On Kernel Density Estimation with Univariate Application

BY

SILOKO, Israel Uzuazor

Department of Mathematics/ICT, Edo University Iyamho, Edo State, Nigeria.

A Seminar Presented at Faculty of Science, Edo University Iyamho, Nigeria.

OUTLINES

1. Abstract
2. Introduction
3. Density Estimators
4. The Kernel Density Estimator
5. Smoothing parameter Selectors
6. Discussions
7. References

ABSTRACT

Kernel density estimation is an important smoothing technique with direct applications such as data exploratory analysis and data visualisation. This review summarizes the most important theoretical aspects of kernel density estimation and provides a description of classical methods for computing the smoothing parameter. The performance of the kernel estimator will be considered based on the classical methods of obtaining the smoothing parameter and this will be fully illustrated by real data examples.

Introduction.

Kernel density estimation is a nonparametric method that estimates the probability density function of a random variable. Estimating probability density functions with kernel functions has had notable success due to their ease of interpretation and visualization.

Kernel density estimation is a fundamental data smoothing problem where inferences about the population are made based on a finite data sample. In some fields such as signal processing and econometrics, kernel density estimation was also termed as the Parzen-Rosenblatt window method, after Emmanuel Parzen (1962) and Murray Rosenblatt (1956), who are usually credited with independently for creating this method in its current form. The kernel density estimator introduced by Rosenblatt (1956) and Parzen (1962) (in the univariate case), is characterised by two components, the smoothing parameter and the kernel function and is of the form

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right). \qquad (1)$$

The estimator given in (1) above contains some parameters which include the sample size n, the kernel function K and the smoothing parameter h. The process of selecting the smoothing parameter is one of the major difficulties in Kernel Density Estimation. The estimator in Equation (1) must satisfy the conditions in Equation (2)

$$\int_{-\infty}^{\infty} tK(t)dt = 0, \qquad \int_{-\infty}^{\infty} t^2 K(t)dt = \mu_2(k) \neq 0\,, \qquad \int_{-\infty}^{\infty} K(t)dt = 1 \tag{2}$$

The smoothing parameter plays a central role in determining the performance of the kernel density estimators. Performance is measured by the closeness of a kernel density estimate to its target density. Although there are other global measures of discrepancy but the mean integrated squared error (MISE) introduced by Rosenblatt (1956) is the most mathematically amenable to theoretical investigation and calibration in practice (Silverman, 1986). The corresponding asymptotic mean integrated squared error (AMISE) of Equation (1) is of the form

$$\begin{aligned} AMISE &= \frac{R(K)}{nh} + \frac{1}{4}\mathrm{k}_2(\mathrm{K})^2 h^4 \int f''(\mathrm{x})^2 dx \\ &= \frac{R(K)}{nh} + \frac{1}{4}\mathrm{k}_2(\mathrm{K})^2 h^4 R(f'') \end{aligned} \tag{3}$$

The smoothing parameter that minimizes the *AMISE* of Equation (3) is given by

$$h_{AMISE} = \left[\frac{R(K)}{\mathrm{k}_2(\mathrm{K})^2 R(f'')}\right]^{1/5} \times n^{-1/5} \tag{4}$$

A range of kernel functions are commonly used such as Uniform, Triangular, Biweight, Triweight, Epanechnikov, Normal kernel etc. The choice of the kernel function is not crucial since most of the kernel functions are probability density functions (Silverman, 1986) but the choice of the smoothing parameter is very important.

The aim of this seminar is to review the most important aspects of kernel density estimation such as data exploratory analysis and data visualisation using the classical approaches in the selection of smoothing parameter. An interesting comprehensive review of kernel smoothing and its applications can be found in Silverman (1986), Scott (1992), Wand and Jones (1995), Simonoff (1996), Schimek (2000) and Sheather (2004).

Some Smoothing Parameter Selectors.

We compare the performance of the popular cross validation selectors. They are the unbiased cross validation, the unbiased cross validation and the smoothed bootstrap method.

The unbiased cross validation is given by

$$UCV(h_1,\dots,h_d) = \frac{1}{\left(2\sqrt{\pi}\right)^d n h_1,\dots,h_d} + \frac{1}{\left(2\sqrt{\pi}\right)^d n^2 h_1,\dots,h_d} \times \sum_{i=1}^{n}\sum_{j\neq i}\left[exp\left\{-\frac{1}{4}\sum_{k=1}^{d}\Delta_{ijk}^2\right\} - \left(2\times 2^{d/2}\right)exp\left\{-\frac{1}{2}\sum_{k=1}^{d}\Delta_{ijk}^2\right\}\right] \tag{5}$$

where $\Delta_{ijk} = \left(\frac{\mathbf{x}_{ik} - \mathbf{X}_{jk}}{h_k}\right)$

The biased cross validation is given by

$$BCV(h_1,\dots,h_d) = \frac{1}{\left(2\sqrt{\pi}\right)^d n h_1 \dots h_d} + \frac{1}{4n(n-1)h_1 \dots h_d} \times \sum_{i=1}^{n}\sum_{j\neq i}\left[\left(\sum_{k=1}^{d}\Delta_{ijk}^2\right)^2 - (2d+4)\left(\sum_{k=1}^{d}\Delta_{ijk}^2\right) + (d^2+2d)\right]\prod_{k=1}^{d}\phi(\Delta_{ijk}) \tag{6}$$

where ϕ is the standard Normal density.

The smoothed bootstrap is of the form

$$B(h_1, \dots, h_d) = \frac{1}{(2\sqrt{\pi})^d n h_1 \dots h_d} + \frac{1}{(2\sqrt{\pi})^d n^2 h_1 \dots h_d} \times \sum_{i=1}^{n} \sum_{j \neq i} \left[\frac{n-1}{2^{d/2} n} \exp\left\{ -\frac{1}{8} \sum_{k=1}^{d} \Delta_{ijk}^2 \right\} + \exp\left\{ -\frac{1}{4} \sum_{k=1}^{d} \Delta_{ijk}^2 \right\} - \frac{2 \times 2^{d/2}}{3^{d/2}} \exp\left\{ -\frac{1}{6} \sum_{k=1}^{d} \Delta_{ijk}^2 \right\} \right] \qquad (7)$$

Results.

Statistical graphs play three important roles in data analysis.

- Graphs provide an initial look at the data, a step that is skipped at the peril of the data analyst.

- Graphs are also employed during model building and model criticism, particularly in diagnostic methods used to understand the fit of a model.

- Finally, presentation graphics can summarize a fitted model for the benefit of others.

We considered the Smoothed Bootstrap method (SBM), the Unbiased Cross Validation (UCV) and the Biased Cross Validation (BCV) and compared their results in terms of performance using the asymptotic mean integrated squared error (AMISE) as the error criterion function. The data set examined is the Annual Snowfall in Buffalo Scott (1992). The sample size of this data is 63. The kernel estimates are shown in Figure3.1, Figure3.2, Figure3.3 and Figure3.4 using the standard normal kernel.

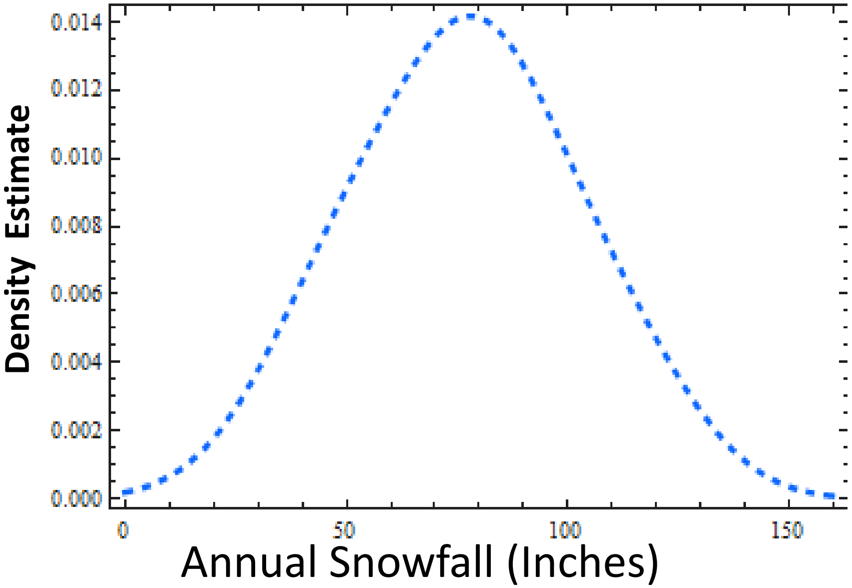

Figure3.1: Kernel Estimate of SBM

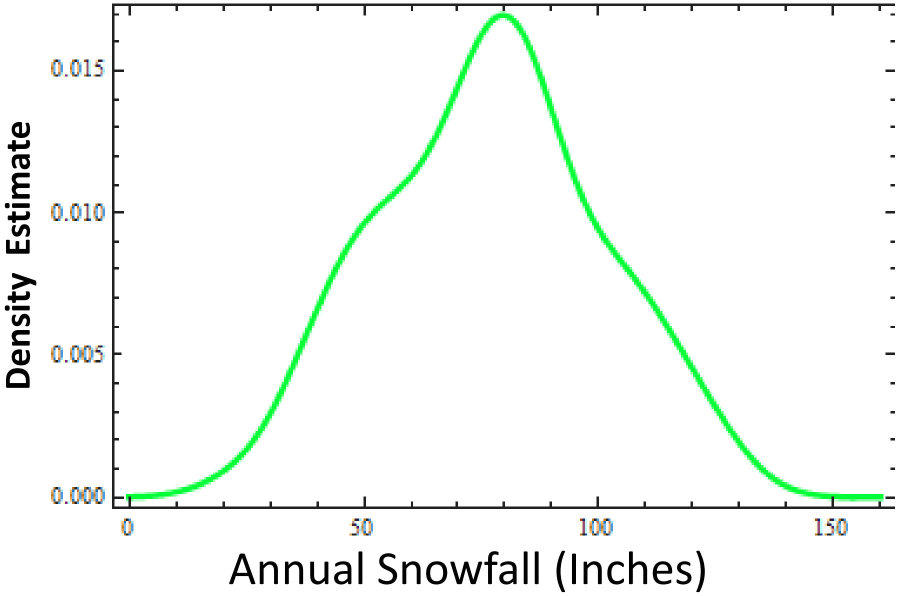

Figure3.2: Kernel Estimate of UCV

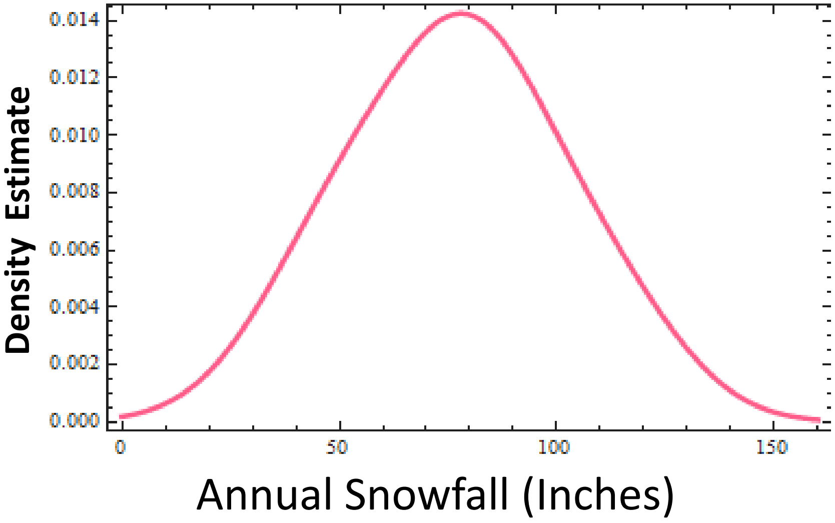

Figure3.3: Kernel Estimate of BCV

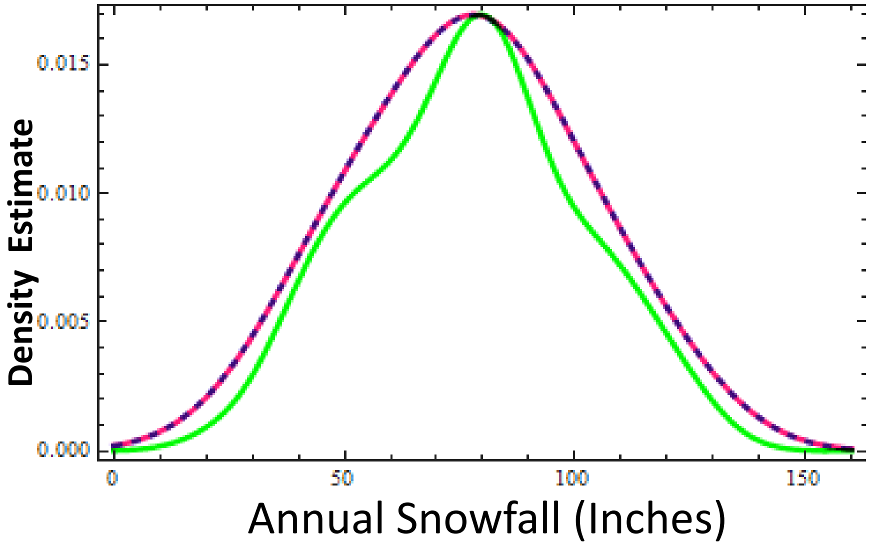

Figure3.4: Kernel Estimate of SBM, UCV and BCV

Table1: Variance, Bias^2 And AMISE For the Data Set.

Methods	hx	Variance	Bias	AMISE
SBM	15.2	0.000294585	0.248053007	0.248347592
UCV	9.01	0.000496969	0.225235149	0.225732118
BCV	15.0	0.000298513	0.247604148	0.247902661

Discussion of Results

The bias cross validation and the bootstrap estimates are similar as can be seen in Figure3.4. The unbiased cross validation is popularly known for selecting smaller values of smoothing parameter. There is a visual evidence of at least two modes in this data as reflected by the UCV estimate. Also, in terms of performance, the UCV yielded the smallest value of the AMISE when compare with the other two smoothing parameter selectors.

REFERENCES.

Parzen, E. (1962). On the Estimation of a Probability Density Function and the Mode. Annals of Math. Statist. **33:** 1065-1076.

Rosenblatt, M. (1956). Remarks on Some Nonparametric Estimates of a Density Function. Annals of Mathematical Statistics. **27:** 832-837.

Schimek, M.J. (2000). A Comparative of Several Smoothing Methods in Density Estimation, Wiley, New York.

Sheather, S. J. (2004). Density Estimation. Statistical Science. **19:** 588–597

Scott, D.W. (1992). Multivariate Density Estimation. Theory, Practice and Visualisation. Wiley, New York.

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

Simonoff, J.S. (1996). Smoothing Methods in Statistics. Springer, New York.

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