Abbasi), mrevan@cu.edu.tr(M.R.Özkale)
Received: 07.03.2023; Accepted: 05.08.2024

The development of the GLM theory is built on the fact that the response variable belongs to the family of exponential distributions. The probability function for a response variable Y belonging to the family of exponential distributions is often expressed as

$$f_{Y_i}(y_i, \theta_i, \phi) = \exp\left[\frac{\theta_i y_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi)\right], i = 1, 2, ..., n \tag{1.1}$$

where θ_i is a canonical parameter, ϕ is a nuisance or dispersion parameter and $a(.)$, $b(.)$ and $c(.)$ are the regular functions pertaining to the type of probability density function being considered.

The most frequently used approach for estimating parameters in GLMs is maximum likelihood (ML) estimation. However, the multicollinearity issue has a significant impact on ML estimations resulting in large standard error and to be far away from the true parameter value. There have been many estimation methods proposed in GLMs and in linear regression to combat multicollinearity. Utilizing previous data in addition to the sample data is one technique to deal with the multicollinearity problem. Applying this information would allow one to choose between stochastic and exact linear restrictions.

Some of the studies carried out within the framework of linear regression under exact linear restrictions or stochastic linear restrictions are as follows; Özkale [24] established the restricted principal component regression (RPCR) estimator with exact linear constraints. Daojiang and Wu [6] developed a stochastic restricted (PCR) estimator in linear models with stochastic linear constraints. Gargi and Chandra [8] introduced a two parameter stochastic restricted principal component estimator to tackle multicollinearity when additional stochastic linear constraints are available. On the other hand, the principal component regression (PCR) estimator has been combined with other estimators in some of the studies that have been proposed like; by merging the ridge and principal component techniques, Baye and Parker [5] proposed the $r-k$ class estimator. Arum and Ugwuowo [3] developed a robust $r-k$ class estimator to deal with multicollinearity. Arum et al. [4] combined the Kibria-Lukman estimator with the PCR estimator in linear regression model. Farghali et al. [7] combined the James-Stein estimator and the PCR estimator in linear regression. Lukman et al. [13] combined the PCR estimator and the modified ridge-type estimator in linear regression. Akram et al. [2] combined the PCR estimator and the ridge estimator in the inverse Gaussian regression model. Under stochastic linear constraints, Shalini and Sarkar [26] suggested an $r-k$ class estimator based on a combination of ridge and PCR estimators. Jianwen and Yang [10] presented the restricted $r-d$ and $r-k$ class estimators under exact linear restrictions.

Some studies have also been done in GLMs using exact or stochastic linear restrictions to solve the issue of multicollinearity, such as; [11,12,19,21,23]. The PCR estimator has also been established in GLMs solely as well as by combining it with other estimators to handle the multicollinearity issue, for example; Smith and Marx [25] presented the generalized ridge and PCR estimators in GLMs. Özkale [20] used the PCR and Liu estimators to create the $r-d$ class estimator in GLMs. Abbasi and Özkale [1] introduced a $r-k$ class estimator in GLMs by merging the PCR and ridge estimators.

Although PCR, Liu and $r-d$ class estimators have been developed to solve the multicollinearity problem, enhancing these estimators is the primary focus of any research. Özkale and Nyquist [21] mentioned that incorporating the stochastic linear constraints into the sample information improves parameter estimation. Furthermore, there is no research in the literature that combines the PCR estimator with the Liu estimator under stochastic linear constraints in GLMs, which is a notable gap in the literature. Thus, in this study a stochastic restricted $r-d$ class estimator is introduced which contains the ML, PCR, and $r-d$ class estimators as special cases. It is hoped that the new estimator will enhance the estimation when there is additional information available in the form of stochastic linear restrictions. In addition, the newly proposed estimator might be applicable to any nonlinear regression models like; logistic, Poisson, gamma and negative binomial regressions etc.

The primary goal of this research is to develop an iterative stochastic restricted $r-d$ class estimator in GLMs by merging the PCR and Liu estimators when stochastic linear restrictions are available as prior information in addition to sample information.

This paper is organized as follows: Sect. 2, presents an iterative stochastic restricted $r-d$ class estimator and some of its special cases. Sect. 3, presents the mean square error (MSE) of the proposed estimator in its first-order approximated (FOA) form. Sect. 4, gives two numerical examples via binomial and Poisson distributed responses. Sect. 5, shows the simulation studies. Sect. 6, concludes the study.

2. Iterative stochastic restricted $r-d$ class estimator in GLMs

Many estimators in GLMs have been developed to improve parameter estimation. However, the most often used estimating technique is the ML estimation, which is based on the iterative re-weighted least squares (IRLS) and defined as follows: $\hat{\beta}^{(t+1)} = \left(X^T\hat{W}^{(t)}X\right)^{-1} X^T\hat{W}^{(t)}\hat{z}^{(t)}$ where t is the iteration step, $\hat{W}^{(t)} = diag(w_{ii}^{(t)})$ is a diagonal weight matrix having weights $w_{ii} = \frac{1}{var(Y_i)}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2 = \frac{1}{var(Y_i)}(g^T(\mu_i))^2$ computed at $\hat{\beta}^{(t)}$ and $\hat{z}^{(t)}$ is a $n\times 1$ working response with elements $z_i^{(t)} = x_i^T\hat{\beta}^{(t)} + \left(y_i - \mu_i^{(t)}\right)\frac{\partial\eta_i^{(t)}}{\partial\mu_i^{(t)}}$, while $\mu_i^{(t)}$ and $\frac{\partial\eta_i^{(t)}}{\partial\mu_i^{(t)}}$ are evaluated at $\hat{\beta}^{(t)}$. The ML estimator has been developed by considering the relationship;

$$g(\mu_i) = x_i^T\beta = \eta_i,$$

where β is the vector of regression coefficients and x_i^T is the i-th row of the explanatory variables matrix X.

Thus, Smith and Marx [25] proposed an iterative PCR estimator in GLMs by considering the linear predictor $\eta = X\beta$ as $\eta = XMM^T\beta = Z\alpha$, where $Z = XM$, $\alpha = M^T\beta$ and $Z^T\hat{W}Z = \Lambda = diag(\lambda_j)$ is a $p\times p$ diagonal matrix containing the eigenvalues of $X^T\hat{W}X$ $(\lambda_1 = \lambda_{\max} \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_p = \lambda_{\min})$ with $\hat{W}$ the weight matrix evaluated at the ML estimator at convergence and $M = [m_1, \cdots, m_p]$ is a $p\times p$ orthogonal matrix and m_j are the eigenvectors corresponding to the eigenvalues λ_j. Therefore, for the i-th element of the Z matrix the linear predictor η_i might be expressed as $\eta_i = x_i^TMM^T\beta = z_i^T\alpha$ where $z_i^T = x_i^TM$ is the row vector of the Z matrix. The use of a reduced set of principal components (PCs) is a very handy option to deal with the multicollinearity problem. Thus, the Z matrix and α vector are partitioned as $Z = \begin{bmatrix} Z_r & Z_{p-r}\end{bmatrix}$ and $\alpha = \begin{bmatrix}\alpha_r^T & \alpha_{p-r}^T\end{bmatrix}$ where $Z_r = XM_r (r \leqslant p)$ consists of PCs with higher eigenvalues that will be kept in the model. Accordingly, the M and Λ matrices can be built as $M = \begin{bmatrix} M_r & M_{p-r}\end{bmatrix}$ and $\Lambda = \begin{bmatrix}\Lambda_r & 0\\ 0 & \Lambda_{p-r}\end{bmatrix}$, where $\Lambda_r = Z_r^T\hat{W}Z_r = M_r^TX^T\hat{W}XM_r$ and $\Lambda_{p-r} = Z_{p-r}^T\hat{W}Z_{p-r} = M_{p-r}^TX^T\hat{W}XM_{p-r}$. Thus, we consider a smaller set of PCs that is $\eta_{ri} = z_{ri}^T\alpha_r$ where z_{ri}^T is the row vector of the matrix Z_r and $\alpha_r = M_r^T\beta$.

By using the reduced set of PCs the PCR and $r-d$ class estimators have been developed respectively by [25] and [20] as

$$\hat{\beta}_r^{(t)} = M_r(M_r^TX^T\hat{W}^{(t)}XM_r)^{-1}M_r^TX^T\hat{W}^{(t)}u_r^{(t)}.$$

$$\hat{\beta}_{r-d}^{(t)} = M_r(M_r^TX^T\hat{W}^{(t)}XM_r + I_r)^{-1}(M_r^TX^T\hat{W}^{(t)}u_r^{(t)} + dM_r^T\hat{\beta}). \tag{2.1}$$

Now we consider the stochastic restrictions on the parameters in order to develop the stochastic restricted $r-d$ class estimator in GLMs.

We consider the stochastic restrictions on the parameter as a prior information in order to develop the stochastic restricted $r-d$ class estimator in GLMs. Assume that the prior information on β is such that $h = H\beta + \phi^*, \phi^* \sim N(0, a(\phi)\Sigma)$ where where h is a $q \times 1$ random vector independent with the sample information and H is a characterized $q \times p$ known matrix with $rank(H) = q$ having a set of q linearly independent constraints on the parameters:

$$H = \begin{bmatrix} H_1 \\ H_2 \\ \vdots \\ H_q \end{bmatrix},$$

where $H_i = [H_{i1}, H_{i2}, \cdots, H_{iq}]$.

If we are working in a subspace of parameters, this stochastic constraint must also be handled appropriately to that subspace by the form for the reduced model

$$h = H_r\alpha_r + \phi^*, \phi^* \sim N(0, a(\phi)\Sigma), \tag{2.2}$$

where h is a random vector of size $q \times 1$, $\alpha_r = M_r^T\beta$ and $H_r = HM_r$ denotes a $q \times p$ matrix having $rank(H_r) = q$.

By assembling the sample and prior information, we consider the following objective function to obtain the stochastic restricted $r-d$ class estimator:

$$\begin{aligned} F(\alpha_r;y,h,d) &= l(\alpha_r) - \frac{1}{2a(\phi)}(\alpha_r - d\hat{\alpha}_r)^T(\alpha_r - d\hat{\alpha}_r) - \frac{q}{2}\ln(2\pi) - \frac{q}{2}\ln|a(\phi)\Sigma| \\ &\quad - \frac{1}{2a(\phi)}(h - H_r\alpha_r)^T\Sigma^{-1}(h - H_r\alpha_r), \end{aligned} \tag{2.3}$$

where

$$l(\alpha_r) = \sum_{i=1}^{n}\left\{\frac{\theta_i y_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi)\right\},$$

is the log-likelihood function of the reduced model and d is a basing parameter and $\hat{\alpha}_r = M_r^T\hat{\beta}$ is the ML estimator of the reduced model.

By considering Eq. (2.3) to estimate α_r with elements α_{r_j}, $j = 1, 2\cdots, r$ taking derivatives of $F(\alpha_r; y, h, d)$ with respect to α_{rj} and using the chain rule we obtain,

$$\frac{\partial F(\alpha_r; y, h, d)}{\partial \alpha_{rj}} = \sum_{i=1}^{n}\frac{(y_i - \mu_i)}{a(\phi)b^T(\theta_i)}\frac{1}{g^T(\mu_i)}z_{r,ij} - \frac{\sum_{j=1}^{p}(\alpha_{rj} - d\hat{\alpha}_{rj})}{a(\phi)} + \frac{1}{a(\phi)}T_j \tag{2.4}$$

where T_j are the components of p vector $T = H_r^T\Sigma^{-1}(h - H_r\alpha_r)$ while $z_{r,ij}$ is the ij-th element of Z_r.

In matrix notation, Eq. (2.4) can be written as:

$$S(\alpha_r, d) = \frac{1}{a(\phi)}[Z_r^TWD^{-1}(y - \mu) - \alpha_r + d\hat{\alpha}_r + H_r^T\Sigma^{-1}(h - H_r\alpha_r)]$$

where $D = diag(1/g^T(\mu_i))$ and $W = diag(w_{ii}^{-1})$.

Now taking the derivative of Eq. (2.4) with respect to α_{rv}, we get:

$$\begin{aligned} \frac{\partial^2 F(\alpha_r; y, h, d)}{\partial \alpha_{r,j}\partial \alpha_{r,v}} &= \frac{1}{a(\phi)}\sum_{i=1}^{n}(y_i - \mu_i)\frac{\partial}{\partial \alpha_{rv}}\frac{1}{w_{ii}}g^T(\mu_i)z_{r,ij} - \sum_{i=1}^{n}\frac{z_{r,ij}z_{r,iv}}{a(\phi)w_{ii}} \\ &\quad - \frac{1}{a(\phi)}\delta_{jv} - \frac{1}{a(\phi)}u_{jv}, \end{aligned} \tag{2.5}$$

since $\delta_{jv} = 1$ if $j = v$ and 0 otherwise and u_{jv} are the components of $p \times p$ matrix $U = H_r^T \Sigma^{-1} H_r$. Minus times the expected value of the second-order derivative gives:

$$Q_{jv}(\alpha_{rj}, d) = -E\left[\frac{\partial^2}{\partial\alpha_{rj}\partial\alpha_{rv}} F(\alpha_r; y, h, d)\right] = \frac{1}{a(\phi)}\left[\sum_{i=1}^{n} \frac{z_{r,ij} z_{r,iv}}{w_{ii}} + I_r + u_{jv}\right].$$

It can be expressed in matrix form as:

$$Q(\alpha_r, d) = \frac{1}{a(\phi)}[(Z_r^T W Z_r + I_r) + H_r^T \Sigma^{-1} H_r].$$

By using the technique of Fisher's scoring algorithm, we have:

$$\hat{\alpha}_{SR-rd}^{(t+1)} = \hat{\alpha}_{SR-rd}^{(t)} + \left\{[Q(\alpha_r, d)]^{-1}[S(\alpha_r, d)]\right\}_{\alpha_r = \hat{\alpha}_{SR-rd}^{(t)}},$$

where t represents the iteration step. Premultiplying both sides by $Q(\alpha_r, d)$ gives us:

$$[Q(\alpha_r, d)]_{\alpha_r = \hat{\alpha}_{r-d}^{(t)}} \hat{\alpha}_{SR-rd}^{(t+1)} = [Q(\alpha_r, d)]_{\alpha_r = \hat{\alpha}_{r-d}^{(t)}} \hat{\alpha}_{SR-rd}^{(t)} + [S(\alpha_r, d)]_{\alpha_r = \hat{\alpha}_{r-d}^{(t)}}.$$

We proceed with the computation by replacing the values of the Q and S matrices as follows:

$$\begin{aligned}[(Z_r^T \hat{W} Z_r + I_r) + H_r^T \Sigma^{-1} H_r]\hat{\alpha}_{SR-rd}^{(t+1)} &= [(Z_r^T \hat{W} Z_r + I_r) + H_r^T \Sigma^{-1} H_r]\hat{\alpha}_{SR-rd}^{(t)} \\ &\quad + [Z_r^T \hat{W} D^{-1}(y - \hat{\mu}) - \alpha_r + d\hat{\alpha}_r \\ &\quad + H_r^T \Sigma^{-1}(h - H_r \hat{\alpha}_{SR-rd}^{(t)})],\end{aligned}$$

where $\hat{W}$ is the weight matrix calculated at the ML estimator. This results in

$$\begin{aligned}\hat{\alpha}_{SR-rd}^{(t+1)} &= [(Z_r^T \hat{W} Z_r + I_r) + H_r^T \Sigma^{-1} H_r]^{-1}\{Z_r^T \hat{W} Z_r \hat{\alpha}_{SR-rd}^{(t)} \\ &\quad + Z_r^T \hat{W} D^{-1}(y - \hat{\mu}^{(t)}) + d\hat{\alpha}_r + H_r^T \Sigma^{-1} h\}.\end{aligned} \tag{2.6}$$

By applying the inverse formula, we get:

$$\begin{aligned}[(Z_r^T \hat{W} Z_r + I_r) + H_r^T \Sigma^{-1} H_r]^{-1} &= (Z_r^T \hat{W} Z_r + I_r)^{-1} - (Z_r^T \hat{W} Z_r + I_r)^{-1} H_r^T \\ &\quad \times [\Sigma + H_r (Z_r^T \hat{W} Z_r + I_r)^{-1} H_r^T]^{-1} \\ &\quad \times H_r (Z_r^T \hat{W} Z_r + I_r)^{-1}.\end{aligned}$$

Solving:

$$\begin{aligned}&\{(Z_r^T \hat{W} Z_r + I_r)^{-1} - (Z_r^T \hat{W} Z_r + kI_r)^{-1} H_r^T [\Sigma + H_r (Z_r^T \hat{W} Z_r + I_r)^{-1} H_r^T]^{-1} \\ &\times H_r (Z_r^T \hat{W} Z_r + I_r)\} H_r^T \Sigma^{-1} h \\ = \; &(Z_r^T \hat{W} Z_r + I_r)^{-1} H_r^T \{I - [\Sigma + H_r (Z_r^T \hat{W} Z_r + I_r)^{-1} H_r^T]^{-1} \\ &\times H_r (Z_r^T \hat{W} Z_r + I_r) H_r^T\} \Sigma^{-1} h \\ = \; &(Z_r^T \hat{W} Z_r + I_r)^{-1} H_r^T [\Sigma + H_r (Z_r^T \hat{W} Z_r + I_r)^{-1} H_r^T]^{-1} h,\end{aligned}$$

transforms Eq. (2.6) to:

$$\begin{aligned}\hat{\alpha}_{SR-rd}^{(t+1)} &= (Z_r^T \hat{W} Z_r + I_r)^{-1}\{Z_r^T \hat{W} Z_r \hat{\alpha}_{SR-rd}^{(t)} + Z_r^T \hat{W} D^{-1}(y - \hat{\mu}^{(t)}) + d\hat{\alpha}_r\} \\ &\quad - (Z_r^T \hat{W} Z_r + I_r)^{-1} H_r^T [\Sigma + (Z_r^T \hat{W} Z_r + I_r)^{-1} H_r^T]^{-1} \\ &\quad \times H_r (Z_r^T \hat{W} Z_r + I_r)^{-1}\{Z_r^T \hat{W} Z_r \hat{\alpha}_{SR-rd}^{(t)} + Z_r^T \hat{W} D^{-1}(y - \hat{\mu}^{(t)}) + d\hat{\alpha}_r\} \\ &\quad + (Z_r^T \hat{W} Z_r + I_r)^{-1} H_r^T [\Sigma + H_r (Z_r^T \hat{W} Z_r + I_r)^{-1} H_r^T]^{-1} h.\end{aligned}$$

On further simplification, we get:

$$\begin{aligned}\hat{\alpha}^{(t+1)}_{SR-rd} &= (Z_r^T\hat{W}Z_r+I_r)^{-1}\left[Z_r^T\hat{W}Z_r\hat{\alpha}^{(t)}_{SR-rd}+Z_r^T\hat{W}D^{-1}(y-\hat{\mu}^{(t)})+d\hat{\alpha}_r\right]\\ &\quad -(Z_r^T\hat{W}Z_r+I_r)^{-1}H_r^T[\Sigma+H_r(Z_r^T\hat{W}Z_r+I_r)^{-1}H_r^T]^{-1}\\ &\quad \times\{H_r(Z_r^T\hat{W}Z_r+I_r)^{-1}[Z_r^T\hat{W}Z_r\hat{\alpha}^{(t)}_{SR-rd}+Z_r^T\hat{W}D^{-1}(y-\hat{\mu}^{(t)})+d\hat{\alpha}_r]-h\},\end{aligned}$$

where $\hat{\mu}^{(t)}$ and D are evaluated at $\hat{\alpha}^{(t)}_{SR-rd}$.

Transforming back to the original parameters, we obtained an iterative stochastic restricted $r-d$ class estimator in GLMs and denoted it by $\hat{\beta}^{(t+1)}_{SR-rd}$ leads to:

$$\begin{aligned}\hat{\beta}^{(t+1)}_{SR-rd} &= M_r(M_r^TX^T\hat{W}XM_r+I_r)^{-1}[M_r^TX^T\hat{W}XM_rM_r^T\hat{\beta}^{(t)}_{SR-rd}\\ &\quad +M_r^TX^T\hat{W}D^{-1}(y-\hat{\mu}^{(t)})+dM_r^T\hat{\beta}]-M_r(M_r^TX^T\hat{W}XM_r+I_r)^{-1}\\ &\quad \times H_r^T[\Sigma+H_r(M_r^TX^T\hat{W}XM_r+I_r)^{-1}H_r^T]^{-1}\\ &\quad \times\{H_r(M_r^TX^T\hat{W}XM_r+I_r)^{-1}[M_r^TX^T\hat{W}XM_rM_r^T\hat{\beta}^{(t)}_{SR-rd}\\ &\quad +M_r^TX^T\hat{W}D^{-1}(y-\hat{\mu}^{(t)})+dM_r^T\hat{\beta}]-h\},\end{aligned}$$

where $\hat{\beta}$ is the ML estimator. On further simplification we get:

$$\begin{aligned}\hat{\beta}^{(t+1)}_{SR-rd} &= M_r(M_r^TX^T\hat{W}XM_r+I_r)^{-1}(M_r^TX^T\hat{W}u_r^{(t)}+dM_r^T\hat{\beta})-M_r\\ &\quad \times(M_r^TX^T\hat{W}XM_r+I_r)^{-1}H_r^T[\Sigma+H_r(M_r^TX^T\hat{W}XM_r+I_r)^{-1}H_r^T]^{-1}\\ &\quad \times[H_r(M_r^TX^T\hat{W}^{(t)}XM_r+I_r)^{-1}M_r^TX^T\hat{W}^{(t)}u_r^{(t)}-h],\end{aligned} \tag{2.7}$$

where $u_r^{(t)}=XM_rM_r^T\hat{\beta}^{(t)}_{SR-rd}+D^{-1}(y-\hat{\mu}^{(t)})$.

Let us denote $\hat{\beta}^{(t)}_{r-d}=M_r(M_r^TX^T\hat{W}XM_r+I_r)^{-1}(M_r^TX^T\hat{W}u_r^{(t)}+dM_r^T\hat{\beta})$ as in the from of iterative $r-d$ class estimator proposed by [20] in GLMs given by Eq. (2.1). Considering this notation Eq. (2.7) is in the form of a stochastic restricted estimator as:

$$\begin{aligned}\hat{\beta}^{(t+1)}_{SR-rd} &= \hat{\beta}^{(t)}_{r-d}-M_r(M_r^TX^T\hat{W}XM_r+I_r)^{-1}M_r^TH^T\\ &\quad \times[\Sigma+HM_r(M_r^TX^T\hat{W}XM_r+I_r)^{-1}M_r^TH^T]^{-1}(H\hat{\beta}^{(t)}_{r-d}-h).\end{aligned} \tag{2.8}$$

Thus, we obtained an iterative stochastic restricted $r-d$ class estimator in GLMs.

We write the FOA form of the iterative stochastic restricted $r-d$ class estimator as:

$$\begin{aligned}\hat{\beta}^{(1)}_{SR-rd} &= \hat{\beta}^{(1)}_{r-d}-M_r(M_r^TX^T\hat{W}^{(0)}XM_r+I_r)^{-1}M_r^TH^T\\ &\quad \times[\Sigma+HM_r(M_r^TX^T\hat{W}^{(0)}XM_r+I_r)^{-1}M_r^TH^T]^{-1}\\ &\quad \times(H\hat{\beta}^{(1)}_{r-d}-h),\end{aligned} \tag{2.9}$$

where $\beta^{(1)}_{r-d}=M_r(M_r^TX^T\hat{W}^{(0)}XM_r+I_r)^{-1}(M_r^TX^T\hat{W}^{(0)}u^{(0)}+dM_r^T\hat{\beta^{(1)}})$ is the FOA $r-d$ class estimator given by [20] and $u^{(0)}=XM_rM_r^T\beta^{(0)}+D^{-1}(y-\mu^{(0)})$ is the initial working response.

Since the initial working response $u_r^{(0)}=XM_rM_r^T\beta^{(0)}+D^{-1}(y-\mu^{(0)})$ and weight matrix $\hat{W}^{(0)}$ of the estimators are same in the FOA form, the $\hat{\beta}^{(1)}_{SR-rd}$ in Eq. (2.9) is a general estimator which contains a different estimators under particular conditions.

1) If $H=0$ that is no prior information, we get the FOA $r-d$ class estimator in GLMs which is given by [20]

$$\hat{\beta}^{(1)}_{r-d}=M_r(M_r^TX^T\hat{W}^{(0)}XM_r+I_r)^{-1}(M_r^TX^T\hat{W}^{(0)}u^{(0)}+dM_r^T\hat{\beta^{(1)}}). \tag{2.10}$$

2) If $d=0$ and $H=0$ we get the FOA PCR estimator proposed by [25]

$$\hat{\beta}_r^{(1)} = M_r(M_r^T X^T \hat{W}^{(0)} X M_r)^{-1} M_r^T X^T \hat{W}^{(0)} u^{(0)}. \tag{2.11}$$

3) If $d=0$, $r=p$ and $H=0$ we get the FOA ML estimator

$$\hat{\beta}^{(1)} = (X^T \hat{W}^{(0)} X)^{-1} X^T \hat{W}^{(0)} u^{(0)}. \tag{2.12}$$

3. MSE of the FOA stochastic restricted $r-d$ class estimator

This section demonstrates the MSE of the FOA stochastic restricted $r-d$ class estimator. In this regard, we may write the FOA form of the SR-rd class estimator in more simplified form in terms of α and β.

In view of $\hat{\beta}_{SR-rd}^{(1)}$ given by Eq. (2.9) and writing $\hat{\beta}_{r-d}^{(1)}$ by ([20]) as:

$$\begin{aligned}\hat{\beta}_{r-d}^{(1)} &= M_r(M_r^T X^T \hat{W}^{(0)} X M_r + I_r)^{-1}(M_r^T X^T \hat{W}^{(0)} X M_r + dI_r) \\ &\quad \times M_r^T (X^T \hat{W}^{(0)} X)^{-1} X^T \hat{W}^{(0)} u.\end{aligned} \tag{3.1}$$

we can easily get the bias and the variance of the SR-rd class estimator:

$$\begin{aligned}E(\hat{\beta}_{SR-rd}^{(1)}) &= M_r S_r(1)^{-1} S_r(d) M_r^T \beta - M_r S_r(1)^{-1} M_r^T H^T P_r^{-1} \\ &\quad \times [M_r S_r(1)^{-1} S_r(d) M_r^T \beta - H M_r M_r^T \beta].\end{aligned} \tag{3.2}$$

where $S_r(1) = (\Lambda_r^{(0)} + I_r)$, $S_r(d) = (M_r^T X^T \hat{W}^{(0)} X M_r + dI_r)$ and $P_r = \Sigma + H_r S_r(1)^{-1} H_r^T$ and the bias of $\hat{\beta}_{SR-rd}^{(1)}$ can be found as:

$$\begin{aligned}Bias(\hat{\beta}_{SR-rd}^{(1)}) &= E(\hat{\beta}_{SR-rd}^{(1)}) - \beta \\ &= \{(d-1) M_r S_r(1)^{-1} M_r^T - M_{p-r} M_{p-r}^T \\ &\quad - M_r S_r(1)^{-1} M_r^T H^T P_r^{-1} [M_r S_r(1)^{-1} S_r(d) M_r^T - H M_r M_r^T]\}\beta.\end{aligned}$$

The variance of the SR-rd class estimator can be expressed as:

$$\begin{aligned}var(\hat{\beta}_{SR-rd}^{(1)}) &= a(\phi)\{[M_r S_r(1)^{-1} S_r(d) M_r^T - M_r S_r(1)^{-1} M_r^T H^T P_r^{-1} H M_r S_r(1)^{-1} S_r(d) M_r^T] \\ &\quad \times \Lambda_r^{-1} [M_r S_r(1)^{-1} S_r(d) M_r^T - M_r S_r(1)^{-1} M_r^T H^T P_r^{-1} H M_r S_r(1)^{-1} S_r(d) M_r^T]^T \\ &\quad + M_r S_r(1)^{-1} M_r^T H^T P_r^{-1} \Sigma P_r^{-1} H M_r S_r(1)^{-1} M_r^T\}.\end{aligned}$$

Thus, the MSE of the $\hat{\beta}_{SR-rd}^{(1)}$ is given by

$$MSE(\hat{\beta}_{SR-rd}^{(1)}) = var(\hat{\beta}_{SR-rd}^{(1)}) + [Bias(\hat{\beta}_{SR-rd}^{(1)})][Bias(\hat{\beta}_{SR-rd}^{(1)})]^T.$$

4. Numerical illustrations

4.1. Logistic regression

This section shows the application of the estimators via a numerical illustration. An apple juice data set is considered which has been used by [28] as well as [22] in order to develop a logistic regression model. This data set contains the variables pH (x_1), nisin concentration (x_2) (IU/ml), temperature (x_3), and soluble solids concentration (Brix) (x_4). Alicyclobacillus acidoterrestris growing in the apple juice is a response variable where 1 shows growth, 0 shows no growth. More detail about variables can be found from [28]. The SR-rd class estimator is compared with ML, PCR and $r-d$ class estimators, respectively. We used the MATLAB programming language to evaluate the results.

Before computing the results, we standardize the explanatory variables by means of unit length scaling and then add the intercept term in the model. Thus, we consider the logistic regression model as:

$$\pi_i = \frac{1}{1+\exp(-x_i^T\beta)} = \frac{1}{1+\exp-(\beta_0+\beta_1 x_{i1}+\cdots+\beta_4 x_{i4})}, i=1,\cdots,n=37$$

where x_{ij} $j=1,\cdots,4$ denotes the i-th observation of the j-th explanatory variable. An iterative procedure is adopted to evaluate the results and the ordinary least square (OLS) estimator is used as an initial estimate to further computations. The iterative process is repeated until the desired convergence criterion is reached, such as the norm of the difference in parameter estimates between iterations being smaller than 1×10^{-6} is achieved. At the final iteration, the eigenvalues of $X^T\hat{W}X$ are computed as $\lambda_1 = 4.2143$, $\lambda_2 = 0.1774$, $\lambda_3 = 0.1145$, $\lambda_4 = 0.0718$ and $\lambda_5 = 0.0303$ and the condition number (CN) (see [22]) is calculated as $\kappa = \frac{\lambda_{\max}}{\lambda_{\min}} = 138.8583$ which shows that there is a multicollinearity problem in this data set. Therefore, to cope with the multicollinearity problem we proceed with the biased estimation methods.

We impose the stochastic restrictions on the parameters such as: $\beta_0+\beta_1+\beta_2+\beta_3+\beta_4+\phi^*=0$ where $\phi^*\sim(0,a(\phi)\Sigma)$ this leads to the H matrix $H=\left[\begin{array}{ccccc}1&1&1&1&1\end{array}\right]$ and $h=0$ while an estimated value of the $var(\phi^*)=a(\phi)\Sigma=\sigma^2$ is arbitrarily taken as 0.5.

To select the number of PCs, percentage of total variation (PTV) method is used which is computed by the following formula:

$$PTV = \frac{\sum_{j=1}^{r}\lambda_j}{\sum_{j=1}^{q}\lambda_j}\times 100$$

where r represents the number of PCs that will retained in the model. There is no hard and fast rule to choose the PTV value it is arbitrarily chosen as 0.95 and it gives $r=2$.

As it is obvious that the $r-d$ class estimator is a special case of the SR-rd class estimator when $H=0$. Therefore, for the convenience we find the optimum value of the shrinkage parameter d by following [20] which gives:

$$d_{opt} = 1-\left(\sum_{j=1}^{r}\frac{a(\phi)}{\lambda_j(\lambda_j+1)}\Big/\sum_{j=1}^{r}\frac{\lambda_j(\alpha_j^0)^2+a(\phi)}{\lambda_j(\lambda_j+1)^2}\right)$$

where $a(\phi)$ is unknown and can be estimated by using the pearson method for the reduced model such as

$$\hat{\phi}_r = (n-r)^{-1}\sum_{i=1}^{n}(y_i-\hat{\mu}_i)^2.$$

Since $a(\phi)=1$ for the binomial distribution implies that $\hat{\phi}_r=1$. The results of iteratively obtained estimates and the scalar mean square error (SMSE) values of the FOA estimators are given in the Table 1. Table 1 shows the performance of the estimators in terms of the SMSE criterion. It is obvious that the performance of all the biased estimators is better than the ML estimator as the SMSE value of the ML estimator is greater than the other estimators. Nevertheless, the SR-rd class estimator outperforms all other estimators in terms of SMSE criterion as it acquires the smallest SMSE value compared to its counterparts. This shows that the stochastic restrictions significantly improve the estimators' performance.

Table 1 shows the results for only one value of the shrinkage parameter d while Fig.1 is established to see the performance for further values of d. From Fig.1 it is clear that the SR-rd class estimator outperforms its counterparts completely for all the values of d.

Table 1. The estimated coefficients and the SMSE values when $d = 0.8207$ for apple juice data.

Coefficients	ML	PCR	$r-d$	$SR-rd$
β_0	-1.3159	-1.0344	-0.5517	-0.4120
β_1	8.9941	-0.4661	-0.4052	0.3540
β_2	-10.7939	-0.3545	-0.3314	0.3492
β_3	6.1903	-0.2585	-0.2194	0.1782
β_4	-5.8053	-0.1958	-0.1910	0.2203
SMSE	17.5909	7.5100	12.7124	1.2977

Then, the performance of the $r-d$ class estimator is better than the ML estimator for all d values and the PCR estimator when the value of d is less than 0.6 approximately. This numerical example illustrates that incorporating the stochastic restrictions with the sample information considerably improves the estimators' performance. That is, the SR-rd class estimator dominates its counterparts regardless of value of d for the apple juice data.

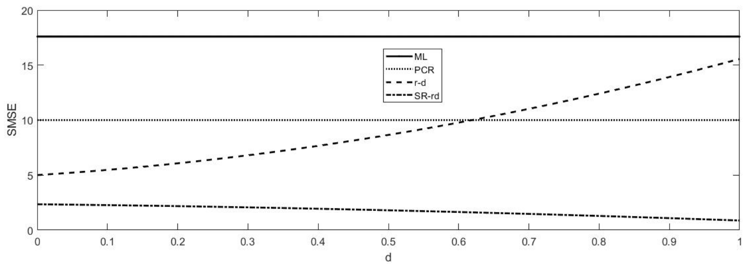

Figure 1. Graph of the SMSE values of the ML, PCR, SR-rd and $r-d$ class estimators for apple juice data.

4.2. Poisson regression

In this section, we consider a real-life data set to figure out the estimators' performances in Poisson regression. The original data set was provided by [16], which was also used by [17]. Furthermore, this data set has also been utilized by [11].

Myers [16] reports 44 observations on mines in the Appalachian coal regions of western Virginia. The data set comprises four explanatory variables: (x_1) represents the inner burden thickness (measured in feet), (x_2) the percentage of extraction of the lower previously mined seam, (x_3) the lower seam height (measured in feet), and (x_4) the duration (years) that the mine has been open. The number of fractures or injuries (y) that occur in the upper seams of the mines was investigated to determine if these factors were related to it. Using a log link function, the Poisson regression model has been applied to this data set by [11, 16, 17],: $log(\mu_i)=\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i}$ where μ_i is the estimated number of upper seam injuries or fractures at the i-th coal mining the location.

The MATLAB programming language is used to get our results. The intercept term is included to the model after the explanatory variables have been normalized using unit length scaling prior to computing the results. The findings are evaluated iteratively, and the OLS estimator is employed as a starting estimate for additional calculations. The iterative procedure is carried out repeatedly until the intended convergence condition is fulfilled, e.g., the norm of the difference in parameter estimations between iterations being less than 1×10^{-6}. The eigenvalues of $X^T\hat{W}X$ are found as $\lambda_1 = 98.6908, \lambda_2 = 2.2452, \lambda_3 = 1.6254, \lambda_4 = 1.2299, \lambda_5 = 0.9730$. Consequently, the condition number (CN) is computed as $CN = \lambda_{max}/\lambda_{min} = 101.4284$, indicating the presence of multicollinearity.

Table 2. The estimated coefficients and the SMSE values when $d = 0.2383$ for the mine data.

Coefficients	ML	PCR	$r-d$	$SR-rd$
β_0	0.5646	2.0298	0.9744	0.9771
β_1	-1.5241	-0.4849	-0.2072	-0.1637
β_2	4.6499	-0.0941	-0.0307	-0.0057
β_3	-0.3114	-0.7174	-0.3040	-0.2352
β_4	-1.5417	-0.5658	-0.2412	-0.1896
SMSE	2.9115	25.9473	24.3663	1.7179

We apply the same stochastic constraints on the parameters as we did in the binomial regression, such as: $\beta_0 + \beta_1 + \beta_2 + \beta_3 + \beta_4 + \phi^* = 0$ where $\phi^* \sim (0, a(\phi)\Sigma)$ yields the matrices $H = [11111]$ and $h = 0$, and an approximated value of 0.5 is randomly selected for $var(\phi^*) = a(\phi)\Sigma = \sigma^2$ where $a(\phi)$ is 1 in Poisson regression. We employed the PTV method with cut off value 0.95 yields in the number of PCs as $r = 2$ and determined the optimum value of the shrinkage parameter d as 0.2383 where the formula given in Section 4.1 is applied. Table 2 displays the obtained results which are explained as follows:

Table 2 shows the results of estimated coefficients and their SMSE values. Table 2 makes it evident that, in comparison to all of its competitors, the SR-rd class estimator outperforms its competitors with the lowest SMSE value, indicating that the estimator's performance is greatly enhanced by using stochastic restrictions. Then, in contrast to PCR and r-d class estimators, the ML estimator obtains the lesser SMSE value when $r = 2$. Since Table 2 displays the estimators' findings for the specific values of the shrinkage parameter $d = 0.2383$ so, Fig. 2 is provided to view the performance for the remaining d values.

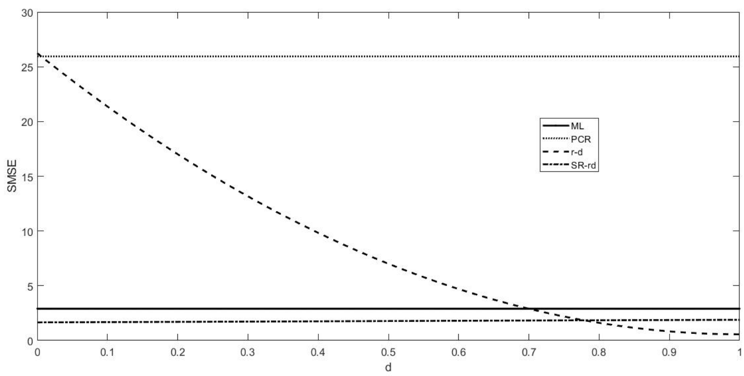

Figure 2. Graph of the SMSE values of the ML, PCR, SR-rd and $r-d$ class estimators for mine data.

It is evident from Fig.2 that when d is between 0 and 0.8 the SR-rd class estimator outperforms the other estimators. The r-d class estimator outperforms all other estimators when d increases beyond 0.8. This suggests that a significant factor in determining the estimators' performance is the shrinkage parameter d. Consequently, this numerical example indicates that for the specific values of d, the SR-rd class estimator performs the best out of the ML, PCR, and $r-d$ class estimators.

5. Simulation studies

In this section, simulation experiments are performed when a response variable belongs to Poisson and negative binomial distributions, respectively in order to figure out the performance of the estimators which are the ML, PCR, $r-d$ class and SR-rd class estimators, respectively. The performance of the estimators is assessed using the estimated mean square error (EMSE) criterion which is computed as:

$$EMSE(\tilde{\beta}) = \frac{1}{MCN} \sum_{s=1}^{MCN} \left(\tilde{\beta}_{(s)} - \beta\right)^T \left(\tilde{\beta}_{(s)} - \beta\right),$$

where the subscript s denotes the s-th replication of the simulation experiment and $\tilde{\beta}_{(s)}$ is the estimates of β and MCN represents the number of replications in the Monte Carlo simulation experiment which is replicated up to 500 times. Our results are evaluated by MATLAB programming language. The remaining steps of conducting the Monte Carlo simulation experiment are illustrated as:

1. The sample size are considered as $n = 25$, 50, 200, 400, 800 and a number of explanatory variables used is $p = 4$.

2. The explanatory variables are produced by following [15] such as:

$$x_{ij} = (1-\rho^2)^{1/2} \upsilon_{ij} + \rho \upsilon_{i,p+1}, i = 1, \cdots, n, j = 1, \cdots, p.$$

where ρ^2 denotes the degree of multicollinearity between any two explanatory variables and υ_{ij} are independent standard normal pseudo-random numbers. The explanatory variables are standardized by unit length standardization approach before computing the results.

3. The values of ρ^2 are considered to be $\rho^2 = 0.80$, 0.90, and 0.99.

4. The parameter vector β is calculated as a normalized eigenvector correspond to the largest eigenvalue of the $X^T X$ matrix so that $\beta^T \beta = 1$.

5. By following [9, 30, 31], the stochastic restriction for $p = 4$ is considered as:

$$H = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}, h = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \ \Sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

6. The values of d are considered as $d = 0, 0.10, 0.30, 0.40, 0.50, 0.70, 0.80, 0.99$.

7. To determine the number of PCs, we applied the percentage of total variation (PTV) approach as described in Section 4.1.

8. The response variable from the Poisson distribution is generated as $y_i \sim P(\mu_i)$ having the log-link function $\mu_i = \exp(\beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_p x_{ip})$, and for the negative binomial (NB) distribution it is produced as $y_i \sim NB(\mu_i, \mu_i + \eta\mu_i^2)$ where we choose $\eta = 1/3$ (see[29] †) and β_j, $j = 1, \ldots, p$ is considered as defined in step 4.

9. The OLS estimator $\hat{\beta}_{ols} = \beta^{(0)} = \left(X^T X\right)^{-1} X^T y$ is taken as the initial estimate of β.

The key findings are as follows, and the results are displayed in Tables 3-14 in Appendix A:

i) For the Poisson response it is seen that when the sample size is small $n = 25$ and $n = 50$ for all the values of d and $\rho^2 = 0.99$ then, SR-rd class estimator performs better than its competitors. However, when $\rho^2 = 0.80, 0.90$ then, it performs better for the d values greater than 0.10.

ii) When $n = 200$ the performance of the SR-rd class estimator is better for $d > 0.10$ and $\rho^2 = 0.80, 0.90$. While it outperforms its counterparts at all degrees of multicollinearity and the values of d except $d = 0, 0.10$ for $n = 400$ and $n = 800$.

iii) For the optimum d values it is seen that SR-rd class estimators acquires smallest EMSE values as compared to its counterparts and outperforms them for almost all the sample sizes and degrees of multicollinearity except $n = 50, 200$ and $\rho^2 = 0.99$.

iv) For the negative binomial response it is observed that the SR-rd class outclasses its competitors when $n = 200$ regardless of the values of d and multicollinearity degrees

†We can also choose $\eta = 1, 2$ by following [14]

ρ^2. Nevertheless, for the remaining sample sizes it performs better when the d values are greater than 0.10 and degrees of multicollinearity are 0.80 and 0.90, respectively.

v) For the optimum d values, the performance of the SR-rd class estimator is superior than its counterparts regardless of the sample size and degrees of multicollinearity when the response variable follows negative binomial distribution.

In general, the EMSE of the ML estimator increases as the degree of multicollinearity increases. The EMSE of the PCR likewise rises when ρ^2 rises from 0.80 to 0.90, then falls at $\rho^2 = 0.99$. The EMSE of the $r-d$ and SR-rd class estimators on the other hand, decreases with increasing degrees of multicollinearity.

6. Conclusion

In this study, we proposed the stochastic restricted $r-d$ class in GLMs and its performance is compared with different estimators existing in the literature. The performance of the estimators is evaluated via numerical examples and simulation studies when a response variable belongs to binomial, Poisson, and negative binomial distributions, respectively. The results revealed that the proposed estimator outperforms its counterparts for the particular values of the shrinkage parameter d.

7. Acknowledgements

The authors thank the referees for their valuable comments and suggestions, which have significantly improved the quality of this article.

Author Contribution. Both authors contributed equally to the preparation of this manuscript.

Conflict of interest. The authors declare that they have no conflicts of interest or competing financial interests.

Funding. The authors declare that no funds or grants were received during the preparation of this manuscript.

Data availability. The data used in the numerical examples were previously employed in [22, 28] and [11, 16, 17] and can be found from these references.

References

[1] A. Abbasi and M.R. Özkale, *The r-k class estimator in generalized linear models applicable with simulation and empirical study using a Poisson and Gamma responses*, Hacet. J. Math. Stat. **50** (2), 594-611, 2021.

[2] M.N. Akram, M. Amin, A.F. Lukman, and S. Afzal, *Principal component ridge type estimator for the inverse Gaussian regression model.* J. Stat. Comput. Simul. **92** (10), 2060-2089, 2022.

[3] K.C. Arum and F.I. Ugwuowo, *Combining principal component and robust ridge estimators in linear regression model with multicollinearity and outlier*, Concurr. Comput. Pract. Exp. **34** (10), 6803, 2022.

[4] K.C. Arum, F.I. Ugwuowo, H.E. Oranye, T.O. Alakija, T.E. Ugah, and O.C. Asogwa, *Combating outliers and multicollinearity in linear regression model using robust Kibria Lukman mixed with principal component estimator, simulation and computation.* Sci. Afr. **19** (17), 2023.

[5] M.R. Baye and D.F. Parker, *Combining ridge and principal component regression: A money demand illustration*, Commun. Stat. Theory Methods **13** (2), 197-205, 1984.

[6] H. Daojiang, and Y. Wu *A stochastic restricted principal components regression estimator in the linear model* Sci. World J. **84** (1), 2014.

[7] R.A. Farghali, A.F. Lukman and A. Ogunleye, *Enhancing model predictions through the fusion of Stein estimator and principal component regression.* J. Stat. Comput. Simul. **94** (8), 1760-1775, 2024.

[8] T. Gargi and S. Chandra *Two-parameter stochastic restricted principal component estimator in linear regression model*, Pak. J. Stat. **35** (2), 127-154, 2019.

[9] Y.E. Gawdat, *A stochastic restricted mixed Liu-Type estimator in logistic regression model*, Appl. Math. Sci. **7**, 311-322, 2020.

[10] X. Jianwen, and H.u. Yang, *On the restricted $r-k$ class estimator and the restricted $r-d$ class estimator in linear regression* J. Stat. Comput. Simul. **81** (6), 679-691, 2011.

[11] F. Kurtoğlu and M.R. Özkale, *Restricted ridge estimator in generalized linear models: Monte Carlo simulation studies on Poisson and binomial distributed responses*, Commun. Stat. Simul. Comput. **48** (4), 1-28, 2017.

[12] F. Kurtoğlu and M.R. Özkale, *Restricted Liu estimator in generalized linear models: Monte Carlo simulation studies on gamma and Poisson distributed responses*, Hacet. J. Math. Stat. **48** (4), 1191-1218, 2019.

[13] A.F. Lukman, K. Ayinde, O. Oludoun, and C. Onate, *Combining modified ridge type and principal component regression estimators.* Sci. Afr. **9**, e00536, 2020.

[14] K. Månsson, B.M.G. Kibria and G. Sukur, *On ridge estimators for the negative binomial regression model. Econ. Model. 29(2), 178-184 (2012)*, Econ. Model. **29** (4), 1483-1488, 2012.

[15] G.C. McDonald and D.I. Galarneau, *A monte carlo evaluation of some ridge-type estimators*, J.Am.Stat.Assoc. **70** (350), 407-416, 1975.

[16] R.H. Myers, *Classical and modern regression with applications.* Belmont, CA: Duxbury press, 1990.

[17] B.D. Marx, *A continuum of principal component generalized linear regressions.* Comput. Statist. Data Anal. **13** (4), 385-393, 1992.

[18] J.A. Nelder and R. W. M. Wedderburn, *Generalized Linear Models*, J.R. Statist.Soc.A **135** (3), 370-384, 1972.

[19] H. Nyquist, *Restricted Estimation of Generalized Linear Models*, J. R. Stat. Soc.Ser.C. **40** (1), 133-141, 1991.

[20] M.R. Özkale, *The r-d class estimator in generalized linear models: applications on gamma, Poisson and binomial distributed responses*, J. Stat. Comput. Simul. **89** (4), 615-640, 2019.

[21] M.R. Özkale and H. Nyquist, *The stochastic restricted ridge estimator in generalized linear models*, Stat. Pap. **62** (3) 1421-1460 (2021) 2019.

[22] M.R. Özkale, *Iterative algorithms of biased estimation methods in binary logistic regression*, Stat. Pap. **57**, 991-1016, 2016.

[23] M.R. Özkale and A.Abbasi *Iterative restricted OK estimator in generalized linear models and the selection of tuning parameters via MSE and genetic algorithm*, Stat. Pap., 1-62, 2022.

[24] M.R. Özkale *Principal components regression estimator and a test for the restrictions*, Statistics **36** (15), 43(6), 541-551, 2009.

[25] E.P. Smith and B.D. Marx, *Ill-conditioned information matrices, generalized linear models and estimation of the effects of acid rain*, Environmetrics **1** (1), 57-71, 1990.

[26] C. Shalini, and N. Sarkar, *A restricted r-k class estimator in the mixed regression model with autocorrelated disturbances*, Stat. Pap. **57** (2), 429-449, 2016.

[27] N. Varathan, and P. Wijekoon, *Liu-Type logistic estimator under stochastic linear restrictions*, Stat. Pap. Ceylon J. Sci.47(1), 21-34, 2018.

[28] P. Wel, De Massaguer P.R., A.D., Zuniga, S.H. Saraiva, *Modeling the growth limit of Alicyclobacillus acidoterrestris CRA7152 in apple juice: effect of pH, Brix, temperature and nisin concentration*, J. Food Process. Preserv. **35** (4) 509-517, 2011.

[29] P. Walter W, *Maximum likelihood estimation for the negative binomial dispersion parameter*, Biometrics 863-867, 1990.

[30] J. Wu and Y. Asar, *On the stochastic restricted Liu-type maximum likelihood estimator in logistic regression model. Commun. Fac. Sci. Univ. Ank. Ser. A1 Math. Stat.* **68** (1), 643-653, 2019.
[31] Z. Weibing, and Y. Li, *A new stochastic restricted Liu estimator for the logistic regression model* Open J. Stat. **8** (1), 25-37, 2018.

Appendix A: Simulation Results

Table 3. The EMSE values of the estimators for different d for the Poisson response when $n = 25$.

Est.	$d = 0$	$d = 0.1$	$d = 0.30$	$d = 0.40$	d=0.50	$d = 0.70$	d=0.80	d=0.99
				$\rho^2 = 0.80$				
ML	7.1864	6.8942	6.4895	7.2262	7.1971	6.8918	7.2738	7.1419
PCR	10.3118	10.1647	10.5890	9.8074	12.6956	11.3482	12.7170	10.1192
$r-d$	**0.9217**	**1.1637**	1.8435	2.5617	3.1010	4.3014	5.3566	6.8559
SR-rd	1.2562	1.3095	**1.4839**	**1.6199**	**1.6706**	**2.0941**	**2.3954**	**2.6803**
				$\rho^2 = 0.90$				
ML	12.3916	11.9163	12.2511	11.7175	12.6646	12.2279	12.1153	12.6999
PCR	26.9213	20.0241	19.1447	12.2667	16.0788	21.8360	11.9529	10.5482
$r-d$	**0.5993**	**0.8709**	1.8325	2.1362	3.0019	4.4474	5.5376	7.7362
SR-rd	1.2386	1.2916	**1.4740**	**1.4348**	**1.6204**	**1.9925**	**2.2104**	**2.7113**
				$\rho^2 = 0.99$				
ML	112.7106	114.2049	114.2242	121.1617	109.2537	113.7359	112.0079	121.2008
PCR	0.4087	0.4511	0.2856	1.9056	0.2468	0.2631	0.3277	0.5905
$r-d$	0.2097	0.1831	0.1901	0.1928	0.2019	0.2074	0.1967	**0.2323**
SR-rd	**0.2094**	**0.1828**	**0.1899**	**0.1925**	**0.2017**	**0.2072**	**0.1966**	**0.2323**

Table 4. The EMSE values of the estimators for different d for the Poisson response when $n = 50$.

Est.	$d = 0$	$d = 0.1$	$d = 0.30$	$d = 0.40$	d=0.50	$d = 0.70$	$d = 0.80$	$d = 0.99$
				$\rho^2 = 0.80$				
ML	7.5491	7.3446	7.2994	7.3580	7.0511	7.4228	7.4040	7.2944
PCR	7.7548	8.8601	8.0291	8.0121	7.9239	8.8167	7.7557	8.4945
$r-d$	**0.9159**	**1.1829**	2.0764	2.6363	3.1106	4.7303	5.5560	7.1968
SR-rd	1.3137	1.3387	**1.5998**	**1.7313**	**1.9649**	**2.3268**	**2.5403**	**3.0925**
				$\rho^2 = 0.90$				
ML	12.4470	15.0687	14.9511	13.8893	14.3763	14.0801	14.6561	13.1432
PCR	8.2959	11.0925	16.6163	9.4651	11.5372	10.1830	10.9422	10.1010
$r-d$	**0.5352**	**0.9489**	2.1798	2.6779	3.5692	5.7381	7.0860	8.9264
SR-rd	1.1896	1.2397	**1.4316**	**1.4566**	**1.6023**	**1.9995**	**2.0666**	**2.4745**
				$\rho^2 = 0.99$				
ML	134.0537	125.8559	131.2492	138.3114	127.3479	125.5206	129.1854	131.9482
PCR	0.2490	0.3323	0.2452	0.3105	1.0323	0.4646	0.1998	0.2516
$r-d$	0.1939	0.1911	0.2075	0.1937	0.2023	0.2036	**0.1839**	0.1996
SR-rd	**0.1936**	**0.1910**	**0.2073**	**0.1933**	**0.2020**	**0.2033**	**0.1839**	**0.1992**

Table 5. The EMSE values of the estimators for different d for the Poisson response when $n = 200$.

Est.	$d=0$	$d=0.1$	$d=0.30$	$d=0.40$	d=0.50	$d=0.70$	d=0.80	d=0.99
				$\rho^2=0.80$				
ML	7.7709	7.5349	7.4537	7.7803	7.6163	7.6842	7.4035	7.8267
PCR	7.7666	7.5375	7.4524	7.7766	7.6160	7.6813	7.4047	7.8213
$r-d$	**0.8821**	**1.1979**	2.0692	2.7114	3.3117	4.8480	5.5153	7.7193
SR-rd	1.3259	1.3900	**1.6079**	**1.7700**	**1.9841**	**2.3428**	**2.5752**	**3.1532**
				$\rho^2=0.90$				
ML	13.1614	13.4129	14.1682	12.9232	12.7735	14.8008	13.1424	13.3069
PCR	9.2911	9.6255	9.9910	8.8974	8.8431	10.1850	9.1417	9.1452
$r-d$	**0.5470**	**0.8935**	2.0444	2.5212	3.2721	5.9328	6.4946	9.0262
SR-rd	1.2783	1.3238	**1.5310**	**1.6337**	**1.7751**	**2.2022**	**2.4755**	**2.8103**
				$\rho^2=0.99$				
ML	129.2784	125.9164	123.7971	123.0112	123.5068	131.3514	128.4290	130.2244
PCR	0.2615	0.2298	0.2214	0.2520	0.2303	0.2572	0.2570	0.2444
$r-d$	**0.2077**	**0.1868**	**0.1872**	**0.2150**	**0.1959**	**0.2314**	**0.2425**	**0.2444**
SR-rd	0.2079	0.1870	0.1876	0.2154	0.1964	0.2320	0.2432	0.2451

Table 6. The EMSE values of the estimators for different d for the Poisson response when $n = 400$.

Est.	$d=0$	$d=0.1$	$d=0.30$	$d=0.40$	d=0.50	$d=0.70$	d=0.80	d=0.99
				$\rho^2=0.80$				
ML	8.5098	8.1723	8.3823	8.6628	8.6903	8.9931	8.6131	8.8280
PCR	8.5100	8.1684	8.3811	8.6604	8.6900	8.9917	8.6130	8.8253
$r-d$	**0.8028**	**1.1413**	2.1242	2.8298	3.6112	5.5471	6.3181	8.7020
SR-rd	1.2789	1.3286	**1.6290**	**1.7462**	**1.9953**	**2.4537**	**2.6856**	**3.2944**
				$\rho^2=0.90$				
ML	17.0741	16.1766	18.0327	16.7381	17.0349	16.2191	16.2744	15.4757
PCR	11.8496	10.5941	12.0584	11.5623	11.4313	10.6250	11.3391	11.0271
$r-d$	**0.5138**	**0.8579**	2.2587	3.0157	4.0017	6.0674	7.9343	10.8547
SR-rd	1.1221	1.2319	**1.4818**	**1.6193**	**1.8323**	**2.4000**	**2.8600**	**3.5565**
				$\rho^2=0.99$				
ML	149.7117	156.0796	158.5484	159.3846	152.9927	155.0625	157.3454	147.5700
PCR	0.2497	0.2402	0.2632	0.2102	0.2373	0.2642	0.2278	0.2230
$r-d$	**0.2102**	0.2001	0.2131	0.1769	0.2048	**0.2387**	**0.2133**	0.2228
SR-rd	**0.2101**	**0.2000**	**0.2130**	**0.1768**	**0.2047**	0.2388	0.2134	**0.2227**

Table 7. The EMSE values of the estimators for different d for the Poisson response when $n = 800$.

Est.	$d = 0$	$d = 0.1$	$d = 0.30$	$d = 0.40$	d=0.50	$d = 0.70$	d=0.80	d=0.99
				$\rho^2 = 0.80$				
ML	9.1431	8.7199	8.1083	8.1535	8.4806	8.5529	8.5042	8.6198
PCR	9.1432	8.7199	8.1077	8.1531	8.4801	8.5530	8.5044	8.6201
$r-d$	**0.8865**	**1.2152**	2.1023	2.7393	3.5475	5.3128	6.2652	8.4976
SR-rd	1.3148	1.3742	**1.5745**	**1.7764**	**2.0481**	**2.4902**	**2.7501**	**3.3014**
				$\rho^2 = 0.90$				
ML	15.7550	16.0459	16.2482	16.8045	15.1216	14.7729	15.2173	15.5289
PCR	10.1507	10.2880	10.6108	11.2034	10.5730	9.9752	9.7348	9.9664
$r-d$	**0.5119**	**0.8672**	2.0334	2.9853	3.7489	5.7212	6.8460	9.8132
SR-rd	1.0150	1.0638	**1.2921**	**1.4436**	**1.6030**	**2.0073**	**2.2431**	**2.6353**
				$\rho^2 = 0.99$				
ML	144.3698	132.3865	149.2753	145.3130	139.6298	148.6552	145.4144	146.5951
PCR	0.2747	0.2627	0.2447	0.2459	0.2458	0.2937	0.2458	0.2774
$r-d$	**0.2238**	0.2156	0.2032	0.2058	0.2105	0.2673	0.2294	0.2765
SR-rd	0.2239	**0.2155**	**0.2031**	**0.2057**	**0.2104**	**0.2672**	**0.2293**	**0.2763**

Table 8. EMSE values of the estimators for optimal d values for Poisson response.

n	ρ^2	ML	PCR	$r-d$	$SR-rd$
25	0.80	6.9453	11.1457	4.3463	**1.6836**
	0.90	12.0440	10.5317	5.1277	**1.3135**
	0.99	170.5235	0.4939	0.3753	**0.3751**
50	0.80	7.3804	7.6847	4.8033	**1.7647**
	0.90	13.6653	9.3467	7.0684	**1.2203**
	0.99	131.0450	**0.3856**	0.3862	0.3861
200	0.80	8.1127	8.1157	5.1726	**1.9037**
	0.90	13.5521	9.3269	6.8171	**1.5113**
	0.99	189.2925	**0.3454**	0.4942	0.4941
400	0.80	9.0093	9.0095	5.7968	**2.0766**
	0.90	16.6859	11.4681	8.5690	**2.8098**
	0.99	162.8223	0.2429	0.2192	**0.2190**
800	0.80	8.7608	8.7598	5.7504	**2.1190**
	0.90	15.6775	10.6188	7.6886	**1.7937**
	0.99	145.5790	0.2611	0.2324	**0.2224**

Table 9. The EMSE values of the estimators for different d for the Negative binomial response when $n=25$.

Est.	$d=0$	$d=0.1$	$d=0.30$	$d=0.40$	d=0.50	$d=0.70$	d=0.80	d=0.99
				$\rho^2=0.80$				
ML	9.1074	9.5980	10.2565	9.3882	9.2744	9.7675	9.5896	10.0516
PCR	10.3984	9.7418	13.3112	9.9233	14.0423	9.2381	9.8496	11.2955
$r-d$	**1.3511**	1.8146	2.6144	2.9083	3.3213	4.4299	5.1637	6.4985
$SR-rd$	1.4718	**1.6200**	**1.9156**	**1.9034**	**2.0268**	**2.2932**	**2.5101**	**2.9338**
				$\rho^2=0.90$				
ML	18.7953	18.0598	18.1788	17.2055	18.4284	18.7703	17.2301	18.0383
PCR	10.4657	13.1702	11.2751	10.6238	15.0920	12.0882	11.8814	10.0369
$r-d$	**1.1165**	1.3828	2.3382	2.7239	3.6722	5.0914	5.5334	7.0487
$SR-rd$	1.2537	**1.3399**	**1.4343**	**1.4384**	**1.8079**	**2.0821**	**2.2591**	**2.5252**
				$\rho^2=0.99$				
ML	165.6990	168.0916	149.4495	180.7034	166.2091	162.0064	169.8203	173.8080
PCR	0.7346	0.5770	0.5056	1.5061	0.5583	0.5480	0.5590	0.6631
$r-d$	**0.2706**	**0.2735**	**0.2756**	**0.2926**	**0.2807**	**0.2829**	**0.2695**	**0.2542**
$SR-rd$	0.2709	0.2739	0.2763	0.2933	0.2812	0.2830	0.2701	0.2544

Table 10. The EMSE values of the estimators for different d for the Negative binomial response when $n=50$.

Est.	$d=0$	$d=0.1$	$d=0.30$	$d=0.40$	d=0.50	$d=0.70$	d=0.80	d=0.99
				$\rho^2=0.80$				
ML	10.5030	10.3884	10.7881	10.5483	10.3613	10.0554	11.2145	11.0879
PCR	17.9434	16.1505	13.7538	15.3950	15.6535	13.7255	15.5429	13.7572
$r-d$	**1.5207**	2.0625	3.4077	4.1778	4.8762	6.5419	8.5867	10.7112
$SR-rd$	1.5427	**1.6873**	**1.9485**	**2.2529**	**2.4436**	**2.8334**	**3.5297**	**4.0185**
				$\rho^2=0.90$				
ML	20.4313	20.7371	19.8677	19.3987	20.8570	19.7703	19.9500	20.5702
PCR	19.6605	20.6764	17.7198	17.7910	18.1055	18.5602	19.4569	17.3411
$r-d$	**1.0240**	1.7584	3.4428	4.4565	6.2904	8.9836	10.8558	15.3569
$SR-rd$	1.2124	**1.3347**	**1.6427**	**1.7863**	**2.1132**	**2.6199**	**3.0307**	**3.8591**
				$\rho^2=0.99$				
ML	197.0314	198.5683	197.6993	187.0360	190.7589	189.7423	184.9493	188.2342
PCR	0.3635	0.3042	0.3204	0.3701	0.3207	0.3307	0.3940	0.3382
$r-d$	**0.2615**	**0.2570**	**0.2813**	**0.3255**	**0.2921**	**0.2629**	**0.3520**	**0.3228**
$SR-rd$	0.2623	0.2579	0.2824	0.3268	0.2933	0.2639	0.3537	0.3244

Table 11. The EMSE values of the estimators for different d for the Negative binomial response when $n = 200$.

Est.	$d = 0$	$d = 0.1$	$d = 0.30$	$d = 0.40$	d=0.50	$d = 0.70$	d=0.80	d=0.99
				$\rho^2 = 0.80$				
ML	11.9762	11.7232	11.6577	10.3722	10.8209	10.6775	11.3011	11.8537
PCR	12.4096	12.1101	12.0132	10.5775	11.0865	10.9622	11.5914	12.1446
$r-d$	1.5530	2.1244	3.5471	3.8962	4.9689	6.9564	8.5797	11.6924
$SR-rd$	**1.4731**	**1.6716**	**2.0705**	**2.1088**	**2.3945**	**2.8884**	**3.3597**	**4.0225**
				$\rho^2 = 0.90$				
ML	22.8148	22.5053	21.3322	21.0611	21.3735	19.8828	20.6017	21.2106
PCR	16.3249	15.7344	14.8500	15.1424	15.2462	14.1287	14.2302	15.2927
$r-d$	1.0131	1.6024	3.2456	4.4134	5.7295	8.3257	10.1359	14.8702
$SR-rd$	**0.9105**	**1.0935**	**1.5750**	**1.8400**	**2.3393**	**2.9300**	**3.5470**	**4.8129**
				$\rho^2 = 0.99$				
ML	198.4872	200.4537	207.8937	203.8673	192.6462	188.9706	182.1771	191.5636
PCR	0.3540	0.3486	0.3605	0.3270	0.3062	0.3442	0.3633	0.3351
$r-d$	0.2703	0.2819	0.2955	0.2781	0.2646	0.3160	0.3435	0.3331
$SR-rd$	**0.2702**	**0.2818**	**0.2954**	**0.2780**	**0.2645**	**0.3156**	**0.3432**	**0.3329**

Table 12. The EMSE values of the estimators for different d for the Negative binomial response when $n = 400$.

Est.	$d = 0$	$d = 0.1$	$d = 0.30$	$d = 0.40$	d=0.50	$d = 0.70$	d=0.80	d=0.99
				$\rho^2 = 0.80$				
ML	11.0976	11.3574	11.6036	11.3010	10.8555	11.5872	11.1093	11.8080
PCR	11.2209	11.5945	11.7012	11.4839	11.0031	11.6575	11.2613	11.9380
$r-d$	1.4866	2.0765	3.5506	4.2886	4.9936	7.5366	8.4470	11.6574
$SR-rd$	**1.4309**	**1.6169**	**2.0479**	**2.2022**	**2.4014**	**3.0856**	**3.4616**	**4.3530**
				$\rho^2 = 0.90$				
ML	19.5736	20.6923	20.5379	21.4430	20.2524	20.1159	21.7899	21.4349
PCR	13.9206	14.4655	14.0970	15.6375	14.4294	14.2980	15.5049	15.0299
$r-d$	**0.9112**	1.5040	3.0810	4.5492	5.4393	8.4278	11.0457	14.7564
$SR-rd$	1.2683	**1.3572**	**1.5465**	**1.6953**	**1.7744**	**2.2029**	**2.6202**	**3.2482**
				$\rho^2 = 0.99$				
ML	203.5570	194.3169	207.3676	195.4191	196.5150	196.5505	196.4927	208.3822
PCR	0.3401	0.3417	0.3029	0.3172	0.3067	0.3492	0.3210	0.3328
$r-d$	**0.2800**	**0.2793**	**0.2498**	**0.2699**	**0.2745**	**0.3186**	**0.3045**	**0.3304**
$SR-rd$	0.2801	0.2795	0.2501	0.2702	0.2748	0.3190	0.3049	0.3310

Table 13. The EMSE values of the estimators for different d for the Negative binomial response when $n = 800$.

Est.	$d = 0$	$d = 0.1$	$d = 0.30$	$d = 0.40$	d=0.50	$d = 0.70$	d=0.80	d=0.99
				$\rho^2 = 0.80$				
ML	11.5687	11.0544	11.5924	11.0004	11.8775	11.4609	10.8301	11.5768
PCR	11.6330	11.1333	11.6877	11.0719	11.9811	11.5526	10.8761	11.6198
$r-d$	**1.5209**	1.9649	3.4862	4.1322	5.4117	7.4353	8.1915	11.4168
$SR-rd$	1.5273	**1.5938**	**2.0889**	**2.2360**	**2.4765**	**3.0943**	**3.3861**	**4.3231**
				$\rho^2 = 0.90$				
ML	20.8121	21.5817	19.9909	21.4171	21.3386	22.4651	22.2622	21.3557
PCR	14.0105	14.2715	13.6106	14.9334	14.3562	14.9251	15.5999	15.0550
$r-d$	**0.8988**	**1.4890**	2.9870	4.3772	5.4706	8.7919	11.1488	14.7524
$SR-rd$	1.4384	1.5632	**1.7896**	**1.9946**	**2.2314**	**2.6066**	**3.1215**	**3.6559**
				$\rho^2 = 0.99$				
ML	204.0482	196.0830	195.5125	192.6773	193.8712	194.1164	199.5015	203.1150
PCR	0.3215	0.0035	0.3136	0.34761	0.3412	0.3539	0.3511	0.3365
$r-d$	**0.2597**	**0.0027**	**0.2630**	**0.2957**	**0.2996**	**0.3264**	**0.3296**	**0.3352**
$SR-rd$	0.2598	0.0028	0.2631	0.2958	0.2997	0.3265	0.3298	0.3353

Table 14. EMSE values of the estimators when the d value are obtained by optimum method for the Negative binomial response.

n	ρ^2	ML	PCR	$r-d$	$SR-rd$
25	0.80	9.6399	14.8745	5.3299	**1.9546**
	0.90	16.7319	12.3005	6.0829	**1.5213**
	0.99	166.6831	1.3082	0.9484	**0.4636**
50	0.80	10.8162	13.9474	8.3705	**2.9431**
	0.90	20.3491	17.9154	12.8051	**3.1688**
	0.99	195.4379	0.3426	0.4735	**0.4733**
200	0.80	11.6451	11.8982	9.0668	**2.9943**
	0.90	20.8585	14.3496	11.6472	**3.9411**
	0.99	200.3968	0.3469	0.4883	**0.4881**
400	0.80	11.3287	11.5121	8.7592	**3.1583**
	0.90	20.8917	15.1284	12.9137	**1.9832**
	0.99	194.2428	19.2937	2.4783	**1.4782**
800	0.80	11.9741	12.0191	9.5258	**3.2672**
	0.90	20.8163	13.8114	11.3467	**1.9811**
	0.99	200.1923	21.3551	0.4943	**0.4940**