



Nonparametric Estimation Method for the Distribution Function Using Various Types of Ranked Set Sampling

Ramy Saad Ghareeb¹  Rikan A. Ahmed² 

¹Postgraduate Student ,Department of Statistics and Informatics, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq, ²Department of Statistics and Informatics, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq.

Article information

Article history:

Received July 1 ,2024

Revised :December 20,2024

Accepted December 25 ,2024

Available June 1, 2025

Keywords:

Local Polynomial Regression,
Ranked Set Sample,Median
Ranked Set Sample,Mean Square
Error,Cumulative Distribution
Function

Correspondence:

Ramy Saad Ghareeb
ramybsako@gmail.com

Abstract

The purpose of this research is to estimate the cumulative distribution function CDF using the local polynomial regression LPR and compare it to parameter estimation using the method of moments and the maximum likelihood method to calculate both the mean square error and the bias using the ranked sets sample RSS and the median ranked sets sample $MRSS$. As well as RSS frequently produces more exact estimates than simple random sampling SRS for the same sample size. By ranking samples based on some easily measurable characteristic, the variability within each set is decreased, resulting in more accurate estimations. We investigated three different degrees of local polynomial regression: the first, second, and third. The simulation analysis demonstrated that the second degree outperforms the other degrees. Also, when LPR is used to analyze RSS data, it takes advantage of the reduced variability within each ranked set, resulting in more precise and reliable regression function estimates. Following that, we investigated several degrees of bandwidth (0.1, 0.2, ... and 0.9) and discovered that the bandwidth of degree 0.8 is superior to the other degrees based on a simulation study. Finally, we analyzed the relative efficiency of each of the three approaches: LPR , MOM , and MLE , and we discovered that LPR is more efficient than the other methods for estimating the CDF in different kernels (normal (gaussian), epanechnikov). The numerical results provide that the suggested estimator CDF based on LPR is more efficient than other methods, as predicted by the simulation analysis.

DOI <https://doi.org/10.33899/ijqjoss.2025> , ©Authors, 2025, College of Computer Science and Mathematics University of Mosul.

This is an open access article under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cumulative Distribution Function CDF is a strong tool for understanding and evaluating random variables, as well as forecasting future occurrences; it is a foundational idea in probability theory and statistics. Also, establish that the occurrence is likely to take place until a specific point. When we utilize it, we encounter several obstacles. One of the challenges is studying nonparametric analysis in-depth and identifying several population features, such as odds, survival analysis, hazards, etc. Estimating the CDF using ranked sets sample RSS is more efficient than basic random sample SRS because RSS frequently yields more precise estimates with the same sample size, as well studying variables and measuring them is not easy; sometimes it is too expensive or time-consuming, but ranking the variables is easy or has a negligible cost. The first to introduce the rank-set sample was McIntyre (1952) for estimating the paster of yields in Australia and expressed expectations about how to develop the estimator that would be more effective for the paster of yields. Halls and Dell (1966) conducted a field trial evaluating its applicability to the estimation of forage yields in a pine-hardwood forest, the terminology ranked set sampling was, coined by

Halls and Dell. Takahasi and Wakimoto (1968) proved the first theoretical result is that, when ranking is perfect, the ranked set sample mean is an unbiased estimator of the population mean, and the variance of the ranked set sample mean is always smaller than the variance of the mean of a simple random sample of the same size. Research has continued in ranked sets since (1997). Muttlak (1997) suggested studying median ranked sets sampling to estimate the population mean instead of ranked sets sampling, and it is a strategy to minimize the error in ranking. Gulati (2004) studied the empirical distribution function of Stokes and Sagar with smooth estimators and properties using simulation to compare the smooth and empirical estimators. Frey (2012) derived the constraint to estimate the cumulative distribution function with the mean of the population to create a Woodruff-type confidence interval for the population quantile. Al-Saleh and Ahmad (2019) suggested a new technique of ranked sets sampling, which was called Moving extreme ranked sets sampling, to estimate the cumulative distribution function and then compared the proposed estimator with the corresponding estimator based on both. Zamanzade (2020) established two estimators in moving extreme ranked sets sampling with simulation and also showed that the proposed estimators provide a substantial improvement over their competitors and prove that the estimators are utilized to estimate the stress-strength probability. Abdallah and Al-Omari (2022) considered the problem of estimating the cumulative distribution function and the odds measure under moving extreme ranked set sampling.

The paper is structured as follows: Section 2 describes ranked sets sampling and median ranked sets sampling. Section 3 describes local polynomial regression. Section 4 Estimation of cumulative distribution function using the Maximum Likelihood Method, Method of Moments, and local polynomial regression. Section 5 Simulation study and conclusions.

2. Description RSS & MRSS

2-1. Methodology of Ranked Sets Sample (RSS)

McIntyre (1952) was the first one who suggested the ranked sets sampling as a strategy to estimate the paster of yields. In the RSS technique, taking samples is much cheaper than measurement of the variable. We will describe how to select ranked sets sampling in the following steps:

1. Draw randomly m^2 sample units from the population of interest. Divided the m^2 units into m groups each one of the groups has a size of m .
2. Based on the judgment rank the unit sets without actual measurement by eyes or by a bit price method for the variable of interest.
3. From the first set, select the smallest order observation, discard the other units, and then from the second set, select the second smallest order observation, and discard the other units. The process continues until get the m^{th} maximum order observation.
4. From steps 1-3 we can get the ranked sets sample RSS of one group
5. To obtain RSS with size k , we can repeat steps (1-3) r times, where $k = mr$.

Let $\{Y_{(i;j)}, i = 1, 2, \dots, m \text{ and } j = 1, 2, \dots, r\}$ be the RSS element set, $Y_{(i;j)}$ be the judgment order statistics of the i^{th} sample of size m , and the j^{th} cycle of the r repeated. You should notice that we utilize square brackets $[\cdot]$, when the ranked sets sample is imperfect ranking it means there is an error in ranking, if there is no error in ranking it means that the ranked sets sample is perfect ranking we use the round brackets (\cdot) , it's very important to note that for each i $\{Y_{(i;1)}, Y_{(i;2)}, \dots, Y_{(i;r)}\}$ are independent and identically distributed (*iid*). And for each j $\{Y_{(1;j)}, Y_{(2;j)}, \dots, Y_{(m;j)}\}$ are just independent.

2-2. Methodology of Median Ranked Sets Sample (MRSS)

Muttlak (1997) suggested a new strategy of ranked sets sampling, which is called median ranked sets sampling MRSS, to minimize errors in the process of ranking units within groups and to increase the efficiency of the estimator in the presence of errors in ranking, also to increase the efficiency over RSS with perfect ranking. The following summarizes the MRSS procedure for drawing a sample of size k . We will describe how to select median ranked sets sampling MRSS in the following steps:

- 1) Draw randomly m^2 sample units from the population of interest.
- 2) Divided the m^2 units into m groups each one of the groups has a size of m .
- 3) If the sample size of m group is odd, the odd will be measured by rule $\frac{m+1}{2}$ it is equal to the units in the medial of the groups, if the sample size of the m group is even, the even will be measured the first half group units with rule $\frac{m}{2}$ and the second half with rule $\frac{m+2}{2}$.
- 4) Steps 1-3 can be replayed r times, if necessary to get MRSS of size $k = mr$.

If the m groups are odd, the median ranked sets sample odd, symbolized $MRSS_o$, where the units of the $MRSS_o$ for variable Y , is described as follows:

$\left\{Y_{i(\frac{m+1}{2})j}, i = 1, 2, \dots, m, j = 1, 2, \dots, r\right\}$ Be the judgment order statistics of the i^{th} sample of size m and the j^{th} cycle of the r repeated.

If the m groups are even, the median ranked sets sample even, symbolized $MRSS_e$, Also $\left\{\frac{m}{2}\right\}$ is the units of the $MRSS_e$ for variable Y , is described as follows:

Where $\left\{Y_{i(\frac{m}{2})j}, Y_{\frac{m}{2}+i(\frac{m+2}{2})j}, i = 1, 2, \dots, \frac{m}{2}, j = 1, 2, \dots, r\right\}$ be the judgment order statistics of the i^{th} sample of size $\left\{\frac{m}{2}\right\}$ and the j^{th} cycle of the r repeated.

3. Local Polynomial Regression (LPR)

Local polynomial regression is a nonparametric technique used to generalize kernel regression, also used to model functions and smoothing one of the statistics plots, which is called the scatter plot. One of the most important uses is to find the relationship between the dependent variable and the independent variable, LPR is better than other types of regression for having a good performance near the boundary. For each point of z_o a WLS which is low order weighted least square regression is fit at each point of z . By the Fan and Gijbels (1996), (Z_i, Y_i) are defined according to the fixed model in equation (1).

$$Y_i = m(Z_i) + \sigma(Z_i)\varepsilon_i \quad i = 1, \dots, n \quad (1)$$

Where $Z_i = \frac{i}{n}$, $\sigma(Z_i)$ is the variance of Y_i at point Z_i , ε_i is a residual error with normal distribution with mean zero and variance σ^2 . For estimating $m(Z_i)$ we use a Tylor series.

$$m(z_i) \approx m(z) + m^{(1)}(z_i)(z_i - z) + \dots + \frac{m^{(P)}(z_i)(z_i - z)^{(P)}}{P!} \quad (2)$$

We need the point in the area of z because it gives us a higher weight than the other point remaining, we can estimate the unknown parameters in equation (2) by using WLS weighted least square, depending on the following formula:

$$Z = \begin{bmatrix} 1 & (z_1 - z) & \dots & (z_1 - z)^P \\ 1 & (z_2 - z) & \dots & (z_2 - z)^P \\ \vdots & \vdots & \ddots & \vdots \\ 1 & (z_n - z) & \dots & (z_n - z)^P \end{bmatrix}_{n \times m}; Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{n \times 1} \quad \hat{\beta} = (Z^T W Z)^{-1} Z^T W Y$$

$$W = \begin{bmatrix} \frac{1}{h} k\left(\frac{(z_1 - z)}{h}\right) & 0 & \dots & 0 \\ 0 & \frac{1}{h} k\left(\frac{(z_2 - z)}{h}\right) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{h} k\left(\frac{(z_n - z)}{h}\right) \end{bmatrix}_{n \times m}$$

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_n \end{bmatrix}_{n \times 1}, e_1 = \{1 \ 0 \ 0 \ \dots \ 0\} \text{ and } w = \text{diag}\left\{\frac{1}{h} k\left(\frac{(z_i - z)}{h}\right)\right\} \quad i = 1, 2, \dots, m; j = 1, 2, \dots, r$$

$k(\cdot)$ represents kernel function, h represents bandwidth. W represents the diagonal elements matrix of weight, where $e_1 = \{1 \ 0 \ 0 \ \dots \ 0\}$ is the $(P + 1)$ times 1 with 1 in the first entry and 0 elsewhere.

4. Estimation of Cumulative Distribution Function (CDF)

4.1. Estimation of (CDF) Using The Maximum Likelihood Method (MLE):

4.1.1. Based on Ranked Sets Sampling (RSS).

Al-Saleh and Ahmad (2019) suggested and proved CDF using MLE based on

$\{Y_{(i;j)}, i = 1, 2, \dots, m \text{ and } j = 1, 2, \dots, r\}$ the represent RSS that we selected from the population with $Pdf f(\cdot)$ and $CDF F(\cdot)$, and then we will use the maximum likelihood estimation MLE for estimating $CDF F(Y_{RSS})$ depending on the RSS. They assumed that the variable $Y_i = \sum_{j=1}^r Y_{(i;j)}$; $i = 1, 2, \dots, m$ is distributed according to a binomial distribution with mass parameter m and success probability $F(Y_{RSS})$. Therefore, the estimator of the probability distribution function is defined according to the following relationship:

$$\hat{F}_{MLE}(Y_{RSS}) = \frac{\sum_{j=1}^r \sum_{i=1}^m Y_{(i,j)}}{mr} = \frac{\sum_{j=1}^r \sum_{i=1}^m Y_{(i,j)}}{k} = \bar{Y}_{RSS} \quad (3)$$

\bar{Y}_{RSS} is the mean obtained by estimating the CDF based on MLE by using RSS.

4.1.2. Based on Odd Median Ranked Sets Sampling (MRSS_o).

Let $\left\{Y_{i(\frac{m+1}{2})j}, i = 1, \dots, m, j = 1, \dots, r\right\}$, be median ranked sets sample odd MRSS_o of size $(k = mr)$, that we selected from the population with Pdf $f(\cdot)$ and CDF $F(\cdot)$, and then we will use the maximum likelihood estimation MLE for estimating CDF $F(Y_{MRSS_o})$ depending on the MRSS_o, note that for each i , $\left\{I\left(Y_{i(\frac{m+1}{2})1} \leq Y_{MRSS_o}\right), I\left(Y_{i(\frac{m+1}{2})2} \leq Y_{MRSS_o}\right), \dots, I\left(Y_{i(\frac{m+1}{2})r} \leq Y_{MRSS_o}\right)\right\}$ are independent and identically distributed (iid) each unit distributed Bernoulli distribution with probability of success $F(Y_{MRSS_o})$, and $I(\cdot)$ represent indicator.

Let $\left\{Y_{iMRSS_o} = \sum_{j=1}^r I\left(Y_{i(\frac{m+1}{2})j} \leq Y_{MRSS_o}\right), i = 1, \dots, m\right\}$, then variable Y_{iMRSS_o} distributed binomial with mass parameter m and success probability $F(Y_{MRSS_o})$.

The likelihood function is determined according to equation (4):

$$g\left(Y_{iMRSS_o} \middle| m, F(Y_{MRSS_o})\right) = \prod_{j=1}^r \binom{m}{Y_{iMRSS_o}} \left(F(Y_{MRSS_o})\right)^{Y_{iMRSS_o}} \left(1 - F(Y_{MRSS_o})\right)^{m-Y_{iMRSS_o}} \quad (4)$$

Therefore, the estimator of the probability distribution function is defined according to the following relationship:

$$\hat{F}_{MLE}(Y_{MRSS_o}) = \frac{\sum_{j=1}^r \sum_{i=1}^m Y_{i(\frac{m+1}{2})j}}{mr} \Rightarrow \frac{\sum_{j=1}^r \sum_{i=1}^m Y_{i(\frac{m+1}{2})j}}{k} = \bar{Y}_{MRSS_o} \quad (5)$$

\bar{Y}_{MRSS_o} is the mean obtained by estimating the CDF based on MLE by using MRSS_o.

4.1.3. Based on Even Median Ranked Sets Sampling (MRSS_e).

The MRSS_e element set is $\left\{Y_{i(\frac{m}{2})j}, Y_{\frac{m}{2}+i(\frac{m+2}{2})j}, i = 1, 2, \dots, \frac{m}{2} \text{ and } j = 1, 2, \dots, r\right\}$ of size $(k = mr)$, that we selected from the population with Pdf $f(\cdot)$ and CDF $F(\cdot)$, and then we will use the maximum likelihood estimation MLE for estimating CDF $F(Y_{MRSS_e})$ depending on the MRSS_e, depending on the withdrawal method of median ranked sets, sample even note that for each i

$\left\{I\left(Y_{i(\frac{m}{2})1}, Y_{\frac{m}{2}+i(\frac{m+2}{2})1} \leq Y_{MRSS_e}\right), \dots, I\left(Y_{i(\frac{m}{2})r}, Y_{\frac{m}{2}+i(\frac{m+2}{2})r} \leq Y_{MRSS_e}\right)\right\}$ are independent and identically distributed (iid) each unit distributed Bernoulli distribution with a probability of success $F(Y_{MRSS_e})$, and $I(\cdot)$ represent indicator.

Let $\left\{Y_{iMRSS_e} = \sum_{j=1}^r I\left(Y_{i(\frac{m}{2})j}, Y_{\frac{m}{2}+i(\frac{m+2}{2})j} \leq Y_{MRSS_e}\right), i = 1, \dots, \frac{m}{2}\right\}$, then variable Y_{iMRSS_e} distributed binomial with mass parameter m and success probability $F(Y_{MRSS_e})$.

The likelihood function is determined according to equation (6):

$$w\left(Y_{iMRSS_e} \middle| m, F(Y_{MRSS_e})\right) = \prod_{j=1}^r \binom{\frac{m}{2}}{Y_{iMRSS_e}} \left(F(Y_{MRSS_e})\right)^{Y_{iMRSS_e}} \left(1 - F(Y_{MRSS_e})\right)^{\frac{m}{2}-Y_{iMRSS_e}} \times \binom{m}{Y_{\frac{m}{2}+iMRSS_e}} \left(F(Y_{MRSS_e})\right)^{Y_{\frac{m}{2}+iMRSS_e}} \left(1 - F(Y_{MRSS_e})\right)^{m-Y_{\frac{m}{2}+iMRSS_e}} \quad (6)$$

Therefore, the estimator of the probability distribution function is defined according to the following relationship:

$$\hat{F}_{MLE}(Y_{MRSS_e}) = \frac{\sum_{j=1}^r \left\{ \sum_{i=1}^{\frac{m}{2}} Y_{i(\frac{m}{2})j} + \sum_{i=\frac{m}{2}+1}^m Y_{i(\frac{m+2}{2})j} \right\}}{k} = \bar{Y}_{MRSS_e} \quad (7)$$

\bar{Y}_{MRSS_e} is the mean obtained by estimating the CDF based on MLE by using MRSS_e.

4.2. Estimation of (CDF) Using the Method of Moments

4.2.1 Based on RSS:

Stokes and Sager (1988) suggested a new estimator, which is called the method of moments, for estimating the distribution function $F(Y_{RSS})$, by using the ranked sets sample, the equation (8) shows us the suggested estimator.

$$\hat{F}_{MOM}(Y_{RSS}) = \frac{\sum_{j=1}^r \sum_{i=1}^m I(Y_{(i,j)} \leq Y_{RSS})}{k} \quad (8)$$

They proved that CDF ranked set sample $\hat{F}_{RSS}(Y_{RSS})$ is an unbiased estimator for $F(Y_{RSS})$ and also has better efficiency than CDF of simple random sample $\hat{F}_{SRS}(Y_{SRS})$.

4.2.2. Based on MRSS₀:

Depending on the description of $MRSS_0$ in the subsection 4.1.2, the distribution

$\{Y_{i(\frac{m+1}{2})j}, i = 1, 2, \dots, m \text{ and } j = 1, 2, \dots, r\}$ for r independent binomial distribution with each mass parameter m and success probability P_{MRSS_0} , can be obtained in equation (9)

$$g(Y_{iMRSS_0} | m, P_{MRSS_0}) = r \binom{m}{Y_{iMRSS_0}} (P_{MRSS_0})^{Y_{iMRSS_0}} (1 - P_{MRSS_0})^{m - Y_{iMRSS_0}} \quad (9)$$

We take the expectation for equation (9) for the purpose of calculating the distribution parameter estimate P_{MRSS_0} , which represents the CDF $F(Y_{MRSS_0})$ and probability of success, based on the method of moments, through equation (10) we will get the mean:

$$E(Y_{iMRSS_0}) = k P_{MRSS_0} = \sum_{j=1}^r \sum_{i=1}^m Y_{i(\frac{m+1}{2})j} \quad (10)$$

$$\hat{F}_{MOM}(Y_{MRSS_0}) = \frac{\sum_{j=1}^r \sum_{i=1}^m Y_{i(\frac{m+1}{2})j}}{k} = \bar{Y}_{MRSS_0} \quad (11)$$

\bar{Y}_{MRSS_0} is the mean obtained by estimating the CDF based on MOM by using $MRSS_0$

4.2.3. Based on MRSS_e:

Depending on the description of $MRSS_e$ in the subsection 4.1.3, the distribution $\{Y_{i(\frac{m}{2})j}, Y_{\frac{m}{2}+i(\frac{m+2}{2})j}, i = 1, 2, \dots, \frac{m}{2}; \text{ and } j = 1, 2, \dots, r\}$ for r independent binomial distribution with mass parameter m and success probability P_{MRSS_e} , which can be obtained in equation (12)

$$w(Y_{iMRSS_e} | k, P_{MRSS_e}) = r \left\{ \binom{\frac{m}{2}}{Y_{iMRSS_e}} (P_{MRSS_e})^{Y_{iMRSS_e}} (1 - P_{MRSS_e})^{\frac{m}{2} - Y_{iMRSS_e}} + \right. \\ \left. \binom{m}{Y_{\frac{m}{2}+iMRSS_e}} (P_{MRSS_e})^{Y_{\frac{m}{2}+iMRSS_e}} (1 - P_{MRSS_e})^{m - Y_{\frac{m}{2}+iMRSS_e}} \right\} \quad (12)$$

To calculate the distribution parameter estimate P_{MRSS_e} , Where $P_{MRSS_e} = F(Y_{MRSS_e})$ which is the probability of success and represents the CDF of the function, based on the method of moments, the expectation of the $P_{MRSS_e} = F(Y_{MRSS_e})$ is defined according to the following:

$$E(Y_{iMRSS_e}) = k P_{MRSS_e} = \sum_{j=1}^r \left\{ \sum_{i=1}^{\frac{m}{2}} Y_{i(\frac{m}{2})j} + \sum_{i=\frac{m}{2}+1}^m Y_{i(\frac{m+2}{2})j} \right\} \quad (13)$$

$$\hat{F}_{MOM}(Y_{MRSS_e}) = \frac{\sum_{j=1}^r \left\{ \sum_{i=1}^{\frac{m}{2}} Y_{i(\frac{m}{2})j} + \sum_{i=\frac{m}{2}+1}^m Y_{i(\frac{m+2}{2})j} \right\}}{k} = \bar{Y}_{MRSS_e} \quad (14)$$

\bar{Y}_{MRSS_e} is the mean obtained by estimating the CDF based on MOM by using $MRSS_e$

4.3. Estimation of (CDF) Using Local Polynomial Regression (LPR)

4.3.1. LPR based on RSS:

Estimation of cumulative distribution function CDF using local polynomial regression LPR based on ranked sets sampling RSS , $F_{LPR}(Y_{RSS})$ can be linearized by using a Tylor series, depending on section (3) equation (2) that we defined LPR with P degree. Consider a fixed regression model:

$$F(Y_{RSS}) = F(Y_{(i,j)}) + \sigma(Y_{(i,j)})\varepsilon_{ij}; i = 1, 2, \dots, m; j = 1, 2, \dots, r \quad (15)$$

$\sigma(Y_{(i,j)})$ is the variance of Y_{RSS} at point $Y_{(i,j)}$, ε_{ij} is a residual error with normal distribution with mean zero and variance σ^2 .

$$F_{LPR}(Y_{RSS}) \approx F(Y) + F^{(1)}(Y)(Y_{RSS} - Y) + \frac{F^{(2)}(Y)}{2!}(Y_{RSS} - Y)^2 + \dots + \frac{F^{(p)}(Y)}{p!}(Y_{RSS} - Y)^p \quad (16)$$

Where \approx represent the approximate equality, $F^{(i)}(Y) = \frac{\partial^{(i)} F_{LPR}(Y_{RSS})}{\partial Y_{RSS}^{(i)}} \Big|_{Y_{RSS}=Y}$ $i = 1, 2, \dots, P$

Where P is the degree of the local polynomial range and Y is an observation from a data neighborhood around Y_{RSS} . If the CDF of the ranked sets sample $F_{LPR}(Y_{RSS})$ is unknown, the equation (16) is as follows:

$$F_{LPR}(Y_{(i,j)}) \approx F(Y_{RSS}) + F^{(1)}(Y_{(i,j)})(Y_{(i,j)} - Y_{RSS}) + \frac{F^{(2)}(Y_{(i,j)})}{2!}(Y_{(i,j)} - Y_{RSS})^2 + \dots + \frac{F^{(p)}(Y_{(i,j)})}{p!}(Y_{(i,j)} - Y_{RSS})^p \quad (17)$$

$i = 1, 2, \dots, m; j = 1, 2, \dots, r$

By estimating the ranked sets sample units $F_{LPR}(Y_{(i,j)})$, the equation (17) will be as follows:

$$\hat{F}_{LPR}(Y_{(i,j)}) \approx \beta_{0RSS} + \beta_{1RSS}(Y_{(i,j)} - Y_{RSS}) + \beta_{2RSS}(Y_{(i,j)} - Y_{RSS})^2 + \dots + \beta_{pRSS}(Y_{(i,j)} - Y_{RSS})^p \quad (18)$$

Where $\hat{\beta}_{iRSS} = \frac{F^{(i)}(Y_{(i,j)})}{i!}$ $i = 0, 1, 2, \dots, P$

Using ordinary least square to estimate the unknown parameter β_{RSS} . Fan and Gijbles (1996) mentioned that the point in the area of Y_{RSS} is giving us a higher weight than the other point. Rather, we can estimate the unknown parameters in equation (18) by using weighted least square depending on the following formula:

$$\hat{\beta}_{RSS} = \begin{bmatrix} \hat{\beta}_{0RSS} \\ \hat{\beta}_{1RSS} \\ \vdots \\ \hat{\beta}_{pRSS} \end{bmatrix}_{p \times 1}; \hat{\beta}_{RSS} = (Y_{RSS}^* W_{RSS} Y_{RSS}^*)^{-1} Y_{RSS}^* W_{RSS} \hat{F}_{RSS} \quad (19)$$

Where

$$Y_{RSS}^* = \begin{bmatrix} 1 & (Y_{(1,1)} - Y_{RSS}) & \dots & (Y_{(1,1)} - Y_{RSS})^p \\ \vdots & \vdots & & \vdots \\ 1 & (Y_{(m,m)} - Y_{RSS}) & \dots & (Y_{(m,m)} - Y_{RSS})^p \\ \vdots & \vdots & & \vdots \\ 1 & (Y_{(m,m)r} - Y_{RSS}) & \dots & (Y_{(m,m)r} - Y_{RSS})^p \end{bmatrix}; \hat{F}_{RSS} = \begin{bmatrix} \hat{F}(Y_{(1,1)}) \\ \vdots \\ \hat{F}(Y_{(m,m)}) \\ \vdots \\ \hat{F}(Y_{(m,m)r}) \end{bmatrix}_{k \times 1}$$

$$\& W_{RSS} = \text{diag} \left\{ \frac{1}{h_{RSS}} k_{RSS} \left(\frac{(Y_{(i,j)} - Y_{RSS})}{h_{RSS}} \right) \right\}, i = 1, \dots, m, j = 1, \dots, r \quad (20)$$

Since $k_{RSS}(\cdot)$ represent kernel function, h_{RSS} represent bandwidth that manages the size of the point in the zone of Y_{RSS} . W_{RSS} represents diagonal elements matrix of weight, since the estimator of $\hat{\beta}_{0RSS}$ is the CDF of local polynomial regression in ranked sets sample, where e_1 is the vector has 1 in the first entry and somewhere else is zero.

$$\hat{F}_{LPR}(Y_{RSS}) = e_1 \times \hat{\beta}_{RSS} = e_1 \times (Y_{RSS}^* W_{RSS} Y_{RSS}^*)^{-1} Y_{RSS}^* W_{RSS} \hat{F}_{RSS} = \hat{\beta}_{0RSS}; e_1 = \{1 \ 0 \ 0 \ \dots \ 0\} \quad (21)$$

We have concluded in equation (21) that the estimator $\hat{\beta}_{0RSS}$ is equal likely to the $\hat{F}_{LPR}(Y_{RSS})$ which is the estimator of the CDF based on LPR with degree P , this indicates that the estimator $\hat{\beta}_{0RSS}$ is equal to the estimator $\hat{F}_{LPR(0)}(Y_{RSS})$ which is the estimator of the CDF based on LPR with degree $P = 0$ in RSS , with the value of the vector $e_1^T = \{1\}$ in equation (22).

$$\hat{F}_{LPR(0)}(Y_{RSS}) = \frac{\sum_{j=1}^r \sum_{i=1}^m k_{RSS}(Y_{(i,j)} - Y_{RSS}) \hat{F}_{RSS}}{\sum_{j=1}^r \sum_{i=1}^m k_{RSS}(Y_{(i,j)} - Y_{RSS})} \quad (22)$$

We have proved in equation (21) that the estimator $\hat{\beta}_{o_{RSS}}$ is equal likely to the $\hat{F}_{LPR}(Y_{RSS})$, and also the estimator $\hat{\beta}_{o_{RSS}}$ is equal to the $\hat{F}_{LPR(1)}(Y_{RSS})$ which is the estimator of the CDF based on LPR with degree $P = 1$ in RSS, with the value of vector $e_1^T = \{1, 0\}$, in equation (23).

$$\hat{F}_{LPR(1)}(Y_{RSS}) = k_{RSS}^{-1} \frac{S_0(Y_{(i,l)j}, h_{RSS}) - S_1(Y_{(i,l)j}, h_{RSS})(Y_{(i,l)j} - Y_{RSS})}{S_0(Y_{(i,l)j}, h_{RSS})S_2(Y_{(i,l)j}, h_{RSS}) - S_1(Y_{(i,l)j}, h_{RSS})^2} \quad (23)$$

Where $S_L(Y_{(i,l)j}, h_{RSS}) = k_{RSS}^{-1} \sum_{j=1}^r \sum_{i=1}^m (Y_{(i,l)j}, Y_{RSS})^L k_{RSS_h}(Y_{(i,l)j}, Y_{RSS})$

The properties of CDF will be studied based on local polynomial using ranked sets sampling depending on these conditions:

- 1- The function $F^{(2)}(Y)$ and σ each continuous on $[0, 1]$.
- 2- The kernel k_{RSS} is symmetric about zero and is reinforced on $[-1, 1]$.
- 3- The bandwidth $h_{RSS} = h_{RSS_k}$ it is a satisfactory sequence, $h_{RSS} \rightarrow 0$ and $kh_{RSS} \rightarrow \infty$ as $k \rightarrow \infty$.
- 4- The point $Y_{(i,l)j}$ at which the estimate is made satisfactory $h_{RSS} < Y_{(i,l)j} < 1 - h_{RSS}$ for all $k \geq k_0$, where k_0 is fixed.

In this subsection, we will make a derivative for the bias and MSE of the estimate CDF based on local polynomial.

$$E(\hat{F}_{LPR(1)}(Y_{RSS})) = E\{e_1^T (Y_{RSS}^*{}^T W_{RSS} Y_{RSS}^*)^{-1} Y_{RSS}^*{}^T W_{RSS} \hat{F}_{RSS}\}$$

$$E(\hat{F}_{LPR(1)}(Y_{RSS})) = F_{RSS} + \frac{1}{2} h_{RSS}^{(2)} F_{RSS}^{(2)} \mu_2(k_{RSS}) + O(h_{RSS}^{(2)} k_{RSS}^{-1}) \quad (24)$$

$$Bias(\hat{F}_{LPR(1)}(Y_{RSS})) = \frac{1}{2} h_{RSS}^{(2)} F_{RSS}^{(2)} \mu_2(k_{RSS}) + O(h_{RSS}^{(2)} k_{RSS}^{-1}) \quad (25)$$

$$MSE(\hat{F}_{LPR(1)}(Y_{RSS})) = e_1^T (Y_{RSS}^*{}^T W_{RSS} Y_{RSS}^*)^{-1} Y_{RSS}^*{}^T W_{RSS} V(F(Y_{RSS})) W_{RSS} Y_{RSS}^* (Y_{RSS}^*{}^T W_{RSS} Y_{RSS}^*)^{-1} e_1 \quad (26)$$

$$MSE(\hat{F}_{LPR(1)}(Y_{RSS})) = (kh_{RSS})^{-1} R(k_{RSS}) V(F(Y_{RSS})) + O\{(kh_{RSS})^{-1}\} \quad (27)$$

In the applied aspect, in Table 1.1, we use the R-Program to obtain values of the empirical MSE and Bias of CDF based on LPR by using RSS, with four level sets size ($m = 2, 3, 4, 5$) and number of cycles ($r = 3, 5, 7$). We take degree $P = 2$ of LPR with bandwidth 0.8 because they give a good result compared to other bandwidths and degrees and use kernel (normal, epanechnikov) with three values of probability ($P_{RSS} = 0.25, 0.50, 0.75$).

Table 1.1: The empirical mean square error and bias of CDF based on LPR by Using RSS

Kernel / Normal					
	r = 3	m = 2, k = 6	m = 3, k = 9	m = 4, k = 12	m = 5, k = 15
$P_{RSS} = 0.25$		Mse = 0.000107486 Bias = 0.0004636507	Mse = 0.0001035949 Bias = 0.0004551812	Mse = 0.0001014252 Bias = 0.0004503893	Mse = 7.986923e-05 Bias = 0.0003996729
$P_{RSS} = 0.50$		946986e-05 Bias = 0.00034	Mse = 5.383707e-05 Bias = 0.0003281374	4.66936e-05 Bias = 0.000304	4.70322e-05 Bias = 0.000304
$P_{RSS} = 0.75$		Mse = 2.168546e-05 Bias = 0.0002082569	Mse = 2.295053e-05 Bias = 0.0002142453	Mse = 4.179461e-05 Bias = 0.000289118	Mse = 2.598391e-05 Bias = 0.0002279645
	r = 5	m = 2, k = 10	m = 3, k = 15	m = 4, k = 20	m = 5, k = 25
$P_{RSS} = 0.25$		Mse = 3.894282e-05 Bias = 0.00027908	Mse = 2.321515e-05 Bias = 0.0002154769	Mse = 3.138624e-05 Bias = 0.0002505444	Mse = 2.607705e-05 Bias = 0.0002283727
$P_{RSS} = 0.50$		158395e-05 Bias = 0.00024	Mse = 2.049009e-05 Bias = 0.0002024356	1.98626e-05 Bias = 0.000199	1.083776e-05 Bias = 0.000204
$P_{RSS} = 0.75$		Mse = 1.459885e-05 Bias = 0.0001708733	Mse = 1.734523e-05 Bias = 0.0001862537	Mse = 1.692903e-05 Bias = 0.0001840056	Mse = 1.845636e-05 Bias = 0.0001921268
	r = 7	m = 2, k = 14	m = 3, k = 21	m = 4, k = 28	m = 5, k = 35
$P_{RSS} = 0.25$		Mse = 1.264575e-05 Bias = 0.000159033	Mse = 1.352716e-05 Bias = 0.000164482	Mse = 1.15039e-05 Bias = 0.0001516832	Mse = 1.099064e-05 Bias = 0.0001482609
$P_{RSS} = 0.50$		211694e-05 Bias = 0.00015	Mse = 1.238188e-05 Bias = 0.0001573651	3.98311e-05 Bias = 0.00016	Mse = 9.6613e-06 bias = 0.0001390058
$P_{RSS} = 0.75$		Mse = 7.667563e-06 Bias = 0.0001238351	Mse = 9.508394e-06 Bias = 0.0001379014	Mse = 1.167593e-05 Bias = 0.0001528131	Mse = 6.628009e-06 Bias = 0.0001151348
Kernel / Epanechnikov					
	r = 3	m = 2, k = 6	m = 3, k = 9	m = 4, k = 12	m = 5, k = 15
$P_{RSS} = 0.25$		Mse = 0.0001206227 Bias = 0.0004911674	Mse = 8.56113e-05 Bias = 0.0004137905	Mse = 6.935339e-05 Bias = 0.0003724336	Mse = 8.86133e-05 Bias = 0.0004209829
$P_{RSS} = 0.50$		661436e-05 Bias = 0.00033	Mse = 5.640625e-05 Bias = 0.0003358757	6.636432e-05 Bias = 0.00033	7.04077e-05 Bias = 0.00033

	$P_{RSS} = 0.75$	Mse = 1.99026e-05 Bias = 0.0001995124	Mse = 2.287863e-05 Bias = 0.0002139095	Mse = 3.196004e-05 Bias = 0.0002528242	Mse = 4.15294e-05 Bias = 0.0002881992
	$r = 5$	$m = 2, k = 10$	$m = 3, k = 15$	$m = 4, k = 20$	$m = 5, k = 25$
	$P_{RSS} = 0.25$	Mse = 3.369704e-05 Bias = 0.0002596037	Mse = 2.682286e-05 Bias = 0.0002316154	Mse = 1.952314e-05 Bias = 0.0001976013	Mse = 2.263933e-05 Bias = 0.0002127878
	$P_{RSS} = 0.50$	Mse = 2.034624e-05 Bias = 0.0002017238	Mse = 1.930976e-05 Bias = 0.0001965185	Mse = 1.933795e-05 Bias = 0.0001966619	Mse = 2.062103e-05 Bias = 0.0002030814
	$P_{RSS} = 0.75$	Mse = 7.79924e-06 Bias = 0.0001248939	Mse = 1.103365e-05 Bias = 0.0001485506	Mse = 1.969276e-05 Bias = 0.0001984578	Mse = 1.909245e-05 Bias = 0.0001954096
	$r = 7$	$m = 2, k = 14$	$m = 3, k = 21$	$m = 4, k = 28$	$m = 5, k = 35$
	$P_{RSS} = 0.25$	Mse = 1.560244e-05 Bias = 0.000176649	Mse = 1.161882e-05 Bias = 0.000152439	Mse = 1.186293e-05 Bias = 0.000154032	Mse = 1.385473e-05 Bias = 0.0001664616
	$P_{RSS} = 0.50$	Mse = 1.03374e-05 bias = 0.00014	Mse = 1.377892e-05 Bias = 0.0001660056	Mse = 1.074807e-05 Bias = 0.0001466156	Mse = 1.121305e-05 Bias = 0.0001497535
	$P_{RSS} = 0.75$	Mse = 8.002706e-06 Bias = 0.0001265125	Mse = 6.681595e-06 Bias = 0.0001155993	Mse = 6.970129e-06 Bias = 0.0001180689	Mse = 1.073552e-05 Bias = 0.00014653

From the results of table 1.1, *figure 1*, and *figure 2*, we can observe the following:

1- For fixed values P_{RSS} and m , when the number of cycles is increasing $r = 3, 5$ and 7 , the values of Mse and Bias are decreasing. For more evidence, we can see *Figure 1* when the blue line gradually curves as the number of cycles increases.

2- For fixed values m and sample size k , when the probability is increasing $P_{RSS} = 0.25, 0.50$, and 0.75 , the blue line gradually curves; see *Figure 2*, which means that values of Mse and Bias are decreasing.

3- Kernel epanechnikov gives us a better result than the gaussian kernel; for evidence, the result of empirical Mse and Bias in epanchnikov are less than gaussian.

4.3.2. LPR based on MRSS₀:

Estimation of cumulative distribution function *CDF* using local polynomial regression *LPR* based on median ranked sets sampling odd $MRSS_0$, we take P degree of local polynomial regression, and using a Tylor series for linearizing $F_{LPR}(Y_{MRSS_0})$ as follows:

Consider a fixed regression model:

$$F(Y_{MRSS_0}) = F\left(Y_{i(\frac{m+1}{2})j}\right) + \sigma\left(Y_{i(\frac{m+1}{2})j}\right)\varepsilon_{ij}; i = 1, 2, \dots, m; j = 1, 2, \dots, r \quad (28)$$

$\sigma\left(Y_{i(\frac{m+1}{2})j}\right)$ is the variance of Y_{MRSS_0} at point $Y_{i(\frac{m+1}{2})j}$, ε_{ij} is a residual error with normal distribution with mean zero and variance σ^2 .

$$F_{LPR}(Y_{MRSS_0}) \approx F(Y) + F^{(1)}(Y)(Y_{MRSS_0} - Y) + \frac{F^{(2)}(Y)}{2!}(Y_{MRSS_0} - Y)^2 + \dots + \frac{F^{(P)}(Y)}{P!}(Y_{MRSS_0} - Y)^P \quad (29)$$

Where $F^{(i)}(Y) = \frac{\partial^i F(Y_{MRSS_0})}{\partial Y_{MRSS_0}^i} \Big|_{Y_{MRSS_0}=Y}$ $i = 1, 2, \dots, P$; Y is an observation from a data neighborhood around Y_{MRSS_0} , equation

(29) is reformulated as follows:

$$F_{LPR}\left(Y_{i(\frac{m+1}{2})j}\right) \approx F(Y_{MRSS_0}) + F^{(1)}\left(Y_{i(\frac{m+1}{2})j}\right)\left(Y_{i(\frac{m+1}{2})j} - Y_{MRSS_0}\right) \quad (30)$$

$$+ \frac{F^{(2)}\left(Y_{i(\frac{m+1}{2})j}\right)}{2!}\left(Y_{i(\frac{m+1}{2})j} - Y_{MRSS_0}\right)^2 + \dots + \frac{F^{(P)}\left(Y_{i(\frac{m+1}{2})j}\right)}{P!}\left(Y_{i(\frac{m+1}{2})j} - Y_{MRSS_0}\right)^P$$

$$i = 1, 2, \dots, m; j = 1, 2, \dots, r$$

By estimating the ranked sets, sample units $F_{LPR}\left(Y_{i\left(\frac{m+1}{2}\right)j}\right)$ in equation (30) will be as follows:

$$\begin{aligned} \hat{F}_{LPR}\left(Y_{i\left(\frac{m+1}{2}\right)j}\right) &\approx \beta_{0MRSSO} + \beta_{1MRSSO}\left(Y_{i\left(\frac{m+1}{2}\right)j} - Y_{MRSSO}\right) + \beta_{2MRSSO}\left(Y_{i\left(\frac{m+1}{2}\right)j} - Y_{MRSSO}\right)^2 \\ &+ \dots + \beta_{PMRSSO}\left(Y_{i\left(\frac{m+1}{2}\right)j} - Y_{MRSSO}\right)^P \end{aligned} \quad (31)$$

Where $\beta_{iMRSSO} = \frac{F^{(i)}\left(Y_{i\left(\frac{m+1}{2}\right)j}\right)}{i!}$ $i = 0, 1, 2, \dots, P$

$$\hat{\beta}_{MRSSO} = \begin{bmatrix} \hat{\beta}_{0MRSSO} \\ \hat{\beta}_{1MRSSO} \\ \vdots \\ \hat{\beta}_{PMRSSO} \end{bmatrix}_{p \times 1}; \hat{\beta}_{MRSSO} = (Y_{MRSSO}^{*T} W_{MRSSO} Y_{MRSSO}^*)^{-1} Y_{MRSSO}^{*T} W_{MRSSO} \hat{F}_{MRSSO} \quad (32)$$

$$Y_{MRSSO}^* = \begin{bmatrix} 1 & \left(Y_{1\left(\frac{m+1}{2}\right)1} - Y_{MRSSO}\right) & \dots & \left(Y_{1\left(\frac{m+1}{2}\right)1} - Y_{MRSSO}\right)^P \\ \vdots & \vdots & & \vdots \\ 1 & \left(Y_{m\left(\frac{m+1}{2}\right)1} - Y_{MRSSO}\right) & \dots & \left(Y_{m\left(\frac{m+1}{2}\right)1} - Y_{MRSSO}\right)^P \\ \vdots & \vdots & & \vdots \\ 1 & \left(Y_{m\left(\frac{m+1}{2}\right)r} - Y_{MRSSO}\right) & \dots & \left(Y_{m\left(\frac{m+1}{2}\right)r} - Y_{MRSSO}\right)^P \end{bmatrix}_{p \times p} \quad \hat{F}_{MRSSO} = \begin{bmatrix} \hat{F}\left(y_{1\left(\frac{m+1}{2}\right)1}\right) \\ \vdots \\ \hat{F}\left(y_{m\left(\frac{m+1}{2}\right)1}\right) \\ \vdots \\ \hat{F}\left(y_{m\left(\frac{m+1}{2}\right)r}\right) \end{bmatrix}_{p \times 1}$$

$$W_{MRSSO} = \text{diag} \left\{ \frac{1}{h_{MRSSO}} k_{MRSSO} \left(\frac{\left(Y_{i\left(\frac{m+1}{2}\right)j} - Y_{MRSSO}\right)}{h_{MRSSO}} \right) \right\}, i = 1, \dots, m, j = 1, \dots, r \quad (33)$$

$k_{MRSSO}(\cdot)$ represents the kernel function of $MRSSO$, and h_{MRSSO} represents the bandwidth that manages the size of the point in the zone of Y_{MRSSO} . W_{MRSSO} represents the diagonal element matrix of weight since the estimator of $\hat{\beta}_{0MRSSO}$ is the *CDF* of local polynomial regression in median ranked sets sample odd, where e_1 is the vector with 1 in the first entry and elsewhere is zero.

$$\hat{F}_{LPR}(Y_{MRSSO}) = e_1 \times \hat{\beta}_{MRSSO} = e_1 \times (Y_{MRSSO}^{*T} W_{MRSSO} Y_{MRSSO}^*)^{-1} Y_{MRSSO}^{*T} W_{MRSSO} \hat{F}_{MRSSO} = \hat{\beta}_{0MRSSO} \quad (34)$$

Estimation of the *CDF* based on *LPR* with degree $P = 0$, with the value of vector $e_1^T = \{1\}$, indicates that the estimator $\hat{\beta}_{0MRSSO}$ is equally likely to the $\hat{F}_{LPR}(Y_{MRSSO})$, which is equal to the estimator $\hat{F}_{LPR(0)}(Y_{MRSSO})$ in equation (35).

$$\hat{F}_{LPR(0)}(Y_{MRSSO}) = \frac{\sum_{j=1}^r \sum_{i=1}^m k_{MRSSO} \left(Y_{i\left(\frac{m+1}{2}\right)j} - Y_{MRSSO} \right) \hat{F}_{MRSSO}}{\sum_{j=1}^r \sum_{i=1}^m k_{MRSSO} \left(Y_{i\left(\frac{m+1}{2}\right)j} - Y_{MRSSO} \right)} \quad (35)$$

We have proved in equation (34) that the estimator $\hat{\beta}_{0MRSSO}$ is equally likely to the estimator $\hat{F}_{LPR}(Y_{MRSSO})$, which is equal to the estimator $\hat{F}_{LPR(1)}(Y_{MRSSO})$, which is the estimator of the *CDF* based on *LPR* with degree $P = 1$, with the value of the vector being $e_1^T = \{1, 0\}$, in equation (36).

$$\hat{F}_{LPR(1)}(Y_{MRSSO}) = k_{MRSSO}^{-1} \frac{S_0\left(Y_{i(\frac{m+1}{2})j}, h_{MRSSO}\right) - S_1\left(Y_{i(\frac{m+1}{2})j}, h_{MRSSO}\right)\left(Y_{i(\frac{m+1}{2})j} - Y_{MRSSO}\right)}{S_0\left(Y_{i(\frac{m+1}{2})j}, h_{MRSSO}\right)S_2\left(Y_{i(\frac{m+1}{2})j}, h_{MRSSO}\right) - S_1\left(Y_{i(\frac{m+1}{2})j}, h_{MRSSO}\right)^2} \quad (36)$$

$$\text{Where } S_L\left(Y_{i(\frac{m+1}{2})j}, h\right) = k_{MRSSO}^{-1} \sum_{j=1}^r \sum_{i=1}^m \left(Y_{i(\frac{m+1}{2})j}, Y_{MRSSO}\right)^L k_{MRSSO_h}\left(Y_{i(\frac{m+1}{2})j}, Y_{MRSSO}\right)$$

The properties of *CDF* will be studied based on local polynomial using median ranked sets sampling odd, depending on these conditions:

- 1- The function $F^{(2)}(Y)$ and σ each continuous on $[0,1]$.
- 2- The kernel k_{MRSSO} is symmetric about zero and is reinforced on $[-1,1]$.
- 3- The bandwidth $h_{MRSSO} = h_{MRSSO_k}$ it is a satisfactory sequence, $h_{MRSSO} \rightarrow 0$ and $kh_{MRSSO} \rightarrow \infty$ as $k \rightarrow \infty$.
- 4- The point $Y_{i(\frac{m+1}{2})j}$ at which the estimate is made satisfactory $h_{MRSSO} < Y_{i(\frac{m+1}{2})j} < 1 - h_{MRSSO}$ for all $k \geq k_0$, where k_0 is fixed.

In this subsection, we will make a derivative for the Bias and *MSE* of the estimate *CDF* based on local polynomial.

$$E\left(\hat{F}_{LPR(1)}(Y_{MRSSO})\right) = E\left\{e_1^T(Y_{MRSSO}^*{}^T W_{MRSSO} Y_{MRSSO}^*)^{-1} Y_{MRSSO}^*{}^T W_{MRSSO} \hat{F}_{MRSSO}\right\}$$

$$E\left(\hat{F}_{LPR(1)}(Y_{MRSSO})\right) = F_{MRSSO} + \frac{1}{2} h_{MRSSO}^{(2)} F_{MRSSO}^{(2)} \mu_2(k_{MRSSO}) + O(h_{MRSSO}^{(2)} k_{MRSSO}^{-1}) \quad (37)$$

$$\text{Bias}\left(\hat{F}_{LPR(1)}(Y_{MRSSO})\right) = \frac{1}{2} h_{MRSSO}^{(2)} F_{MRSSO}^{(2)} \mu_2(k_{MRSSO}) + O(h_{MRSSO}^{(2)} k_{MRSSO}^{-1}) \quad (38)$$

$$\text{MSE}\left(\hat{F}_{LPR(1)}(Y_{MRSSO})\right) = e_1^T(Y_{MRSSO}^*{}^T W_{MRSSO} Y_{MRSSO}^*)^{-1} Y_{MRSSO}^*{}^T W_{MRSSO} V(F(Y_{MRSSO})) \quad (39)$$

$$\text{MSE}\left(\hat{F}_{LPR(1)}(Y_{MRSSO})\right) = (k_{MRSSO} h_{MRSSO})^{-1} R(k_{MRSSO}) V(F(Y_{MRSSO})) + O\left\{(k_{MRSSO} h_{MRSSO})^{-1}\right\} \quad (40)$$

In the Table 1.2, we use the R-Program to achieve the result of the empirical mean square error and bias of *CDF* based on *LPR* by using *MRSS odd* with four levels of set size ($m = 3, 5, 7, 9$), with three values of probability ($P_{MRSSO} = 0.25, 0.50, 0.75$), and with the same number of cycles r , degree of kernel, type of kernel, and bandwidth as in the previous table 1.1.

Table 1.2: The empirical mean square error and bias of CDF based on LPR by Using MRSS ODD

Kernel / Normal					
	r = 3	m = 3, k = 9	m = 5, k = 15	m = 7, k = 21	m = 9, k = 27
	$P_{MRSSO} = 0.25$	Mse = 8.512535e-05 Bias = 0.0004126145	Mse = 7.691912e-05 Bias = 0.0003922222	Mse = 8.174425e-05 Bias = 0.0004043371	Mse = 7.924077e-05 Bias = 0.0003980974
	$P_{MRSSO} = 0.50$	865291e-05 Bias = 0.0003	Mse = 5.529911e-05 Bias = 0.0003325631	3.341845e-05 Bias = 0.00035	3.342218e-05 Bias = 0.00035
	$P_{MRSSO} = 0.75$	Mse = 2.326374e-05 Bias = 0.0002157023	Mse = 3.806345e-05 Bias = 0.000275911	Mse = 4.086824e-05 Bias = 0.0002858959	Mse = 3.77983e-05 Bias = 0.0002749484
	r = 5	m = 3, k = 15	m = 5, k = 25	m = 7, k = 35	m = 9, k = 45
	$P_{MRSSO} = 0.25$	Mse = 2.637953e-05 Bias = 0.0002296934	Mse = 2.60296e-05 Bias = 0.0002281649	Mse = 2.278614e-05 Bias = 0.0002134767	Mse = 2.114632e-05 Bias = 0.0002056517
	$P_{MRSSO} = 0.50$	524377e-05 Bias = 0.0001	Mse = 1.809422e-05 Bias = 0.0001902326	1.96922e-05 Bias = 0.00019	1.632715e-05 Bias = 0.00018
	$P_{MRSSO} = 0.75$	Mse = 1.107483e-05 Bias = 0.0001488276	Mse = 1.713709e-05 Bias = 0.0001851329	Mse = 1.3506e-05 Bias = 0.0001643533	Mse = 1.482615e-05 Bias = 0.0001721984
	r = 7	m = 3, k = 21	m = 5, k = 35	m = 7, k = 49	m = 9, k = 63
	$P_{MRSSO} = 0.25$	Mse = 1.56476e-05 Bias = 0.0001769045	Mse = 9.443144e-06 Bias = 0.0001374274	Mse = 1.042319e-05 Bias = 0.0001443828	Mse = 1.071913e-05 Bias = 0.0001464181

	$P_{MRSS_0} = 0.50$	674636e-06 Bias = 0.0001	Mse = 9.210139e-06 Bias = 0.0001357213	.552546e-06 Bias = 0.00013	7.716423e-06 Bias = 0.0001
	$P_{MRSS_0} = 0.75$	Mse = 9.022209e-06 Bias = 0.0001343295	Mse = 6.555129e-06 Bias = 0.0001145	Mse = 9.568994e-06 Bias = 0.0001383401	Mse = 8.611258e-06 Bias = 0.0001312346
Kernel / Epanechnikov					
	r = 3	m = 3, k = 9	m = 5, k = 15	m = 7, k = 21	m = 9, k = 27
	$P_{MRSS_0} = 0.25$	Mse = 0.0001048237 Bias = 0.0004578727	Mse = 7.279169e-05 Bias = 0.0003815539	Mse = 6.217946e-05 Bias = 0.0003526456	Mse = 6.378918e-05 Bias = 0.0003571811
	$P_{MRSS_0} = 0.50$	Mse = 5.373519e-05 Bias = 0.0003278267	Mse = 5.708709e-05 Bias = 0.0003378967	Mse = 7.544543e-05 Bias = 0.0003884467	Mse = 5.500072e-05 Bias = 0.0003316647
	$P_{MRSS_0} = 0.75$	Mse = 3.294451e-05 Bias = 0.0002566886	Mse = 3.944943e-05 Bias = 0.0002808894	Mse = 3.64872e-05 Bias = 0.0002701377	Mse = 3.562654e-05 Bias = 0.0002669327
	r = 5	m = 3, k = 15	m = 5, k = 25	m = 7, k = 35	m = 9, k = 45
	$P_{MRSS_0} = 0.25$	Mse = 2.706822e-05 Bias = 0.0002326724	Mse = 2.142626e-05 Bias = 0.0002070085	Mse = 2.363208e-05 Bias = 0.0002174032	Mse = 1.916575e-05 Bias = 0.0001957843
	$P_{MRSS_0} = 0.50$	Mse = 2.02618e-05 Bias = 0.0002013047	Mse = 1.732828e-05 Bias = 0.0001861627	Mse = 2.406087e-05 Bias = 0.0002193667	Mse = 1.673458e-05 Bias = 0.0001829458
	$P_{MRSS_0} = 0.75$	Mse = 1.15668e-05 Bias = 0.0001520973	Mse = 1.023598e-05 Bias = 0.0001430803	Mse = 1.64224e-05 Bias = 0.0001812313	Mse = 1.666829e-05 Bias = 0.0001825831
	r = 7	m = 3, k = 21	m = 5, k = 35	m = 7, k = 49	m = 9, k = 63
	$P_{MRSS_0} = 0.25$	Mse = 1.297796e-05 Bias = 0.0001611084	Mse = 1.355451e-05 Bias = 0.0001646481	Mse = 1.365714e-05 Bias = 0.0001652703	Mse = 1.144232e-05 Bias = 0.0001512767
	$P_{MRSS_0} = 0.50$.17379e-05 bias = 0.00015	Mse = 1.000622e-05 Bias = 0.0001414653	Mse = 1.265157e-05 Bias = 0.0001590696	Mse = 1.054311e-05 Bias = 0.000145211
	$P_{MRSS_0} = 0.75$	Mse = 9.345159e-06 Bias = 0.0001367125	Mse = 8.477178e-06 Bias = 0.0001302089	Mse = 8.385006e-06 Bias = 0.0001294991	Mse = 1.065344e-05 Bias = 0.0001459688

The following steps describe table 1.2, Figure 3, and Figure 4:

- 1- When the number of cycles is increasing $r = 3, 5$ and 7 , for fixed values P_{MRSS_0} and m , the values of Mse and Bias are decreasing, and also Figure 3 proves that the blue line is curving gradually when we increase the cycles r .
- 2- In Figure 4, we notice that the blue line continues to curve as the probability increases ($P_{MRSS_0} = 0.25, 0.50, 0.75$), and this indicates a decrease in the values of Mse and bias for fixed values m and sample size k .
- 3- In table 1.2, we notice that the kernel epanechnikov is better than the kernel gaussian. Evidence of this is that the experimental values in Mse and Bias are lower than what they are in the kernel gaussian.

4.3.3. LPR based on $MRSS_e$:

Estimation of cumulative distribution function CDF using local polynomial regression LPR based on median ranked sets sampling even $MRSS_e$, we can linearize the CDF , $F(Y_{MRSS_e})$, by using a Tylor series and with P degree of local polynomial regression.

The regression fixed model based on median ranked sets sample even:

$$F(Y_{MRSS_e}) = F\left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j}\right) + \sigma\left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j}\right)\varepsilon_{ij} \quad (39)$$

$$i = 1, 2, \dots, \frac{m}{2}; j = 1, 2, \dots, r$$

$$F_{LPR}(Y_{MRSS_e}) \approx F(Y) + F^{(1)}(Y)(Y_{MRSS_e} - Y) + \frac{F^{(2)}(Y)}{2!}(Y_{MRSS_e} - Y)^2 + \dots + \frac{F^{(P)}(Y)}{P!}(Y_{MRSS_e} - Y)^P \quad (40)$$

where $F^{(i)}(Y) = \frac{\partial^{(i)} F(Y_{MRSS_e})}{\partial Y_{MRSS_e}^{(i)}} \Big|_{Y_{MRSS_e}=Y}$, $i = 1, 2, \dots, P$; Y is an observation from the data neighborhood around Y_{MRSS_e} ; if the

CDF of ranked sets sample $F_{LPR}(Y_{MRSS_e})$ is unknown, the equation (40) will be as follows:

$$F_{LPR}\left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j}\right) \approx F(Y_e) + F^{(1)}\left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j}\right)\left(\left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j}\right) - Y_{MRSS_e}\right) + \quad (41)$$

$$+ \frac{F^{(2)}\left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j}\right)}{2!} \left(\left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j} \right) - Y_{MRSS_e} \right)^2 + \dots$$

$$+ \frac{F^{(P)}\left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j}\right)}{P!} \left(\left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j} \right) - Y_{MRSS_e} \right)^P$$

$i = 1, 2, \dots, \frac{m}{2} ; j = 1, 2, \dots, r, \text{ where } \approx \left\{ \frac{m}{2} + \frac{m}{2} = m \right\}$

By estimating the ranked sets sample units $F_{LPR}\left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j}\right)$, the equation (41) will be:

$$\hat{F}_{LPR}\left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j}\right) \approx \beta_{0MRSS_e} + \beta_{1MRSS_e} \left(\left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j} \right) - Y_{MRSS_e} \right) + \quad (42)$$

$$\beta_{2MRSS_e} \left(\left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j} \right) - Y_{MRSS_e} \right)^2 + \dots + \beta_{PMRSS_e} \left(\left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j} \right) - Y_{MRSS_e} \right)^P$$

Where $\beta_{iMRSS_e} = \frac{F^{(i)}\left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j}\right)}{i!} \quad i = 0, 1, 2, \dots, P$

$$\hat{\beta}_{MRSS_e} = \begin{bmatrix} \hat{\beta}_{0MRSS_e} \\ \hat{\beta}_{1MRSS_e} \\ \vdots \\ \hat{\beta}_{PMRSS_e} \end{bmatrix}_{p \times 1} ; \hat{\beta}_{MRSS_e} = (Y_{MRSS_e}^{*T} W_{MRSS_e} Y_{MRSS_e}^*)^{-1} Y_{MRSS_e}^{*T} W_{MRSS_e} \hat{F}_{MRSS_e} \quad (43)$$

$$Y_{MRSS_e}^* = \begin{bmatrix} 1 & \left(\left(Y_{1(\frac{m}{2})1} + Y_{\frac{m}{2}+1(\frac{m+2}{2})1} \right) - Y_{MRSS_e} \right) & \dots & \left(\left(Y_{1(\frac{m}{2})1} + Y_{\frac{m}{2}+1(\frac{m+2}{2})1} \right) - Y_{MRSS_e} \right)^P \\ \vdots & \vdots & & \vdots \\ 1 & \left(\left(Y_{\frac{m}{2}(\frac{m}{2})1} + Y_{m(\frac{m+2}{2})1} \right) - Y_{MRSS_e} \right) & \dots & \left(\left(Y_{\frac{m}{2}(\frac{m}{2})1} + Y_{m(\frac{m+2}{2})1} \right) - Y_{MRSS_e} \right)^P \\ \vdots & \vdots & & \vdots \\ 1 & \left(\left(Y_{\frac{m}{2}(\frac{m}{2})r} + Y_{m(\frac{m+2}{2})r} \right) - Y_{MRSS_e} \right) & \dots & \left(\left(Y_{\frac{m}{2}(\frac{m}{2})r} + Y_{m(\frac{m+2}{2})r} \right) - Y_{MRSS_e} \right)^P \end{bmatrix}_{P \times P}$$

$$\hat{F}_{MRSS_e} = \begin{bmatrix} \hat{F}\left(Y_{1(\frac{m}{2})1} + Y_{\frac{m}{2}+1(\frac{m+2}{2})1}\right) \\ \vdots \\ \hat{F}\left(Y_{\frac{m}{2}(\frac{m}{2})1} + Y_{m(\frac{m+2}{2})1}\right) \\ \vdots \\ \hat{F}\left(Y_{\frac{m}{2}(\frac{m}{2})r} + Y_{m(\frac{m+2}{2})r}\right) \end{bmatrix}_{k \times 1}$$

$$W_{MRSS_e} = \text{diag} \left\{ \frac{1}{h_{MRSS_e}} k_{MRSS_e} \left(\frac{\left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j} \right) - Y_{MRSS_e}}{h_{MRSS_e}} \right) \right\}, i = 1, \dots, \frac{m}{2}, j = 1, \dots, r \quad (44)$$

$k_{MRSS_e}(\cdot)$ represents the kernel function of $MRSS_e$, and h_{MRSS_e} represents the bandwidth that manages the size of the point in the zone of Y_{MRSS_e} . W_{MRSS_e} represents the diagonal element matrix of weight since the estimator of $\hat{\beta}_{0MRSS_e}$ is the CDF of local polynomial regression in median ranked sets sample even, where e_1 is the vector with 1 in the first entry and elsewhere is zero.

$$\hat{F}(Y_{MRSS_e}) = e_1 \times \hat{\beta}_{MRSS_e} = e_1 \times (Y_{MRSS_e}^T W_{MRSS_e} Y_{MRSS_e})^{-1} Y_{MRSS_e}^T W_{MRSS_e} \hat{F}_{MRSS_e} = \hat{\beta}_{0MRSS_e} \quad (45)$$

In equation (45), we get the estimator $\hat{\beta}_{0MRSS_e}$, which is equally likely to the $\hat{F}_{LPR}(Y_{MRSS_e})$, which is equal to the $\hat{F}_{LPR(0)}(Y_{MRSS_e})$ the estimator of the CDF based on LPR with degree $P = 0$, with the value of vector $e_1^T = \{1\}$, in equation (46).

$$\hat{F}_{LPR(0)}(Y_{MRSS_e}) = \frac{\sum_{j=1}^r \sum_{i=1}^{\frac{m}{2}} k_{MRSS_e} \left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j} - Y_{MRSS_e} \right) \hat{F}_{MRSS_e}}{\sum_{j=1}^r \sum_{i=1}^{\frac{m}{2}} k_{MRSS_e} \left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j} - Y_{MRSS_e} \right)} \quad (46)$$

In the equation (47) we prove $\hat{F}_{LPR(1)}(Y_{MRSS_e})$, which is the estimator of the CDF based on LPR with degree $P = 1$, with the value of the vector being $e_1^T = \{1, 0\}$, this indicates that the estimation of $\hat{\beta}_{0MRSS_e}$ is equally likely to the $\hat{F}_{LPR}(Y_{MRSS_e})$, which is equal to the $\hat{F}_{LPR(1)}(Y_{MRSS_e})$ in equation (47).

$$\hat{F}_{LPR(1)}(Y_{MRSS_e}) = \frac{k_{MRSS_e}^{-1} \left(S_0 \left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j}, h_{MRSS_e} \right) - S_1 \left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j}, h_{MRSS_e} \right) \left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j} - Y_{MRSS_e} \right) \right)}{S_0 \left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j}, h_{MRSS_e} \right) S_2 \left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j}, h_{MRSS_e} \right) - S_1 \left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j}, h_{MRSS_e} \right)} \quad (47)$$

Where $S_L \left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j}, h_{MRSS_e} \right) =$

$$k_{MRSS_e}^{-1} \sum_{j=1}^r \sum_{i=1}^{\frac{m}{2}} \left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j}, Y_{MRSS_e} \right)^L k_{MRSS_e h} \left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j}, Y_{MRSS_e} \right) \quad (48)$$

The properties of CDF will be studied based on local polynomial using median ranked sets sampling even depending on these conditions:

- 1- The function $F^{(2)}(Y)$ and σ each continuous on $[0, 1]$.
 - 2- The kernel k_{MRSS_e} is symmetric about zero and is reinforced on $[-1, 1]$.
 - 3- The bandwidth $h_{MRSS_e} = h_{MRSS_e k}$ it is a satisfactory sequence, $h_{MRSS_e} \rightarrow 0$ and $kh_{MRSS_e} \rightarrow \infty$ as $k \rightarrow \infty$.
 - 4- The point $\left(Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j} \right)$ at which the estimate is made satisfactory.
- $h_{MRSS_e} < Y_{i(\frac{m}{2})j} + Y_{\frac{m}{2}+i(\frac{m+2}{2})j} < 1 - h_{MRSS_e}$ for all $k \geq k_0$, where k_0 is fixed.

In this subsection, we will make a derivative for the bias and MSE of the estimate CDF based on local polynomial.

$$E \left(\hat{F}_{LPR(1)}(Y_{MRSS_e}) \right) = E \left\{ e_1^T (Y_{MRSS_e}^* W_{MRSS_e} Y_{MRSS_e}^*)^{-1} Y_{MRSS_e}^* W_{MRSS_e} \hat{F}_{MRSS_e} \right\}$$

$$E \left(\hat{F}_{LPR(1)}(Y_{MRSS_e}) \right) = F_{MRSS_e} + \frac{1}{2} h_{MRSS_e}^{(2)} F_{MRSS_e}^{(2)} \mu_2(k_{MRSS_e}) + O(h_{MRSS_e}^{(2)} k_{MRSS_e}^{-1}) \quad (49)$$

$$Bias \left(\hat{F}_{LPR(1)}(Y_{MRSS_e}) \right) = \frac{1}{2} h_{MRSS_e}^{(2)} F_{MRSS_e}^{(2)} \mu_2(k_{MRSS_e}) + O(h_{MRSS_e}^{(2)} k_{MRSS_e}^{-1}) \quad (50)$$

$$MSE \left(\hat{F}_{LPR(1)}(Y_{MRSS_e}) \right) = e_1^T (Y_{MRSS_e}^* W_{MRSS_e} Y_{MRSS_e}^*)^{-1} Y_{MRSS_e}^* W_{MRSS_e} V \left(F(Y_{MRSS_e}) \right) \quad (51)$$

$$MSE \left(\hat{F}_{LPR(1)}(Y_{MRSS_e}) \right) = (k_{MRSS_e} h_{MRSS_e})^{-1} R(k_{MRSS_e}) V \left(F(Y_{MRSS_e}) \right) + O \left\{ (k_{MRSS_e} h_{MRSS_e})^{-1} \right\} \quad (52)$$

In Table 1.3, we will solve the empirical mean square error and bias of CDF based on LPR by using MRSS even with four levels of set size ($m = 2, 4, 6, 8$), with three levels of probability ($P_{MRSS_e} = 0.25, 0.50, 0.75$) with the same number of cycles r , degree of kernel, type of kernel, and bandwidth as in the previous tables.

Table 1.3: The empirical mean square error and bias of CDF based on LPR by Using MRSS EVEN

Kernel / Normal					
	r = 3	m = 2, k = 6	m = 4, k = 12	m = 6, k = 18	m = 8, k = 24
	$P_{MRSS_e} = 0.25$	Mse = 0.0001086189 Bias = 0.0004660877	Mse = 9.811098e-05 Bias = 0.0004429695	Mse = 6.557998e-05 Bias = 0.0003621601	Mse = 8.739888e-05 Bias = 0.0004180882
	$P_{MRSS_e} = 0.50$	418607e-05 Bias = 0.0003	Mse = 5.512962e-05 Bias = 0.000332053	664477e-05 Bias = 0.0003	312043e-05 Bias = 0.0003
	$P_{MRSS_e} = 0.75$	Mse = 1.983807e-05 Bias = 0.0001991887	Mse = 2.367623e-05 Bias = 0.0002176062	Mse = 3.517974e-05 Bias = 0.0002652536	Mse = 3.618193e-05 Bias = 0.0002690053
	r = 5	m = 2, k = 10	m = 4, k = 20	m = 6, k = 30	m = 8, k = 40
	$P_{MRSS_e} = 0.25$	Mse = 3.73182e-05 Bias = 0.0002731966	Mse = 2.330108e-05 Bias = 0.0002158753	Mse = 2.754449e-05 Bias = 0.0002347104	Mse = 2.789499e-05 Bias = 0.000236199

	$P_{MRSS_e} = 0.50$.611155e-05 Bias = 0.0001	Mse = 2.056189e-05 Bias = 0.00020279	.56869e-05 Bias = 0.00017	.545758e-05 Bias = 0.0001
	$P_{MRSS_e} = 0.75$	Mse = 1.09177e-05 Bias = 0.000147768	Mse = 1.606776e-05 Bias = 0.0001792639	Mse = 1.418493e-05 Bias = 0.0001684336	Mse = 1.940943e-05 Bias = 0.000197025
	$r = 7$	$m = 2, k = 14$	$m = 4, k = 28$	$m = 6, k = 42$	$m = 8, k = 56$
	$P_{MRSS_e} = 0.25$	Mse = 1.541827e-05 Bias = 0.0001756034	Mse = 1.022398e-05 Bias = 0.0001429964	Mse = 1.08598e-05 Bias = 0.0001473757	Mse = 1.15695e-05 Bias = 0.0001521151
	$P_{MRSS_e} = 0.50$.078183e-05 Bias = 0.0001	Mse = 1.026643e-05 Bias = 0.000143292	.229568e-05 Bias = 0.0001	.829758e-06 Bias = 0.0001
	$P_{MRSS_e} = 0.75$	Mse = 6.130213e-06 Bias = 0.0001107268	Mse = 7.483774e-06 Bias = 0.0001223419	Mse = 8.35094e-06 Bias = 0.0001292358	Mse = 8.344404e-06 Bias = 0.0001291852
Kernel / Epanechnikov					
	$r = 3$	$m = 2, k = 6$	$m = 4, k = 12$	$m = 6, k = 18$	$m = 8, k = 24$
	$P_{MRSS_e} = 0.25$	Mse = 0.0001040585 Bias = 0.0004561985	Mse = 0.0001081452 Bias = 0.0004650703	Mse = 7.688896e-05 Bias = 0.0003921453	Mse = 8.177441e-05 Bias = 0.0004044117
	$P_{MRSS_e} = 0.50$	Mse = 5.506826e-05 Bias = 0.0003318682	Mse = 4.512977e-05 Bias = 0.0003004323	Mse = 4.609768e-05 Bias = 0.0003036369	Mse = 6.321508e-05 Bias = 0.0003555702
	$P_{MRSS_e} = 0.75$	Mse = 2.04245e-05 Bias = 0.0002021114	Mse = 3.123355e-05 Bias = 0.0002499342	Mse = 3.560996e-05 Bias = 0.0002668706	Mse = 4.591781e-05 Bias = 0.0003030439
	$r = 3$	$m = 2, k = 10$	$m = 4, k = 20$	$m = 6, k = 30$	$m = 8, k = 40$
	$P_{MRSS_e} = 0.25$	Mse = 3.372501e-05 Bias = 0.0002597114	Mse = 2.574094e-05 Bias = 0.0002268962	Mse = 2.332808e-05 Bias = 0.0002160004	Mse = 2.738825e-05 Bias = 0.0002340438
	$P_{MRSS_e} = 0.50$	Mse = 2.510962e-05 Bias = 0.0002240965	Mse = 1.965092e-05 Bias = 0.0001982469	Mse = 1.50904e-05 Bias = 0.0001737262	Mse = 1.491986e-05 Bias = 0.0001727418
	$P_{MRSS_e} = 0.75$	Mse = 7.975208e-06 Bias = 0.000126295	Mse = 1.677073e-05 Bias = 0.0001831433	Mse = 1.190957e-05 Bias = 0.0001543345	Mse = 1.722819e-05 Bias = 0.0001856243
	$r = 7$	$m = 2, k = 14$	$m = 4, k = 28$	$m = 6, k = 42$	$m = 8, k = 56$
	$P_{MRSS_e} = 0.25$	Mse = 1.558296e-05 Bias = 0.0001765387	Mse = 1.510541e-05 Bias = 0.0001738126	Mse = 9.30583e-06 Bias = 0.0001364246	Mse = 1.232422e-05 Bias = 0.0001569982
	$P_{MRSS_e} = 0.50$	Mse = 1.284562e-05 Bias = 0.0001602849	Mse = 9.537822e-06 Bias = 0.0001381146	Mse = 9.276924e-06 Bias = 0.0001362125	Mse = 1.013299e-05 Bias = 0.0001423586
	$P_{MRSS_e} = 0.75$	Mse = 7.110055e-06 Bias = 0.0001192481	Mse = 8.564457e-06 Bias = 0.0001308775	Mse = 8.637282e-06 Bias = 0.0001314327	Mse = 8.434095e-06 Bias = 0.0001298776

Table 1.3, Figure 5, and Figure 6, were interpreted as follows:

- 1- The values of P_{MRSS_e} and m are immutability, Mse and Bias are decreasing, such as the blue line curving in the figure 5, this happens when the repetitions are increasing, $r = 3, 5, 7$.
- 2- As the probability $P_{MRSS_e} = 0.25, 0.50$, and 0.75 increases, the value of both Mse and Bias decreases if the sample size k and m are stable. Figure 6 proves that through the curvature of the blue line.
- 3- On the basis of the result empirical Mse and Bias, we confirmed that the kernel epanechnikov is better than the kernel gaussian.

5. Simulation Study And Conclusion

5.1. Monte Carlo Comparisons

In this part, a comparison is made between the *CDF* estimator for the three methods (moment, maximum likelihood, and local polynomial regression) based on the relative efficiency *RE*, which is defined according to the following relationship:

$$RE_i(CDF) = \frac{MSE(\hat{F}_i(Y_k))}{MSE(\hat{F}(Y_k))} \quad i = MOM, MLE ; k = RSS, MRSS \quad (53)$$

To generate *RSS* and *MRSS*, we assume that the ranking process is done using imperfect ranking mode, with a binomial distribution when simulation $n = 500$ and sample size $k = m * r$ in four levels of sets size $m = 2, 3, 4, 5$, in *RSS*. And in *MRSS* $m = 3, 5, 7, 9$ for type *odd*, also $m = 2, 4, 6, 8$ for type *even*, cycles $r = 3, 5, 7$. With three degrees of probability (0.25, 0.50, 0.75), we used two types of kernels (normal, epanechnikov). We use bandwidth 0.8 with degree 2 for *LPR* because they give a good result compared to other bandwidths and degrees. In Tables 2.1, 2.2, and 2.3, we use R-programing to obtain the result of the relative efficiency *RE*.

Table 2.1: The Relative Efficiency (RE) Of the Method Of Moment And MLE To LPR based on RSS

Kernel / Normal					
r = 3		m = 2, k = 6	m = 3, k = 9	m = 4, k = 12	m = 5, k = 15
MOM & LPR	$P_{RSS} = 0.25$	1.135750702	1.16406889	1.17344999	1.471932307
MLE & LPR	$P_{RSS} = 0.25$	1.137516514	1.162210688	1.174450728	1.474069551
MOM & LPR	$P_{RSS} = 0.50$	8.206516713	8.963245957	10.20831334	10.00206242
MLE & LPR	$P_{RSS} = 0.50$	8.208776681	8.952738327	10.19476759	10.00882799
MOM & LPR	$P_{RSS} = 0.75$	5.063775451	47.27995388	25.63634402	40.74663898
MLE & LPR	$P_{RSS} = 0.75$	5.063065298	47.27842886	25.64026797	40.71708223
r = 5		m = 2, k = 10	m = 3, k = 15	m = 4, k = 20	m = 5, k = 25
MOM & LPR	$P_{RSS} = 0.25$	3.081517979	5.066695671	3.676639827	4.322041795
MLE & LPR	$P_{RSS} = 0.25$	3.089373086	5.06586001	3.67284517	4.334577723
MOM & LPR	$P_{RSS} = 0.50$	2.225527765	22.94889871	23.22915932	21.62112914
MLE & LPR	$P_{RSS} = 0.50$	2.226000338	22.96821049	23.25431716	21.65576338
MOM & LPR	$P_{RSS} = 0.75$	7.401329557	61.04156589	61.22146396	55.01398976
MLE & LPR	$P_{RSS} = 0.75$	7.400843217	60.98252949	61.2500539	54.99610974
r = 7		m = 2, k = 14	m = 3, k = 21	m = 4, k = 28	m = 5, k = 35
MOM & LPR	$P_{RSS} = 0.25$	9.350327185	8.481033713	9.664600701	9.828317550
MLE & LPR	$P_{RSS} = 0.25$	9.314370441	8.501444501	9.677066038	9.871335973
MOM & LPR	$P_{RSS} = 0.50$	3.896041410	37.03106475	31.81991703	44.73795452
MLE & LPR	$P_{RSS} = 0.50$	3.899763472	37.01806995	31.86743149	44.7348804
MOM & LPR	$P_{RSS} = 0.75$	1.386600932	108.5673353	85.88318018	146.7357995
MLE & LPR	$P_{RSS} = 0.75$	1.385961876	108.5636544	85.8482365	146.9243931
Kernel / Epanechnikov					
r = 3		m = 2, k = 6	m = 3, k = 9	m = 4, k = 12	m = 5, k = 15
MOM & LPR	$P_{RSS} = 0.25$	1.012059090	1.408594426	1.716100684	1.326686852
MLE & LPR	$P_{RSS} = 0.25$	1.013632592	1.406345891	1.717564203	1.328613199
MOM & LPR	$P_{RSS} = 0.50$	8.620434815	8.554989917	8.456819846	8.24706609
MLE & LPR	$P_{RSS} = 0.50$	8.622808771	8.544960886	8.445598208	8.252644556
MOM & LPR	$P_{RSS} = 0.75$	5.517384663	47.42853921	33.52502062	25.49415595
MLE & LPR	$P_{RSS} = 0.75$	5.516610895	47.4270094	33.53015203	25.47566302
r = 5		m = 2, k = 10	m = 3, k = 15	m = 4, k = 20	m = 5, k = 25
MOM & LPR	$P_{RSS} = 0.25$	3.561232678	4.385218429	5.910724402	4.978331956
MLE & LPR	$P_{RSS} = 0.25$	3.570310627	4.384495166	5.904623949	4.992771429
MOM & LPR	$P_{RSS} = 0.50$	2.360911893	24.35167501	23.85938013	21.84837033
MLE & LPR	$P_{RSS} = 0.50$	2.361413214	24.37216724	23.88522051	21.88336858
MOM & LPR	$P_{RSS} = 0.75$	1.385402937	95.95917942	52.62949429	53.18112657
MLE & LPR	$P_{RSS} = 0.75$	1.385311902	95.86637242	52.65407185	53.16384225
r = 7		m = 2, k = 14	m = 3, k = 21	m = 4, k = 28	m = 5, k = 35
MOM & LPR	$P_{RSS} = 0.25$	7.578423631	9.874006138	9.372102845	7.796579219
MLE & LPR	$P_{RSS} = 0.25$	7.549280754	9.897769309	9.384190921	7.830704748
MOM & LPR	$P_{RSS} = 0.50$	4.566728578	33.27649772	41.39732994	38.54676471
MLE & LPR	$P_{RSS} = 0.50$	4.571091377	33.26482046	41.45914569	38.54411601
MOM & LPR	$P_{RSS} = 0.75$	1.328531874	154.4991877	143.8662039	90.59330149
MLE & LPR	$P_{RSS} = 0.75$	1.327919581	154.4939494	143.8076684	90.7097374

Table 2.2: The Relative Efficiency (RE) Of the Method Of Moment And MLE To LPR Of MRSS ODD

Kernel / Normal					
r = 3		m = 3, k = 9	m = 5, k = 15	m = 7, k = 21	m = 9, k = 27
MOM & LPR	$P_{MRSS_0} = 0.25$	1.415389188	1.527881494	1.402094949	1.406503243
MLE & LPR	$P_{MRSS_0} = 0.25$	1.414098151	1.531513881	1.405189967	1.412945634
MOM & LPR	$P_{MRSS_0} = 0.50$	8.226359101	8.515932716	7.231069192	7.052961598
MLE & LPR	$P_{MRSS_0} = 0.50$	8.222205855	8.499512922	7.225271195	7.053201262
MOM & LPR	$P_{MRSS_0} = 0.75$	4.664795944	27.80029136	25.26318726	26.65122506
MLE & LPR	$P_{MRSS_0} = 0.75$	4.662367272	27.79661329	25.27047898	26.64191247
r = 5		m = 3, k = 15	m = 5, k = 25	m = 7, k = 35	m = 9, k = 45
MOM & LPR	$P_{MRSS_0} = 0.25$	4.458475947	4.336927959	4.756751253	4.901259415
MLE & LPR	$P_{MRSS_0} = 0.25$	4.466569344	4.335760058	4.73894657	4.898157221

MOM & LPR	$P_{MRSS_0} = 0.50$	3.088215710	24.9128672	21.95669351	25.34578294
MLE & LPR	$P_{MRSS_0} = 0.50$	3.088062861	24.93722305	21.95971501	25.32617144
MOM & LPR	$P_{MRSS_0} = 0.75$	9.552986366	59.23718671	72.01075818	62.79158109
MLE & LPR	$P_{MRSS_0} = 0.75$	9.559189622	59.21191988	72.07851325	62.82913636
r = 7		m = 3, k = 21	m = 5, k = 35	m = 7, k = 49	m = 9, k = 63
MOM & LPR	$P_{MRSS_0} = 0.25$	7.326612388	11.49452979	9.720133663	8.919277964
MLE & LPR	$P_{MRSS_0} = 0.25$	1.322835909	11.41768038	9.745375456	8.955212783
MOM & LPR	$P_{MRSS_0} = 0.50$	5.287188996	47.05310094	47.58948973	49.52667836
MLE & LPR	$P_{MRSS_0} = 0.50$	5.287012619	46.91815183	47.52536847	49.5257712
MOM & LPR	$P_{MRSS_0} = 0.75$	1.145476679	148.3637927	95.61809737	99.81082903
MLE & LPR	$P_{MRSS_0} = 0.75$	1.144674214	148.3842194	95.57852163	99.64525508
Kernel / Epanechnikov					
r = 3		m = 3, k = 9	m = 5, k = 15	m = 7, k = 21	m = 9, k = 27
MOM & LPR	$P_{MRSS_0} = 0.25$	1.149410868	1.614515338	1.843264641	1.747199133
MLE & LPR	$P_{MRSS_0} = 0.25$	1.148362441	1.618353688	1.847333509	1.755202058
MOM & LPR	$P_{MRSS_0} = 0.50$	8.979216413	8.249211862	6.078342982	8.132878988
MLE & LPR	$P_{MRSS_0} = 0.50$	8.974683071	8.233306339	6.073469261	8.133155348
MOM & LPR	$P_{MRSS_0} = 0.75$	3.294042012	26.82358148	28.29655331	28.27585839
MLE & LPR	$P_{MRSS_0} = 0.75$	3.292327007	26.82003263	28.30472056	28.26597812
r = 5		m = 3, k = 15	m = 5, k = 25	m = 7, k = 35	m = 9, k = 45
MOM & LPR	$P_{MRSS_0} = 0.25$	4.345040051	5.268698317	4.586477365	5.407750805
MLE & LPR	$P_{MRSS_0} = 0.25$	4.352927529	5.267279497	4.569310023	5.404328033
MOM & LPR	$P_{MRSS_0} = 0.50$	2.323389334	26.0140591	17.9700734	24.7286995
MLE & LPR	$P_{MRSS_0} = 0.50$	2.323274339	26.03949151	17.9725463	24.70956546
MOM & LPR	$P_{MRSS_0} = 0.75$	9.146669779	99.17496908	59.22260449	55.85200401
MLE & LPR	$P_{MRSS_0} = 0.75$	9.152609192	99.13266732	59.27832716	55.88540876
r = 7		m = 3, k = 21	m = 5, k = 35	m = 7, k = 49	m = 9, k = 63
MOM & LPR	$P_{MRSS_0} = 0.25$	8.833738122	8.007998814	7.418449251	8.355552021
MLE & LPR	$P_{MRSS_0} = 0.25$	8.842005986	7.954459438	7.437713899	8.389215649
MOM & LPR	$P_{MRSS_0} = 0.50$	3.907380366	43.30962142	32.17081358	36.2482038
MLE & LPR	$P_{MRSS_0} = 0.50$	3.907250019	43.18540868	32.12746718	36.24753986
MOM & LPR	$P_{MRSS_0} = 0.75$	1.105891296	114.7249474	109.1196595	80.67786555
MLE & LPR	$P_{MRSS_0} = 0.75$	1.105116564	114.7407427	109.0744956	80.54403085

Table 2.3: The Relative Efficiency (RE) Of the Method Of Moment And MLE To LPR based on MRSS EVEN

Kernel / Normal					
r = 3		m = 2, k = 6	m = 4, k = 12	m = 6, k = 18	m = 8, k = 24
MOM & LPR	$P_{MRSS_e} = 0.25$	1.121467811	1.213447261	1.764840428	1.297525781
MLE & LPR	$P_{MRSS_e} = 0.25$	1.124377986	1.211776704	1.770773642	1.297350721
MOM & LPR	$P_{MRSS_e} = 0.50$	9.008062404	8.634465465	9.953317382	7.173374769
MLE & LPR	$P_{MRSS_e} = 0.50$	9.009959571	8.638287367	9.958276137	7.185803392
MOM & LPR	$P_{MRSS_e} = 0.75$	5.533058407	45.2334261	29.71497231	28.20200581
MLE & LPR	$P_{MRSS_e} = 0.75$	5.534016162	45.28242039	29.74194806	28.18686013
r = 5		m = 2, k = 10	m = 4, k = 20	m = 6, k = 30	m = 8, k = 40
MOM & LPR	$P_{MRSS_e} = 0.25$	3.217770418	4.952221099	4.012744473	3.795039898
MLE & LPR	$P_{MRSS_e} = 0.25$	3.225021571	4.94776637	4.016705337	3.786597522
MOM & LPR	$P_{MRSS_e} = 0.50$	2.979935512	22.41561938	28.1837648	27.35911443
MLE & LPR	$P_{MRSS_e} = 0.50$	2.980518944	22.41636834	28.1842429	27.38291505
MOM & LPR	$P_{MRSS_e} = 0.75$	9.897176145	64.53177045	70.13118147	49.04326917
MLE & LPR	$P_{MRSS_e} = 0.75$	9.894483270	64.54179052	70.10334207	49.09162196
r = 7		m = 2, k = 14	m = 4, k = 28	m = 6, k = 42	m = 8, k = 56
MOM & LPR	$P_{MRSS_e} = 0.25$	7.653310002	10.89775215	9.644330466	8.469671982
MLE & LPR	$P_{MRSS_e} = 0.25$	7.665166066	10.89591333	9.640288035	8.519949004

MOM & LPR	$P_{MRSS_e} = 0.50$	4.383895869	43.43432917	34.10624707	57.80866613
MLE & LPR	$P_{MRSS_e} = 0.50$	4.379929010	43.36149957	34.12770176	57.71607135
MOM & LPR	$P_{MRSS_e} = 0.75$	1.733270932	133.91933	112.9850412	106.2509198
MLE & LPR	$P_{MRSS_e} = 0.75$	1.734050677	133.9057005	113.0363767	106.5025735
Kernel / Epanechnikov					
r = 3		m = 2, k = 6	m = 4, k = 12	m = 6, k = 18	m = 8, k = 24
MOM & LPR	$P_{MRSS_e} = 0.25$	1.170616528	1.100857921	1.505264215	1.386770018
MLE & LPR	$P_{MRSS_e} = 0.25$	1.173654243	1.099342366	1.510324759	1.386582918
MOM & LPR	$P_{MRSS_e} = 0.50$	8.863753821	10.54768947	10.07144394	7.162634295
MLE & LPR	$P_{MRSS_e} = 0.50$	8.865620595	10.55235823	10.07646155	7.175044309
MOM & LPR	$P_{MRSS_e} = 0.75$	5.374192759	34.28867356	29.35597232	22.22237951
MLE & LPR	$P_{MRSS_e} = 0.75$	5.375123014	34.32581311	29.38262217	22.21044514
r = 5		m = 2, k = 10	m = 4, k = 20	m = 6, k = 30	m = 8, k = 40
MOM & LPR	$P_{MRSS_e} = 0.25$	3.560603837	4.48282386	4.738023875	3.865256086
MLE & LPR	$P_{MRSS_e} = 0.25$	3.568627556	4.478791373	4.742700642	3.856657508
MOM & LPR	$P_{MRSS_e} = 0.50$	1.912071150	23.45475428	29.29782511	28.34515203
MLE & LPR	$P_{MRSS_e} = 0.50$	1.912445509	23.45553796	29.29832211	28.36981044
MOM & LPR	$P_{MRSS_e} = 0.75$	1.354878769	61.82682567	83.52995952	55.25257732
MLE & LPR	$P_{MRSS_e} = 0.75$	1.354510127	61.83642573	83.49680131	55.30705199
r = 7		m = 2, k = 14	m = 4, k = 28	m = 6, k = 42	m = 8, k = 56
MOM & LPR	$P_{MRSS_e} = 0.25$	7.572425265	7.376059306	11.25482628	7.950999739
MLE & LPR	$P_{MRSS_e} = 0.25$	7.584156027	7.374814719	11.2501088	7.998197858
MOM & LPR	$P_{MRSS_e} = 0.50$	3.679574828	46.75234031	45.20458505	38.96374121
MLE & LPR	$P_{MRSS_e} = 0.50$	3.676245288	46.67394715	45.2330212	38.9013312
MOM & LPR	$P_{MRSS_e} = 0.75$	1.494407568	117.0210791	109.2393765	105.1210118
MLE & LPR	$P_{MRSS_e} = 0.75$	1.495079855	117.0091694	109.2890101	105.3699893

The next five steps are a description and explanation of relative efficiency RE tables :

- 1- LPR was more efficient than the method of moment and maximum likelihood at the same set size and cycles for estimating CDF .
- 2- In different kernels, LPR remains more efficient.
- 3- By increasing the number of cycles, which is r , the efficiency of LPR will increase.
- 4- Evidently, using LPR significantly improves the behavior of the method of moment and MLE . It is apparent that all the proposed procedures based on LPR can be the best choice.
- 5- The relative efficiency in the case of $MRSS$ odd, even has a higher efficiency result with the LPR method compared to the RSS .

5.2 Conclusion

This article is concerned with estimating the cumulative distribution function CDF based on the local polynomial regression LPR depending on RSS and $MRSS$. A new CDF estimator depends on LPR is derived. The resulting proposed estimator is used to introduce three ways of estimating CDF (the method of moments, the maximum likelihood method, and LPR), based on RSS and $MRSS$. The method of moments and the maximum likelihood method based on RSS were suggested by Al-Saleh and Ahmad (2019). In this study we get the same result of CDF estimate for the method of moments and the maximum likelihood method, CDF estimator based on LPR can have some advantages over their competitors for fixed or non-fixed samples; on the empirical side, we use kernel (normal, epanechnikov) with bandwidth (0.1, 0.2, ..., 0.9) and three levels of degree of kernel, We have concluded that kernel epanechnikov is a little better than normal, with bandwidth 0.8 giving the best result. Compared to the others, bandwidth and degree 2 are also like that. We have concluded that CDF of LPR based on $MRSS$ is better than CDF of LPR based on RSS because the data in $MRSS$ is stable and is less prone to ranking errors for fixed or non-fixed samples. Depending on the relative efficiency, we get that there is no big difference between both kernels, nonetheless we conclude that relative efficiency depends on whether the $MRSS$ is better than the RSS and has more efficiency. We recommend using LPR based on estimators.

Fig. 1: The mean square error and bias of CDF based on RSS with $P = 0.25$.

$$m = 2, r = 3, k = 6$$

$$m = 2, r = 5, k = 10$$

$$m = 2, r = 7, k = 14$$

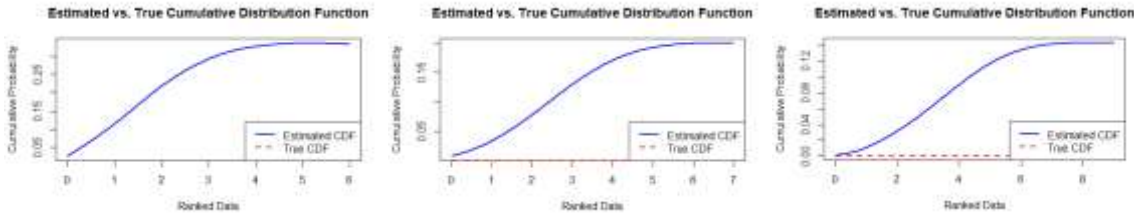


Fig. 2: The mean square error and bias of CDF based on RSS with $m = 5, r = 3, k = 15$

$$P = 0.25$$

$$P = 0.5$$

$$P = 0.75$$

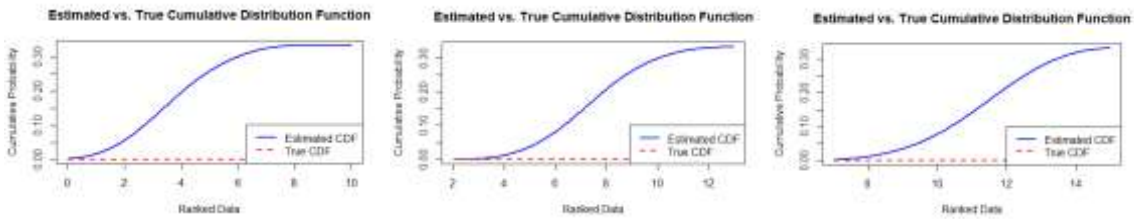


Fig. 3: The mean square error and bias of CDF based on MRSS ODD with $m = 5, P = 0.5$.

$$m = 5, r = 3, k = 15$$

$$m = 5, r = 5, k = 25$$

$$m = 5, r = 7, k = 35$$

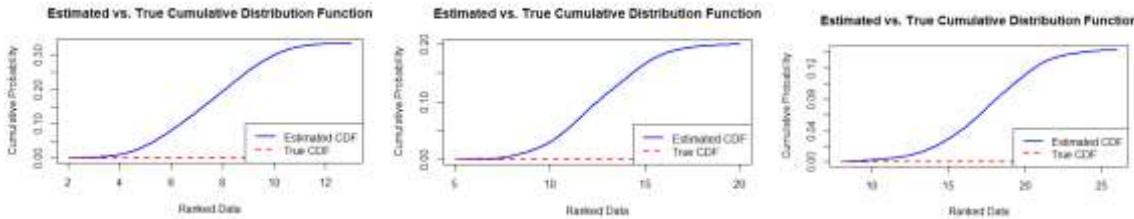


Fig. 4: The mean square error and bias of CDF based on MRSS ODD with $m = 9, r = 5, k = 45$

$$P = 0.25$$

$$P = 0.5$$

$$P = 0.75$$

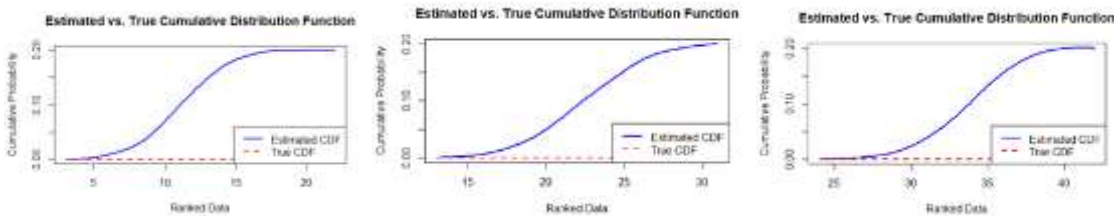


Fig. 5: The mean square error and bias of CDF based on MRSS EVEN with $m = 6, P = 0.75$.

$$m = 6, r = 3, k = 18$$

$$m = 6, r = 5, k = 35$$

$$m = 6, r = 7, k = 42$$

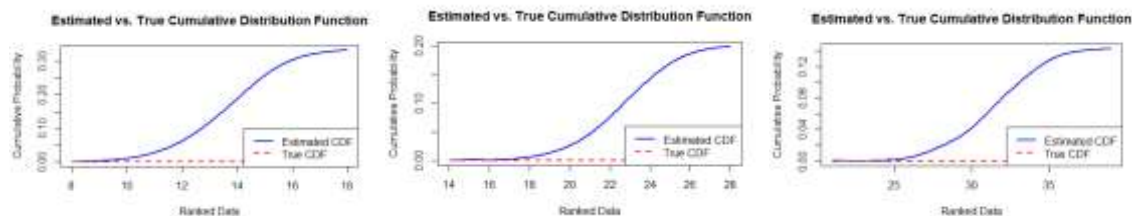
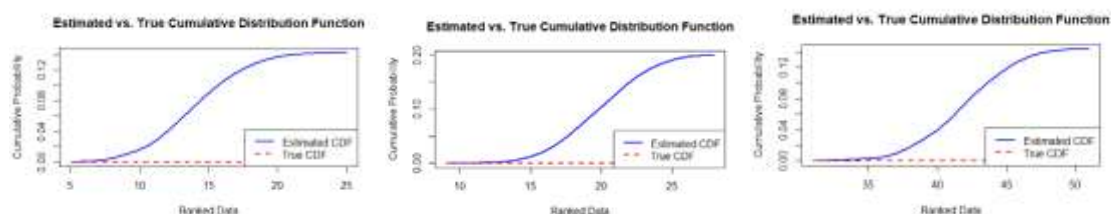


Fig. 6: The mean square error and bias of CDF based on MRSS EVEN with $m = 8, r = 7, k = 56$

$P = 0.25$

$P = 0.5$

$P = 0.75$



References

1. Al-Saleh, M. F., & Ahmad, D. M. R. (2019). Estimation of the distribution function using moving extreme ranked set sampling (MERSS). In *Ranked Set Sampling* (pp. 43-58). Academic Press.
2. Abdallah, M. S., & Al-Omari, A. I. (2022). ON THE NONPARAMETRIC ESTIMATION OF THE ODDS AND DISTRIBUTION FUNCTION USING MOVING EXTREME RANKED SET SAMPLING. *Investigación Operacional*, 43(1), 90-102.
3. AL_Rahman, R., & Mohammad, S. (2022). Generalized ratio-cum-product type exponential estimation of the population mean in median ranked set sampling. *Iraqi Journal of Statistical Sciences*, 19(1), 54-66.
4. Frey, J. (2012). Constrained nonparametric estimation of the mean and the CDF using ranked-set sampling with a covariate. *Annals of the Institute of Statistical Mathematics*, 64(2), 439-456.
5. Fan, J., Gijbels, I., Hu, T. C., & Huang, L. S. (1996). A study of variable bandwidth selection for local polynomial regression. *Statistica Sinica*, 113-127.
6. Gulati, S. (2004). Smooth non-parametric estimation of the distribution function from balanced ranked set samples. *Environmetrics*, 15(5), 529-539.
7. HA, M. (1997). Median ranked set sampling. *J Appl Stat Sci*, 6, 245-255.
8. Halls, L. K., & Dell, T. R. (1966). Trial of ranked-set sampling for forage yields. *Forest Science*, 12(1), 22-26.
9. McIntyre, G. A. (1952). A method for unbiased selective sampling, using ranked sets. *Australian journal of agricultural research*, 3(4), 385-390.
10. Stokes, S. L., & Sager, T. W. (1988). Characterization of a ranked-set sample with application to estimating distribution functions. *Journal of the American Statistical Association*, 83(402), 374-381.
11. Takahasi, K., & Wakimoto, K. (1968). On unbiased estimates of the population mean based on the sample stratified by means of ordering. *Annals of the institute of statistical mathematics*, 20(1), 1-31.
12. Zamanzade, E., Mahdizadeh, M., & Samawi, H. M. (2020). Efficient estimation of cumulative distribution function using moving extreme ranked set sampling with application to reliability. *AstA Advances in Statistical Analysis*, 104(3), 485-502.

طريقة التقدير اللامعلمية لدالة التوزيع باستخدام أنواع مختلفة من مجموعات العينات المرتبة

رامي سعد غريب¹ و ريسان عبد العزيز احمد²

¹ طالب ماجستير، قسم الإحصاء والمعلوماتية، كلية علوم الحاسوب والرياضيات، جامعة الموصل، موصل، العراق.

² قسم الإحصاء والمعلوماتية، كلية علوم الحاسوب والرياضيات، جامعة الموصل، موصل، العراق.

الخلاصة: الغرض من هذا البحث هو تقدير دالة التوزيع التراكمي CDF باستخدام الانحدار متعدد الحدود المحلي LPR ومقارنته مع الطرق المعلمية وهي كل من طريقة العزوم وطريقة الامكان الأعظم لغرض حساب متوسط مربع الخطأ والتحيز بالأعتماد على منهجية مجموعة عينات المرتبة RSS ومنهجية مجموعة عينات المرتبة الوسطى $MRSS$. بالإضافة إلى أن RSS غالباً تنتج تقديرات أكثر دقة من عينة عشوائية بسيطة SRS و لنفس حجم العينة، ويتم ذلك من خلال ترتيب العينات بناءً على بعض الخصائص التي يمكن قياسها بسهولة، يتم تقليل التباين داخل كل مجموعة، مما يؤدي إلى تقديرات أكثر دقة. لقد قمنا بدراسة ثلاث درجات مختلفة من الانحدار متعدد الحدود المحلي: الدرجة الأولى والثانية والثالثة، أظهر تحليل المحاكاة أن الدرجة الثانية تتفوق على الدرجات الأخرى، وعندما يتم استخدام LPR لتحليل بيانات RSS ، فإنه يستفيد من التباين المنخفض داخل كل مجموعة مرتبة، وهذا ما يؤدي إلى تقديرات أكثر دقة وموثوقية لدالة الانحدار، إضافة إلى ذلك، قمنا بدراسة درجات مختلفة من النطاق الترددي (0.1, 0.2, ..., 0.9) وتبيننا لنا من خلال دراسة المحاكاة أن النطاق الترددي للدرجة 0.8 يتفوق على الدرجات الأخرى. بعد ذلك، قمنا بتحليل الكفاءة النسبية لكل من الأساليب الثلاثة: LPR و MOM و MLE ، و تبيننا لنا أن طريقة LPR هي أكثر كفاءة من الطرق الأخرى لتقدير CDF في نواة مختلفة (epanechinkov, (gaussian) normal)، و من ثم أظهرت دراسة المحاكاة أن مقدر CDF المقترح بناءً على LPR أكثر كفاءة من الطرق المعلمية.

الكلمات المفتاحية: الانحدار متعدد الحدود المحلي، عينة مجموعة مرتبة، متوسط عينة مجموعة مرتبة، خطأ مربع دالة التوزيع التراكمي