International Journal of Mathematics and Statistics Invention (IJMSI)
E-ISSN: 2321 – 4767 P-ISSN: 2321 - 4759
www.ijmsi.org Volume 12 Issue 4 || Jul.-Aug. 2024 || PP-15-21

Quantile Based Slope Estimators in Simple Linear Regression

Sharada V. Bhat and *Shrinath M. Bijjargi
Department of Statistics, Karnatak University, Dharwad - 580 003, India.
Corresponding Author: Shrinath M. Bijjargi
Email: shrinathphd@gmail.com

***ABSTRACT**: In this paper, some estimators are proposed for slope parameter of simple linear regression (SLR). These estimators are based on the quantiles of slopes obtained using various kinds of distances among predictor observations. The variances of the proposed estimators under various symmetric error distributions are obtained. The optimality of the quantile estimators is established by minimizing their variances. The performance of the estimators is evaluated. The estimators are extended to multiple responses for a single predictor variable.*

***KEYWORDS**: Quantile based estimator, Quasi range, Robust, Slope parameter, Simple linear regression and optimality.*

Date of Submission: 05-08-2024 Date of acceptance: 17-08-2024

I. INTRODUCTION

Regression analysis serves as a fundamental tool in statistical modeling offering insights into relationships between variables through the estimation of parameters. One of the simplest forms in regression is SLR model. Let, $Y_i, i = 1, \dots, n$ be the response variable with distributions

$$P(Y_i \le y) = F_i(y) = F(y - \alpha - \beta x_i) = F(e_i). \quad (1)$$

Here, $F(\cdot)$ is continuous cumulative distribution function (cdf), $x_1, x_2, \dots, x_n$ are known constants, α and β are the intercept and slope parameters respectively. The error, e_i capturing deviations from the linear relationship are independent and identically distributed (iid) from continuous distribution with cdf $F(e_i)$ having zero mean and finite variance σ^2.

The estimation of β is crucial in SLR model by (1) for quantifying the rate of change in y with respect to x. The least squares (LS) method which provides best linear unbiased estimator (BLUE) is sensitive to outliers and non-normal data. Hence, robust estimation techniques have been developed in literature by focusing on estimators of slopes based on various measures of central tendency obtained using different types of distances viz. half ranges and quasi ranges among $x_i s$. Let $x_{(i)}$ be the i^{th} order statistic and y_i^* be the corresponding y value. For $n = 2m$, the i^{th} half range is defined as

$$h_i = \left(x_{(m+i)} - x_{(i)}\right), \; i = 1, 2, \cdots m, \quad (2)$$

i^{th} quasi range is defined as

$$q_i = x_{(n-i)} - x_{(i+1)}, \; i = 1, 2, \cdots m - 1 \quad (3)$$

and $q_0 = x_{(n)} - x_{(1)}$ is the range of n observations.

The complete method due to [3] assumes all $x_{(i)}s$ are distinct and yields a robust estimator for β based on the median of slopes $b_i = \frac{y_j^* - y_i^*}{x_{(j)} - x_{(i)}}, 1 \le i < j \le n$ resulting in $\frac{n(n-1)}{2}$ slopes. The incomplete method due to [3] is based on the median of slopes $b_i = \frac{y_{m+i}^* - y_i^*}{h_i}$ which make use of half ranges resulting in $\frac{n}{2}$ slopes. [8] extended Theil's complete method to cases where not all $x_{(i)}s$ are distinct, leading to the well-known Theil-Sen estimator. [11] proposed an alternative estimator based on median of subset of slopes given by

$$b_i = \frac{y_{i+m}^* - y_i^*}{x_{(i+m)} - x_{(i)}} \text{ and } b_{i+1} = \frac{y_{i+m+1}^* - y_i^*}{x_{(i+m+1)} - x_{(i)}}$$

for $m = \frac{n}{2}$ if n is even and $m = \frac{n-1}{2}$ if n is odd. [2] suggested an estimator based on location estimator due to [4] instead of median. [1] and [7] assuming equidistant $x_{(i)}s$, utilized median and Hodges-Lehmann estimator in Theil's incomplete method and generalized the regression model with multiple responses at $x_{(i)}$. [6] proposed an estimator for β based on average of successive slopes $b_i = \frac{y_i - y_{i-1}}{x_i - x_{i-1}}, i = 2, \dots, n$. Subsequently, [5] proposed an estimator where the average was replaced by the median of successive slopes. [12] proposed a class of estimators based on weighted averages of slopes $b_i = \frac{y^*_{m+i} - y^*_{m-i+1}}{q_{m-i}}$ with weights i^k and found its optimality at $k = 2$.

In this paper, we propose the estimators based on quantiles of the slopes obtained from $h_i s$ and $q_i s$. These estimators provide flexible and robust method for estimating the slope parameter and is a generalization of estimators based on median. They also capture a broader range of data characteristics and improve the performance of estimators under various distributions.

The proposed estimators are given in section 2, their distributional properties are derived in section 3 and section 4 contains optimization of quantile based slope estimators. The proposed estimators are illustrated through examples in section 5 and their extension in case of multiple responses are given in section 6. The conclusions are provided in section 7.

II. PROPOSED QUANTILE BASED ESTIMATORS

In this section, we propose quantile based estimators. The usage of quantiles in the context of estimation of β, allows for flexibility in estimating β by using various partition values of the distribution of the slopes. By considering different quantiles, we can develop estimators that are adaptable to different nature of data and error distribution leading to robust and efficient estimators. In regression modeling, when we encounter situations where the error distribution is not normal, the robust estimators become preferable.

Quantiles are statistical measures that divide a data set or probability distribution into intervals with equal probabilities. The p^{th} sample quantile is the value below which, a fraction p of the data falls. The asymptotic distribution of p^{th} sample quantile, $\hat{\xi}_p$ is given by

$$N\left(\xi_p, \frac{p(1-p)}{nf^2(\xi_p)}\right) \tag{4}$$

where, f is probability density function evaluated at ξ_p.

We propose quantile based estimators from the sample slopes obtained through different kinds of distances. The slope b_i can be generally expressed as

$$b_i = \frac{D_{y^*_i}}{D_{x_{(i)}}}, \quad i = 1, \dots, m \tag{5}$$

where, $D_{x_{(i)}}$ represent either h_i or q_{m-i} of variable x and $D_{y^*_i}$ is difference of corresponding y values. The proposed estimators of β are given by

$$\hat{\beta}_p = p^{th}\ quantile\ of\ b_i = b_{(r)}, \quad 0 < p < 1, \tag{6}$$

$$r = \begin{cases} mp & , if\ mp\ is\ an\ integer \\ [mp+1]\ or\ [mp]+1 & , if\ mp\ is\ not\ an\ integer \end{cases},$$

where, $[t]$ is largest integer $\leq t$. The specific choice of p can be determined based on the desired properties of the estimator, such as minimizing the variance of $\hat{\beta}_p$ or the distribution $F(\cdot)$.

III. DISTRIBUTIONAL PROPERTIES OF $\hat{\boldsymbol{\beta}}_p$

In this section, we deal with the distributional properties of the proposed quantile based slope estimators. We obtain mean and variance of these estimators under different error distributions.

Let
$$b_i = \frac{D_{y^*_i}}{D_{x_{(i)}}} = \beta + \frac{D_{e_i}}{D_{x_{(i)}}}$$
$$= \beta + u_i, \tag{7}$$

where $u_i = \frac{D_{e_i}}{D_{x_{(i)}}}$ is continuous random variable and symmetric around zero. Hence, it follows that, $E(b_i) = \beta\ \forall\ i = 1, \dots, m$ and b_i is symmetric around β. Therefore, the proposed estimator which is any quantile of b_i, will be an unbiased estimator of β. Hence,

$$\sqrt{m}(\hat{\beta}_p - \beta) \to N\left(0, \frac{p(1-p)}{f^2(\xi_p)}\right) \tag{8}$$

where $f(\xi_p)$ is the density of b_i evaluated at ξ_p. The distribution of b_i is influenced by the distribution of u_i which is depending on distribution of D_{e_i}. We derive the variances of $\hat{\beta}_p$ when e_i has uniform, normal, Laplace and Cauchy distributions with mean zero and variance σ^2.

- **Uniform distribution**

  When $e_i \sim U(-\sqrt{3}\sigma, \sqrt{3}\sigma)$, the probability density function of u_i is given by

$$f(u_i)=\begin{cases}\frac{D_{x_{(i)}}\left(D_{x_{(i)}}u_i+2\sqrt{3}\sigma\right)}{12\sigma^2} & ;\ -\frac{2\sqrt{3}\sigma}{D_{x_{(i)}}}<u_i\le 0\\ \frac{D_{x_{(i)}}\left(2\sqrt{3}\sigma-D_{x_{(i)}}u_i\right)}{12\sigma^2} & ;\ 0<u_i\le\frac{2\sqrt{3}\sigma}{D_{x_{(i)}}}\end{cases} \tag{9}$$

and $b_i\sim Tr\left(\beta-\frac{2\sqrt{3}\sigma}{D_{x_{(i)}}},\beta+\frac{2\sqrt{3}\sigma}{D_{x_{(i)}}},\beta\right)$ which is symmetric around β. The density of b_i is given by

$$f(b_i)=\begin{cases}\frac{D^2_{x_{(i)}}\left(b_i-\beta+\frac{2\sqrt{3}\sigma}{D_{x_{(i)}}}\right)}{12\sigma^2} & ;\ \beta-\frac{2\sqrt{3}\sigma}{D_{x_{(i)}}}<b_i\le\beta\\ \frac{D^2_{x_{(i)}}\left(-b_i+\beta+\frac{2\sqrt{3}\sigma}{D_{x_{(i)}}}\right)}{12\sigma^2} & ;\beta<b_i\le\beta+\frac{2\sqrt{3}\sigma}{D_{x_{(i)}}}\end{cases}. \tag{10}$$

To obtain variance of $\hat{\beta}_p$, ξ_p is obtained by solving $F_{b_i}(\xi_p)=p$, that is,

$$F_{b_i}(\xi_p)=\begin{cases}\int_{\beta-\frac{2\sqrt{3}\sigma}{D_{x_{(i)}}}}^{\xi_P}\frac{D^2_{x_{(i)}}\left(b_i-\beta+\frac{2\sqrt{3}\sigma}{D_{x_{(i)}}}\right)}{12\sigma^2}\,db_i=p\ ;\ \beta-\frac{2\sqrt{3}\sigma}{D_{x_{(i)}}}<b_i\le\beta\\ \frac{1}{2}+\int_{\beta}^{\xi_P}\frac{D^2_{x_{(i)}}\left(-b_i+\beta+\frac{2\sqrt{3}\sigma}{D_{x_{(i)}}}\right)}{12\sigma^2}db_i=p\quad ;\beta<b_i\le\beta+\frac{2\sqrt{3}\sigma}{D_{x_{(i)}}}\end{cases}. \tag{11}$$

Solving (11), we get

$$\xi_p=\begin{cases}\beta-\frac{2\sqrt{3}\sigma}{D_x}+\sqrt{p\left(\frac{24\sigma^2}{D_x^2}\right)} & ;\beta-\frac{2\sqrt{3}\sigma}{D_x}<b_i\le\beta\\ \beta+\frac{2\sqrt{3}\sigma}{D_x}-\sqrt{(1-p)\left(\frac{24\sigma^2}{D_x^2}\right)} & ;\ \beta<b_i\le\beta+\frac{2\sqrt{3}\sigma}{D_x}\end{cases}$$

where, D_x is h_i or q_{m-i} of the p^{th} quantile of b_is and $f^2(\xi_p)$ is given by

$$f^2(\xi_p)=\begin{cases}\frac{p\,D_x^2}{6\sigma^2} & ;0<p\le\frac{1}{2}\\ \frac{(1-p)\,D_x^2}{6\sigma^2} & ;\ \frac{1}{2}<p<1\end{cases}.$$

Using (8), on substituting $f^2(\xi_p)$, the variance of $\hat{\beta}_p$ is given by

$$V(\hat{\beta}_p)=\begin{cases}\frac{6(1-p)\sigma^2}{mD_x^2} & ;\ 0<p\le\frac{1}{2}\\ \frac{6p\sigma^2}{mD_x^2} & ;\ \frac{1}{2}<p<1\end{cases}. \tag{12}$$

Proceeding on similar lines, we obtain $V(\hat{\beta}_p)$ when e_i has normal, Laplace and Cauchy distributions.

- **Normal distribution**

When $e_i\sim N(0,\sigma^2)$, $b_i\sim N\left(\beta,\frac{2\sigma^2}{D^2_{x_{(i)}}}\right)$ and $\xi_p=\beta+\frac{2\sigma}{D_x}\,erf^{-1}(2p-1)$, where $erf(x)$ is error function defined as $erf(x)=\frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt.$

Hence, $$f^2(\xi_p)=\frac{D_x^2}{4\pi\sigma^2}e^{-\left(\sqrt{2}\,erf^{-1}(2p-1)\right)^2}$$

and $$V(\hat{\beta}_p)=\frac{4p(1-p)\pi\sigma^2}{mD_x^2e^{-\left(\sqrt{2}\,erf^{-1}(2p-1)\right)^2}}\ . \tag{13}$$

- **Laplace distribution**

When $e_i\sim L\left(0,\frac{\sqrt{2}}{\sigma}\right)$, then b_i has pdf given by

$$f(b_i)=\frac{\sigma D_{x_{(i)}}}{4\sqrt{2}}\left(1+\frac{\sigma D_{x_{(i)}}}{\sqrt{2}}|b_i-\beta|\right)e^{-\frac{\sigma D_{x_{(i)}}}{\sqrt{2}}|b_i-\beta|};\ -\infty<b_i<\infty,\sigma>0\ . \tag{14}$$

Then,

$$F(\xi_p)=\begin{cases}\frac{1}{4}\left(2-\frac{\sigma D_{x_{(i)}}}{\sqrt{2}}(\xi_p-\beta)\right)e^{\frac{\sigma D_{x_{(i)}}}{\sqrt{2}}(\xi_p-\beta)} & ;\ b_i<\beta\\ 1-\frac{1}{4}\left(2-\frac{\sigma D_{x_{(i)}}}{\sqrt{2}}(\xi_p-\beta)\right)e^{-\frac{\sigma D_{x_{(i)}}}{\sqrt{2}}(\xi_p-\beta)} & ;\ b_i>\beta\end{cases}. \tag{15}$$

As solving (15) directly for ξ_p is complicated, we use Lambert W function discussed in [13]. It is denoted as $W(z)$ which is used in solving equations of the form $te^t=C$. The ξ_p is given by

$$\xi_p = \begin{cases} \beta + \frac{\sqrt{2}\left(2+W\left(\frac{-4p}{e^2}\right)\right)}{\sigma D_x} & ;\ b_i < \beta \\ \beta + \frac{\sqrt{2}\left(2-W\left(\frac{4(1-p)}{e^{-2}}\right)\right)}{\sigma D_x} & ;\ b_i < \beta \end{cases}. \tag{16}$$

Hence,

$$f^2(\xi_p) = \begin{cases} \frac{\sigma^2 D_x^2}{32}\left(-1 - W\left(\frac{-4p}{e^2}\right)\right)^2 e^{4+2W\left(\frac{-4p}{e^2}\right)} & ; 0 < p \le \frac{1}{2} \\ \frac{\sigma^2 D_x^2}{32}\left(3 - W\left(\frac{4(1-p)}{e^{-2}}\right)\right)^2 e^{4-2W\left(\frac{4(1-p)}{e^{-2}}\right)} & ; \frac{1}{2} < p < 1 \end{cases} \tag{17}$$

and

$$V(\hat{\beta}_p) = \begin{cases} \frac{32p(1-p)}{m\sigma^2 D_x^2\left(-1-W\left(\frac{-4p}{e^2}\right)\right)^2 e^{4+2W\left(\frac{-4p}{e^2}\right)}} & ;\ 0 < p \le \frac{1}{2} \\ \frac{32p(1-p)}{m\sigma^2 D_x^2\left(3-W\left(\frac{4(1-p)}{e^{-2}}\right)\right)^2 e^{4-2W\left(\frac{4(1-p)}{e^{-2}}\right)}} & ;\ \frac{1}{2} < p < 1 \end{cases}. \tag{18}$$

- **Cauchy distribution**

When $e_i \sim C(0, \lambda)$, $b_i \sim C\left(\beta, \frac{2\lambda}{D_{x_{(i)}}}\right)$ and $\xi_P = \beta + \frac{2\lambda}{D_x}\tan\left(\pi p - \frac{\pi}{2}\right)$.

Hence,

$$f^2(\xi_p) = \frac{D_x^2}{4\lambda^2\pi^2\left(1+tan^2\left(\pi p - \frac{\pi}{2}\right)\right)^2}$$

and

$$V(\hat{\beta}_p) = \frac{4p(1-p)\lambda^2\pi^2\left(1+tan^2\left(\pi p-\frac{\pi}{2}\right)\right)^2}{m\ D_x^2}. \tag{19}$$

IV. OPTIMIZATION OF QUANTILE BASED SLOPE ESTIMATORS

In this section, we find the optimal quantile slope estimator obtained at minimum $V(\hat{\beta}_p)$ for different error distributions using numerical method. Also, we evaluate the performance of the proposed estimators.

We use a numerical optimization technique due to [10] to achieve minimization for the variance expressions given by (12), (13), (18) and (19). It is a root-finding algorithm which does not require the derivative of the function. By applying Brent's method, we observe that, for uniform, normal and Cauchy distributions, the $V(\hat{\beta}_p)$ is minimum at $p = \frac{1}{2}$, which corresponds to the median of b_i. This suggests that the median is the optimal quantile for uniform, normal and Cauchy error distributions. However, for Laplace distribution, as p increases, the variance decreases, reaching its minimum at $p = 0.99$. Thus, for Laplace distribution, 99th percentile yields minimum $V(\hat{\beta}_p)$. In figure 1, we plot the $V(\hat{\beta}_p)$ for $0 < p < 1$, $D_x = 1$, $m = 100$, $\lambda = 1$ for Cauchy distribution and $\sigma = 1$ for other distributions under consideration.

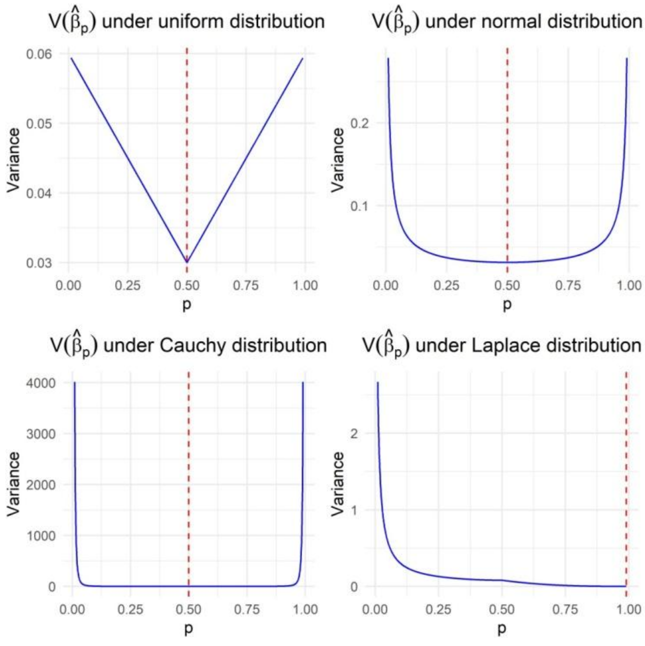

Figure 1. $V(\hat{\beta}_p)$ **under various distributions**

The Figure 1 reveals that, the optimal quantiles obtained from $V(\hat{\beta}_p)$ plotted against p and obtained from Brent's method are the same. These results highlight the importance of selecting an appropriate quantile estimator based on distribution of e_i. The robustness of these optimal quantiles to varying sample sizes enhance their practical utility.

The performance of the estimators is also evaluated by considering relative efficiency (RE) which is the ratio of the variances of two estimators, indicating how much more efficient one estimator is compared to another. Suppose, $\hat{\beta}_{\theta_1}$ and $\hat{\beta}_{\theta_2}$ are two unbiased estimators, then

$$RE(\hat{\beta}_{\theta_1}, \hat{\beta}_{\theta_2}) = \frac{V(\hat{\beta}_{\theta_2})}{V(\hat{\beta}_{\theta_1})} \tag{20}$$

where θ_1 and θ_2 are any two quantiles. The $V(\hat{\beta}_p)$ are computed for various values of p and error distributions in Table 1.

Table 1. $V(\hat{\beta}_p)$ for various error distributions

Error distribution	p						
	0.01	0.10	0.25	0.50	0.75	0.90	0.99
Uniform	0.059400	0.054000	0.045000	0.030000	0.045000	0.054000	0.059400
Normal	0.002787	0.000584	0.000371	0.000314	0.000371	0.000584	0.002787
Laplace	2.671198	0.301563	0.128939	0.080000	0.011886	0.001107	0.000012
Cauchy	4014.959828	3.896483	0.296088	0.098696	0.296088	3.896483	4014.959828

From Table 1, it is observed that, for uniform, normal and Cauchy distributions, $RE(\hat{\beta}_{0.5}, \hat{\beta}_{0.25}) = RE(\hat{\beta}_{0.5}, \hat{\beta}_{0.75}) > 1$ which shows $\hat{\beta}_{0.5}$ is better estimator than $\hat{\beta}_{0.25}$ and $\hat{\beta}_{0.75}$. However, RE of $(\hat{\beta}_{0.5}, \hat{\beta}_{0.75})$ is 1.5, 1.19 and 2.99 indicating 50%, 19% and 199% efficiency of median as estimator compared to other quartiles respectively for uniform, normal and Cauchy distributions. For Laplace error distribution, $RE(\hat{\beta}_{0.5}, \hat{\beta}_{0.25}) = 1.6 > 1$ and $RE(\hat{\beta}_{0.5}, \hat{\beta}_{0.75}) = 0.1486 < 1$. This shows that, as p increases, efficiency of $\hat{\beta}_p$ increases. However, while the efficiency increases as p approaches 0.99, the robustness of the estimator may relatively decrease. Therefore, for Laplace distribution, any quantile with $p > \frac{1}{2}$ is more efficient than the median and the choice of p may be balanced between efficiency and robustness based on the data characteristics.

V. ILLUSTRATION

In this section, we illustrate the proposed estimator using data from [9]. The dataset contains 12 observations from a Belgian insurance company, which gives details of monthly payments made in 1979 due to the expiration of life-insurance contracts. The payments are expressed as a percentage of the total annual amount as per the company's reporting standards. Notably, the December payments exhibit a significant spike, attributed to one extraordinarily high supplementary pension payout, introducing an outlier effect in the data. The data is given by

Table 2. Monthly payment data due to [9]

Month (x)	Payment (y)
1	3.22
2	9.62
3	4.50
4	4.94
5	4.02
6	4.20
7	11.24
8	4.53
9	3.05
10	3.76
11	4.23
12	42.69

After examining the distribution of the data using R software, it is found that, the data follows Cauchy distribution. The computed values of $\hat{\beta}_p$ and $V(\hat{\beta}_p)$ under Cauchy distribution using h_i and q_{m-i} are presented in Table 3 and regression lines are plotted in Figure 2.

Table 3. Computations of $\hat{\beta}_p$ and $V(\hat{\beta}_p)$

p	$D_{x_{(i)}} = h_i$		$D_{x_{(i)}} = q_{m-i}$	
	$\hat{\beta}_p$	$V(\hat{\beta}_p)$	$\hat{\beta}_p$	$V(\hat{\beta}_p)$
0.10	-0.84833	$1.803927\lambda^2$	-0.58889	$0.801745\lambda^2$
0.25	-0.24167	$0.137078\lambda^2$	-0.37800	$0.197392\lambda^2$
0.50	-0.19667	$0.045693\lambda^2$	-0.10571	$0.033570\lambda^2$
0.75	1.33667	$0.137078\lambda^2$	3.58818	$0.040783\lambda^2$
0.90	6.41500	$1.803927\lambda^2$	7.04000	$64.94138\lambda^2$

To fit the regression model given in (1), α is estimated using $\hat{\alpha} = \tilde{y} - \hat{\beta}_p \tilde{x}$ where, $\tilde{y}$ and $\tilde{x}$ are median of y and x respectively.

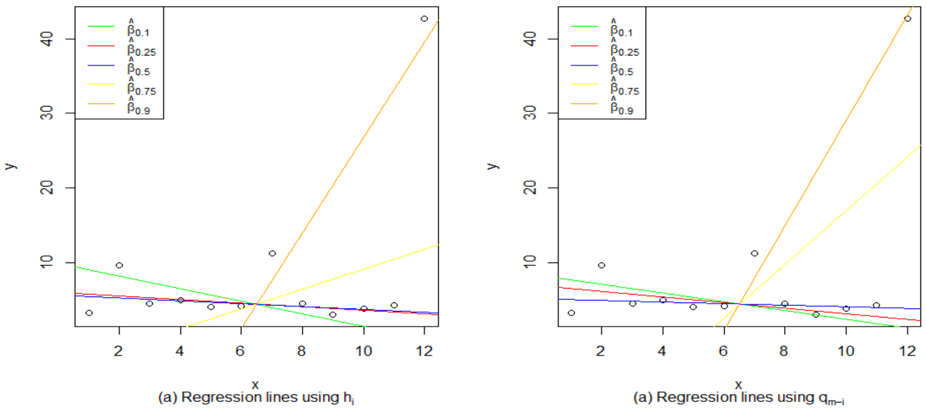

Figure 2. The fitted regression lines

From Table 3 and Figure 2, it is observed that, $\hat{\beta}_{0.5}$ has minimum variance and shows better fit to the data. Also, the estimators obtained using q_{m-i} have lower variances than those obtained from h_i for $\hat{\beta}_{0.10}$, $\hat{\beta}_{0.5}$ and $\hat{\beta}_{0.75}$.

VI. EXTENSION OF PROPOSED ESTIMATORS

In this section, we extend our proposed estimators for β to accommodate situations where multiple responses, y_{jk} are observed for each predictor x_j. We encounter two kinds of situations where in the estimator is proposed considering slopes of all the responses and estimator is proposed using the function of responses. In the first situation, the model is formulated as

$$y_{jk} = \alpha + \beta x_j + e_{jk}, \quad j = 1, \dots, n, \; k = 1, \dots, c \, . \tag{21}$$

where y_{jk} represents the response for the j^{th}observation at the k^{th} instance. The corresponding errors, e_{jk} are symmetrically distributed around zero with variance σ^2. Arrange the data pairs (y_{jk}, x_j) in ascending order based on x_j, with y^*_{jk} representing the y_{jk} observations corresponding to the $x^{th}_{(j)}$ order statistics. The y^*_{jk} values are random and sorted according to their occurrence. When $D_{x_{(i)}} = h_i$, we get $mc = N$ slopes given by

$$b_{ij} = \frac{y^*_{(m+i)j} - y^*_{ij}}{h_i}; \quad i = 1, 2, \dots, m, j = 1, 2, \dots, c \tag{22}$$

and when $D_{x_{(i)}} = q_{m-i}$, we get $mc = N$ slopes given by

$$b_{ij} = \frac{y^*_{(m+i)j} - y^*_{(m-i+1)j}}{q_{m-i}}; \quad i = 1, 2, \dots, m, j = 1, 2, \dots, c \tag{23}$$

Hence, the proposed estimator given in (6) is generalized for N slopes by replacing $\hat{\beta}_p$ by $\hat{\beta}^*_p$, b_i by b_{ij} and m by N. $\hat{\beta}^*_p$ is also an unbiased estimator of β. The variance of $\hat{\beta}^*_p$ can be obtained by making appropriate changes in (12), (13), (18) and (19) for uniform, normal, Laplace and Cauchy error distributions respectively. In the second situation, the model can be taken considering various functions of y_{jk} such as the maximum, minimum, average and median of y_{jk}. That is,

$$g(y_{j\cdot}) = \alpha + \beta x_j + g(e_{j\cdot}), \qquad j = 1, 2, \dots, n \tag{24}$$

where $g(\cdot)$ and $g(e_{j\cdot})$ are functions of $y_{j\cdot}$ and $e_{j\cdot}$ respectively. Also, $g(e_{j\cdot})$ are iid, symmetric around zero with constant variance σ^2_g. Here, we get m slopes given by

$$b'_i = \frac{D_{g(y^*_{i\cdot})}}{D_{x_{(i)}}} = \beta + \frac{D_{g(e_{i\cdot})}}{D_{x_{(i)}}}, i = 1, 2, \dots, m$$
$$= \beta + u'_i$$

where $u_i' = \frac{D_{g(e_{i.})}}{D_{x_{(i)}}}$ is symmetric around zero. Hence, $E(b_i') = \beta \ \forall \ i = 1,2,\dots,m$. The proposed estimators are given by replacing $\hat{\beta}_p$ by $\hat{\beta}_p'$ and b_i by b_i' in (6). The $\hat{\beta}_p'$ is an unbiased estimator of β and $V(\hat{\beta}_p') = \frac{p(1-p)}{mf'^2(\xi_p)}$, where, $f'^2(\cdot)$ is pdf of b_i' evaluated at ξ_p.

VII. CONCLUSIONS

- The estimators based on p^{th} quantile of slopes obtained using two kinds of distances among $x_i s$ are proposed for slope parameter β in SLR.
- The distributions of slopes are derived when e_i follows uniform, normal, Laplace and Cauchy distributions.
- The proposed estimators are consistent and unbiased.
- The efficiency of the estimators is established through deriving their variances and identifying optimal quantiles (p) for all the distributions under consideration.
- The optimal p is obtained using Brent's numerical method, by minimizing the variances. It is found that, for uniform, normal and Cauchy distributions, the estimator based on median $(p = 0.5)$ remains optimal, while for Laplace distribution, the estimator based on 99^{th} percentile $(p = 0.99)$ is optimal.
- Among the proposed estimators, the estimator using the quasi ranges of predictor variables is befitting than those using half ranges.
- The proposed estimators are extended to include scenarios such as occurrence of multiple responses.

REFERENCES

[1]. A. P. Gore and K. M. Rao, Nonparametric tests for slope in linear regression problems, Biometrical Journal, 24(3), 1982, 229-237.
[2]. H. Nadia and A. A. Mohammad, Model of robust regression with parametric and nonparametric methods, Mathematical Theory and Modeling, 3(5), 2013, 27-39.
[3]. H. Theil, A rank-invariant method of linear and polynomial regression analysis, Indagationes mathematicae, 12(85), (1950), 85-173.
[4]. J. L. Gastwirth, On robust procedures, Journal of the American Statistical Association, 61(316), 1966, 929-948.
[5]. J. Singthongchai, N. Thongmual and N. Nitisuk, An Improved Simple Averaging Approach for Estimating Parameters in Simple Linear Regression Model, Mathematics and Statistics, 9(6), 2021, 939 - 946.
[6]. K. R. Cliff and K. M. Billy, Estimation of the parameters of a linear regression system using the simple averaging method, Global Journal of Pure and Applied Mathematics, 13(11), 2017, 7749-7758.
[7]. M. K. Rao, Contribution to nonparametric inference in regression and other linear models, PhD thesis, 1982, Savitribai Phule Pune University, Pune, India.
[8]. P. K. Sen, Estimates of the regression coefficient based on Kendall's tau, Journal of the American statistical association, 63(324), 1968, 1379-1389.
[9]. P. Rousseeuw, B. Daniels, and A. Leroy, Applying robust regression to insurance, Insurance: Mathematics and Economics, 3(1), (1984), 67-72.
[10]. R. P. Brent, An algorithm with guaranteed convergence for finding a zero of a function, The computer journal, 14(4), 1971, 422-425.
[11]. S. S. Hussain and P. Sprent, Non- parametric regression. Journal of the Royal Statistical Society: Series A (General), 146(2), 1983, 182-191.
[12]. S. V. Bhat and S. M. Bijjargi, Classes of slope estimators based on quasi ranges in simple linear regression, International Journal of Statistics and Systems. 19(1), 2024, 1-11.
[13]. T. P. Dence, A brief look into the Lambert W function, Applied mathematics, 4, 2013, 887-892.