

Artículo de investigación

Bootstrap versus Jackknife: Confidence intervals, hypothesis testing, density estimation, and kernel regression

Laura Cruz¹, Jessica Blanco¹, Ramón Giraldo²✉

¹Estudiante Maestría en Estadística. Universidad Nacional de Colombia, Sede Bogotá.

²Universidad Nacional de Colombia, Sede Bogotá, Departamento de Estadística, Grupo de investigación en Estadística Espacial.

Recepción: 08-abril-2023 **Aceptado:** 15-abril-2024 **Publicado:** 23-jul-2024

Cómo citar: Cruz, L., Blanco, J., & Giraldo, R. (2024). Bootstrap versus Jackknife: Confidence intervals, hypothesis testing, density estimation, and kernel regression. *Ciencia En Desarrollo*, 15(2). <https://doi.org/10.19053/01217488.v15.n2.2024.15900>

Abstract

Bootstrap and Jackknife methods are compared in various statistical contexts. Initially, these are evaluated using estimates of coefficients of variation obtained from samples of different probability models (Normal, Gamma, Binomial, and Poisson) generated by Monte Carlo simulation. With the results, the bias and variance of the estimators are evaluated. The performance of the two inferential procedures considered in one-sample problems, density estimation, and kernel regression, is also studied. The results show that in the case of the Jackknife coefficient of variation, it has a lower bias but a higher standard error. Bootstrap is a more powerful estimator in this context. Both methodologies produce similar results regarding density estimation (histogram and kernel). In regression Kernel, it is observed that Jackknife allows obtaining estimates of the regression bandwidth closer to the classical ones than those found with Bootstrap. The corresponding confidence intervals with Jackknife are shorter than those established with Bootstrap.

Keywords: Bootstrap, confidence intervals, Jackknife, kernel regression, local regression, power of the test.

Resumen

Se comparan los métodos Bootstrap y Jackknife en varios contextos estadísticos. Inicialmente usando estimaciones de coeficientes de variación obtenidos a partir de muestras de varios modelos de probabilidad (Normal, Gama, Binomial y Poisson) generadas por simulación de Monte Carlo. Con los resultados se evalúa sesgo y varianza los estimadores. También se estudia el desempeño de los dos procedimientos inferenciales considerados en problemas de una muestra, estimación de la densidad y regresión kernel. Los resultados muestran que en el caso del coeficiente de variación Jackknife tiene menor sesgo pero mayor error estándar. Bootstrap es más potente en este contexto. En lo referente a la estimación de la densidad (histograma y Kernel) y la estimación del ancho de banda en la estimación de la función de regresión Jackknife produce estimaciones más cercanas a las clásicas que las halladas con Bootstrap. Los correspondientes intervalos de confianza con Jackknife son más cortos que los establecidos con Bootstrap.

Palabras Clave: Bootstrap, estimación kernel de la densidad, intervalos de confianza, Jackknife, potencia de la prueba, regresión kernel.

1 Introduction

Two popular resampling methods widely used in real data analysis are Bootstrap [6, 10] and Jackknife [24, 25, 31]. Both are nonparametric statistical methods. These computer-intensive techniques can be used to estimate bias and standard errors of non-traditional estimators, especially useful when the sampling distribution of an estimator is unknown or cannot be defined mathematically so that classical statistical analysis methods are not available. Samples from the observed data allow us to draw conclusions about the population of interest. Nowadays, these approaches are feasible because of the availability of high-speed computing. Confidence intervals based on Bootstrap and other resampling methods (for example, Jackknife) should be used whenever there is cause to doubt the assumptions of parametric confidence intervals. When the underlying distribution of some statistic of interest is unknown, these strategies can be beneficial.

Bootstrap and Jackknife methods are powerful techniques used in statistics to estimate the variability of a statistic or to assess the goodness of fit of a statistical model. While the Bootstrap resamples with replacement, the Jackknife method systematically leaves out one observation at a time. The choice of method depends on the problem, and both methods can be useful in different scenarios. Bootstrap and Jackknife have been used and compared in several statistical scenarios. Among others in linear regression [11, 33], quantile regression [15, 19], analysis of variance [7, 8], and generalized linear models [21]. Bootstrap and Jackknife can be used in kernel density estimation and kernel regression to estimate the variability of the density and regression functions and, consequently, define confidence intervals. In this work, we explore the applicability of these methods in the estimation of the bandwidth in both scenarios, kernel density, and kernel regression estimation. The analysis is carried out using simulated data in R [26].

The article is organized as follows: Initially, we show in Section 2 a review of the bias, standard error, and confidence intervals using Bootstrap and Jackknife. An illustration based on the coefficient of variation is also shown. In Sections 3 and 3.2, Bootstrap and Jackknife are compared in the context of hypothesis testing for one sample problems. Notably, using Monte Carlo simulations, the power of the tests based on these two strategies is estimated when testing hypotheses about the coefficient of variation are carried out. In Sections 4 and 5, we show a comparison of these methodologies in the context of kernel density estimation [5] and kernel regression [3].

2 Background: Bootstrap and Jackknife

Here we give an overview of Bootstrap and Jackknife methods. The estimations on bias and standard errors (consequently the respective confidence intervals) by both approaches are presented. Assume Y_1, \dots, Y_n a random sample of $Y \sim f(y, \Theta)$ with Θ a parameters vector that defines the probability model of interest (for example $\Theta = (\mu, \sigma)$ in the case of a normal distribution or $\Theta = (\alpha, \beta)$ for a Gamma distribution). Suppose we want to estimate θ a particular parameter (or a function of parameters) of the distribution. If the distribution of the estimator $\hat{\theta}$ is unknown, Bootstrap and Jackknife procedures (sections 2.1 and 2.2) can be used for obtaining a CI for θ .

2.1 Bias, standard error, and confidence intervals using Bootstrap

Based on the sample, we can obtain B size n samples with replacement (Table 1) denoted as Y_{bj}^* , with $b = 1, \dots, B$ and $j = 1, \dots, n$. At each case, the estimator $(\hat{\theta}_b^*)$ of the parameter of interest is calculated.

Using the Bootstrap samples and the estimations $\hat{\theta}_b^*, b = 1, \dots, B$ in

Table 1: Representation of the Bootstrap random samples: Asterisk indicates that a sample size n with replacement is obtained from Y_1, \dots, Y_n .

	j			
b	1	...	n	$\hat{\theta}$
1	Y_{11}^*	...	Y_{1n}^*	$\hat{\theta}_1^*$
2	Y_{21}^*	...	Y_{2n}^*	$\hat{\theta}_2^*$
...
B	Y_{B1}^*	...	Y_{Bn}^*	$\hat{\theta}_B^*$

Table 1, the Bootstrap estimator, its bias and variance are given by

$$\hat{\theta}_{boot} = \frac{1}{B} \sum_{b=1}^B \hat{\theta}_b^* \tag{1}$$

$$\begin{aligned} \text{bias}_{boot} &= (\hat{E}(\hat{\theta}) - \hat{\theta}) \\ &= \frac{1}{B} \sum_{b=1}^B \hat{\theta}_b^* - \hat{\theta} \end{aligned} \tag{2}$$

$$V(\hat{\theta})_{boot} = \frac{1}{B} \sum_{b=1}^B \left(\hat{\theta}_b^* - \frac{1}{B} \sum_{b=1}^B \hat{\theta}_b^* \right)^2 \tag{3}$$

Assuming normality a $100(1 - \alpha)\%$ CI to θ is given by

$$\left(\frac{1}{B} \sum_{b=1}^B \hat{\theta}_b^* \pm Z_{1-\frac{\alpha}{2}} \sqrt{V(\hat{\theta})_{boot}} \right)$$

In general the CI can be obtained as

$$\left(\hat{\theta}_{\frac{\alpha}{2}}^*, \hat{\theta}_{1-\frac{\alpha}{2}}^* \right), \tag{4}$$

with $\hat{\theta}_{\frac{\alpha}{2}}^*$ and $\hat{\theta}_{1-\frac{\alpha}{2}}^*$ percentiles obtained from $\hat{\theta}_b^*, b = 1, \dots, B$.

2.2 Bias, standard error, and confidence intervals using Jackknife

As in the previous section, assume that we have a sample Y_1, \dots, Y_n be a random sample of $Y \sim f(y, \theta)$, θ is the parameter of interest and $\hat{\theta} = g(Y_1, \dots, Y_n)$ its estimator. The bias $(E(\hat{\theta}) - \theta)$ and the variance $V(\hat{\theta})$ are unknown. Let $\hat{\theta}_{-i}$ the estimator obtained after deleting $Y_i, i = 1, \dots, n$. The Jackknife estimator of θ is defined as

$$\hat{\theta}_{jack} = \frac{1}{n} \sum_{i=1}^n \tilde{\theta}_i, \tag{5}$$

where $\tilde{\theta}_i = n\hat{\theta} - (n-1)\hat{\theta}_{-i}$. $\tilde{\theta}_i, i = 1, \dots, n$, are called Tukey's pseudovalues. Alternatively we have

$$\begin{aligned} \hat{\theta}_{jack} &= \frac{1}{n} \sum_{i=1}^n \tilde{\theta}_i \\ &= \frac{1}{n} \sum_{i=1}^n (n\hat{\theta} - (n-1)\hat{\theta}_{-i}) \\ &= \frac{1}{n} \sum_{i=1}^n n\hat{\theta} - (n-1) \frac{1}{n} \sum_{i=1}^n \hat{\theta}_{-i} \\ &= n\hat{\theta} - (n-1)\hat{\theta}_n, \end{aligned}$$

with $\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^n \hat{\theta}_{-i}$.

In order to estimate the bias of the Jackknife estimator, $E(\hat{\theta})$ and θ are replaced by $\hat{\theta}$ and $\hat{\theta}_{\text{jack}}$ respectively. Specifically

$$\begin{aligned} \text{bias}_{\text{jack}} &= \hat{\theta} - \hat{\theta}_{\text{jack}} \\ &= \hat{\theta} - (n\hat{\theta} - (n-1)\hat{\theta}_n) \\ &= \hat{\theta} - n\hat{\theta} + (n-1)\hat{\theta}_n \\ &= (n-1)\hat{\theta}_n - (n-1)\hat{\theta} \\ &= (n-1)(\hat{\theta}_n - \hat{\theta}). \end{aligned}$$

Let Y_1, \dots, Y_n be a random sample and \bar{Y} the sample mean. The variance of this statistic can be approximated using the sample variance as

$$V(\bar{Y}) = \frac{\sigma^2}{n} \rightarrow \hat{V}(\bar{Y}) = \frac{S^2}{n}$$

and consequently

$$\hat{V}(\bar{Y}) = \frac{1}{n(n-1)} \sum_{i=1}^n (Y_i - \bar{Y})^2. \quad (6)$$

Adapting the equation (6) to the pseudo-values $\hat{\theta}_1, \dots, \hat{\theta}_n$ we have

$$\hat{V}(\hat{\theta})_{\text{jack}} = \frac{1}{n(n-1)} \sum_{i=1}^n \left(\hat{\theta}_i - \frac{1}{n} \sum_{j=1}^n \hat{\theta}_j \right)^2. \quad (7)$$

It can be shown that [9]

$$T = \frac{\hat{\theta}_{\text{jack}} - \theta}{\sqrt{\hat{V}(\hat{\theta}_{\text{jack}})}} \sim t\text{-student}_{n-1}.$$

Then a $100(1 - \alpha)\%$ CI for θ is

$$\left(\hat{\theta}_{\text{jack}} \pm t_{n-1, 1-\frac{\alpha}{2}} \sqrt{\hat{V}(\hat{\theta}_{\text{jack}})} \right).$$

3 Inference on the coefficient of variation using Bootstrap and Jackknife

In this Section, we show, using Monte Carlo simulation, the behavior of Bootstrap and Jackknife in statistical inference (estimation and hypothesis testing) on the coefficient of variation.

3.1 Estimation of the coefficient of variation

Assume $Y \sim f(y, \Theta)$, $\mu = \mathbb{E}(Y)$, and $\sigma^2 = \mathbb{V}(Y)$, and we want to estimate the coefficient of variation $\theta = \text{CV} = \frac{\sigma}{\mu}$. Based on Monte Carlo simulation [4], we compare Bootstrap and Jackknife in terms of bias and standard errors of estimation. [33] conducted a similar study based on Normal data. We extend that work using simulations from four probability models (Normal, Gamma, Poisson, and Binomial) obtained using R [26]. In Table 2, we show the expressions of the CV of the four models considered.

To estimate the CV using Bootstrap, we first take a random sample size n with replacement from the original dataset. We then calculate the CV of this bootstrap sample. We repeat this process B times to obtain B bootstrap samples and CV estimates. The standard deviation of these B estimates can be used as an estimate of the standard error of the CV. The confidence interval for the CV can then be obtained using the standard error and the desired level of confidence. To estimate the CV using Jackknife, we create n subsamples by leaving out one observation from the original dataset

Table 2: Probability models considered to study the consistency and power of the tests on the coefficient of variation

Model	Parameters (Θ)	Mean	Variance	CV
Normal	$(\mu = 10, \sigma = 2)$	μ	σ^2	$\frac{\sigma}{\mu} = 0.200$
Gamma	$(\alpha = 2, \beta = 2)$	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$	$\frac{\sqrt{\alpha}}{\alpha} = 0.707$
Poisson	$\lambda = 5$	λ	λ	$\frac{\sqrt{\lambda}}{\lambda} = 0.447$
Binomial	$(n = 10, p = 0.4)$	np	$np(1-p)$	$\frac{\sqrt{np(1-p)}}{np} = 0.387$

Table 3: Summary of Bootstrap estimation of the coefficient of variation

General case	Coefficient of variation
$\hat{\theta}$	$\widehat{\text{CV}} = S/\bar{X}$
$\hat{\theta}_{\text{boot}} = \frac{1}{B} \sum_{b=1}^B \hat{\theta}_b^*$	$\widehat{\text{CV}}_{\text{boot}} = \frac{1}{B} \sum_{b=1}^B \widehat{\text{CV}}_b^*$
$\text{bias}_{\text{boot}} = \hat{\theta}_{\text{boot}} - \hat{\theta}$	$\text{bias}_{\text{boot}} = \widehat{\text{CV}}_{\text{boot}} - \widehat{\text{CV}}$
$V(\hat{\theta})_{\text{boot}} = \frac{1}{B} \sum_{b=1}^B \left(\hat{\theta}_b^* - \frac{1}{B} \sum_{b=1}^B \hat{\theta}_b^* \right)^2$	$V(\widehat{\text{CV}})_{\text{boot}} = \frac{1}{B} \sum_{b=1}^B \left(\widehat{\text{CV}}_b^* - \frac{1}{B} \sum_{b=1}^B \widehat{\text{CV}}_b^* \right)^2$

at a time. We then calculate the CV of each subsample. The mean of these n estimates is used as an estimate of the CV. To facilitate the interpretation of the results in Tables 3 and 4 are presented the expressions of Bootstrap and Jackknife estimation defined in Sections 2.1 and 2.2 corresponding to the coefficient of variation. We consider samples size $n = 5, 10, 15, \dots, 200$. The results are presented in Figure 1. Various aspects are remarkable in this Figure. The Jackknife bias is less than the Bootstrap one, and its performance increases with continuous distributions (Normal and Gamma). Bootstrap underestimates the CV in all cases considered. We note that the greater the sample size, the better the Bootstrap estimation (less bias). For n values close to 200, the estimations by these methodologies are very similar. The methods produce similar standard error estimations from relatively small sample sizes. The results in Figure 1 suggest that Jackknife can be a better option for estimating bias and standard deviation of the coefficient of variation, particularly when the sample size is small.

3.2 Hypothesis testing on the coefficient of variation

The power of a test using Bootstrap or Jackknife depends on the number of resamples or Jackknife samples used, as well as the characteristics of the original dataset. Generally, increasing the number of resamples or Jackknife samples will increase the power of the test, but at the cost of computational time. An essential factor that can affect the power of the test is the underlying distribution of the data. Bootstrap and Jackknife methods can perform well if the data are normally distributed. Here, we empirically compare (using simulated data) the power of tests based on Bootstrap and Jackknife. For this purpose, we simulate samples from the distributions in Table 2. Specifically, we test the hypothesis

Table 4: Summary of Jackknife estimation of the coefficient of variation

General case	Coefficient of variation
$\hat{\theta}$	$\widehat{\text{CV}} = \frac{S}{\bar{X}}$
$\hat{\theta}_{-i}$	$\widehat{\text{CV}}_{-i} = \frac{S_{-i}}{\bar{X}_{-i}}$
$\hat{\theta}_n = \frac{\sum_{i=1}^n \hat{\theta}_{-i}}{n}$	$\widehat{\text{CV}}_n = \frac{\sum_{i=1}^n \widehat{\text{CV}}_{-i}}{n}$
$\hat{\theta}_i = n\hat{\theta} - (n-1)\hat{\theta}_n$	$\widehat{\text{CV}}_i = n\widehat{\text{CV}} - (n-1)\widehat{\text{CV}}_n$
$\hat{\theta}_{\text{jack}} = \frac{\sum_{i=1}^n \hat{\theta}_i}{n}$	$\widehat{\text{CV}}_{\text{jack}} = \frac{\sum_{i=1}^n \widehat{\text{CV}}_i}{n}$
$\text{bias}_{\text{jack}} = \hat{\theta} - \hat{\theta}_{\text{jack}}$	$\text{bias}_{\text{jack}} = \widehat{\text{CV}} - \widehat{\text{CV}}_{\text{jack}}$
$\hat{V}(\hat{\theta}_{\text{jack}}) = \frac{\sum_{i=1}^n (\hat{\theta}_i - \frac{1}{n} \sum_{j=1}^n \hat{\theta}_j)^2}{n(n-1)}$	$\hat{V}(\widehat{\text{CV}}_{\text{jack}}) = \frac{\sum_{i=1}^n (\widehat{\text{CV}}_i - \frac{1}{n} \sum_{j=1}^n \widehat{\text{CV}}_j)^2}{n(n-1)}$

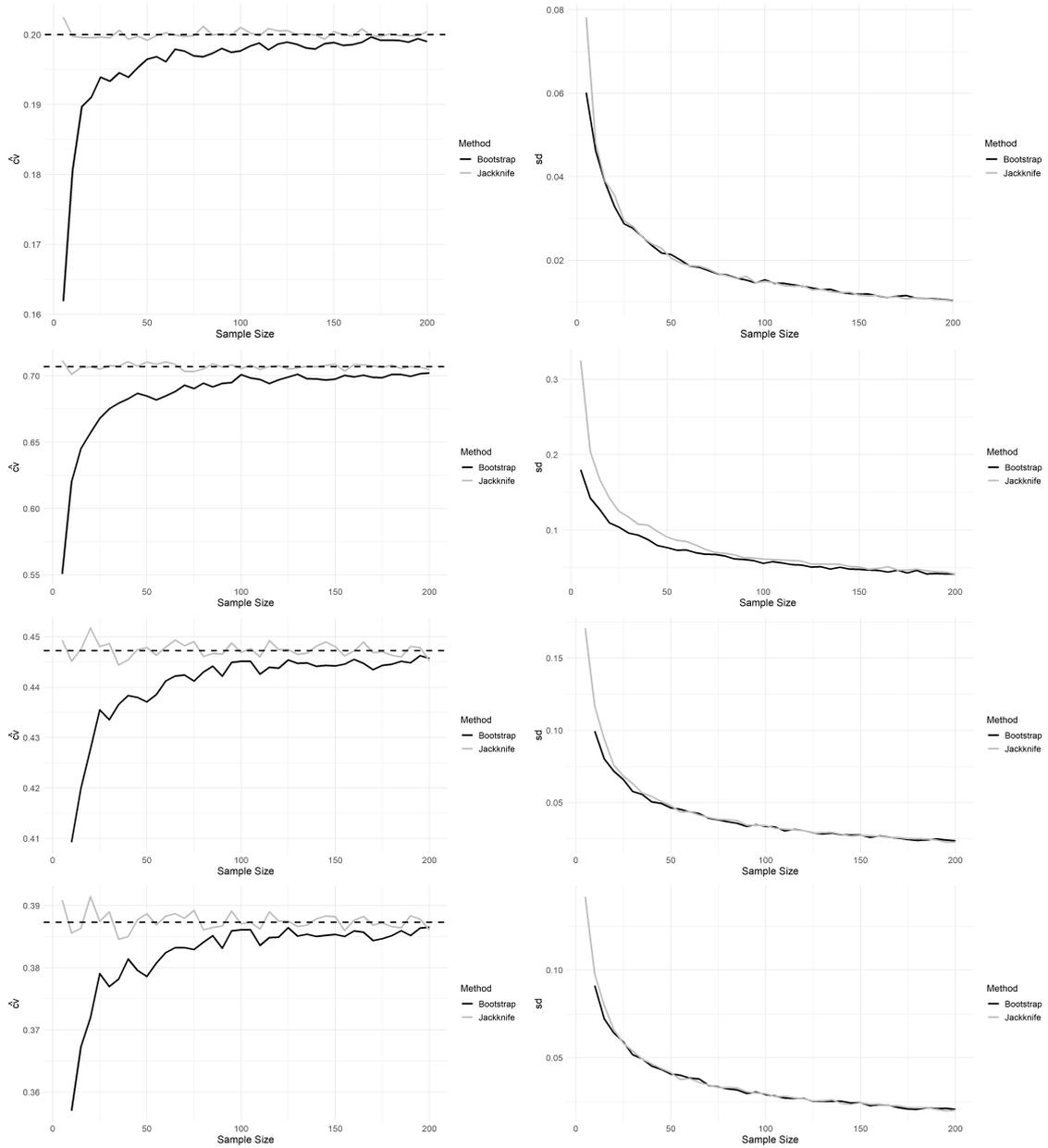


Figure 1: Estimation (left) and standard error (right) of the coefficient of variation according to the sample size (grey and black curves correspond to Jackknife and Bootstrap, respectively). The dashed line in the left panel corresponds to the CV of reference. From top to bottom, we have the results for Normal, Gamma, Poisson, and Binomial distributions, respectively.

$$\begin{aligned} H_0 &: CV = CV_0 \\ H_a &: CV > CV_0 \end{aligned} \tag{8}$$

with CV_0 defined with the expressions in the last column of the Table 2. The steps to calculate the power curves for each one of the probability models considered are the following

1. We fix a sample size $n = 100$ and a significance level $\alpha = 5\%$.
2. One sample size n is simulated under the null hypothesis.
3. We generate many Bootstrap and Jackknife samples by re-sampling with replacement or deleting one observation at a time, respectively.
4. For each Bootstrap or Jackknife sample, the CV is calculated and used to test whether it is significantly different from the CV under the null model (given in Table 2).
5. The proportion of times the null hypothesis is rejected across all Bootstrap or Jackknife samples is calculated. This proportion is an estimate of the power of the test.
6. We repeat steps 3-5 for many values of μ, α, λ , and p , respectively, under the alternative hypothesis.

The results obtained are presented in Figure 2. These suggest that in all cases (four probability models), the tests based on Bootstrap are more powerful than those found with Jackknife.

4 Review of Bootstrap and Jackknife in kernel density estimation

Kernel density estimation is widely used in several applied contexts, including, among others, marine biology [22], chemistry [20], and econometrics [34]. A fundamental aspect of kernel estimation is bandwidth selection. Usually, leave-one-out and k-fold cross-validation are used for establishing the optimal bandwidth [3]. Here we compare the performance of Bootstrap and Jackknife in this scenario. Bootstrap and Jackknife can be used to estimate the variability of the kernel density estimator, but they differ in how they generate the resamples. Bootstrap requires random sampling with replacement, while Jackknife involves leaving out one observation at a time. In this section, we compare these strategies according to their performances in both histogram density estimation (Section 4.1) and kernel density estimation (Section 4.2). Specifically, we estimate the uncertainty of the estimated density function in kernel density estimation.

4.1 Bootstrap and Jackknife in bandwidth histogram estimation

The histogram is one of the most broadly used graphical tools in descriptive data analysis [29]. This tool is a kernel density estimator where the underlying kernel is uniform [3]. Although there are better options for estimating the density, in this work, we consider the histogram given its extensive use in real data analysis. Specifically, it is established which of the two resampling methodologies (Bootstrap or Jackknife) performs best in estimating the amplitude of the class intervals.

Given an observed sample x_1, \dots, x_n , the histogram estimator of the density function $f(x)$ is defined as

$$\begin{aligned} \hat{f}_H(x) &= \frac{1}{nh} \sum_{i=1}^n \frac{1}{2} I_{[x_i-h, x_i+h]} \left(\frac{x_i - x}{h} \right) \\ &= \frac{1}{nh} \sum_{i=1}^n K_U(u), \end{aligned}$$

with

$$\begin{aligned} u &= \left(\frac{x_i - x}{h} \right), \\ K_U(u) &= \begin{cases} \frac{1}{2} & -1 < u < 1 \\ 0 & \text{other case} \end{cases} \end{aligned}$$

The optimal bandwidth h is the one that minimizes the asymptotic mean integrated squared error (AMISE), defined as [3]

$$AMISE(\hat{f}_h) = \frac{1}{nh} + \frac{h^2}{12} R(f'), \text{ with } R(f') = \int_{-\infty}^{\infty} (f'(x))^2 dx, \tag{9}$$

where $f(x)$ is the population density. To estimate h is usually considered $f(x)$ as a Gaussian distribution with mean and standard deviation estimated from the sample. Taking derivative of the AMISE with respect to h and equating to zero yields

$$h = \left(\frac{6}{nR(f')} \right)^{1/3}.$$

Under the Gaussianity assumption

$$\begin{aligned} R(f') &= \frac{1}{4\sqrt{\pi}\sigma^3}, \\ h &= 3.491\sigma n^{-1/3}, \text{ and therefore} \\ \hat{h} &= 3.491\hat{\sigma} n^{-1/3}, \text{ with } \hat{\sigma} = \frac{IQR}{1.35} \end{aligned} \tag{10}$$

4.2 Bootstrap and Jackknife in bandwidth kernel density estimation

The problem of estimating h depends only on the σ estimation. Here we evaluate the performance of Bootstrap and Jackknife in the estimation of h . The estimations obtained are compared with the classical estimation given in Equation (10). The Tables 5 and 6 show in parallel the general definitions of Bootstrap and Jackknife estimators and the corresponding expressions for estimating the bandwidth h by means of these approaches. In order to compare the

Table 5: Summary of Bootstrap estimation of the bandwidth h in kernel density estimation. Assume that \hat{h}_b^* is the estimation of the bandwidth h based on a Bootstrap sample

General case	Bandwidth h
$\hat{\theta}$	$\hat{h} = 3.491\hat{\sigma} n^{-1/3}$
$\hat{\theta}_{boot} = \frac{1}{B} \sum_{b=1}^B \hat{\theta}_b^*$	$\hat{h}_{boot} = \frac{1}{B} \sum_{b=1}^B \hat{h}_b^*$
$bias_{boot} = \hat{\theta}_{boot} - \hat{\theta}$	$bias_{boot} = \hat{h}_{boot} - \hat{h}$
$V(\hat{\theta}_{boot}) = \frac{1}{B} \sum_{b=1}^B \left(\hat{\theta}_b^* - \frac{1}{B} \sum_{b=1}^B \hat{\theta}_b^* \right)^2$	$V(\hat{h}_{boot}) = \frac{1}{B} \sum_{b=1}^B \left(\hat{h}_b^* - \frac{1}{B} \sum_{b=1}^B \hat{h}_b^* \right)^2$

methodologies, we conduct a simulation study. Suppose $X \sim N(\mu = 10, \sigma = 3)$. Using R software [26], we simulate random samples size $n = 10, 20, 30, \dots, 200$ from X , and at each case, we estimate the parameter h by using the estimator in Equation (10). We also carry out an estimation of h through Bootstrap and Jackknife (see Tables 5 and 6). The simulation results are shown in Table 7. There are several remarkable aspects of this Table.

With small samples, the estimations by the three methods (classical, Bootstrap, and Jackknife) are very similar, and for large samples ($n \geq 60$), the three estimations coincide. In all cases, the estimations are very close to the value of the parameter h . This result indicates that any of them can be used. However, one advantage of Bootstrap and Jackknife over classical estimation is that these

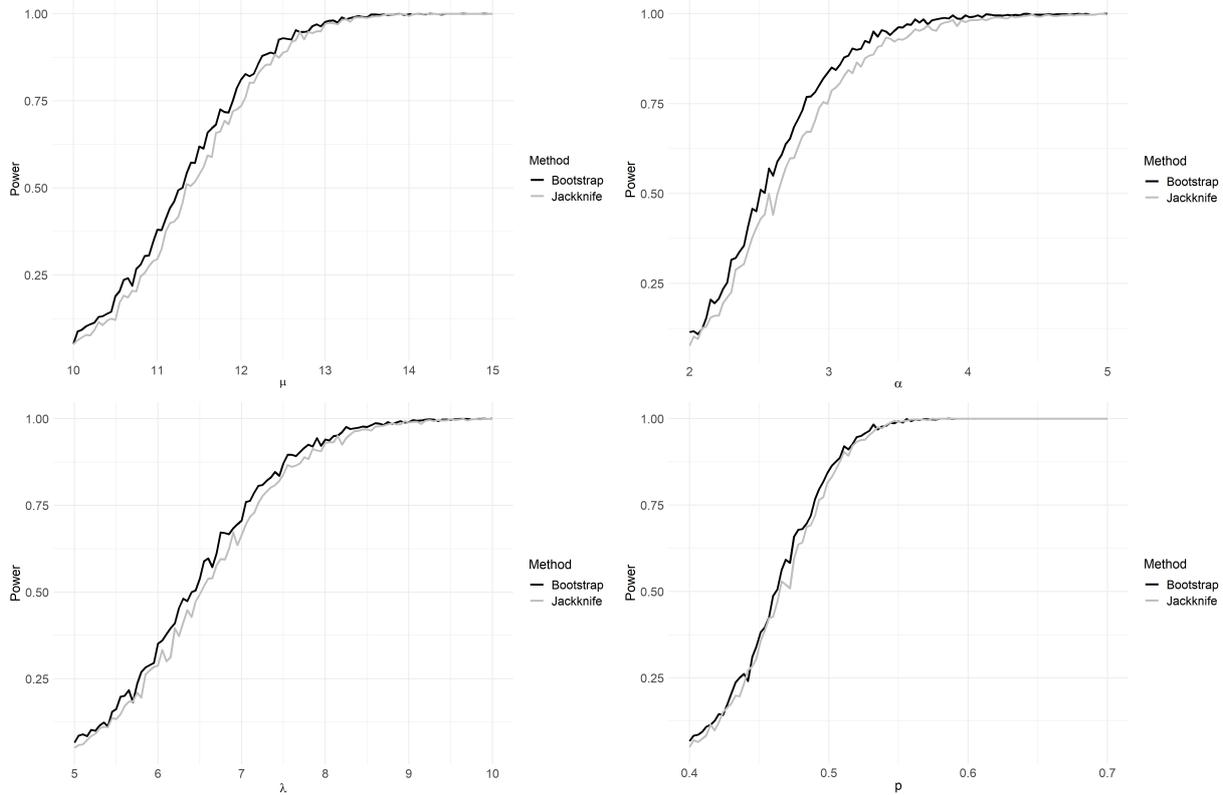


Figure 2: Power curves based on simulated data. One side hypothesis of the coefficients of variation obtained with four probability models. Normal ($\mu = 10, \sigma = 2$) (top left), Gamma ($\alpha = 2, \beta = 2$) (top right), Poisson ($\lambda = 5$) (bottom left), Binomial ($n = 10, p = 0.4$)

Table 6: Summary of Jackknife estimation of the bandwidth h in kernel density estimation. Assume that $\hat{\sigma}_{-i}$ is the estimation of standard deviation based on a sample $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$.

General case	Bandwidth h
$\hat{\theta}$	$\hat{h} = 3.491\hat{\sigma}n^{-1/3}$
$\hat{\theta}_{-i}$	$\hat{h}_{-i} = 3.491\hat{\sigma}_{-i}(n-1)^{-1/3}$
$\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^n \hat{\theta}_{-i}$	$\hat{h}_n = \frac{1}{n} \sum_{i=1}^n \hat{h}_{-i}$
$\hat{\theta}_i = n\hat{\theta} - (n-1)\hat{\theta}_n$	$\hat{h}_i = nh - (n-1)\hat{h}_n$
$\hat{\theta}_{jack} = \frac{1}{n} \sum_{i=1}^n \hat{\theta}_i$	$\hat{h}_{jack} = \frac{1}{n} \sum_{i=1}^n \hat{h}_i$
$bias_{jack} = \hat{\theta} - \hat{\theta}_{jack}$	$bias_{jack} = h - \hat{h}_{jack}$
$V(\hat{\theta}_{jack}) = \frac{\sum_{i=1}^n (\hat{\theta}_i - \frac{1}{n} \sum_{j=1}^n \hat{\theta}_j)^2}{n(n-1)}$	$V(\hat{h}_{jack}) = \frac{\sum_{i=1}^n (\hat{h}_i - \frac{1}{n} \sum_{j=1}^n \hat{h}_j)^2}{n(n-1)}$

Table 7: Assume $X \sim N(\mu, \sigma)$. h : optimal amplitude in histogram density estimation. \hat{h} : classical estimation. \hat{h}_{boot} and \hat{h}_{jack} are the approaches based on Bootstrap and Jackknife, respectively. In these cases we also include 95% confidence intervals.

n	μ	σ	$\hat{\sigma}$	h	\hat{h}	\hat{h}_{boot}	CI(\hat{h}_{boot})	\hat{h}_{jack}	CI(\hat{h}_{jack})
10	10	3	2.1	4.9	5.1	4.7	(2.9, 6.4)	5.3	(2.8, 7.7)
20	10	3	2.4	3.9	3.1	3.0	(2.1, 3.7)	3.1	(2.2, 4.0)
30	10	3	2.6	3.4	2.9	2.8	(2.1, 3.5)	3.0	(2.2, 3.7)
40	10	3	2.7	3.1	2.7	2.7	(2.2, 3.2)	2.7	(2.2, 3.3)
50	10	3	2.6	2.8	2.5	2.4	(2.0, 2.8)	2.5	(2.1, 2.9)
60	10	3	2.7	2.7	2.4	2.4	(2.0, 2.8)	2.4	(2.0, 2.9)
70	10	3	3.0	2.5	2.3	2.3	(1.9, 2.6)	2.3	(2.0, 2.7)
80	10	3	3.0	2.4	2.3	2.3	(1.9, 2.5)	2.3	(2.0, 2.6)
90	10	3	3.0	2.3	2.2	2.2	(1.9, 2.4)	2.2	(1.9, 2.5)
100	10	3	3.0	2.3	2.1	2.1	(1.9, 2.3)	2.1	(1.9, 2.4)
120	10	3	3.0	2.1	2.0	2.0	(1.8, 2.2)	2.0	(1.9, 2.2)
150	10	3	3.0	2.0	1.9	1.9	(1.7, 2.0)	1.9	(1.7, 2.0)
200	10	3	3.0	1.8	1.7	1.7	(1.6, 1.8)	1.7	(1.6, 1.8)

allow for assessing the uncertainty of the estimations (employing the corresponding confidence intervals). As expected with both methodologies (Bootstrap and Jackknife), a large sample size provides narrower confidence intervals. In all cases, Bootstrap and Jackknife confidence intervals contain the corresponding parameter h . The results indicate that Bootstrap and Jackknife are valid and valuable alternatives for estimating the interval bandwidth in the histogram density estimation. These are preferable to the classical estimation since they allow obtaining, in addition to the point estimation, a measure of variability in the estimation. In the case of small samples, the Bootstrap intervals are slightly narrower than those obtained with Jackknife. This point suggests that Bootstrap might be more suitable when n values are small. Let x_1, \dots, x_n a sample size n of a population with unknown density $f(x)$. The kernel density

estimator of $f(x)$ is given by

$$\hat{f}_k(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{x-x_i}{h}\right)$$

where K is a kernel function (Gaussian, Epanechnikov, triangular, biweight, etc.) and h is the bandwidth. The optimal h is the value that minimizes the $AMISE(\hat{f}_k(x))$ defined as

$$AMISE = \frac{R(k)}{nh} + \frac{h^4 \sigma_k^4}{4} R(f''), \text{ where}$$

$$R(k) = \int K^2(u)du \text{ and } R(f'') = \int (f''(x))^2 dx.$$

Table 8: Assume $X \sim N(\mu, \sigma)$. h : optimal amplitude in kernel density estimation. \hat{h} : classical estimation. \hat{h}_{boot} and \hat{h}_{jack} are the approaches based on Bootstrap and Jackknife, respectively. In these cases we also include 95% confidence intervals.

n	μ	σ	$\hat{\sigma}$	h	\hat{h}	\hat{h}_{boot}	CI(\hat{h}_{boot})	\hat{h}_{jack}	CI(\hat{h}_{jack})
10	10	3	2.1	2.0	2.1	2.0	(1.2, 2.7)	2.1	(1.2, 3.1)
20	10	3	2.4	1.7	1.4	1.3	(1.0, 1.7)	1.4	(1.0, 1.8)
30	10	3	2.6	1.6	1.4	1.4	(1.0, 1.7)	1.4	(1.0, 1.8)
40	10	3	2.7	1.5	1.4	1.3	(1.1, 1.6)	1.4	(1.0, 1.6)
50	10	3	2.6	1.5	1.3	1.2	(1.0, 1.4)	1.3	(1.0, 1.5)
60	10	3	2.7	1.4	1.3	1.2	(1.0, 1.5)	1.3	(1.0, 1.5)
70	10	3	3.0	1.4	1.2	1.2	(1.0, 1.5)	1.2	(1.0, 1.4)
80	10	3	3.0	1.3	1.2	1.2	(1.1, 1.4)	1.2	(1.0, 1.4)
90	10	3	3.0	1.3	1.2	1.2	(1.0, 1.3)	1.2	(1.0, 1.4)
100	10	3	3.0	1.3	1.2	1.2	(1.0, 1.3)	1.2	(1.0, 1.3)
120	10	3	3.0	1.2	1.1	1.1	(1.0, 1.3)	1.2	(1.0, 1.3)
150	10	3	3.0	1.2	1.1	1.1	(1.0, 1.3)	1.1	(1.0, 1.2)
200	10	3	3.0	1.1	1.0	1.0	(1.0, 1.1)	1.1	(1.0, 1.1)

Differentiating the AMISE with respect to h we have

$$\frac{\partial AMISE}{\partial h} = -\frac{R(k)}{nh^2} + \frac{4h^3 \sigma_k^4 R(f'')}{4}$$

Equating to zero

$$h^3 \sigma_k^4 R(f'') = \frac{R(k)}{nh^2}$$

$$h^5 = \left[\frac{R(k)}{n \sigma_k^4 R(f'')} \right]$$

$$h = \left(\frac{R(k)}{\sigma_k^4 R(f'')} \right)^{1/5} \frac{1}{n^{1/5}}$$

The optimal bandwidth h is obtained assuming $f(x)$ as a Gaussian density. Hence, after some calculations, we have that the optimal bandwidth in kernel density estimation is obtained by

$$\hat{h} = 1.059 \hat{\sigma} n^{-1/5}. \tag{11}$$

Changing $\hat{h} = 3.491 \hat{\sigma} n^{-1/3}$ by Equation (11) in Tables 5 and 6, we obtain the corresponding expressions to do the estimation of the optimal bandwidth in kernel density estimation by using Bootstrap and Jackknife. In Table 8, we show the estimations of the optimal bandwidth in kernel density estimation based on the same samples simulated to generate the results in Table 7. As in the particular case of the histogram, the bandwidth estimations shown in Table 8 indicate that Bootstrap and Jackknife can be favorable alternatives to carry out the estimation of the density using the general methodology based on kernel. We present the results using a Gaussian kernel; however, we obtained similar results with others. As in the case of the histogram estimation, in this section, we can conclude that using Bootstrap or Jackknife can be preferable because these approaches allow having an estimation of the variability for the estimator. Particularly with small sample sizes is helpful to know the uncertainty in the estimation. As in the case of the histogram estimation, we note that Bootstrap can be preferable with small samples because narrower confidence intervals are obtained. The results in Tables 7 and 8 show that the estimators \hat{h}_{boot} and \hat{h}_{jack} are consistent, i.e, $\lim_{n \rightarrow \infty} \hat{h}_{boot} = h$ and $\lim_{n \rightarrow \infty} \hat{h}_{jack} = h$.

5 Bootstrap and Jackknife in bandwidth kernel regression estimation

Bootstrap and Jackknife have been compared under various regression contexts. Among others in linear [2], generalized linear [12, 30, 33], and logistic regression [14]. This paper compares the efficiency of Bootstrap and Jackknife in estimating the bandwidth in kernel regression. This parameter, denoted as h (as in Sections 4.1 and 4.2 but now in a regression scenario), determines the width of the kernel function, which in turn affects the smoothness of the

estimated regression function. The choice of the bandwidth is critical, as it can greatly affect the bias and variance of the regression estimator. Suppose we have an observed bivariate random sample $(y_i, x_i), i = 1, \dots, n$ with y and x the response and predictor variables, respectively. We want to estimate the regression model.

$$y_i = m(x_i) + e_i, \quad i = 1, \dots, n. \tag{12}$$

Let $f(x)$ and $f(x, y)$ be the univariate and bivariate density functions of the variable X and the random vector (Y, X) . The estimation of a kernel regression model based on the observed sample $(x_i, y_i), i = 1, \dots, n$ is given by [3]

$$\begin{aligned} \hat{m}(x) &= \frac{1}{\hat{f}(x)} \int y f(y, x) dy \\ &= \frac{1}{\frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{x-x_i}{h}\right)} \int y f(y, x) dy \\ &= \frac{\frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{x-x_i}{h}\right) y_i}{\frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{x-x_i}{h}\right)} \\ &= \frac{\sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) y_i}{\sum_{i=1}^n K\left(\frac{x-x_i}{h}\right)} \\ &= \sum_{i=1}^n w_i y_i, \quad \text{with } w_i = \frac{K\left(\frac{x-x_i}{h}\right)}{\sum_{i=1}^n K\left(\frac{x-x_i}{h}\right)}, \end{aligned}$$

where K is a kernel (Gaussian, rectangular, triangular, Epanechnikov, etc) and $h > 0$ is the bandwidth that controls the amount of smoothing. For a particular $i, i = 1, \dots, n$, we have

$$\begin{aligned} \hat{y}_i &= \hat{m}(x_i) \\ &= \frac{\sum_{j=1}^n K\left(\frac{x_i-x_j}{h}\right) y_j}{\sum_{m=1}^n K\left(\frac{x_i-x_m}{h}\right)} \\ &= \sum_{j=1}^n w(x_i, x_j) y_j. \end{aligned} \tag{13}$$

In matrix notation, the estimates at the n sampling points are calculated as

$$\begin{pmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{pmatrix} = \begin{pmatrix} w(x_1, x_1) & w(x_1, x_2) & \cdots & w(x_1, x_n) \\ w(x_2, x_1) & w(x_2, x_2) & \cdots & w(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ w(x_n, x_1) & w(x_n, x_2) & \cdots & w(x_n, x_n) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

A key point in kernel regression is to determine the bandwidth h . The usual strategy is based on choosing the h value that minimizes the mean squared error (MSE) defined as [18]

$$MSE(h) = \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - \hat{y}_i}{1 - v/n} \right)^2, \tag{14}$$

with

$$v = \text{trace}(W) = \sum_{i=1}^n w(x_i, x_i). \tag{15}$$

In general, a small bandwidth will result in a high degree of local smoothing, while a large bandwidth will lead to less local smoothing and more global effects. A bandwidth that is too small may generate overfit of the data, while a bandwidth that is too large may produce an over-smoothing of the data and miss significant local trends. Various methods can be used to determine the optimal bandwidth, such

as cross-validation or minimizing a particular criterion (e.g., mean squared error or Akaike's information criterion). Cross-validation involves partitioning the data into training and validation sets and evaluating the performance of the model with different bandwidth values. The bandwidth that results in the best behavior on the validation set is then chosen as the optimal bandwidth. In this Section, we compare Bootstrap and Jackknife to establish their efficiency in estimating the optimal bandwidth h in kernel regression according to the criterion in Equation (14). The Jackknife method in kernel regression involves repeatedly fitting the kernel regression estimator to a subset of the data, leaving out one observation each time. The estimation of the kernel regression function is then calculated using all the observations and each subset of the data. The Jackknife estimation of the bandwidth is obtained using the pseudovalues are calculated as

$$\begin{aligned} \tilde{h}_i &= n\hat{h} - (n-1)\hat{h}_{-i}, \text{ with} \\ \hat{h}_n &= \frac{\sum_{i=1}^n \tilde{h}_i}{n}, \end{aligned}$$

where \hat{h} and \hat{h}_{-i} are defined using the criterion in Equation (14), with and without considering, respectively, the i -th observation $(y_i, x_i), i = 1, \dots, n$. The pseudo-values can be used to estimate h , the variance of the kernel regression estimator \hat{h} and a confidence interval for h . Specifically we have

$$\begin{aligned} \hat{h}_{\text{jack}} &= \frac{1}{n} \sum_{i=1}^n \tilde{h}_i, \\ \hat{V}(\hat{h}_{\text{jack}}) &= \frac{1}{n(n-1)} \sum_{i=1}^n \left(\tilde{h}_i - \frac{1}{n} \sum_{i=1}^n \tilde{h}_i \right)^2, \\ &\left(\hat{h}_{\text{jack}} \pm t_{n-1, 1-\frac{\alpha}{2}} \sqrt{\hat{V}(\hat{h}_{\text{jack}})} \right). \end{aligned}$$

Using the quantiles $\frac{\alpha}{2}$ and $(1 - \frac{\alpha}{2})$ from the pseudovalues, an approximate confidence interval for the bandwidth of the kernel regression can be obtained as

$$\left(\tilde{h}_{\frac{\alpha}{2}}, \tilde{h}_{1-\frac{\alpha}{2}} \right).$$

On the other hand, the Bootstrap method can also be used to obtain many estimations (using resampling with replacement) and, consequently, a variability measure and a confidence interval for the bandwidth h calculating the quantiles of the distribution of the Bootstrap estimations. Let $(x_{11}^*, y_{11}^*), \dots, (x_{1n}^*, y_{1n}^*), \dots, (x_{B1}^*, y_{B1}^*), \dots, (x_{Bn}^*, y_{Bn}^*)$ be B Bootstrap samples size n taken from $(x_i, y_i), i = 1, \dots, n$. Based on each one of these is obtained an estimation of h according to the criterion in Equation (14), i.e., we find $(\hat{h}_1^*, \dots, \hat{h}_B^*)$. The Bootstrap estimator of the bandwidth, its variance, and the corresponding confidence interval are given by

$$\begin{aligned} \hat{h}_{\text{boot}} &= \frac{1}{B} \sum_{b=1}^B \hat{h}_b^*, \\ V(\hat{h}_{\text{boot}}) &= \frac{1}{B} \sum_{b=1}^B \left(\hat{h}_b^* - \frac{1}{B} \sum_{b=1}^B \hat{h}_b^* \right)^2, \\ &\left(\hat{h}_{\text{boot}} \pm Z_{1-\frac{\alpha}{2}} \sqrt{\hat{V}(\hat{h}_{\text{boot}})} \right) \end{aligned}$$

Using the quantiles of the Bootstrap estimations $(\hat{h}_1^*, \dots, \hat{h}_B^*)$ a confidence interval for the bandwidth of the kernel regression estimator can be also obtained as

$$\left(\hat{h}_{\frac{\alpha}{2}}^*, \hat{h}_{1-\frac{\alpha}{2}}^* \right).$$

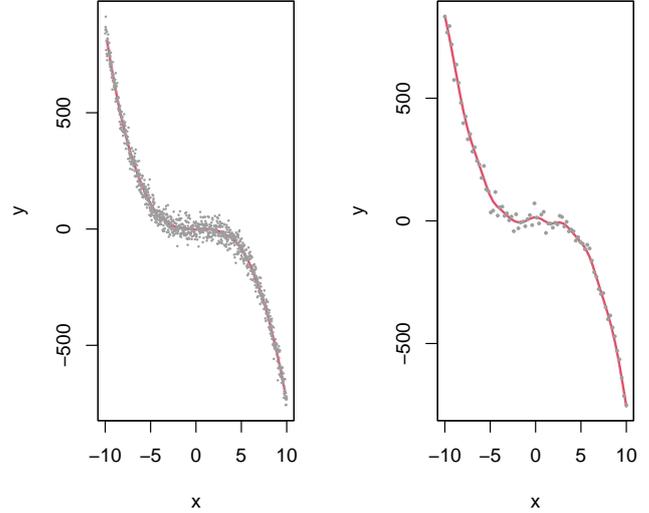


Figure 3: Left panel: Scatterplot of $(x_i, y_i), i = 1, \dots, 1000$ points simulated from the model $Y = 0.2x + 0.5x^2 - 0.8x^3 + \epsilon, \epsilon \sim N(0, 30)$ which is assumed as the population data and fitted kernel regression model with $h = 0.37$ (red curve). Right panel: Simulation of $n = 80$ data from the population data (gray points in the right panel) and estimated kernel regression model (red curve) with $\hat{h} = 0.65$.

5.1 Simulation study

Here we present some simulation results comparing Bootstrap and Jackknife in estimating the bandwidth in kernel regression. For this purpose we assume $Y = 0.2x + 0.5x^2 - 0.8x^3 + \epsilon, \epsilon \sim N(0, 30)$ as the population model (Figure 3). The red line corresponds to the fit of a kernel regression model with bandwidth $h = 0.37$, which was calculated according to the criterion in Equation (14). The comparisons are made by taking samples of size $n = 20, 40, 60, \dots, 200$ from this model. For illustrative purposes, in the right panel is presented an estimated kernel regression model with a sample size $n = 80$ of the population model. In this case, the bandwidth estimation is $\hat{h} = 0.64$. Varying the sample size are obtained the estimations $\hat{h}, \hat{h}_{\text{boot}}$ and \hat{h}_{jack} (Table 9). Note that in practice given a dataset $(x_i, y_i), i = 1, \dots, n$ we have just one estimation of \hat{h} and several estimations of \hat{h}_{boot} and \hat{h}_{jack} obtained by resampling the data recorded. In this real case, we only could determine confidence intervals employing the approaches based on Bootstrap and Jackknife. For this reason, in order to compare the three methodologies (classical, Bootstrap, and Jackknife), we calculate confidence intervals based on the percentiles 5% and 95% obtained from 300 sets of values of $\hat{h}, \hat{h}_{\text{boot}}$ and \hat{h}_{jack} generated with an equal number of samples size $n = 20, 40, 60, \dots, 200$ simulated from the population model (Table 9). The results suggest that the tree estimators are consistent, i.e., as we collect more and more data (when n increases), the difference between the estimated value and the "true value of the parameter" (assumed equal to 0.37) will become smaller and smaller. According to the results in Table 9, the Jackknife estimations are very close to the obtained with the classical approach. We also note (see last row of the table) that the estimates by Jackknife tend more quickly to the fixed reference value ($h = 0.37$) than those obtained by Bootstrap. These results indicate that Jackknife may be a more appropriate and efficient option than Bootstrap to estimate the bandwidth in Kernel regression. In general is accepted that in regression problems, as long as the data set is reasonably large, Bootstrap is often acceptable. However, the estimations \hat{h}_{boot} and the corresponding confidence

Table 9: Assume $Y = 0.2x + 0.5x^2 - 0.8x^3 + \epsilon$ a regression function with $\epsilon \sim N(0, 30)$. The optimal bandwidth (see Section 5) in a kernel regression model calculated with 1000 data simulated from the population model is $h = 0.36$. \hat{h} , \hat{h}_{boot} , and \hat{h}_{jack} identify the approaches classical and based on Bootstrap and Jackknife, respectively, which are calculated for each sample size as the mean of fifty simulations. We also include 95% confidence obtained as percentiles 5% and 95% of the fifty estimations.

n	\hat{h}	CI(\hat{h})	\hat{h}_{boot}	CI(\hat{h}_{boot})	\hat{h}_{jack}	CI(\hat{h}_{jack})
20	0.74	(0.40, 1.22)	0.70	(0.50, 0.96)	0.74	(0.42, 1.20)
40	0.73	(0.39, 0.87)	0.68	(0.49, 0.92)	0.72	(0.45, 0.99)
60	0.68	(0.39, 0.85)	0.66	(0.48, 0.91)	0.70	(0.43, 0.87)
80	0.64	(0.38, 0.80)	0.65	(0.48, 0.88)	0.65	(0.41, 0.84)
100	0.62	(0.37, 0.75)	0.65	(0.47, 0.84)	0.63	(0.39, 0.80)
120	0.57	(0.37, 0.74)	0.64	(0.47, 0.83)	0.57	(0.37, 0.76)
140	0.56	(0.36, 0.73)	0.63	(0.46, 0.78)	0.55	(0.37, 0.75)
160	0.54	(0.36, 0.71)	0.61	(0.46, 0.74)	0.52	(0.36, 0.74)
180	0.55	(0.36, 0.70)	0.60	(0.45, 0.72)	0.50	(0.36, 0.72)
200	0.49	(0.36, 0.65)	0.59	(0.44, 0.71)	0.49	(0.36, 0.70)

Table 10: Results based on Bioluminescence data. Assume $Y = m(x) + e$, with $m(x)$ a kernel regression function. \hat{h} : optimal estimation based on lest square. \hat{h}_{boot} and \hat{h}_{jack} are the approaches based on Bootstrap and Jackknife, respectively. In these cases, we also include 95% confidence intervals.

n	\hat{h}	\hat{h}_{boot}	CI(\hat{h}_{boot})	\hat{h}_{jack}	CI(\hat{h}_{jack})
51	373	371	(349, 385)	372	(371, 376)

intervals $CI(\hat{h}_{boot})$ in Table 9 suggest that in the context of Kernel regression, Bootstrap is not the best option to establish the variability of the bandwidth estimator \hat{h} .

5.2 Application to bioluminescence data

Bioluminescence is the emission of light by living organisms, typically due to chemical reactions involving luciferin and luciferase enzymes [16]. Overall, analyzing bioluminescence data can provide valuable insights into the biological processes and respect for the environmental and experimental factors affecting bioluminescence activity [27]. Bioluminescence is very common in the ocean, at least in the pelagic zone. Bioluminescent creatures occur in all oceans at all depths, with the greatest numbers found in the upper 1000 m of the vast open ocean [32]. The relationship between pelagic bioluminescence and depth has been considered from various perspectives. Among additive models [13], [17], and additive mixed models [36] have been used in this context. This work applies a Kernel regression model to a pelagic bioluminescence dataset taken from [35]. Based on these data, the aim is to illustrate how to define confidence intervals for the bandwidth h parameter. The scatterplot and the fitted Kernel regression model with bandwidth $h = 373$ are shown in Figure 4

From Table 10, we can establish that the optimal bandwidth h that minimizes the criterion in Equation (5) is $\hat{h} = 373$. We also note that the estimations by Bootstrap and Jackknife are very close to this value; however, Jackknife is a better option to define the confidence interval for h because, on the one hand, it includes the \hat{h} value and, on the other hand, is shorter than the obtained by Bootstrap. The results in this Section confirm that Jackknife is a good alternative in kernel regression to define the bandwidth estimation uncertainty.

6 Conclusion and further research

The results in this work suggest that in the particular case of the coefficient of variation, Jackknife tends to produce less biased estimates, but it may have lower power than Bootstrap. In the case of the histogram and kernel density estimation, both methodologies produce similar results. Finally, in the context of Kernel regression, Jackknife is a better option. Resampling methods are helpful nowadays in many contexts. Power studies comparing these approaches

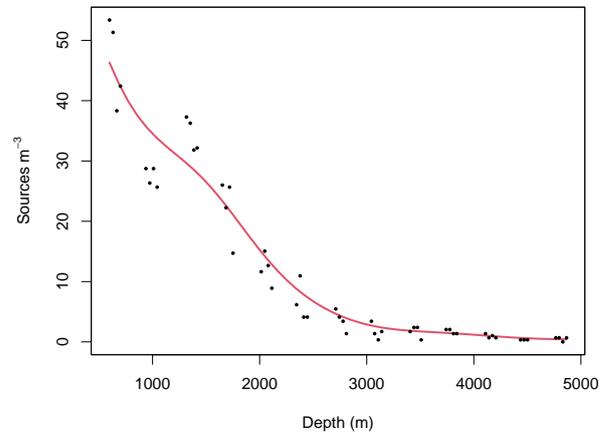


Figure 4: Scatterplot (dots) of pelagic bioluminescence along a depth gradient in the northeast Atlantic Ocean for a particular station. Data taken from [35]. Red curve corresponds to a Kernel regression model with an optimal bandwidth $\hat{h} = 373$.

in many statistical areas are required. For example, evaluating its performance in splines regression and neural networks can be valuable.

References

- [1] Behme, A. Theory of Stochastic Objects: Probability, Stochastic Processes and Inference. Chapman & Hall/CRC Press, 2018
- [2] Bose, A., Chatterjee, S. Comparison of bootstrap and jackknife variance estimators in linear regression: Second order results. Statistica Sinica. 2002, 575-598.
- [3] Bowman, A., Azzalini, A. Applied smoothing techniques for data analysis: the kernel approach with S-Plus illustrations. OUP Oxford, 1973.
- [4] Buckland, S. Monte Carlo confidence intervals. Biometrics. 1984; 811-817.
- [5] Chen, Y. A tutorial on kernel density estimation and recent advances. Biostatistics & Epidemiology. 2017; 1(1), 161-187.
- [6] DiCiccio, T., Efron, B. Bootstrap confidence intervals. Statistical Science. 1996; 11(3): 189-228.
- [7] Donegani, M. Bootstrap adaptive test for two way analysis of variance. Biometrical Journal. 1992; 34(2): 141-146
- [8] Donnelly-Makowecki, L., Moore, R. Hierarchical testing of three rainfall-runoff models in small forested catchments. Journal of Hydrology. 1999; 219(3-4): 136-152.
- [9] Efron, B., Stein, C. The Jackknife estimate of variance. The Annals of Statistics. 1981; 586-596.
- [10] Efron B., Tibshirani R. An introduction to the bootstrap. 1994; CRC Press.
- [11] Freedman D., Peters S. Bootstrapping a regression equation: Some empirical results. Journal of Statistical Software. 2008; 25(7):1-37.
- [12] Garay, A., Hashimoto, E., Ortega, E., Lachos, V. On estimation and influence diagnostics for zero-inflated negative binomial

- regression models. *Computational Statistics & Data Analysis*, 2011; 55(3): 1304-1318.
- [13] Gillibrand E., Jamieson A., Bagley, P., Zuur, A., Priede, I. Seasonal development of a deep pelagic bioluminescent layer in the temperate NE Atlantic Ocean. *Marine Ecology Progress Series*, 2007, 341: 37-43.
- [14] Gong, G. Cross-validation, the jackknife, and the bootstrap: Excess error estimation in forward logistic regression. *Journal of the American Statistical Association*. 1986; 81 (393): 108-113.
- [15] Hahn, J. Bootstrapping quantile regression estimators. *Econometric Theory*. 1995; 11(1): 105-121.
- [16] Han, J., Li, G., Liu, H., Hu, H., Zhang, X. Stimulation of bioluminescence in *Noctiluca* sp. using controlled temperature changes. *Luminescence*. 2013; 28(5): 742-744.
- [17] Heger, A., Ieno, E., King, N., Morris, K., Bagley, P., Priede, I. Deep-sea pelagic bioluminescence over the Mid-Atlantic Ridge. *Deep Sea Research Part II: Topical Studies in Oceanography*. 2008, 55(1-2), 126-136.
- [18] Kohler, M., Schindler, A., Sperlich, S. 2014. A review and comparison of bandwidth selection methods for kernel regression. *International Statistical Review*. 2014; 82(2): 243-274.
- [19] Lu, X., Su, L. Jackknife model averaging for quantile regressions. *Journal of Econometrics*. 2015; 188(1): 40-58.
- [20] McCabe, P., Korb, O., Cole, J. Kernel density estimation applied to bond length, bond angle, and torsion angle distributions. *Journal of Chemical Information and Modeling*, 2014, 54(5): 1284-1288.
- [21] Moulton, L., Zeger, S. Bootstrapping generalized linear models. *Computational Statistics & Data Analysis*. 1991; 11: 53-63.
- [22] O'Brien, S., Webb, A., Brewer, M., Reid, J. Use of kernel density estimation and maximum curvature to set marine protected area boundaries: Identifying a special protection area for wintering red-throated divers in the UK. *Biological Conservation*. 2012, 156: 15-21.
- [23] Oyeyemi, G. Comparison of bootstrap and jackknife methods of resampling in estimating population parameters. *Global Journal of Pure and Applied Sciences*. 2008; 1(2):217-220.
- [24] Quenouille, M. Approximate test of correlation in time series. *Journal of the Royal Statistical Society, Series B*. 1949; 11:68-84
- [25] Quenouille, M. Notes on bias in estimation. *Biometrika*. 1956; 43(3): 353-360.
- [26] R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. <https://www.R-project.org/>. 2022.
- [27] Roda, A., Pasini, P., Mirasoli, M., Michelini, E., Guardigli, M. Biotechnological applications of bioluminescence and chemiluminescence. *TRENDS in Biotechnology*. 2004, 22(6), 295-303.
- [28] Rudemo, M. Empirical choice of histograms and kernel density estimators. *Scandinavian Journal of Statistics*. 1982, 65-78.
- [29] Scott, D. Histogram. *Wiley Interdisciplinary Reviews: Computational Statistics*. 2010, 2(1): 44-48.
- [30] Sherman, M., Cessie, S. A comparison between bootstrap methods and generalized estimating equations for correlated outcomes in generalized linear models. *Communications in Statistics-Simulation and Computation*. 1997; 26(3): 901-925.
- [31] Tukey, J. Bias and confidence in not quite large samples. *Ann. Math. Statist*. 1958; 29: 614.
- [32] Widder, E. Marine bioluminescence. Why do so many animals in the open ocean make light?. *Science*. 2010; 328: 704-708.
- [33] Wu, C. Jackknife, Bootstrap and other resampling methods in regression analysis. *The Annals of Statistics*. 1986; 14(4): 1261-1295.
- [34] Zambom, A., Dias, R. A review of kernel density estimation with applications to econometrics. *International Econometric Review*. 2013; 5(1): 20-42.
- [35] Zuur, A., Ieno, E. Smith, G. *Analysing ecological data*. Springer; 2007.
- [36] Zuur, A. Additive mixed modelling applied on deep-sea pelagic bioluminescent organisms. In *Mixed effects models and extensions in ecology with R*. 2009; 399-422.