JIRSS (2023)
Vol. 22, No. 02, pp 97-117
DOI: 10.22034/jirss.2024.2004038.1024

Bayesian Bandwidth Estimation based on Ranked Set Sampling

Ali Najafi Majidabadi [1], **Nader Nematollahi** [1]

[1] Department of Statistics, Allameh Tabataba'i University, Tehran, Iran.

Received: 05/06/2023, Accepted: 27/04/2024, Published online: 30/10/2024

Abstract. In the estimation of a probability density function (PDF) by kernel method, two inherent problems are the choice of sampling methods and the selection of a bandwidth. In this article, we use the balanced and unbalanced ranked set sampling (RSS) methods and Bayesian bandwidth to estimate a PDF by kernel method. To compare our method with existing methods, we use an extensive simulation study to compare the RSS with simple random sampling (SRS) PDF estimator and also Bayesian bandwidth with other existing bandwidths. As an application, we use the household expenses and income data of the Statistical Center of Iran in 2021 to estimate the PDF of the total expenses of the households in Tehran province.

Keywords. Bayesian Bandwidth, Ranked Set Sampling, Probability Density Function Estimation, Simple Random Sampling.

MSC: 62D05, 62F15, 62G07.

1 Introduction

The probability density function (PDF) is a very basic and fundamental concept in statistics. The problem of the PDF estimation of the study variable in the underlying population appears naturally in many areas such as medicine, seismology, hydrology, economics, and environmental sciences. For example, in economics it is widely used to examine the distribution of income, wages, and poverty (Minoiu and Reddy, 2014). Most of the studies on PDF estimation are based on kernel density estimation and under simple random sampling (SRS) design. For an overview on PDF estimation under SRS,

Ali Najafi Majidabadi (najafiali920@gmail.com)
Corresponding Author: Nader Nematollahi (nematollahi@atu.ac.ir)

see Rosenblatt (1956), Parzen (1962), Silverman (1986), Wand and Jones (1995), and Wasserman (2006).

In Practice, SRS design is not necessarily efficient and one can use more efficient sampling procedures to obtain better samples from the underlying population and possibly improve the efficiency of the estimators. One of these sampling designs is ranked set sampling (RSS) design which can achieve efficiency and reduce cost in certain situations where the measurement of the variable of interest is costly or time consuming but the ranking of a cheaply related auxiliary variable can be easily done by judgment without actual measurement. Chen (1999) used RSS to estimate the PDF and showed that the mean integrated squared error (MISE) of this estimator is less than or equal to the MISE of the SRS estimator. Barabesi and Fattorini (2002) studied the kernel density estimator of the unbalanced RSS. Lim et al. (2014) obtained an optimal bandwidth for kernel density estimation under RSS. Samawi et al. (2018) investigated some of the asymptotic properties of kernel density-based mode estimating using RSS.

One of the most important problems in kernel density estimation is selecting the bandwidth. Classical bandwidth selection methods, such as cross-validation, require symmetric kernels to perform well. Also, the bandwidth used in these methods gives an asymptotic hyper-local value, which may perform poorly in some intervals of the density support, especially in density functions with a positive support (Kulasekera and Padgett, 2006). Gangopadhyay and Cheung (2002) proposed a method to calculate the bandwidth values based on the data under SRS so that the bandwidth can be controlled using the study data. Their main idea was to assume that the bandwidth has its own distribution. In this method, a priori distribution was considered for the bandwidth in the neighborhood of the point where the density is estimated. This type of bandwidth is called the Bayesian bandwidth. They showed by a simulation study that the performance of the PDF estimator based on a reasonable prior distribution for bandwidth compares favorably to other commonly used local and global estimators.

In this paper we use Bayesian bandwidth to estimate a PDF by kernel method based on balanced and unbalanced RSS designs. To this end in Section 2, we state some preliminary results in kernel density estimation under SRS, balanced RSS, and unbalanced RSS designs. In section 3, the Bayesian estimate of bandwidth under balanced and unbalanced RSS is presented. To compare the different sampling designs and different bandwidths in estimation of a PDF, an extensive simulation study is conducted in Section 4. In section 5, we apply the proposed method to estimate the PDF of the total expenses of the households in Tehran province based on the household expenses and income data of Statistical Center of Iran in 2021. Finally, concluding remarks are given in section 6.

2 Preliminaries

In this section, we review some properties of kernel density estimation under SRS and balanced and unbalanced RSS designs.

2.1 Kernel Density Estimation under SRS

let $X_1, ..., X_n$ be a simple random sample of size $n(= Ht)$ from a population with unknown PDF $f(.)$. The kernel density estimator of $f(.)$ at a point x is defined as (Silverman, 1986)

$$\hat{f}_{SRS}(x) = \frac{1}{n\lambda}\sum_{i=1}^{n} K(\frac{x - X_i}{\lambda}) = \frac{1}{n}\sum_{i=1}^{n} k_{\lambda}(x - X_i), \tag{2.1}$$

where λ is the bandwidth and $K(.)$ is a kernel function that has the following properties

$$\int_{-\infty}^{+\infty} K(x)dx = 1, \int_{-\infty}^{+\infty} xK(x)dx = 0, \int_{-\infty}^{+\infty} x^2K(x)dx \neq 0.$$

The second condition may not be zero for asymmetric kernels. The mean and variance of $\hat{f}_{SRS}(x)$ is given by (Silverman, 1986)

$$E(\hat{f}_{SRS}(x)) = \int_{-\infty}^{\infty} \frac{1}{\lambda} K\left(\frac{x-y}{\lambda}\right) f(y)dy,$$

$$Var(\hat{f}_{SRS}(x)) = \frac{1}{n}\left\{\int_{-\infty}^{\infty} \frac{1}{\lambda^2} K^2\left(\frac{x-y}{\lambda}\right) f(y)dy - \left(\int_{-\infty}^{\infty} \frac{1}{\lambda} K\left(\frac{x-y}{\lambda}\right) f(y)dy\right)^2\right\}.$$

Let $i_l(f) = \int x^l f(x)dx$, then Wand and Jones (1995) showed that the approximate mean and variance of $\hat{f}(x)$ are

$$E(\hat{f}_{SRS}(x)) \approx f(x) + \frac{\lambda^2}{2} f''(x) i_2(K) + O(\lambda^2), \tag{2.2}$$

$$Var(\hat{f}_{SRS}(x)) \approx \frac{1}{n\lambda} f(x) i_0(K^2) + O(\frac{\lambda^2}{n}), \tag{2.3}$$

and hence

$$MSE(\hat{f}_{SRS}(x)) \approx \frac{1}{n\lambda} f(x) i_0(K^2) + \frac{\lambda^4}{4} f''^2(x) i_2^2(K).$$

Therefore the $MISE$ of $\hat{f}_{SRS}(x)$ is given by

$$MISE(\hat{f}_{SRS}(x)) = \int MSE(\hat{f}_{SRS}(x))dx \approx \frac{1}{n\lambda} i_0(K^2) + \frac{\lambda^4}{4} i_2^2(K) \int f''^2(x)dx. \tag{2.4}$$

2.2 Kernel Density Estimation under Balanced RSS

To select a sample of size H from the underlying population with PDF $f(.)$ based on RSS, H independent sample of size H are selected from the population. Then, the units in each sample are ranked from smallest to largest by judgment or some auxiliary variable that does not require actual measurement. Now, from the first sample of size H, the unit with the smallest rank, from the second sample the unit with the second rank, and in the same way from the Hth sample of size H, the unit with the largest rank is selected and measured. The H measured units constitute a balanced RSS with set size H. In some situations, ranking H units for large H is difficult, so we select a RSS with a small H and then repeat the sampling scheme t times (t cycles) to obtain a balanced RSS of size $n = tH$. It should be noted that this sampling method needs to collect Ht^2 units, but at the end only $n = tH$ units are measured. Let $X_{[h]i}$ be the hth judgment order statistic of a sample of size H from the ith sampling cycle ($h = 1, ..., H, i = 1, ..., t$). Then for fixed h, $X_{[h]1}, ..., X_{[h]t}$ are independent and identically distributed with PDF $f_{[h]}(.)$. Chen (1999) proposed the following estimator of $f_{[h]}(.)$ based on kernel method and balanced RSS,

$$\hat{f}_{[h]}(x) = \frac{1}{t\lambda} \sum_{i=1}^{t} K(\frac{x - X_{[h]i}}{\lambda}).$$

It can be shown that in a balanced RSS with size H from a population with PDF $f(.)$, $f(x) = \frac{1}{H} \sum_{h=1}^{H} f_{[h]}(x)$ (Chen, 1999). Therefore the kernel density estimator of $f(x)$ is given by

$$\hat{f}_{RSS}(x) = \frac{1}{H} \sum_{h=1}^{H} \hat{f}_{[h]}(x) = \frac{1}{n\lambda} \sum_{h=1}^{H} \sum_{i=1}^{t} K(\frac{x - X_{[h]i}}{\lambda}). \tag{2.5}$$

Chen (1999) showed that the expectation and variance of $\hat{f}_{RSS}(x)$ is given by

$$E(\hat{f}_{RSS}(x)) = \frac{1}{H} \sum_{h=1}^{H} \int \frac{1}{\lambda} K(\frac{x - y}{\lambda}) f_{[h]}(y) dy = \int \frac{1}{\lambda} K(\frac{x - y}{\lambda}) f(y) dy = E(\hat{f}_{SRS}(x)),$$

$$Var(\hat{f}_{RSS}(x)) = Var(\hat{f}_{SRS}(x)) + \frac{1}{Ht} \left(\left[E(\frac{1}{\lambda} K(\frac{x - X}{\lambda})) \right]^2 - \frac{1}{H} \sum_{h=1}^{H} \left[E \frac{1}{\lambda} K(\frac{x - X_{[h]}}{\lambda}) \right]^2 \right),$$

and also showed that $Var(\hat{f}_{RSS}(x)) \leq Var(\hat{f}_{SRS}(x))$. Using the same bandwidth for $\hat{f}_{RSS}(x)$ and $\hat{f}_{SRS}(x)$, Chen (1999) showed that for fixed H and large sample size n,

$$MISE(\hat{f}_{RSS}(x)) = MISE(\hat{f}_{SRS}(x)) - \frac{1}{n} \Delta(f, H) + O(\frac{\lambda^2}{n}),$$

where $\Delta(f, H) = \int \left[\frac{1}{H} \sum_{h=1}^{H} f_{[h]}^2(x) - f^2(x) \right] dx.$

2.3 Kernel Density Estimation under Unbalanced RSS

To select an unbalanced RSS of size $n = t_1 + t_2 + \cdots + t_H$ from the underlying population with PDF $f(.)$, for $i = 1, 2, ..., H$, t_i sample of size H is selected. Then, the units in each sample of size H are ranked from smallest to largest by judgement or some auxiliary variable that does not require actual measurement. Finally, from t_1 sample of size H the units with the smallest rank, from t_2 sample of size H the units with the second rank, and in the same way from t_H sample of size H the units with the largest rank are selected and measured. The measured units constitute an unbalanced RSS with size $n = t_1 + t_2 + \cdots + t_H$. Note that $t_1, t_2, \cdots, t_H$ are known values and if $t_1 = t_2 = \cdots = t_H = t$, we have a balanced RSS. To estimate the PDF $f(.)$, assume $\left\{X_{[h]i}, h = 1, ..., H, i = 1, ..., t_h\right\}$ be the unbalanced RSS of size $n = t_1 + t_2 + \cdots + t_H$. Barabesi and Fattorini (2002) proposed the kernel density estimator based on unbalanced RSS as follows,

$$\hat{f}_{URSS}(x) = \frac{1}{H}\sum_{h=1}^{H}\hat{f}_{[h]}(x), \tag{2.6}$$

where $\hat{f}_{[h]}(x) = \frac{1}{\lambda t_h}\sum_{i=1}^{t_h} K(\frac{x - X_{[h]i}}{\lambda})$. Barabesi and Fattorini (2002) showed that the expectation and variance of $\hat{f}_{URSS}(x)$ are given by

$$E(\hat{f}_{URSS}(x)) = E(\hat{f}_{SRS}(x)),$$

$$Var(\hat{f}_{URSS}(x)) = \frac{1}{H^2}\sum_{h=1}^{H}\frac{\sigma^2_{[h]h}(x)}{t_h},$$

where $\sigma^2_{[h]\lambda}(x) = Var\left[\frac{1}{\lambda}K(\frac{x - X_{[h]}}{\lambda})\right]$. They also showed that $Var(\hat{f}_{URSS}(x)) \leq Var(\hat{f}_{SRS}(x))$ and expressed the asymptotic variance and MISE of $\hat{f}_{URSS}(x)$ as follows

$$Var(\hat{f}_{URSS}(x)) = \frac{i_0(K^2)}{H^2\lambda}\sum_{h=1}^{H}\frac{f_{[h]}(x)}{t_h} + O(\frac{\lambda^2}{n}),$$

$$MISE(\hat{f}_{URSS}(x)) = \frac{i_0(K^2)}{H^2\lambda}\sum_{h=1}^{H}\frac{1}{t_h} + \frac{\lambda^4}{4}i_2^2(K)i_0(f''^2) + O(\frac{\lambda^2}{n}). \tag{2.7}$$

3 Bayesian Bandwidth

In this section, we first review some existing methods for calculating bandwidth and then propose a Bayesian bandwidth for density estimation under RSS.

3.1 Bandwidth Selection

Under SRS, Wand and Jones (1995) obtained optimal bandwidth λ by minimizing the MISE given in (2.4) with respect to λ which yields to

$$\lambda_{AMISE} = \left(\frac{i_0(K^2)}{i_0(f''^2) i_2^2(K) n} \right)^{\frac{1}{5}}. \tag{3.1}$$

Since (3.1) depends on unknown PDF $f(x)$, it can not be calculated directly. There are several methods in literature to select λ based on an estimate of $i_0(f'')$. One method is approximate $f(.)$ by a standard normal distribution which yields to normal scale (NS) bandwidth $\lambda_{NS} = 1.06\sigma n^{\frac{-1}{5}}$. Wasserman (2006) proposed the estimate $\hat{\sigma} = \min(s, \frac{IQR}{1.349})$ for σ, where s is the sample standard deviation and IQR is the interquartile range. Therefore

$$\lambda_{NS} = 1.06\hat{\sigma} n^{\frac{-1}{5}}, \ \ \hat{\sigma} = \min(s, \frac{IQR}{1.349}).$$

Another method is to find an upper bound for (3.1). Terrell (1990) showed that $\lambda_{AMISE} \leq \left(\frac{243 i_0(K^2)}{35 i_2^2(K) n} \right)^{\frac{1}{5}} \sigma$. By approximating $f(.)$ with standard normal PDF and plug-in σ by s, the over smoothing bandwidth obtained as $\lambda_{os} = \left(\frac{243}{70 n \sqrt{\pi}} \right)^{\frac{1}{5}} s$. The most common technique for obtaining bandwidth is the least-squares cross-validation (SCV) proposed by Rodemo (1982) and Bowman (1984). The least SCV is the value $\lambda = \hat{\lambda}_{SCV}$ that minimizes the cross-validation function

$$SCV(\lambda) = \int \hat{f}^2(x;\lambda) dx - \frac{2}{n} \sum_{i=1}^{n} \hat{f}_{(-i)}(X_i;\lambda), \tag{3.2}$$

where $\hat{f}^2(x;\lambda)$ is the kernel density estimate of $f(x)$ and $\hat{f}_{(-i)}(X_i;\lambda)$ is the kernel density estimate based on the sample with X_i deleted, which is called leave-one-out density estimator. The above bandwidths are based on a SRS.

Now, suppose that $\left\{ X_{[h]i}, h = 1, ..., H, i = 1, ..., t_h \right\}$ is an unbalanced RSS of size $n = t_1 + t_2 + \cdots + t_H$. Barabesi and Fattorini (2002) obtained optimal bandwidth λ by minimizing MISE given in (2.7) with respect to λ which yields to

$$\lambda_{URSS} = \left(\frac{\sum_{h=1}^{H} \frac{1}{t_h H^2} i_0(K^2)}{i_2^2(K) i_0(f''^2)} \right)^{\frac{1}{5}}. \tag{3.3}$$

For balanced RSS we have $t_1 = t_2 = \cdots = t_H = t$ with $n = tH$. Therefore, in this case (3.3) reduces to (3.1), i.e., $\lambda_{RSS} = \lambda_{SRS}$.

3.2 Bayesian Bandwidth Estimation under RSS

Gangopadhyay and Cheung (2002) proposed the Bayes estimate of bandwidth λ under SRS. In this section, we use their method to find a Bayes estimate of bandwidth λ under balanced and unbalanced RSS. Let $X^*_{URSS} = \left\{X_{[h]i}, h = 1, ..., H, i = 1, ..., t_h\right\}$ be unbalanced RSS of size $n = t_1 + t_2 + \cdots + t_H$. Based on the discussion in Gangopadhyay and Cheung (2002), if $K(.)$ is the kernel and we consider $k_\lambda(y) = \frac{1}{\lambda}K(\frac{y}{\lambda})$ as a PDF, then λ can be considered as the scale parameter of this density and we could estimate it by Bayes method. We would like to estimate the PDF $f(.)$ at a local point x, therefore only the segment of $f(.)$ in an interval around the point x is important. Using (2.6), consider the truncated version of $f(x)$ given by the following convolution

$$\begin{aligned} f_\lambda(x) &= (f * k_\lambda)(x) = \int f(y)k_\lambda(x-y)dy \\ &= \frac{1}{H}\sum_{h=1}^{H}\int f_{[h]}(y)k_\lambda(x-y)dy \\ &= \frac{1}{H}\sum_{h=1}^{H} E_{f_{[h]}}\left[k_\lambda(x-X_{[h]})\right]. \end{aligned} \tag{3.4}$$

For $f_\lambda(x)$, λ appears as a scale parameter, so we consider the prior $\pi(\lambda)$ for λ. Then, the posterior density of λ given $X = x$ is given by

$$\pi(\lambda|x) = \frac{f_\lambda(x)\pi(\lambda)}{\int f_\lambda(x)\pi(\lambda)d\lambda}. \tag{3.5}$$

The posterior density (3.5) depends on unknown $f_\lambda(x)$, so it can not be computed. Therefore we estimate $f_\lambda(x)$ by $\hat{f_\lambda}(x)$ as follows. Based on the unbalanced sample $\left\{X_{[h]i}, h = 1, ..., H, i = 1, ..., t_h\right\}$, a reasonable estimate of $E_{f_{[h]}}\left[k_\lambda(x - X_{[h]}\right]$ is given by

$$\begin{aligned} \hat{E}_{f_{[h]}}\left[k_\lambda(x-X_{[h]})\right] &= \frac{1}{t_h}\sum_{i=1}^{t_h} k_\lambda\left(x - X_{[h]i}\right) \\ &= \frac{1}{t_h\lambda}\sum_{i=1}^{t_h} K\left(\frac{x-X_{[h]i}}{\lambda}\right). \end{aligned} \tag{3.6}$$

By replacing (3.6) in (3.4), we obtain the estimate of $f_\lambda(x)$ as follows

$$\begin{aligned} \hat{f_\lambda}(x) &= \frac{1}{H}\sum_{h=1}^{H}\hat{E}_{f_{[h]}}\left[k_\lambda(x-X_{[h]})\right] \\ &= \frac{1}{H\lambda}\sum_{h=1}^{H}\frac{1}{t_h}\sum_{i=1}^{t}\left(\frac{x-X_{[h]i}}{\lambda}\right) = \hat{f}_{URSS}(x). \end{aligned} \tag{3.7}$$

By replacing (3.7) in (3.5), the estimated posterior density is given by

$$\hat{\pi}(\lambda|x, X^*_{URSS}) = \frac{\hat{f}_{URSS}(x)\pi(\lambda)}{\int \hat{f}_{URSS}(x)\pi(\lambda)d\lambda}. \tag{3.8}$$

Under the squared error loss (SEL) function, the Bayes estimator of the bandwidth λ is given by

$$\lambda^*_{B,URSS} = \lambda^*(x, X^*_{URSS}) = \int \lambda\hat{\pi}(\lambda|x, X^*_{URSS})d\lambda. \tag{3.9}$$

In some situations, the posterior density (3.8) and the Bayes estimator (3.9) can be obtained in a closed form.

Theorem 3.1. *Suppose the kernel function is the standard normal PDF and λ has the following conjugate prior*

$$\pi(\lambda) = \frac{2}{\Gamma(\alpha)\beta^\alpha\lambda^{2\alpha+1}}e^{-\frac{1}{\beta\lambda^2}},\ \lambda > 0. \tag{3.10}$$

Then,

a) Under the SEL function the Bayes estimator of λ is

$$\lambda^*(x, X^*_{URSS}) = \frac{\Gamma(\alpha)\sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{1}{t_h}\beta^{*\alpha}_{hi}}{\Gamma(\alpha+\frac{1}{2})\sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{1}{t_h}\beta^{*(\alpha+\frac{1}{2})}_{hi}},$$

where $\beta^*_{hi} = \left(\frac{(x-X_{[h]i})^2}{2} + \frac{1}{\beta}\right)^{-1}$.

b) Under the weighted SEL function $L(\lambda,\delta) = (\frac{\delta-\lambda}{\lambda})^2$, the Bayes estimator of λ is

$$\lambda^*(x, X^*_{URSS}) = \frac{\Gamma(\alpha+1)\sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{1}{t_h}\beta^{*(\alpha+1)}_{i}}{\Gamma(\alpha+\frac{3}{2})\sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{1}{t_h}\beta^{*(\alpha+\frac{3}{2})}_{i}}.$$

c) Under the entropy loss function $L(\lambda,\delta) = \frac{\delta}{\lambda} - \ln(\frac{\delta}{\lambda}) - 1$, the Bayes estimator of λ is

$$\lambda^*(x, X^*_{URSS}) = \frac{\Gamma(\alpha+\frac{1}{2})\sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{1}{t_h}\beta^{*(\alpha+\frac{1}{2})}_{i}}{\Gamma(\alpha+1)\sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{1}{t_h}\beta^{*(\alpha+1)}_{i}}.$$

Remark 1. Note that the kernel, that we used in Theorem 3.1 is the normal kernel, i.e., $k_\lambda(y) = \frac{1}{\lambda}K(\frac{y}{\lambda}) = \frac{1}{\sqrt{\pi\lambda^2}}e^{-\frac{1}{2\lambda^2}y^2}$. In this case, the conjugate prior distribution for $\tau = \lambda^2$ is the inverted gamma prior with parameters α, β. Therefore $\lambda = \sqrt{\lambda^2} = \sqrt{\tau}$ has the density $\pi(\lambda)$ given by (3.10)

Remark 2. According to the above discussion, λ can be considered as the scale parameter. Therefore, scale invariant loss functions such as entropy or weighted SEL are suitable for estimating λ.

Proof of Theorem 3.1. Suppose $X^*_{URSS} = \left\{X_{[h]i}, h = 1, ..., H, i = 1, ..., t_h\right\}$, then

$$\hat{\pi}(\lambda|x, X^*_{URSS}) = \frac{\hat{f}_{URSS}(x)\pi(\lambda)}{\int \hat{f}_{URSS}(x)\pi(\lambda)d\lambda}$$

$$= \frac{\frac{1}{H\lambda}\sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{1}{t_h\sqrt{2\pi}}e^{-\frac{(x-X_{[h]i})^2}{2\lambda^2}}\frac{2}{\Gamma(\alpha)\beta^{\alpha}\lambda^{2\alpha+1}}e^{-\frac{1}{\beta\lambda^2}}}{\int_0^{\infty}\frac{1}{H\lambda}\sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{1}{t_h\sqrt{2\pi}}e^{-\frac{(x-X_{[h]i})^2}{2\lambda^2}}\frac{2}{\Gamma(\alpha)\beta^{\alpha}\lambda^{2\alpha+1}}e^{-\frac{1}{\beta\lambda^2}}d\lambda}$$

$$= \frac{\sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{1}{t_h}e^{-\frac{1}{\lambda^2}(\frac{(x-X_{[h]i})^2}{2}+\frac{1}{\beta})}\frac{1}{\lambda^{2\alpha+2}}}{\int_0^{\infty}\sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{1}{t_h}e^{-\frac{1}{\lambda^2}(\frac{(x-X_{[h]i})^2}{2}+\frac{1}{\beta})}\frac{1}{\lambda^{2\alpha+2}}d\lambda}$$

$$= \frac{\sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{1}{t_h}e^{-\frac{1}{\lambda^2\beta^*_{hi}}}\frac{1}{\lambda^{2\alpha+2}}}{\int_0^{\infty}\sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{1}{t_h}e^{-\frac{1}{h^2\beta^*_{hi}}}\frac{1}{\lambda^{2\alpha+2}}d\lambda}$$

$$= \frac{\sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{1}{t_h}e^{-\frac{1}{\lambda^2\beta^*_{hi}}}\frac{1}{\lambda^{2(\alpha+\frac{1}{2})+1}}}{\sum_{h=1}^{H}\sum_{i=1}^{t_h}\int_0^{\infty}\frac{1}{t_h}e^{-\frac{1}{\lambda^2\beta^*_{hi}}}\frac{1}{\lambda^{2(\alpha+\frac{1}{2})+1}}d\lambda}.$$

The denominator of the above fraction can be obtained as follows

$$A_{hi} = \sum_{h=1}^{H}\sum_{i=1}^{t_h}\int_0^{\infty}\frac{1}{t_h}e^{-\frac{1}{\lambda^2\beta^*_{hi}}}\frac{1}{\lambda^{2(\alpha+\frac{1}{2})+1}}d\lambda$$

$$= \sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{\Gamma(\alpha+\frac{1}{2})}{2t_h}\beta_{hi}^{*(\alpha+\frac{1}{2})}\underbrace{\int_0^{\infty}\frac{2e^{-\frac{1}{\lambda^2\beta^*_{hi}}}}{\Gamma(\alpha+\frac{1}{2})\beta_{hi}^{*(\alpha+\frac{1}{2})}\lambda^{2(\alpha+\frac{1}{2})+1}}d\lambda}_{1}$$

$$= \sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{\Gamma(\alpha+\frac{1}{2})}{2t_h}\beta_{hi}^{*(\alpha+\frac{1}{2})}.$$

Therefore,

$$\hat{\pi}(\lambda|x, X^*_{URSS}) = \frac{\sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{1}{t_h}e^{-\frac{1}{\lambda^2\beta^*_{hi}}}\frac{1}{\lambda^{2(\alpha+\frac{1}{2})+1}}}{\sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{\Gamma(\alpha+\frac{1}{2})}{2t_h}\beta_{hi}^{*(\alpha+\frac{1}{2})}}.$$

a) Under the SEL function, the Bayesian bandwidth estimator is given by

$$\lambda^*(x, X^*_{URSS}) = E(\lambda|x, X^*_{URSS}) = \frac{\int_0^\infty \sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{1}{t_h}e^{-\frac{1}{\lambda^2\beta^*_{ri}}}\frac{1}{\lambda^{2\alpha+1}}d\lambda}{\sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{\Gamma(\alpha+\frac{1}{2})}{2t_h}\beta_{hi}^{*(\alpha+\frac{1}{2})}}$$

$$= \frac{\sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{1}{t_h}\int_0^\infty \frac{1}{\lambda^{2\alpha+1}}e^{-\frac{1}{\lambda^2\beta^*_{hi}}}dh}{\sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{\Gamma(\alpha+\frac{1}{2})}{2t_h}\beta_{hi}^{*(\alpha+\frac{1}{2})}}$$

$$= \frac{\sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{\Gamma(\alpha)\beta^{*\alpha}_{hi}}{2t_h}\overbrace{\int_0^\infty \frac{2e^{-\frac{1}{\lambda^2\beta^*_{hi}}}}{\Gamma(\alpha)\beta^{*\alpha}_{hi}\lambda^{2\alpha+1}}d\lambda}^{1}}{\sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{\Gamma(\alpha+\frac{1}{2})}{2t_h}\beta_{hi}^{*(\alpha+\frac{1}{2})}}$$

$$= \frac{\Gamma(\alpha)}{\Gamma(\alpha+\frac{1}{2})}\frac{\sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{1}{t_h}\beta^{*\alpha}_{hi}}{\sum_{h=1}^{H}\sum_{i=1}^{t_h}\frac{1}{t_h}\beta_{hi}^{*(\alpha+\frac{1}{2})}},$$

where $\beta^*_{hi} = \left(\frac{(x-X_{[h]i})^2}{2} + \frac{1}{\beta}\right)^{-1}$.

b,c) The proof is similar to the proof of part (a).

Remark 3. The Bayes estimate of bandwidth λ under balanced RSS can be obtained from Theorem 3.1 by replacing t_h by t. Therefore,

a) Under the SEL function the Bayes estimator of λ is

$$\lambda^*(x, X^*_{URSS}) = \frac{\Gamma(\alpha)\sum_{h=1}^{H}\sum_{i=1}^{t}\beta^{*\alpha}_{hi}}{\Gamma(\alpha+\frac{1}{2})\sum_{h=1}^{H}\sum_{i=1}^{t}\beta_{hi}^{*(\alpha+\frac{1}{2})}}.$$

b) Under the weighted SEL function $L(\lambda,\delta) = (\frac{\delta-\lambda}{\lambda})^2$, the Bayes estimator of λ is

$$\lambda^*(x, X^*_{URSS}) = \frac{\Gamma(\alpha+1)\sum_{h=1}^{H}\sum_{i=1}^{t}\beta_i^{*(\alpha+1)}}{\Gamma(\alpha+\frac{3}{2})\sum_{h=1}^{H}\sum_{i=1}^{t}\beta_i^{*(\alpha+\frac{3}{2})}}.$$

c) Under the entropy loss function $L(\lambda,\delta) = \frac{\delta}{\lambda} - \ln(\frac{\delta}{\lambda}) - 1$, the Bayes estimator of λ is

$$\lambda^*(x, X^*_{URSS}) = \frac{\Gamma(\alpha+\frac{1}{2})\sum_{h=1}^{H}\sum_{i=1}^{t}\beta_i^{*(\alpha+\frac{1}{2})}}{\Gamma(\alpha+1)\sum_{h=1}^{H}\sum_{i=1}^{t}\beta_i^{*(\alpha+1)}}.$$

3.3 Optimal Selection of Parameters α and β

In Section 3.2, we use the prior (3.10) for the normal kernel to obtain Bayesian bandwidth $\lambda^* = \lambda^*_{\alpha,\beta}$ which depends on the two parameters α, β of prior distribution. The implementation and performance of Bayes approach completely depend on choice of these parameters. In this section we use the crossed-validation criterion which is similar to the criterion that used by Chaubey et al. (2015) to obtain optimum values of α and β.

Since $\lambda^* = \lambda^*_{\alpha,\beta}$ depends on α, β, we select the best $\lambda^*_{\alpha,\beta}$ through a data-driven approach involving parameters α and β. A widely used measure of accuracy for assessing the fit between a density estimator $\hat{f}(x)$ and the true density $f(x)$ is

$$ISE(\hat{f}, f) = \int_{-\infty}^{\infty} (\hat{f}(x) - f(x))^2 dx = \int_{-\infty}^{\infty} \hat{f}^2(x)dx - 2\int_{-\infty}^{\infty} \hat{f}(x)f(x)dx + \int_{-\infty}^{\infty} f^2(x)dx. \tag{3.11}$$

The goal in bandwidth selection is to minimize (3.11) to achieve the optimal bandwidth size. However, since the true density $f(x)$ remains unknown, (3.11) becomes impractical as an objective function for optimization. To overcome this challenge, we make adjustments. We omit the final constant term in (3.11) and replace the second term with the leave-one-out estimator. This modification gives rise to an objective function suitable for practical optimization, approximating (3.11), and known as the cross-validation (SCV) criterion.

It is important to note that in the Bayesian method, the local bandwidth choice λ depends on x. Once the local bandwidth λ is plunged into (3.2), the SCV becomes a function of parameters α and β. In this scenario, the optimal values for α and β are obtained by minimizing SCV function (3.2), which guarantees the optimal selection of the Bayesian bandwidth under criterion (3.11).

4 Simulation Study

In this section, we conduct simulation studies to compare the MISE of PDF estimators under RSS and SRS for different bandwidths and also to compare Bayesian with classical bandwidth. In these simulation studies, we generate balanced and unbalanced RSS and SRS of different sizes n from $lognormal(0, 1)$ and $t(3)$ distributions to compute the MISE of the PDF estimators. If $\hat{f}(x)$ is the estimator of the PDF $f(x)$ at point x, then

$$MISE(\hat{f}) = E\left(\int \left[\hat{f}(x) - f(x)\right]^2 dx\right) = E(ISE(\hat{f})).$$

To calculate $ISE(\hat{f})$, we use the grid approximation. The grid values are denoted by $x_1, ..., x_k$ which are the 5th to 95th percentiles of each distribution. From these grid values we estimate $\widehat{ISE}(\hat{f})$ by

$$\widehat{ISE}(\hat{f}) = \frac{1}{k}\sum_{i=1}^{k}\left[\hat{f}(x_i) - f(x_i)\right]^2.$$

In each case, we repeat the procedure $N = 5000$ times and calculate the estimated MISE of $\hat{f}$ as

$$\widehat{MISE}(\hat{f}) = \frac{1}{N}\sum_{j=1}^{N}\widehat{ISE}_j(\hat{f}), \tag{4.1}$$

where $\widehat{ISE}_j(\hat{f})$ is the value of $\widehat{ISE}$ in j-th replication and $\hat{f}$ is one of the kernel density estimates under SRS, RSS, or URSS sample and with one of the bandwidths. Then, we obtain the ratio

$$R_{\hat{f},\hat{f}_{SRS}} = \frac{\widehat{MISE}(\hat{f}_{SRS})}{\widehat{MISE}(\hat{f})},$$

where $\hat{f}_{SRS}$ is the SRS estimate of f with NS, OS, and CSV bandwidths, and $\hat{f}$ is the RSS estimate of f with NS, OS, CSV and Bayesian bandwidths. To generate RSS with imperfect ranking, we use the ranking procedure of Dell and Clutter (1972). In this procedure an auxiliary variable Y (ranking variable) is made from the study variable X by the following model

$$Y = \rho(\frac{X-\mu}{\sigma}) + \sqrt{1-\rho^2}Z\,.$$

where μ and σ are the mean and standard deviation of variable X, ρ is the correlation coefficient of X and Y, and Z is a standard normal random variable that is independent of X.

First we use the balanced RSS and compute the $\frac{\widehat{MISE}(\hat{f}_{SRS})}{\widehat{MISE}(\hat{f}_{RSS})}$ and $\frac{\widehat{MISE}(\hat{f}_{SRS})}{\widehat{MISE}(\hat{f}_{RSS.Bayes})}$ for the values $t = 5$, $H = 2, 4, 6$, $\alpha = 3$, and for a range of β values, where $\hat{f}_{SRS}$ and $\hat{f}_{RSS}$ obtain under NS, OS, and CSV bandwidths and $\hat{f}_{RSS.Bayes}$ obtain under Bayesian bandwidth. Figures 1 and 2 show the results. In this figures, we observe that $\widehat{MISE}(\hat{f}_{RSS}) < \widehat{MISE}(\hat{f}_{SRS})$

for the same classical bandwidths. Figure 1 shows that when the underlying population has Lognormal distribution, for moderate values of β ($\beta \in (0.8, 4)$) the $\hat{f}_{RSS}(x)$ with Bayesian bandwidth works better than the other bandwidths. Also, Figure 2 shows that under Student's t(3) distribution, for small values of β ($\beta \in (0, 0.8)$) the $\hat{f}_{RSS}(x)$ with Bayesian bandwidth works better than the other bandwidths.

Furthermore, we use unbalanced RSS and compute the $\frac{\widehat{MISE}(\hat{f}_{SRS})}{\widehat{MISE}(\hat{f}_{URSS})}$ and $\frac{\widehat{MISE}(\hat{f}_{SRS})}{\widehat{MISE}(\hat{f}_{URSS.Bayes})}$ for the two sets of values $t_1 = 4, t_2 = 2, t_3 = 2, t_4 = 2$ and $t_1 = 4, t_2 = 3, t_3 = 3, t_4 = 0$ with $H = 4$, $\alpha = 3$ and for a range of β values, where $\hat{f}_{SRS}$ and $\hat{f}_{RSS}$ obtain under NS, OS, and CSV bandwidths and $\hat{f}_{RSS.Bayes}$ obtain under Bayesian bandwidth. Figure 3 shows that when the underlying population has Lognormal distribution, for moderate values of β ($\beta \in (0.8, 6)$) the $\hat{f}_{URSS}(x)$ with Bayesian bandwidth performs significantly better than the other bandwidths.

We calculate the ratio of *MISE*s for the values of $\alpha = 3.5$ and $\alpha = 4$, and the results closely resemble those shown in Figures 1-3 and we have not included them to save space.

From Figures 1, 2 and 3, we can observe that when the population distribution is asymmetric, the Bayesian bandwidth performs better than the symmetric distribution. Also, when the unbalanced rank set sampling method is used, the situation is much better for the case where some $t_i = 0$.

5 Real Data Application

The household income and expenses survey has been implemented by Statistical Center of Iran since 1963 in rural areas and since 1969 in urban areas. Since 1975, in addition to household income, expense information was also collected. The general goal of the household income and expenses survey is to estimate the average income and expenses of an urban household and a rural household at the level of country and province. The importance of this project lies in the possibility of examining the trend of consumption of goods and services, studying the interrelationships of economic characteristics of households, evaluating the effects of economic policies in the field of ensuring social justice and examining income distribution, the possibility of examining the poverty line of households, etc.

In the following, we will use the data of the income and expenses of urban households collected by Statistical Center of Iran in 2021 to estimate the density function of the total cost (thousand billion Rials) of the households. To avoid the effect of the large population size on the simulation results, in this study, we only use the information from Tehran province. The data can be accessed at https://ssis.sci.org.ir/1400-2-2, and the columns used include household weight, total household expenditure, and household food expenditure. In the household income and expenditure survey, the total cost is derived by summing the cost of all sub-sections. Since the total cost is closely related to food consumption expenses, it is easy and more efficient to use food consumption expenses as the auxiliary variable for ranking the data and then measure the total cost for the selected sample in RSS procedure. In this study, we use the food consumption expenses of the households as an auxiliary variable with a correlation coefficient 0.70 to the total cost, and the density function of the total expenses of the households is estimated by RSS method with Bayesian Bandwidth and classical bandwidth in 2021 for Tehran province. (Note that the households have a weight and the data is multiplied by the weight of the households.)

First, we fit a distribution on the total expenses data. According to Figure 4, the blue dot follows a distribution that is close to the data. It can be seen that the blue point is close to the three Gamma, Lognormal and Exponential distributions. In order to find the most suitable distribution, we draw the histogram and the empirical distribution function of the data along with the density and the distribution function of the above three distributions in Figure 5. It can be seen from Figure 5 that the Exponential distribution is not suitable. To see which one of the Gamma and Lognormal distribution have better fit, we use the AIC, BIC and KS-TEST criteria. Note that the lower the value of KS, AIC, and BIC indicates that the distribution is better fitted. From Table 1, the results of the above criteria show that the Lognormal distribution has better fit to data.

Table 1: AIC, BIC and KS-TEST criteria

criteria	Exponential	Lognormal	Gamma
AIC	7164.7	6746.5	6907.2
BIC	7170.1	6757.1	6917.8
KS-TEST	0.14	0.027	0.075

In the following, assuming that the data comes from a log normal distribution with a mean of 1.137 and a standard deviation of 0.77, we estimate the PDF using SRS and RSS with SCV and Bayesian bandwidth. Note that based on Remark 2, the weighted loss function is used to estimate the Bayesian bandwidth. To obtain the PDF estimate of the total cost using the SRS and RSS methods with different bandwidths, we first produce n samples from each method and then calculate $\hat{f}_{RSS_i}(t)$ and $\hat{f}_{SRS_i}(t)$ with different bandwidths. We repeat this process N = 50 times. Finally, we calculate $\hat{f}_{RSS}(t) = \frac{1}{N}\sum_{i=1}^{N} \hat{f}_{RSS_i}(t)$ and $\hat{f}_{SRS}(t) = \frac{1}{N}\sum_{i=1}^{N} \hat{f}_{SRS_i}(t)$ for different bandwidths. The results are shown in Figure 6.

Figure 6 shows the histogram of the actual data, along with the PDF estimates using RSS and SRS methods with Bayesian bandwidth and SCV bandwidth, where $t = 25$, $H = 2$, $\rho = 0.70$, $n = tH$. When estimating the PDF $f(x)$ using Bayesian bandwidth, the values of hyperparameters α and β at any point x are determined using the procedure introduced in Section 3.3. In this figure, the PDF estimates by RSS method with Bayesian bandwidth and λ_{SCV} bandwidth, and SRS with λ_{SCV} bandwidth have been shown by $\hat{f}^*_{RSS}(t)$, $\hat{f}_{RSS}(t)$, and $\hat{f}_{SRS}(t)$, respectively. From Figure 6, we observe that the PDF estimate using RSS method with Bayesian bandwidth has better fit to this data than the other two methods. This is consistent with the simulation results obtained in Figure 1 for Lognormal distribution.

6 Conclusion and Discussion

In this paper, we used the Bayesian bandwidth to estimate a PDF under RSS method and compare it with SRS counterpart. We used simulation studies to compare these methods. The simulation studies showed that when the population distribution is Student's t and the hyperparameter β takes the values specified in Figure 2, the PDF estimate with Bayesian bandwidth performs much better than the others. Bayesian bandwidth estimate also performs much better when the underlying population is Lognormal for the values of β that are specified in Figure 1. The Bayesian bandwidth estimate performs significantly better when the underlying population is Lognormal, and we utilize the URSS method for bandwidth estimation. The results are applied in estimation of PDF of total expenses and it is observed that the RSS method using Bayesian bandwidth has a better fit to this data than the other mentioned methods.

When the ranking is imperfect and the sampling design is URSS, Ozturk (2007) demonstrated that the statistical inference based on CDFs $F_{[i]}(x), i = 1, \cdots, H$ of the judgment class distributions of a ranked set sample generated from a distribution F, can be improved by using the estimators of $F_{[i]}(x), i = 1, \cdots, H$. This improvement involves ensuring that these estimators adhere to a stochastic order restriction $F_{[1]}(x) \geq F_{[2]}(x) \geq \cdots \geq F_{[H]}(x)$. Another approach for estimating the PDF $f(x)$ involves initially estimating the CDF $F(x)$ through established methods in the literature, followed by estimating the PDF using the relationship between CDF and PDF. Incorporating the stochastic order restriction in estimating the CDF enables the derivation of an improved estimator for the PDF. We are actively considering the integration of this approach into our forthcoming research endeavors.

References

Barabesi. L., and Fattorini, L. (2002), Kernel Estimators of Probability Density Functions by Ranked-Set Sampling. *Commun Statist Theory-Methods*, **31**, 597-610.

Bowman, A. (1984), An Alternative Method of Cross Validation for the Smoothing of Density Estimate. *Biometrika*, **71**, 353-360.

Chen, Z. (1999), Density Estimation Using Ranked Set Sampling Data. *Enviro Ecolog Statist*, **6**, 135-146.

Chaubey, Y. P., Li, J., and Dewan, I. (2015), Smoothing Parameter Selection for Nonparametric Density Estimation for Length-biased Data: A Bayesian Perspective. *Indian Statistical Institute, Delhi Centre.*

Dell, T. R., and Clutter, J. L. (1972), Ranked Set Sampling Theory with Order Statistics Background. *Biometrics*, **28**, 545-555.

Gangopadhyay, A. K., and Cheung, K. N. (2002), Bayesian Approach to the Choice of Smoothing Parameter in Kernel Density Estimation. *Nonparametric Statistics*, **14**, 655-664.

Kulasekera, K. B., and Padgett, W. J. (2006), Bayes Bandwidth Selection in Kernel Density Estimation with Censored Data. *Nonparametric Statistics*, **18**, 129-143.

Lim, J., Chen, M., Park, S., Wang, X., and Stokes, L. (2014), Kernel Density Estimator from Ranked Set Samples. *Commun Statist Theory-Methods*, **43**, 2156-2168.

Minoiu, C., and Reddy, S. G. (2014), Kernel Density Estimation on Grouped Data: the Case of Poverty Assessment. *The Journal of Economic Inequality*, **12**, 163-189.

Ozturk, O. (2007), Statistical Inference Under a Stochastic Ordering Constraint in Ranked Set Sampling. *Journal of Nonparametric Statistics,* **19**, 131-144.

Parzen, E. (1962), On Estimation of a Probability Density Function and Mode. *Ann Math Statist,* 1065-1076.

Rodemo, M. (1982), Empirical Choice of Histograms and Kernel Density Estimation *Scand J Statist*, **9**, 65-78.

Rosenblatt, M. (1956), Remarks on Some Nonparametric Estimates of a Density Function. *Ann Math Statist*, **27**, 832-837.

Samawi, H., Rochani, H., Yin, J., Linder, D., and Vogel, R. (2018), Notes on Kernel Density Based Mode Estimation Using More Efficient Sampling Designs. *Comput Statis*, **33**, 1071-1090.

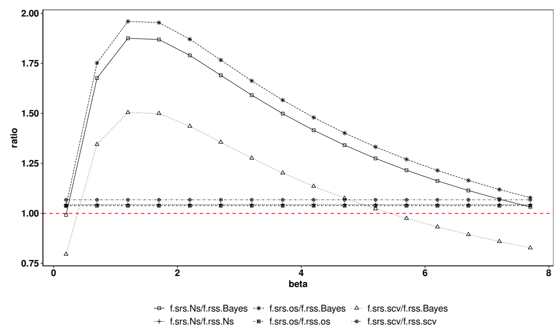
Silverman, B. W. (1986), Density Estimation for Statistics and Data Analysis *London: Chapman and Hall*

Terrell, G. R. (1990), The Maximal Smoothing Principle in Density Estimation. *Journal of the American Statistical Association*, **85**, 470-477.

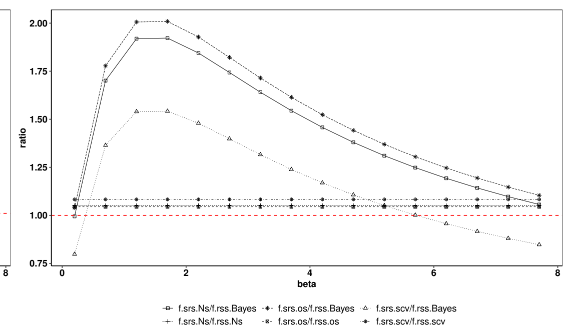
Wand, M. P., and Jones, M. C. (1995), Kernel Smoothing. *London: Chapman and Hall*

Wasserman, L. (2006), *All of Nonparametric Statistics.* New York: Springer

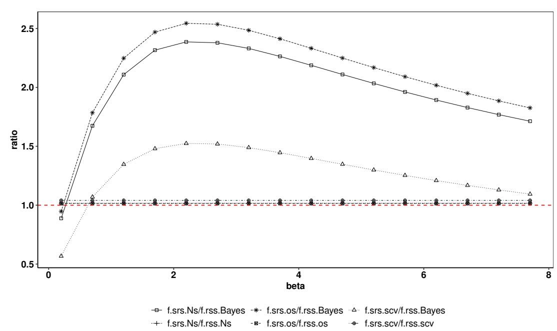
Appendix



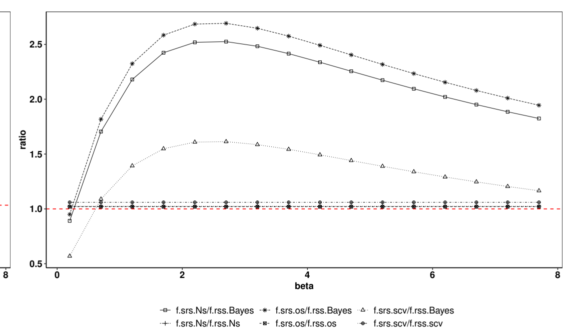
(a) $t = 5$, $H = 2, \rho = 0.5$



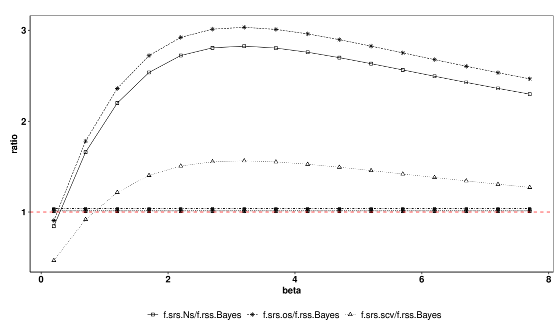
(b) $t = 5$, $H = 2, \rho = 0.9$



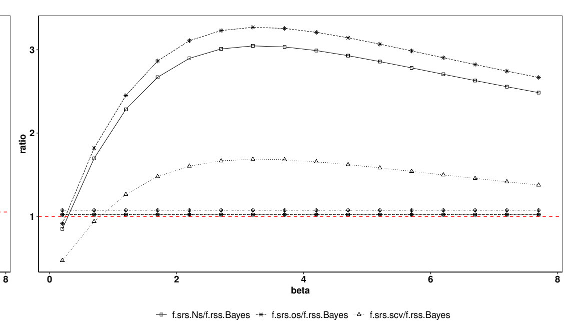
(c) $t = 5$, $H = 4, \rho = 0.5$



(d) $t = 5$, $H = 4, \rho = 0.9$

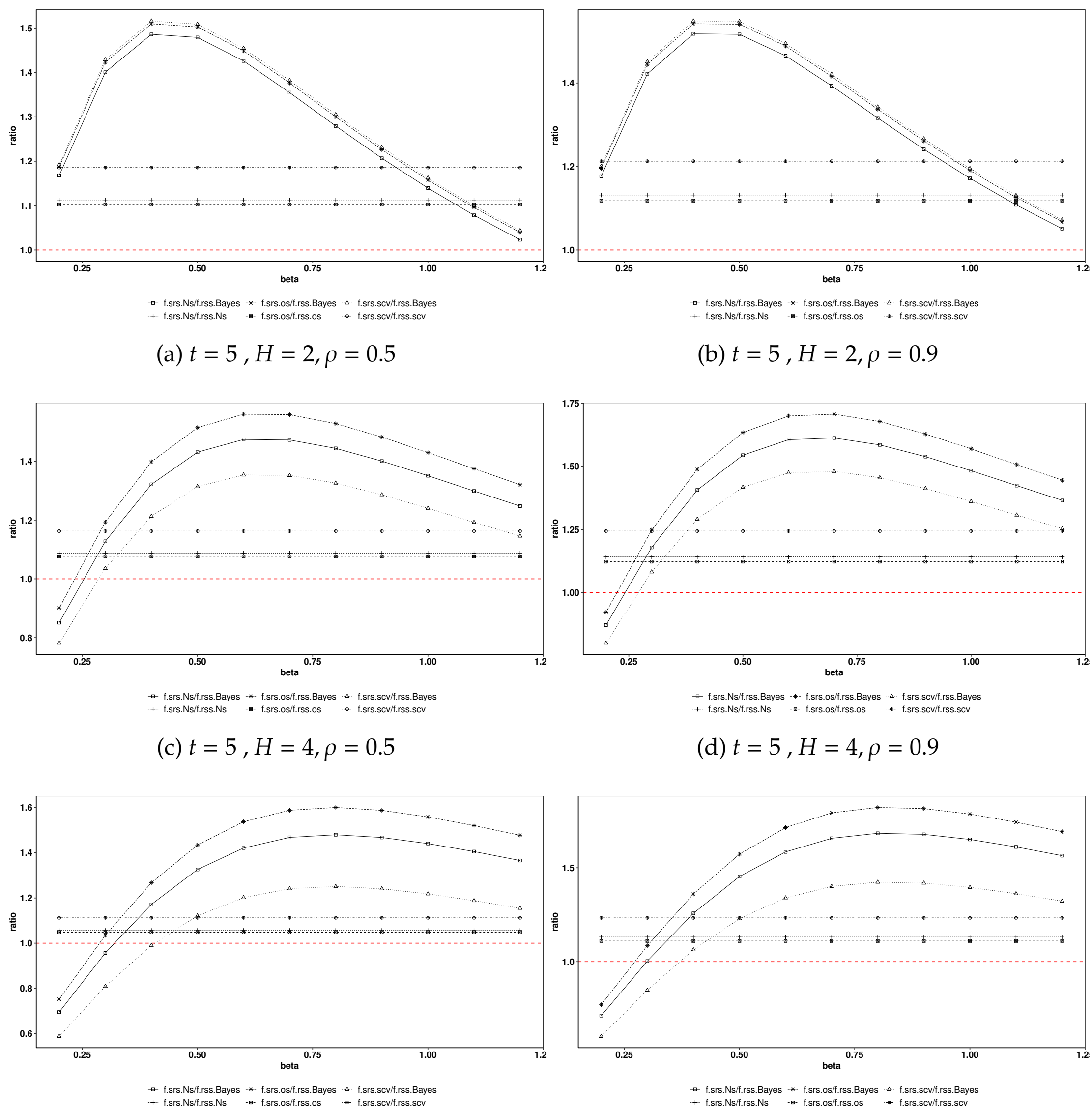


(e) $t = 5$, $H = 6, \rho = 0.5$



(f) $t = 5$, $H = 6, \rho = 0.9$

Figure 1: $\frac{\widehat{MISE}(\hat{f}_{SRS})}{\widehat{MISE}(\hat{f}_{RSS})}$ and $\frac{\widehat{MISE}(\hat{f}_{SRS})}{\widehat{MISE}(\hat{f}_{RSS.Bayes})}$ under $Lognormal(0, 1)$ distribution for $\alpha = 3$, $t = 5$, $H = 2, 4, 6$ and $\rho = 0.5, 0.90$, where $\hat{f}_{SRS}$ and $\hat{f}_{RSS}$ obtain under NS, OS, and CSV bandwidths and $\hat{f}_{RSS.Bayes}$ obtain under Bayesian bandwidth.



(a) $t = 5$, $H = 2, \rho = 0.5$

(b) $t = 5$, $H = 2, \rho = 0.9$

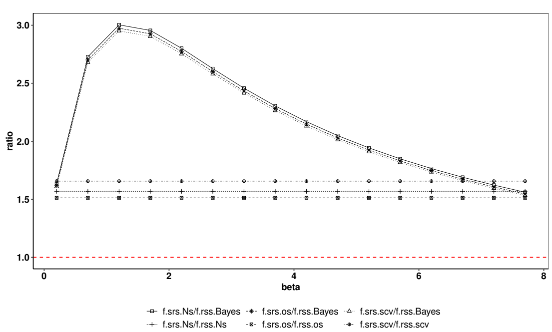
(c) $t = 5$, $H = 4, \rho = 0.5$

(d) $t = 5$, $H = 4, \rho = 0.9$

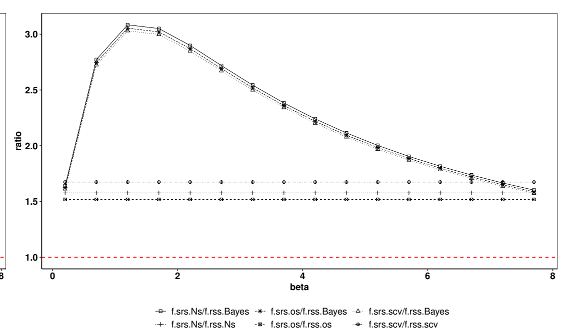
(e) $t = 5$, $H = 6, \rho = 0.5$

(f) $t = 5$, $H = 6, \rho = 0.9$

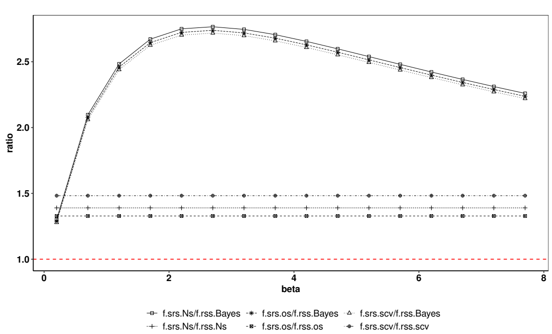
Figure 2: $\frac{\widehat{MISE}(\hat{f}_{SRS})}{\widehat{MISE}(\hat{f}_{RSS})}$ and $\frac{\widehat{MISE}(\hat{f}_{SRS})}{\widehat{MISE}(\hat{f}_{RSS.Bayes})}$ under $t(3)$ distribution for $\alpha = 3$, $t = 5$, $H = 2, 4, 6$ and $\rho = 0.5, 0.90$, where $\hat{f}_{SRS}$ and $\hat{f}_{RSS}$ obtain under NS, OS, and CSV bandwidths and $\hat{f}_{RSS.Bayes}$ obtain under Bayesian bandwidth.



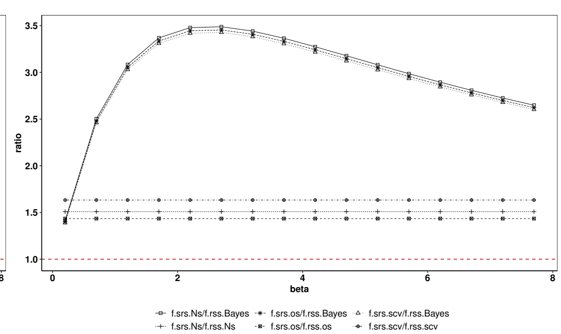
(a) $t_1 = 4, t_2 = 2, t_3 = 2, t_4 = 2$, $\rho = 0.5$



(b) $t_1 = 4, t_2 = 2, t_3 = 2, t_4 = 2$, $\rho = 0.9$



(c) $t_1 = 4, t_2 = 3, t_3 = 3, t_4 = 0$, $\rho = 0.5$



(d) $t_1 = 4, t_2 = 3, t_3 = 3, t_4 = 0$, $\rho = 0.9$

Figure 3: $\frac{\widehat{MISE}(\hat{f}_{SRS})}{\widehat{MISE}(\hat{f}_{URSS})}$ and $\frac{\widehat{MISE}(\hat{f}_{SRS})}{\widehat{MISE}(\hat{f}_{URSS.Bayes})}$ under $Lognormal(0,1)$ distribution for $\alpha = 3$, $t = 5$ and $\rho = 0.5, 0.90$, where $\hat{f}_{SRS}$ and $\hat{f}_{RSS}$ obtain under NS, OS, and CSV bandwidths and $\hat{f}_{RSS.Bayes}$ obtain under Bayesian bandwidth.

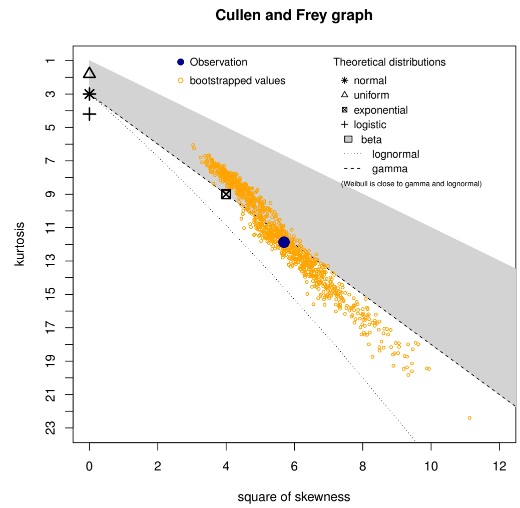


Figure 4: Goodness-of-fit plots to the total cost.

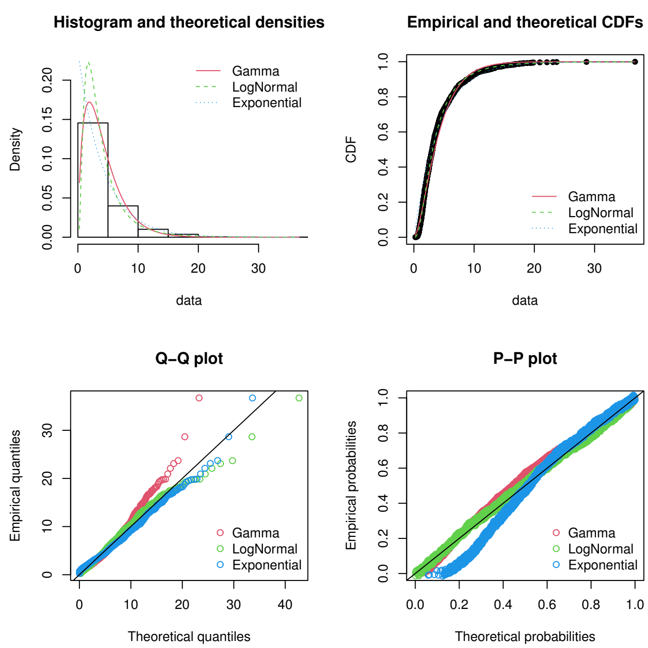


Figure 5: Goodness-of-fit plots of total expenses data.

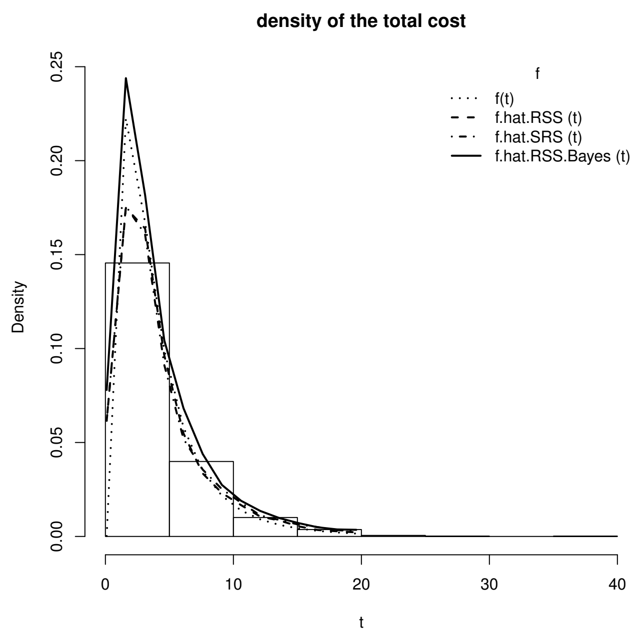


Figure 6: Estimates of the probability density function by $\hat{f}_{RSS}(x)$, $\hat{f}_{SRS}(x)$, and $\hat{f}^{*}_{RSS}(x)$ for $\rho = 0.70$, $t = 25$, $n = tH$, $H = 2$.