*Journal of Statistical Modelling: Theory and Applications*
*Vol. 5, No. 2, 2024, pp. 153-163*
*Yazd University Press 2024*

JSMTA

*Research Paper*

# Bayesian estimation of heteroscedastic skew-normal error regression model

Isiaka Oloyede*, Alfred A. Abiodun
Department of Statistics, University of Ilorin, Nigeria

Received: September 14, 2024/ Revised: May 10, 2025/ Accepted: May 28, 2025

**Abstract:** In statistics, errors are inherent in data and models, particularly heteroscedasticity and skew-normal error structures. These errors were simultaneously generated and infused into the data, leading to uncertainty in parameter estimation. The statistician uses statistical knowledge to elicit information and guide decision-making. Both classical and Bayesian restricted Stein-rule least squares were compared when the data were contaminated with the aforementioned errors. This study proposed an innovative Bayesian generalized restricted Stein-rule least squares method with heteroscedastic skew-normal errors, which was ultimately found to be more efficient compared to non-Bayesian restricted Stein-rule least square estimators. The study observed excellent performance of the Bayesian frameworks, including the Bayes estimate and posterior mean, in comparison to the classical restricted Stein-rule least squares estimators. Therefore, the study recommends Bayesian generalized restricted Stein-rule least squares to analysts and researchers who may encounter such errors in their data.

**Keywords:** Bayesian; Heteroscedasticity; Modeling and least squares; Skew-normal, Simulation.
**Mathematics Subject Classification (2010):** 62C10, 62J05.

## 1 Introduction

Heteroscedasticity in Statistics occurs when a breakdown of assumption of homoscedasticity which resulted from the unequal variances of the diagonal elements of variance covariance of residuals, Harvey (1976) was the proponent of regression model with

*Corresponding author: `oloyede.i@unilorin.edu.ng`

multiplicative heteroscedasticity, Cepeda and Germaman (2000) coined Bayesian Heteroscedastic model and showed the efficiency of the model in their studied, Oloyede et al. (2013) proposed Bayesian generalized heteroscedastic model as efficient method to fit Bayesian heteroscedastic model using Markov Chain Monte Carlo (MCMC) simulation. Obviously, data skewness often induces heteroscedasticity errors, so far, dependent variable follows skew-normal distribution as the error is skewed, there is tendency that such dependent variable that is skew-normal to likely induces heteroscedasticity.

O'Hagan and Leonard (1976) were the proponent of skew-normal distribution which was afterward expanded with the work of Azzalini (1985) who later examined series of univariate skew-normal distributions and their respective properties. Genton et al. (2001) applied multivariate skew-normal model to time series and spatial Statistics whereas Azzalini (1985, 1986) applied such to econometrics, non-linear time series and financial Statistics. Ferreira and Steel (2007) obtained Bayesian regression model from the construction of dual multivariate skew-normal distributions. Lachos et al. (2007) estimated parameters of multivariate skew-normal regression model with EM algorithms. Corrales and Cepeda-Cuervo (2022) examined Bayesian skew-normal regression model where its moments are of normal regression model structures, they were of the opinion that inappropriateness of regression skewness can result into infallible location and scale regression parameters. Moreso, inappropriateness of location and skewness in regression model will produce incorrect inferences.

Though most of the literature considered skew-normal regression model where the distribution of covariates are individually skew-normal but in Azzalini (2005), regression model with skew-normal error was considered. The bane of concern is that skew-normal error or non-normality of error can influence or lead to heteroscedasticity either in heavy tail towards right or left.

Alhamide et al. (2019) carried out study on Bayesian linear regression with three different prior density which includes alpha-skew-normal, normal, and non-informative. The posterior means of the parameters under the alpha-skew-normal prior were found to be more accurate than the comparable posterior means under the normal and non-informative priors. Indacochea (2012) was of the opinion that practitioners will often encounter data-generating processes (DGPs) with error terms that deviate greatly from normality.

In Rubio and Genton (2016), skew-normal error models reflect deviations from the conventional assumption of error normality in terms of heavy tails and asymmetry from their work Bayesian linear regression models with skew-symmetric scale mixtures of normal error distributions. They suggested a noninformative prior structure for these regression models and demonstrated that the resulting posterior distribution is appropriate under moderate conditions. Azzalini (1986) investigated the features of a multivariate skewed normal distribution with scalar skew-normal marginal densities. The random variable $Y$ follows $SN(\mu, \sigma, \delta)$ as in the Azzalini skew-normal distribution, where the mean and variance are as follows:

$$E(Y) = \mu\sqrt{\frac{2}{\pi}}\sigma, \qquad Var(Y) = \sigma^2\left(1 - \frac{2}{\pi}r^2\right).$$

The following are the sections of the study: Section 1 offers an introduction, and Section 2 examines the skew-normal distribution, followed by Section 3 which exemplified heteroscedasticity skew-normal distribution errors while in Section 4 Bayesian heteroscedasticity skew-normal restricted Stein-rule design. Thus, bayesian restricted least squares estimator with skew-normal and heteroscedastic errors were examined in Section 5 which methodically synthesizes the study's statistical foundation. Section 6 considered data generation process and performance metrics. Seventh Section covers data analysis and interpretation, whereas Section 8 deals with the conclusion.

# 2 Skew-normal distribution

The density of the univariate skew-normal distribution, $SN\left(\xi,\omega^2,\alpha\right)$ is $f_{SN}\left(x;\xi,\omega^2,\alpha\right)$ $= 2\omega^{-1}\phi(u)\varphi(\alpha u), x \in \mathbb{R}$ where $u = \omega^{-1}(x-\xi)$, $\xi \in \mathbb{R}$ is a location parameter, $\omega > 0$ is a scale parameter, $\phi(.)$ is the pdf of the univariate standard normal distribution, and $\varphi(.)$ is the cumulative distribution function of the standard normal distribution. Thus, $2\varphi(\alpha u)$ is the skewness factor which is controlled by shape parameter $\alpha \in \mathbb{R}$, the $\alpha$ indicates the direction of the distribution, if $\alpha > 0$ the distribution is skew positively or to the right, if $\alpha < 0$ the distribution is skew negatively or to the left whereas if $\alpha = 0$, the distribution becomes normal distribution.

# 3 Heteroscedasticity skew-normal distribution errors

Let $y = X\beta_R + u_t$ be restricted regression model where $y$ is an $n \times 1$ set of observations on the regressand, $X$ is a set of $n \times p$ full column rank of regressors, $\beta_R$ is $p \times 1$ vectors of unknown restricted parameters while $u$ is an $n \times 1$ vectors of disturbance with heteroscedastic skew-normal distribution errors, where $u_t = \sigma_i\varepsilon_i$, $\varepsilon_i \sim SN(0,1,\alpha)$ and $\sigma_i = \omega \exp\left(X\beta\right)^{\gamma}$. Thus, $\alpha$ denotes the shape parameter that can vary with the skewness of the distribution. The skew-normal regression is a linear regression with errors from the skew-normal distribution $u_i \overset{iid}{\sim} SN\left(0,\ \omega^2,\alpha\right)$, as a result, sample $y$ is assumed to have a skew-normal distribution, $y_i \overset{iid}{\sim} SN\left(\xi_i,\ \omega^2,\alpha\right)$ where $\xi_i = X\beta$. It is of interest that the mean $\mu$ and variance $\omega$ of a skew random variate is not the same as the location $(\xi)$ parameter, $E\left(u_i\right) \neq 0$ except $\alpha = 0$ as in normal linear regression. The mean $E\left(u_i\right) = \sqrt{2/\pi\omega\delta}$, where $\delta = \alpha/\sqrt{1+\alpha^2}$. thus $E\left(y_i\right) = \xi + E\left(u_i\right)$, thus, $s(u) = 2f(u)\varphi(u)$, $u \in \mathbb{R}$.

# 4 Bayesian heteroscedasticity skew-normal restricted Stein-rule design

Let there be $m$ linearly independent restriction that constrain the regression coefficients such that

$$r = R\beta,$$

where $r$ is an $m \times 1$ vector and $R$ is an $m \times p$ matrix of rank $m < p$ , the likelihood function for heteroscedastic skew-normal regression model is expressed as

$$L\left(\beta_R,\sigma^2|X,y\right) = \left(2\pi\omega^2\right)^{-\frac{n}{2}} \exp\left[-\frac{1}{2}\left(y-X\beta_R\right)'\left(\omega^2\Omega\right)^{-1}\left(y-X\beta_R\right)\right]$$

$$\times \exp\left[1_n^{'}\xi_0\left(\alpha^{'}\omega^{-1}\left(y - X\beta_R\right)\right)\right].$$

Take log of both sides and differentiate with respect to $\beta_R$, then skew-normal $\beta_R$ is derived as

$$\tilde{\beta}_R = \left(X'\left(\omega^2\Omega\right)^{-1}X\right)\left(X'\left(\omega^2\Omega\right)^{-1}y - 1_n^{'}\xi_0\alpha'\omega^{-1}X\right).$$

Note: the derivation of $\tilde{\beta}_R$ can be requested from the corresponding author.

The generalized restricted least squares (GRLS) estimates is obtained as follow, adopting the criterion of minimizing the sum of squares $(y - X\beta)\hat{\Omega}\hat{\Omega}(y - X\beta)$ subject to the condition that $R\beta = r$. Then restricted $\beta_R$ is expressed as Oloyede (2023).

$$\hat{\beta}_R = \tilde{\beta} + \left(X'\hat{\Omega}X\right)^{-1}R'\left[R\left(X'\hat{\Omega}X\right)^{-1}R'\right]^{-1}\left(r - R\tilde{\beta}\right),$$

thus $\hat{\beta}_R$ is a constrained estimates, following Chaturvedi et al. (2001), the restricted Stein-rule version of the disturbance errors can be expressed as

$$\hat{\beta}_S = \left[1 - \frac{a}{n}\frac{\left(y - X\tilde{\beta}\right)'\hat{\Omega}\left(y - X\tilde{\beta}\right)}{\tilde{\beta}X\hat{\Omega}X\tilde{\beta}}\right]\tilde{\beta}.$$

Therfore, the restricted Stein-rule heteroscedastic skew-normal estimator is

$$\hat{\beta}_{RS} = \hat{\beta}_S + \left(X'\hat{\Omega}X\right)^{-1}R'\left[R\left(X'\hat{\Omega}X\right)^{-1}R'\right]^{-1}\left(r - R\hat{\beta}_S\right),$$

where $\hat{\Omega} = \Omega(b)$, $b$ is a consistent and efficient estimator of $\beta_{RS}$.

# 5 Bayesian restricted least squares heeroscedatic skew-normal estimator

The posterior distribution for skew-normal restricted $\beta_R$ and is derived by integrating the conjugate of skew-normal-inverse gamma and the restricted likelihood, which belong to the same distribution family. The model is expressed as

$$\begin{aligned} y &= X\beta_R + u_t \\ L\left(\beta_R, \omega^2 | X, y\right) &= \left(2\pi\omega^2\left(1 - \frac{2\delta^2}{\pi}\right)\right)^{-\frac{n}{2}} \exp\left[-\frac{\left(y - X\beta_R\right)'\hat{\Omega}\left(y - X\beta_R\right)}{2\omega^2\left(1 - \frac{2\delta^2}{\pi}\right)}\right] \\ &\quad \times \exp\left[1_n\xi_0\alpha'\omega^{-1}(y - X\beta)\right]. \end{aligned}$$

Let $(y - X\beta_R) = \omega\sqrt{n}$, then we have

$$\begin{aligned} L\left(\beta_R, \omega^2 | X, y\right) &= \left[2\pi\omega^2\left(1 - \frac{2\delta^2}{\pi}\right)\right]^{-\frac{n}{2}} \exp\left[-\frac{\left(y - X\beta_R\right)'\hat{\Omega}\left(y - X\beta_R\right)}{2\omega^2\left(1 - \frac{2\delta^2}{\pi}\right)}\right] \\ &\quad \times \exp\left[1_n\xi_0\alpha'\sqrt{n}\right]. \end{aligned}$$

Simplify further $(2\pi)^{-\frac{n}{2}}$ is excluded because it is a constant which has no significant impact, Oloyede (2023).

$$\begin{aligned} L\left(\beta_R, \omega^2|X, y\right) \propto\ & \omega^{-n}\left(1-\frac{2\delta^2}{\pi}\right)\exp\left[-\frac{y'\hat{\Omega}y-2\beta_R{}'X'\hat{\Omega}y+\beta_R'X'\hat{\Omega}X\beta_R}{2\omega^2\left(1-\frac{2\delta^2}{\pi}\right)}\right] \\ & \times\exp\left[1_n\xi_0\alpha'\sqrt{n}\right] \end{aligned} \tag{1}$$

$$\begin{aligned} L\left(\beta_R, \omega^2|X, y\right) \propto\ & \omega^{-n}\left(1-\frac{2\delta^2}{\pi}\right)\exp\left[-\frac{1}{2\omega^2\left(1-\frac{2\delta^2}{\pi}\right)}\Big(y'\hat{\Omega}y-2\beta_R{}'X'\hat{\Omega}y\right. \\ & +\beta_R{}'X'\hat{\Omega}X\beta_R-2\left(\left(X'\hat{\Omega}X\right)^{-1}X'\hat{\Omega}y\right)'X'\hat{\Omega}y \\ & \left.+2\left(\left(X'\hat{\Omega}X\right)^{-1}X'\hat{\Omega}y\right)'X'\hat{\Omega}X\left(\left(X'\hat{\Omega}X\right)^{-1}X'\hat{\Omega}y\right)\Big)\right] \\ & \times\exp\left[1_n\xi_0\alpha'\sqrt{n}\right]. \end{aligned} \tag{2}$$

Note that (2) was obtained through replacement of $\beta R$ in (1) with $\left(X'\hat{\Omega}X\right)^{-1}X'\hat{\Omega}y)$.

$$\begin{aligned} =&\,\omega^{-n}\left(1-\frac{2\delta^2}{\pi}\right)\exp\left[-\frac{1}{2\omega^2\left(1-\frac{2\delta^2}{\pi}\right)}\left(y-Xb_R\right)'\hat{\Omega}\left(y-Xb_R\right)+b_R{}'X'\hat{\Omega}Xb_R\right. \\ & \left.+\beta_R{}'X'\hat{\Omega}X\beta_R-2\beta'X'\hat{\Omega}Xb\right]\exp\left[1_n\xi_0\alpha'\sqrt{n}\right] \\ =&\,\omega^{-n}\left(1-\frac{2\delta^2}{\pi}\right)\exp\left[-\frac{1}{2\omega^2\left(1-\frac{2\delta^2}{\pi}\right)}\left(\hat{\Omega}(n-k)+\left(\beta_R-b_R\right)'X'\hat{\Omega}X\left(\beta_R-b_R\right)\right)\right] \\ & \times\exp\left[1_n\xi_0\alpha'\sqrt{n}\right] \\ =&\,\omega^{-n}\left(1-\frac{2\delta^2}{\pi}\right)\exp\left[-\frac{1}{2\omega^2\left(1-\frac{2\delta^2}{\pi}\right)}\Big(\hat{\Omega}(n-k)1_n\xi_0\alpha'\sqrt{n}\right. \\ & \left.+\left(\beta_R-b_R\right)'X'\hat{\Omega}X\left(\beta_R-b_R\right)1_n\xi_0\alpha'\sqrt{n}\Big)\right]. \end{aligned}$$

Setting the priors

$$\begin{aligned} p\left(\beta_R|\omega^2\right) &\propto (2\pi)^{-\frac{k}{2}}\left|\hat{\Omega}\right|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}\left(\beta_R-\mathbb{B}\right)'\hat{\Omega}\left(\beta_R-\mathbb{B}\right)\right]\exp\left[1_n\xi_0\alpha'\omega^{-1}\left(\beta_R-\mathbb{B}\right)\right], \\ p\left(\beta_R|\omega^2\right) &\propto (2\pi)^{-\frac{k}{2}}\left|\hat{\Omega}\right|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}\left(\beta_R-\mathbb{B}\right)'\hat{\Omega}\left(\beta_R-\mathbb{B}\right)\right]\exp\left[1_n\xi_0\alpha'\sqrt{n}\right], \\ p\left(\omega^2\right) &\propto \omega^{-(a-k)}\exp\left[-\frac{b}{\omega^2\left(1-\frac{2\delta^2}{\pi}\right)}\right]. \end{aligned}$$

Note that the normal-inverse gamma priors, as explained by Oloyede (2023), were selected as conjugate priors because of the similarity between their prior and posterior density distributions.

$$p\left(\beta_R|\omega^2\right)p\left(\omega^2\right)=(2\pi)^{-\frac{k}{2}}\left|\hat{\Omega}\right|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}\left(\beta_R-\mathbb{B}\right)'\hat{\Omega}\left(\beta-\mathbb{B}\right)1_n\xi_0\alpha'\sqrt{n}\right]$$

$$\times\omega^{-(a-k)}\exp\left[-\frac{b}{\omega^2\left(1-\frac{2\delta^2}{\pi}\right)}\right],$$

$$p\left(\beta_R|\omega^2\right)p\left(\omega^2\right)=\omega^{-(a-k)}\exp\left[-\frac{1}{2\omega^2\left(1-\frac{2\delta^2}{\pi}\right)}\left(\beta_R-\mathbb{B}\right)'\hat{\Omega}\left(\beta_R-\mathbb{B}\right)1_n\xi_0\alpha'\sqrt{n}+2b\right].$$

$n\times k$ size of covariates matrix $X$, $\beta_R$ unknown restricted parameters, $\mathbb{B}$ prior $\mu$-vector of $\beta_R$ (true value), $\sigma^2$ prior variance for $\beta_R$, $\hat{\Omega}^2=\dfrac{(y-X\beta_R)\,\hat{\Omega}\,(y-X\beta_R)}{n-k}$, $a-k$ and $b$ are hyper-parameters, and Joint posterior density

$$\begin{aligned}\pi\left(\beta_R,\omega^2|X,y\right)\propto\,&\omega^{-n}\left(1-\frac{2\delta^2}{\pi}\right)\exp\left[-\frac{1}{2\omega^2\left(1-\frac{2\delta^2}{\pi}\right)}\Big(\hat{\Omega}(n-k)1_n\xi_0\alpha'\sqrt{n}\right.\\&\left.+\left(\beta_R-b_R\right)'X'\hat{\Omega}X\left(\beta_R-b_R\right)1_n\xi_0\alpha'\sqrt{n}\Big)\right]\omega^{-(a-k)}\\&\times\exp\left[-\frac{1}{2\omega^2\left(1-\frac{2\delta^2}{\pi}\right)}\left(\beta_R-\mathbb{B}\right)'\hat{\Omega}^{-1}\left(\beta_R-\mathbb{B}\right)1_n\xi_0\alpha'\sqrt{n}+2b\right]\\\propto\,&\omega^{-n-a+k}\left(1-\frac{2\delta^2}{\pi}\right)\exp\left[-\frac{1}{2\sigma^2}\Big(\hat{\Omega}(n-k)1_n\xi_0\alpha'\sqrt{n}\right.\\&+\left(\beta_R-b_R\right)'X'\hat{\Omega}X\left(\beta_R-b_R\right)1_n\xi_0\alpha'\sqrt{n}\Big)\\&\left.+2b+\omega^2\left(1-\frac{2\delta^2}{\pi}\right)\left(\beta_R-\mathbb{B}\right)'\hat{\Omega}\left(\beta_R-\mathbb{B}\right)1_n\xi_0\alpha'\sqrt{n}\right].\end{aligned}$$

To determine the marginal distribution of $\beta_R$, Jacobian transformation is required after $\omega^2\left(1-\frac{2\delta^2}{\pi}\right)$ is substituted with $s$, $J=\left|\frac{d}{ds}\omega\right|=\left|\frac{d}{ds}s^{-\frac{1}{2}}\right|=\frac{1}{2}s^{-\frac{3}{2}}$. Replace $\omega^2\left(1-\frac{2\delta^2}{\pi}\right)$ with $s$.

$$\begin{aligned}\pi\left(\beta_R,\omega^2|X,y\right)\propto&\left(s^{-\frac{1}{2}}\right)^{-n-a+k}\exp\left[-\frac{s}{2}\Big(\hat{\Omega}(n-k)1_n\xi_0\alpha'\sqrt{n}\right.\\&+(\beta_R-b_R)'X'\hat{\Omega}X(\beta_R-b_R)1_n\xi_0\alpha'\sqrt{n}\Big)\\&\left.+2b+s^{-1}\left(\beta_R-\mathbb{B}\right)'\hat{\Omega}\left(\beta_R-\mathbb{B}\right)1_n\xi_0\alpha'\sqrt{n}\right]\left(\frac{1}{2}s^{-\frac{3}{2}}\right).\end{aligned}$$

To obtain $\beta_R$, perform integration with respect to $s$-nuisance parameter

$$\begin{aligned}\beta_R=&\int_0^\infty\frac{1}{2}s^{\frac{n+a-k-3}{2}}\exp\left[-\frac{s}{2}\Big(\hat{\Omega}(n-k)1_n\xi_0\alpha'\sqrt{n}+\left(\beta_R-b_R\right)'X'\hat{\Omega}X\left(\beta_R-b_R\right)\right.\\&\left.\times1_n\xi_0\alpha'\sqrt{n}\Big)+2b+s^{-1}\left(\beta_R-\mathbb{B}\right)'\hat{\Omega}\left(\beta_R-\mathbb{B}\right)1_n\xi_0\alpha'\sqrt{n}\right]ds.\end{aligned}$$

Recall

$$1=\int_0^\infty\frac{q^{p+1}}{\Gamma(p+1)}s^pe^{-qs}ds,\qquad\frac{\Gamma(p+1)}{q^{p+1}}=\int_0^\infty s^pe^{-qs}ds,$$

where $p = \frac{n+a-k-3}{2}$ and

$$\begin{aligned} q &= \frac{1}{2}\Big(\hat{\Omega}(n-k)1_n\xi_0\alpha'\sqrt{n} + (\beta_R - b_R)^{'}\, X^{'}\hat{\Omega}X\,(\beta_R - b_R)\,1_n\xi_0\alpha'\sqrt{n} \\ &\quad +2b + s^{-1}\,(\beta_R{-}\mathbb{B})'\,\hat{\Omega}\,(\beta_R{-}\mathbb{B})\,1_n\xi_0\alpha'\sqrt{n}\Big), \end{aligned}$$

with df as $v = n - k$

$$\pi\,(\beta|p,q) \propto q^{-(p+1)} = q^{-\frac{n}{2}},$$

then replace it

$$\begin{aligned} \pi\,(\beta|p,q) &\propto \frac{1}{2}\Big(\hat{\Omega}(n-k)1_n\xi_0\alpha'\sqrt{n}X^{'}\hat{\Omega}X\,(\beta_R - b_R)\,1_n\xi_0\alpha'\sqrt{n} + 2b \\ &\quad +s^{-1}\,(\beta_R{-}\mathbb{B})'\,\hat{\Omega}^{-1}\,(\beta_R{-}\mathbb{B})\,1_n\xi_0\alpha'\sqrt{n}\Big)^{-\left(\frac{n+a-k-1}{2}\right)}. \end{aligned}$$

In order to derive marginal posterior density of $\omega^2$ we have

$$\begin{aligned} \pi\left(\omega^2|\beta_R, X, y\right) &\propto \int_0^\infty \omega^{-n-a+k}\exp\Big[-\frac{1}{2\sigma^2}\Big(\hat{\Omega}(n-k)1_n\xi_0\alpha'\sqrt{n} \\ &\quad +(\beta_R - b)^{'}\, X^{'}\hat{\Omega}X\,(\beta_R - B)\,1_n\xi_0\alpha'\sqrt{n}\Big) + 2b \\ &\quad +\omega^2\left(1-\frac{2\delta^2}{\pi}\right)(\beta_R{-}\mathbb{B})'\,\hat{\Omega}^{-1}\,(\beta_R{-}\mathbb{B})\,1_n\xi_0\alpha'\sqrt{n}\Big]d\beta_R. \end{aligned}$$

The study simplified further to have

$$\begin{aligned} \pi\left(\omega^2|\hat{\beta}_R, X, y\right) &\propto \sigma^{-n-a+k}\exp\left[-\frac{\hat{\Omega}(n-k)1_n\xi_0\alpha'\sqrt{n}}{2\omega^2\left(1-\frac{2\delta^2}{\pi}\right)}\right]\left(2\pi\omega^2\left(1-\frac{2\delta^2}{\pi}\right)\right)^{\frac{k}{2}}, \\ \pi\left(\omega^2|\hat{\beta}_R, X, y\right) &\propto \left(\omega^2\left(1-\frac{2\delta^2}{\pi}\right)\right)^{-\frac{1}{2}(n+a-2k)}\exp\left[-\frac{\hat{\Omega}(n-k)1_n\xi_0\alpha'\sqrt{n}}{2\omega^2\left(1-\frac{2\delta^2}{\pi}\right)}\right]. \end{aligned}$$

Thus,

$$\hat{\beta}_R \sim MVN\bigg(\hat{\beta}_R, \omega^2\left(1-\frac{2\delta^2}{\pi}\right)(X^{'}\hat{\Omega}X)^{-1}\Big[1-(X^{'}\hat{\Omega}X)^{-1}R^{'}\,\Big[R(X^{'}\hat{\Omega}X)^{-1}R^{'}\Big]^{-1}R\Big]\bigg),$$

$$\hat{\Omega}^2 \sim IG\left(a_1 - \frac{n}{2}, b_1 + \frac{1}{2}\sum_{i=1}^{n}(y_i - X\beta_R)^2 + 2\Big[R(X^{'}\hat{\Omega}X)^{-1}R^{'}\Big]^{-1}(R\beta_R - r)'(r - R\beta_R)\right),$$

$$\check{\beta}_{RS} \sim MVN\bigg(\hat{\beta}_{RS}, \omega^2\left(1-\frac{2\delta^2}{\pi}\right)(X^{'}\hat{\Omega}X)^{-1}\Big[1-(X^{'}\hat{\Omega}X)^{-1}R^{'}\Big[R(X^{'}\hat{\Omega}X)^{-1}R^{'}\Big]^{-1}R\Big]\bigg),$$

for restricted Stein-rule estimator.

# 6 Data generation processes

The MCMC simulation algorithm was adopted to examine the small sample property of the family of skew-normal restricted least squares estimators in classical and

Bayesian frameworks. Data were generated based on the following parameter of the model: $P = 6,\ n = 30,\ y_t = X_t\beta_R + u_t\ t = 1,\dots,30$, where $\varepsilon_t$ assumed to be generated by heteroscedastic skew-normal error as defined in section 3, $\hat{\beta}_R$ was set as (1.2, 2, 0.8, 0.3, 2.1, 1.1) while seed was set at 1234. 10000 iterations were set for both classical and Bayesian paradigm. The skew-normal error was generated for both negatively and positively skew distribution. Heteroscedasticity was of four categories: $\delta = 0$ No heteroscedasticity, $\delta = 0.3$ mild heteroscedasticity, $\delta = 0.6$ moderately heteroscedasticity, and $\delta = 0.9$ severe Heteroscedasticity. The restriction of parameters was set as

$$R = \begin{pmatrix} 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \qquad r = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix},$$

where $\beta_1 - \beta_3 = 0$, $\beta_2 + \beta_4 = 1$ and $\beta_5 = 0$.

The relative efficiency was computed for each estimator. $R(\hat{\beta}_R)/R(\hat{\beta}_{RS})$, the value of relative efficiency exceeding one indicates $\hat{\beta}_{RS}$ is more efficient compared to $\hat{\beta}_R$. All computations were carried out using R Statistical Software (2024). The dataset class contained the posterior sample for the model parameters.

### 6.1 Quadratic loss and risk function

Quadratic weight loss and risk function that incorporated heteroscedastic skew-normal errors were used to evaluate the performances of classical, Bayes estimate and posterior mean. Let $L(\hat{\beta}_R - \beta) = (\hat{\beta}_R - \beta)Q(\hat{\beta}_R - \beta)$ be quadratic or square error loss function where $\hat{\beta}_R$ is an estimator of $\beta$ and Q are the $\sum_{i=1}^{\beta_R} \hat{\beta}_R$ weight of loss function. For the purpose of comparison of $\hat{\beta}_R,\ and\ \hat{\beta}_{RS}$ in both classical and Bayesian Paradigm, both were compared with the weighted square error loss function and it's associated relative efficiency, to capture small sample properties.

## 7 Data analysis and discussion

Table 1 above revealed the outcome of the study the error term was either heteroscedasticity skew normal error with positive or negative tail. Classical and Bayesian estimators were considered in two separate occasions (negatively skill normal with heteroscedasticity and positively skill normal with heteroscedasticity). Risk waited precision function was adopted to assess the effectiveness of frameworks. It was discovered that Bayesian paradigm (both posterior mean and Bayes estimates) outperformed the family of restricted Stein-rule least squares in all scale of heteroscedasticity both in positively and negatively skew normal error structures.

Table 2 above revealed the relative efficiency of classical and Bayesian families of restricted least squares estimators, choosing restricted least squares as the baseline estimator with which other estimators are compared with, it was succinctly observed that Bayesian framework due its probabilistic nature were more efficient in comparison with classical estimator when the error structure is simultaneously heteroscedastic and skew-normal.

Table 1: Depicts loss and risk function of Heteroscedastic skew-normal error regression model.

| | | Skew | Classical | | Bayes-OLS | | Bayes-Stein | |
|---|---|---|---|---|---|---|---|---|
| $n$ | $\lambda$ | | OLS | Stein-rule | Post-mean | Bayes-est | Post.mean | Bayes-est. |
| 30 | 0 | -ve Skew | 9.8108 | 11.8857 | 5.9273 | 5.9252 | 6.3821 | 6.3758 |
| | 0.3 | | 10.2759 | 12.1053 | 6.0979 | 6.0958 | 6.5685 | 6.5622 |
| | 0.6 | | 10.6084 | 12.2122 | 6.1399 | 6.1378 | 6.6053 | 6.5987 |
| | 0.9 | | 10.9525 | 12.2980 | 6.1200 | 6.1178 | 6.5787 | 6.5717 |
| 250 | 0 | | 9.1764 | 9.3722 | 6.0370 | 6.0352 | 6.0700 | 6.0685 |
| | 0.3 | | 9.2653 | 9.3930 | 6.0520 | 6.0502 | 6.0851 | 6.0836 |
| | 0.6 | | 9.1575 | 9.2503 | 5.9502 | 5.9484 | 5.9828 | 5.9813 |
| | 0.9 | | 9.2011 | 9.1846 | 5.8919 | 5.8901 | 5.9241 | 5.9227 |
| 30 | 0 | +ve Skew | 19.5136 | 19.4580 | 12.1814 | 12.1780 | 13.1357 | 13.1258 |
| | 0.3 | | 18.8017 | 19.1077 | 11.6334 | 11.6299 | 12.5391 | 12.5288 |
| | 0.6 | | 18.3860 | 19.1046 | 11.1300 | 11.1264 | 11.9935 | 11.9824 |
| | 0.9 | | 19.6369 | 20.0537 | 11.6185 | 11.6148 | 12.5197 | 12.5084 |
| 250 | 0 | | 12.8441 | 12.7247 | 8.4891 | 8.4865 | 8.5349 | 8.5327 |
| | 0.3 | | 12.8383 | 12.8140 | 8.4584 | 8.4558 | 8.5039 | 8.5017 |
| | 0.6 | | 12.8539 | 12.9247 | 8.4175 | 8.4149 | 8.4629 | 8.4607 |
| | 0.9 | | 12.8164 | 12.9998 | 8.3254 | 8.3229 | 8.3705 | 8.3683 |

Table 2: Depicts relative efficiency of Heteroscedastic skew-normal error regression model.

| | | Skew | Classical | Bayes-OLS | | Bayes-Stein | |
|---|---|---|---|---|---|---|---|
| $n$ | $\lambda$ | | Stein-rule | Post-mean | Bayes-est | Post.mean | Bayes-est. |
| 30 | 0 | -ve Skew | 1.211489 | 0.60416 | 0.603944 | 0.650511 | 0.6498 |
| | 0.3 | | 1.178023 | 0.593424 | 0.593216 | 0.639221 | 0.63860 |
| | 0.6 | | 1.151187 | 0.578783 | 0.578582 | 0.622649 | 0.62203 |
| | 0.9 | | 1.122848 | 0.558776 | 0.558582 | 0.600654 | 0.60002 |
| 250 | 0 | | 1.021337 | 0.657883 | 0.657687 | 0.661479 | 0.6613 |
| | 0.3 | | 1.013783 | 0.65319 | 0.652996 | 0.656762 | 0.6566 |
| | 0.6 | | 1.010134 | 0.649762 | 0.649566 | 0.653322 | 0.65315 |
| | 0.9 | | 0.998207 | 0.640347 | 0.640152 | 0.643847 | 0.64369 |
| 30 | 0 | +ve Skew | 0.997151 | 0.624252 | 0.624078 | 0.673156 | 0.67264 |
| | 0.3 | | 1.016275 | 0.618742 | 0.618556 | 0.666913 | 0.66636 |
| | 0.6 | | 1.039084 | 0.605352 | 0.605156 | 0.652317 | 0.65171 |
| | 0.9 | | 1.021225 | 0.591667 | 0.591478 | 0.63756 | 0.63698 |
| 250 | 0 | | 0.990706 | 0.660932 | 0.660733 | 0.664501 | 0.66432 |
| | 0.3 | | 0.998106 | 0.658836 | 0.658639 | 0.662389 | 0.66221 |
| | 0.6 | | 1.005507 | 0.654854 | 0.654657 | 0.658388 | 0.65821 |
| | 0.9 | | 1.014309 | 0.649589 | 0.649395 | 0.653105 | 0.65293 |

# 8 Conclusion

This study has proposed a novelty generalized restricted least squares with heteroscedastic skew-normal error, the skew normal prior was used to conjugate heteroscedastic skew normal likelihood, the skew normal posterior density provide Bayesian generalized restricted least squares with heteroscedastic skew-normal error, the uncertainty of the parameters is embedded in its restriction. Thus, risk weighted precision function was adopted to evaluate the performances of the families of both classical and Bayesian estimators considered in the study. Skew-normal prior was adopted to elicit information about the uncertainty of parameters (state of the nature) in a Bayesian paradigm. The study demonstrated superior performance of Bayesian frameworks (both posterior mean and bayes estimate) in comparison with family of classical restricted least squares estimators. The study therefore recommended Bayesian generalized restricted least squares to the analyst and research who may encounter heteroscedastic skew-normal error in their data.

# References

Alhamide, A.A., Kamarulzaman I., Alodat, M.T. and Wan Z.W.Z. (2019). Bayesian inference for linear regression under alpha-skew-normal prior. *Sains Malaysiana*, **48**(1):227–235.

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, **12**(2):171–178.

Azzalini, A. (1986). Further results on a class of distributions which includes the normal ones. *Statistica*, **46**(6):199–208.

Azzalini, A. (2005). The skew-normal distribution and related multivariate families. *Scandinavian Journal of Statistics*, **32**(2):159–188.

Azzalini, A. and Dalla, V. (1996). The multivariate skew-normal distribution. *Biometrika*, **83**(4):715–726.

Cepeda, E. and Germaman, D. (2000). Bayesian modelling of variance heterogeneity in normal regression models. *Brazil Journal of Probability Statistics*, **14**(2):207–221.

Chaturvedi, A, Wan, A.T.K and Singh, S.P. (2001). Stein-rule restricted regression estimator in a linear regression model with non-spherical disturbances. *Communication in Statistical Theory Methods*, **30**(1):55–68.

Corrales, M.L. and Cepeda-Cuervo, E. (2022). Bayesian modelling of location, scale and shape parameters in skew-normal regression models. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, **15**(1):98–111.

Ferreira, J.T. and Steel, M.F. (2007). A new class of skewed multivariate distributions with applications to regression analysis. *Statistica Sinica*, **17**(2):505–529.

Genton, M.G., He, L. and Liu, X. (2001). Moments of skew-normal random vectors and their quadratic forms. *Statistics & probability letters*, **51**(4):319–325.

Harvey, A.C. (1976). Estimating regression models with multiplicative heteroscedasticity. *Econometrica: journal of the Econometric Society*, **44**(3):461–465.

Lachos, V.H., Bolfarine, H., Arellano-Valle, R.B. and Montenegro, L.C. (2007). Likelihood-based inference for multivariate skew-normal regression models. *Communication Statistics Theory Methods*, **36**(9):1769–1786.

Indacochea, D. (2012). *The Effects of skewness and kurtosis on heteroskedasticity-robust bootstrap methods in finite samples*. Doctoral dissertation, Queen's University.

O'Hagan, A. and Leonard, T. (1976). Bayes estimation subject to uncertainty about parameter constraints. *Biometrika*, **63**(1):201–203.

Oloyede, I. (2023). Bayesian linear restricted least squares estimators. *Journal of the Indian Society for Probability and Statistics*, **24**(1):65–74.

Oloyede, I., Ipinyomi, R.A. and Iyaniwura, J.O. (2013). Bayesian generalized least squares with parametric heteroscedasticity linear model, *Asian Journal of Mathematics and Statistics*, **6**:57–75.

R Core Team (2024). *R: A Language and Environment for Statistical Computing.* Version 4.22, R foundation for Statistical Computing. https://www.R-project.org

Rubio, F.J. and Genton, M.G. (2016). Bayesian linear regression with skew-symmetric error distributions with applications to survival analysis. *Statistics in Medicine*, **35**(14):2441–2454.