Journal of Statistical Research *ISSN 0256 - 422 X*
2008, Vol. 42, No. 2, pp. 181-194
Bangladesh

A GRAVITATIONAL APPROACH TO DENSITY ESTIMATION

Min Jiang
Department of Statistical & Actuarial Sciences
The University of Western Ontario, Canada, London, Ontario, Canada, N6A 5B7
Email: mjiang8@uwo.ca

Serge B. Provost
Department of Statistical & Actuarial Sciences
The University of Western Ontario, Canada, London, Ontario, Canada, N6A 5B7
Email: provost@stats.uwo.ca

summary

It is shown that one can obtain a certain kernel density estimate at a given point x by assuming that the sample points and x are subject to the law of universal gravitation or some variant thereof. Some distributional properties of the resulting kernel are discussed, including the asymptotic normality of a certain rescaled kernel. A two-stage algorithm for the selection of an optimal bandwidth is described. The proposed density estimation technique is applied to three widely used data sets.

Keywords and phrases: Density estimation, Law of universal gravitation, Kernels.

AMS Classification: Primary: 62G07; Secondary: 83C99.

1 Introduction

The methodologies for estimating a density function from a given sample can be classified as parametric or non-parametric. The latter, which include kernel density estimators, are more flexible than the former, which have fixed functional forms. As will be shown in the next section, the proposed gravitational approach to density estimation yields kernel density estimates.

Kernel density estimators smooth out the contribution of each data point over a local neighbourhood of that point by replacing it with a kernel function. The contribution of the point x_i to the estimate at the point x depends on distance between x_i and x, the extent of this contribution depending on the shape of the kernel function and the bandwidth accorded to it. If we denote the kernel function by $\mathcal{K}(\cdot)$ and its bandwidth by δ, the estimated density function at any point x is

$$\hat{f}_\delta(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{\delta}\,\mathcal{K}\Big(\frac{x-x_i}{\delta}\Big)\,.$$

Table 1: Some types of kernels

Kernel	$\mathcal{K}(x)$
Uniform	$\frac{1}{2}\,\mathcal{I}(\lvert x\rvert \le 1)$
Triangle	$(1-\lvert x\rvert)\,\mathcal{I}(\lvert x\rvert \le 1)$
Epanechnikov	$\frac{3}{4}(1-x^2)\,\mathcal{I}(\lvert x\rvert \le 1)$
Biweight	$\frac{15}{16}(1-x^2)^2\,\mathcal{I}(\lvert x\rvert \le 1)$
Triweight	$\frac{35}{32}(1-x^2)^3\,\mathcal{I}(\lvert x\rvert \le 1)$
Cosine	$\frac{\pi}{4}cos(\frac{\pi}{2}x)\,\mathcal{I}(\lvert x\rvert \le 1)$
Gaussian	$\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\;\mathcal{I}(x \in \Re)$

To ensure that the estimates $\hat{f}(x)$ is a bona fide density function, the kernel function $\mathcal{K}(\cdot)$ is usually chosen to be a density function that is symmetric around 0. Several types of kernels are listed in Table 1 where $\mathcal{I}(\cdot)$ denotes the indicator function.

Gravitational approaches have been used in various contexts in the scientific literature. For instance, Huff (1966) assumed that the probability that a user chooses a given service point is directly proportional to its attractivity and inversely proportional to the distance from it. Ottensmann (1995), Sheu (2003), and Stopper and Meyburg (1975) resorted to gravitational models in connection with certain circulation and transportation problems while Fotheringham and O'Kelly (1989) studied migratory flow in terms of such models. Chen and Lee (2001) made use of a gravitation-based algorithm in connection with field data extraction, while Ravi and Gowda (1999) studied the clustering of symbolic objects using a gravitational approach.

The gravitational approach to density estimation is described in Section 2 where it is shown that the resulting estimates are in fact kernel density estimates which are based on a so-called $\mathcal{G}$-kernel. A two-stage plug-in bandwidth selection algorithm is also provided. Some properties of the $\mathcal{G}$-kernel and its connection to other distributions are discussed in Section 3 where a rescaled $\mathcal{G}$-kernel is shown to be asymptotically normally distributed. The proposed density estimates are applied to the Buffalo snowfall, the length of treatment spells and the galaxy velocities data sets.

2 The Density Estimate

2.1 Derivation

The proposed gravitational approach to density estimation is related to a well-known law of physics, namely Newton's law of universal gravitation, which states that every point mass attracts every other point mass by a force directed along the line connecting the two, this force being proportional to the product of the masses and inversely proportional to the square of the distance between the two points. Additional considerations on the law of universal gravitation are discussed for example in Scheurer (1988) and Gregory (2006).

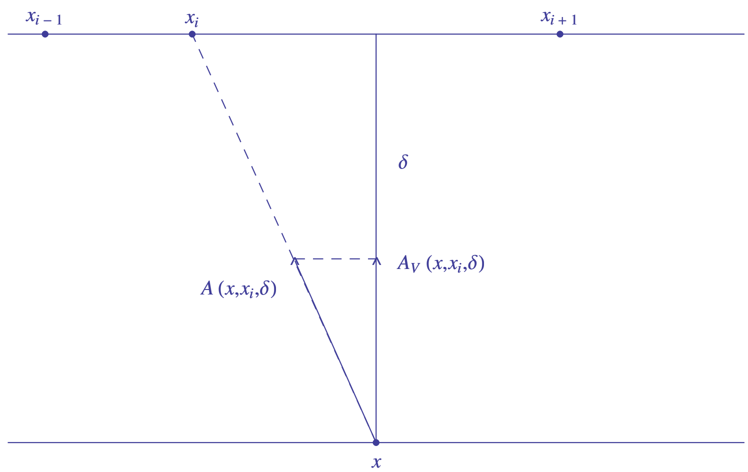

Figure 1: Vertical component of the force of attraction.

Given a random sample $x_1, x_2, \ldots, x_n$, we wish to obtain a density estimate in terms of the combined force of attraction of the sample points. It is assumed that the sample points are lying on a line located at a distance $+\delta$ from the abscissa in the real plane, that is, $(x_i, \delta) \equiv x_i$, $i = 1, \ldots, n$. Letting the force of attraction between x_i and a point $(x, 0) \equiv x$ on the abscissa be denoted by $A(x, x_i, \delta)$, its vertical component, $A_V(x, x_i, \delta)$, clearly satisfies the relationship

$$\frac{A_V(x, x_i, \delta)}{A(x, x_i, \delta)} = \frac{\delta}{\sqrt{\delta^2 + (x - x_i)^2}}\,, \tag{2.1}$$

where $(x - x_i)$ is the horizontal distance between the points x and x_i. This is illustrated in Figure 1. Thus, the combined vertical components of the force of attraction of the points $x_1, \ldots, x_n \equiv \mathbf{x}$ on a point x is given by

$$A_V(x, \mathbf{x}, \delta) = \sum_{i=1}^{n} A_V(x, x_i, \delta)\,,$$

where according to (2.1)

$$A_V(x, x_i, \delta) = \delta\, A(x, x_i, \delta)/\sqrt{\delta^2 + (x - x_i)^2}\,. \tag{2.2}$$

In particular, when the force of attraction is inversely proportional to a power of the square of the distance, one has

$$\frac{k}{(\delta^2 + (x - x_i)^2)^{w'}} \equiv a(x, x_i, \delta) \tag{2.3}$$

where k is a constant. We note that in the physical world, $w' = 1$ and k is the universal gravitational constant. It is assumed that all the points have unit mass. Then, according to (2.1), the vertical component of $a(x, x_i, \delta)$ is

$$a_V(x, x_i, \delta) = \frac{k\,\delta}{(\delta^2 + (x - x_i)^2)^{w'+1/2}} \tag{2.4}$$

and once combined these vertical components of the force of attraction of the points $x_1, x_2, \ldots,\ x_n \equiv \mathbf{x}$ on x add up to

$$a_V(x, \mathbf{x}, \delta) = \sum_{i=1}^{n} a_V(x, x_i, \delta)\,. \tag{2.5}$$

Let us determine the constant k such that $a_V(x, \mathbf{x}, \delta)$ be a *bona fide* density function on the interval $(-\infty, +\infty)$. We observe that in order for $a_V(x, x_i, \delta)$ to be defined, one must have $w' > 0$. Now, letting $w = w' + 1/2$ and noting that

$$\int_{-\infty}^{\infty} \frac{\delta}{(\delta^2 + (x - x_i)^2)^w}\,\mathrm{d}x \;=\; \frac{\sqrt{\pi}\,\Gamma(w - 1/2)}{\Gamma(w)}\,\delta^{2(1-w)}$$

for each x_i, $i = 1, \ldots, n$, one has

$$\begin{aligned} a_V(x, \mathbf{x}, \delta) \;&=\; \frac{1}{n}\sum_{i=1}^{n} \frac{\Gamma(w)}{\delta^{2-2w}\sqrt{\pi}\,\Gamma(-\frac{1}{2} + w)}\,\frac{\delta}{(\delta^2 + (x - x_i)^2)^w} && (2.6)\\ &=\; \frac{1}{n}\sum_{i=1}^{n} \frac{\Gamma(w)}{\sqrt{\pi}\,\Gamma(w - 1/2)}\,\frac{1}{\delta\Big(1 + \big(\frac{x - x_i}{\delta}\big)^2\Big)^w}\,, && (2.7) \end{aligned}$$

which will be denoted by $\hat{f}_\delta(x, w)$. Thus, whenever $w > 1/2$ and

$$k = \Gamma(w)/(n\,\sqrt{\pi}\,\Gamma(w - 1/2))\,, \tag{2.8}$$

$a_V(x, \mathbf{x}, \delta)$ will integrate to one and be positive. Then, on letting

$$g(x, w) = \frac{\Gamma(w)}{\sqrt{\pi}\,\Gamma(w - \frac{1}{2})}\,\frac{1}{(1 + x^2)^w}\;,\quad w > 0.5,\; x \in \Re \tag{2.9}$$

where $g(x, w)$ will be referred to as a $\mathcal{G}$*-kernel* (with parameter w), one has the following representation of the density estimate:

$$\begin{aligned} \hat{f}_\delta(x, w) \;&=\; \frac{1}{n}\sum_{i=1}^{n} \frac{1}{\delta}\, g\Big(\frac{x - x_i}{\delta}, w\Big) && (2.10)\\ &=\; \frac{1}{n}\sum_{i=1}^{n} g_\delta((x - x_i), w) \end{aligned}$$

where

$$g_\delta(x,w) = \frac{\Gamma(w)}{\sqrt{\pi}\Gamma(w-\frac{1}{2})}\frac{1}{\delta(1+(\frac{x}{\delta})^2)^w} \qquad (2.11)$$

will be called a $\mathcal{G}$-*density* (with parameters w and δ). Thus, whenever the force of attraction is as specified in Equation (2.3) with k as given in Equation (2.8), the gravitational approach yields a kernel density estimate wherein the kernel $g(x,w)$ is as specified by Equation (2.9) with $w = w' + 1/2$ and δ corresponds to the bandwidth associated with that kernel.

2.2 Optimal Bandwidth Selection

Given n sample points and the kernel $g(x,w)$ as specified by Equation (2.9), one has to select an appropriate bandwidth δ. The bandwidth in a kernel density estimate should be such that the integrated mean squared error (IMSE), which is defined as

$$IMSE = \int_{-\infty}^{+\infty} E_f[\hat{f}(x) - f(x)]^2 \mathrm{d}x\,, \qquad (2.12)$$

be minimized, see for example Izenman (1991). For any given kernel function $\mathcal{K}(x)$ and unknown density function $f(x)$, the optimal value of δ is

$$\delta_{opt} = k_2^{-2/5}\Big(\int \mathcal{K}(t)^2\mathrm{d}t\Big)^{1/5}\Big(\int f''(x)^2\mathrm{d}x\Big)^{-1/5} n^{-1/5}\,, \qquad (2.13)$$

where $k_2 = \int t^2\mathcal{K}(t)\mathrm{d}t$, see Silverman (1986). It should be noted that the right-hand side of Equation (2.13) cannot be directly evaluated since it involves the unknown density function. However, when the data is approximately normally distributed, the optimal bandwidth can be taken to be

$$\delta_N = k_2^{-2/5}\Big(\int \mathcal{K}(t)^2\Big)^{1/5}\Big(\frac{3}{8}\pi^{-1/2}\sigma^{-5}\Big)^{-1/5} n^{-1/5}\,, \qquad (2.14)$$

where σ is the standard deviation of $f(x)$. When this is not the case, one has to resort to more advanced techniques in order to determine the bandwidth, such as plug-in methodologies or the solution of a bandwidth selector equation, see for instance Sheather and Jones (1991) and Wand and Jones (1995). A version of the two-stage plug-in bandwidth selector is described below.

Two-Stage Plug-In Bandwidth Selection Algorithm

Given a sample $x_1, \ldots, x_n$ and a second-order kernel $\mathcal{K}(x)$, and letting

$$\hat{\psi}_r(g) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{1}{g}\,\mathcal{K}^{(r)}\Big(\frac{x_i - x_j}{g}\Big)$$

be an estimate of $\psi_r = E(f^{(r)}(X))$ where the superscript (r) denotes the r-th derivative, the bandwidth of the kernel density estimate can be determined as follows:

1. Assuming that the true density is normal, estimate ψ_8 by $\hat{\psi}_8^N = 105/(32\sqrt{\pi}\,\hat{\sigma}^9)$, where $\hat{\sigma}$ is the sample standard deviation.

2. Estimate ψ_6 using $\hat{\psi}_6(g_1)$ with

$$g_1 = \Big(\frac{-2\mathcal{K}^{(6)}(0)}{\mu_2(\mathcal{K})\,\psi_8^N\,n}\Big)^{1/9}$$

   where $\mathcal{K}^{(6)}(0)$ is the sixth derivative of kernel function $\mathcal{K}(x)$ evaluated at $x = 0$, $\mu_2(\mathcal{K}) = \int x^2\mathcal{K}(x)\,\mathrm{d}x$ is the second raw moment of the kernel density and n is the size of the sample.

3. Estimate ψ_4 using $\hat{\psi}_4(g_2)$ with

$$g_2 = \Big(\frac{-2\mathcal{K}^{(6)}(0)}{\mu_2(\mathcal{K})\,\psi_6(g_1)\,n}\Big)^{1/7}.$$

4. The two-stage plug-in bandwidth is

$$\hat{\delta}_{PI,2} = \Big(\frac{R(\mathcal{K})}{(\mu_2(\mathcal{K}))^2\,\psi_4(g_2)\,n}\Big)^{1/5} \tag{2.15}$$

   where $R(\mathcal{K}) = \int \mathcal{K}^2(x)\,\mathrm{d}x$.

Note that although the number of stages of functional estimation could be greater than two, two stages appear to provide reasonable density estimates in most cases.

3 Some Properties of the $\mathcal{G}$-kernel

3.1 Basic Distributional Properties

The characteristic function of the $\mathcal{G}$-kernel, $g(x,w)$, as specified by Equation (2.9), which can be obtained by making use of the symbolic computing software *Mathematica* is

$$\phi_X(t) = \frac{2^{3/2-w}\,|t|^{w-1/2}\,\mathrm{BesselK}(1/2-w,|t|)}{\Gamma(w-1/2)}\,,$$

where $\mathrm{BesselK}(n,z)$ denotes a modified Bessel function of the second kind (see for instance Gradshteyn and Ryzhik 2000, p. 703, eqn. 6.649.1). When $w = 1$ (or equivalently $w' = 0.5$), one has

$$g(x,1) = \frac{1}{\pi(1+x^2)}\,,\ x \in \Re\,,$$

which is the density function of a standard Cauchy distribution whose characteristic function is

$$\phi_X(t) = e^{-|t|}\,.$$

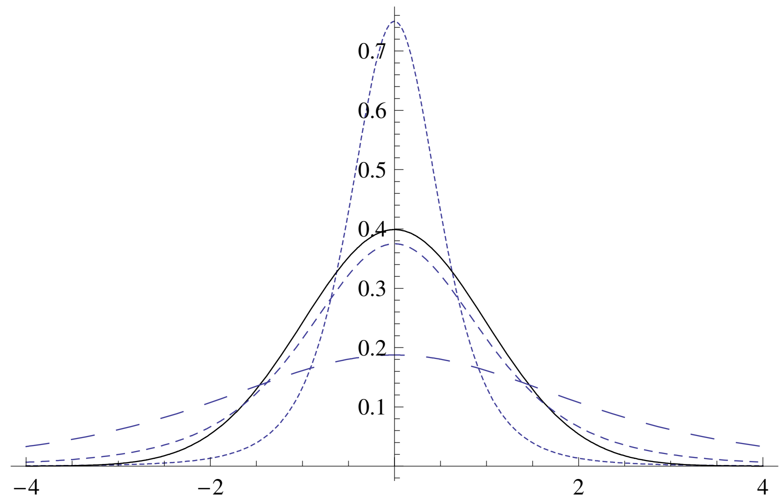

Figure 2: $\mathcal{G}$-densities with $w = 2.5$: δ=1 (short dashes), δ=2 (standard dashes), δ=4 (long dashes); standard normal density (solid line).

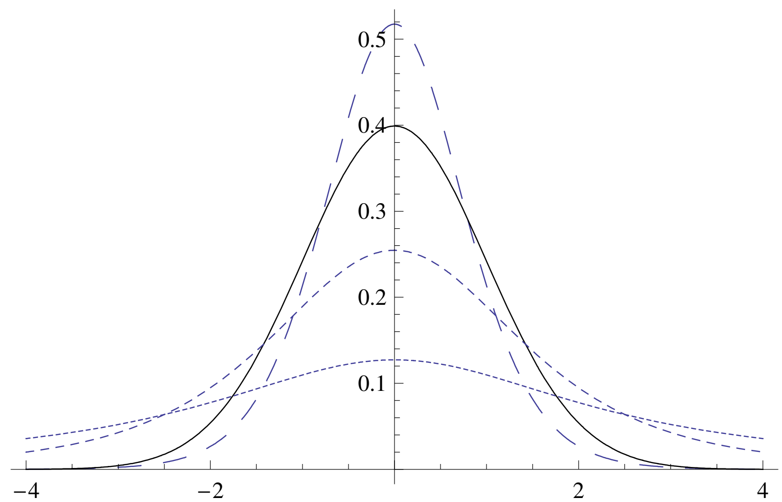

Figure 3: $\mathcal{G}$-densities with $\delta = 2.5$: $w = 1$ (short dashes), $w = 2$ (standard dashes), $w = 6$ (long dashes); standard normal density (solid line).

When $w > \frac{3}{2}$, the mean and variance of the $\mathcal{G}$-kernel are respectively 0 and $\frac{1}{2w-3}$.

Now, consider the $\mathcal{G}$-density, $g_\delta(x, w)$ defined in Equation (2.11) which is a scaled $\mathcal{G}$-kernel. The $\mathcal{G}$-density is plotted in Figure 2 and 3 for various combinations of w and δ to illustrate its behavior with respect to these parameters.

In Figure 2, w is set equal to 2.5 and $\delta = 1$, 2 and 4. It is seen that for a given w, the smaller δ is, the higher the mode. In Figure 3, δ is set equal to 2.5 while $w = 1$, 2 and 6. We note that the smaller w is, the lower the modes and the heavier the tails are.

3.2 Asymptotic Normality of a Certain $\mathcal{G}$-density

In this section, we shall prove that as w goes to $+\infty$, the limiting distribution of the $\mathcal{G}$-kernel, when appropriately rescaled, is that of a standard normal random variable. Let the $\mathcal{G}$-density $g_\delta(x, w) = g(x/\delta, w)/\delta$ be as defined in Equation (2.11). We select δ so that the mode of $g_\delta(x, w)$ coincide with that of a standard normal random variable. Thus, on letting $g_{\delta_N}(0, w) = 1/\sqrt{2\pi}$, we have

$$\frac{\Gamma(w)}{\sqrt{\pi}\Gamma(w-\frac{1}{2})}\frac{1}{\delta_N} = \frac{1}{\sqrt{2\pi}}, \tag{3.1}$$

and so,

$$\delta_N = \frac{\sqrt{2}\,\Gamma(w)}{\Gamma(w-\frac{1}{2})}. \tag{3.2}$$

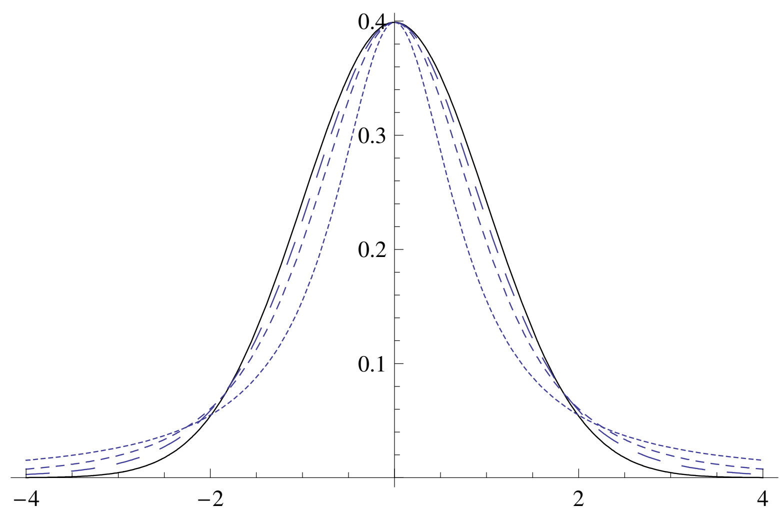

Figure 4: Scaled $\mathcal{G}$-densities: $w = 1$ (short dashes), $w = 2$ (standard dashes), $w = 4$ (long dashes); standard normal density (solid line).

Thus, substituting δ_N into Equation (2.11) yields the following rescaled kernel

$$\begin{aligned} g_{\delta_N}(x, w) &= \frac{\Gamma(w)}{\sqrt{\pi}\Gamma(w-\frac{1}{2})}\frac{1}{\frac{\sqrt{2}\Gamma(w)}{\Gamma(w-\frac{1}{2})}(1+(\frac{x}{\delta_N})^2)^w} \\ &= \frac{1}{\sqrt{2\pi}}\Big(1+\frac{x^2/2}{\big(\Gamma(w)/\Gamma(w-\frac{1}{2})\big)^2}\Big)^{-w}. \end{aligned} \tag{3.3}$$

Using the identity,

$$\frac{\Gamma(n+\frac{1}{2})}{\Gamma(n)} = \sqrt{n}(1 - \frac{1}{8n} + \frac{1}{128n^2} + \cdots\ ),$$

it is seen that

$$\frac{\Gamma(w)}{\Gamma(w-\frac{1}{2})} \to \sqrt{w - 1/2} \quad \text{as} \quad w \to +\infty\,, \tag{3.4}$$

so that

$$\Big(1 + \frac{x^2/2}{(\Gamma(w)/\Gamma(w-\frac{1}{2}))^2}\Big)^w \to e^{\frac{x^2}{2}} \quad \text{as} \quad w \to +\infty\,.$$

Thus

$$g_{\delta_N}(x, w) \to \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad \text{as} \quad w \to +\infty\,.$$

The density function $g_{\delta_N}(x, w)$ is plotted in Figure 4 for $w = 1, 2$ and 4.

3.3 Connection to the Type-2 Beta Distribution

Consider a type-2 beta random variable whose density function is

$$f(x) = \frac{1}{B(\alpha, \beta-\alpha)}\, x^{\alpha-1}(1+x)^{-\beta} \quad \text{for } x > 0\,,$$

where $\alpha > 0$, $\beta - \alpha > 0$ and $B(\alpha, \beta) = \Gamma(\alpha)\,\Gamma(\beta)/\Gamma(\alpha+\beta)$, and apply the mapping $y = x^{1/2}$ for $y > 0$, so that the transformation be one-to-one. Then

$$f(y) = \frac{2}{B(\alpha, \beta-\alpha)}\, y^{2\alpha-1}(1+y^2)^{-\beta},\ y > 0\,,$$

which on letting $\alpha = \frac{1}{2}$ and symmetrizing yields

$$\begin{aligned} g(y) &= \frac{1}{B(\frac{1}{2}, \beta-\frac{1}{2})} \frac{1}{(1+y^2)^\beta} \\ &= \frac{\Gamma(\beta)}{\Gamma(\frac{1}{2})\Gamma(\beta-\frac{1}{2})} \frac{1}{(1+y^2)^\beta}, \quad y \in \Re\,, \end{aligned}$$

that is, the kernel $g(y, \beta)$ as defined in Equation (2.9).

4 Application to Three Data sets

The proposed density estimates $\hat{f}_\delta(x, w)$, as specified by Equation (2.10), is applied to three widely used data set in this section. We observe that w acts as a smoothing parameter: the larger w is, the smoother the resulting density estimate obtained in conjunction with the previously discussed two-stage plug-in bandwidth selection algorithm.

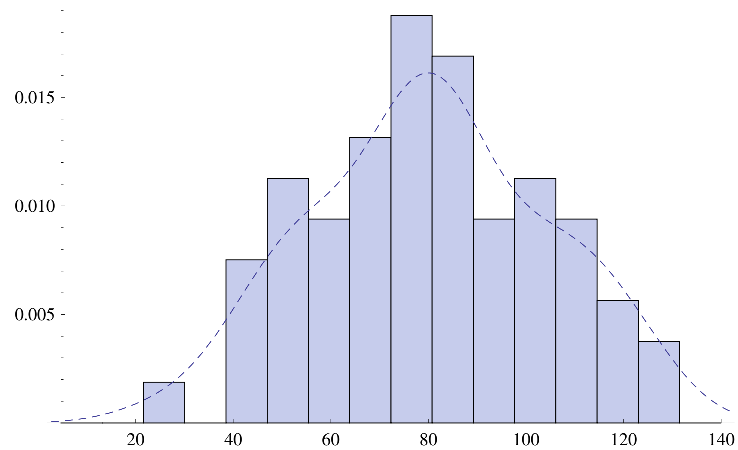

Figure 5: Density estimate $\hat{f}_{69.33}(x, 25)$ and 13-bin histogram (Buffalo snowfall)

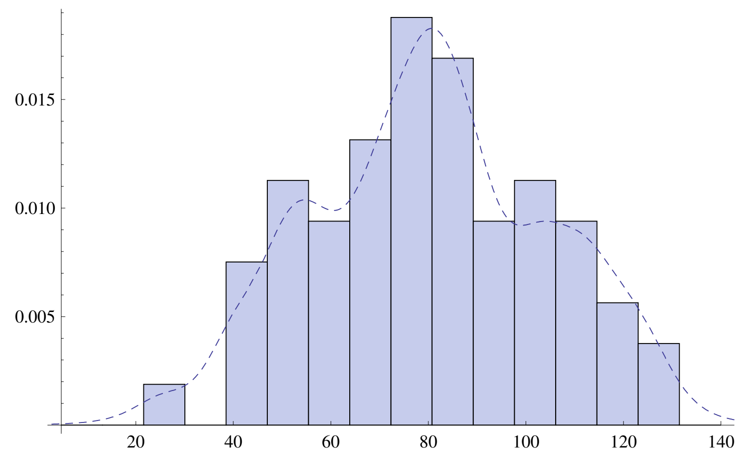

Figure 6: Density estimate $\hat{f}_{12.19}(x, 2.5)$ and 13-bin histogram (Buffalo snowfall)

4.1 The Buffalo Snowfall Data Set

Consider the set of 63 values of annual snowfall precipitations in Buffalo, in inches, for the winters 1910/11 to 1972/73, which is available for instance from the R (or S-Plus) data base. The histogram in Figure 5 suggests that the density function might be trimodal. We made use of the direct two-stage plug-in algorithm described in Section 2.2 to select the optimal bandwidth associated with the $\mathcal{G}$-kernel with parameter $w = 25$, that is, $g(x, 25)$ and determined that $\hat{\delta}_{PI,2} = 69.33$. The resulting density estimate $\hat{f}_{69.33}(x, 25)$ is superimposed on the histogram in Figure 5. A density estimate was similarly obtained assuming that $w = 2.5$. The resulting density function $\hat{f}_{12.19}(x, 2.5)$ is plotted in Figure 6.

4.2 The Length of Treatment Spells Data Set

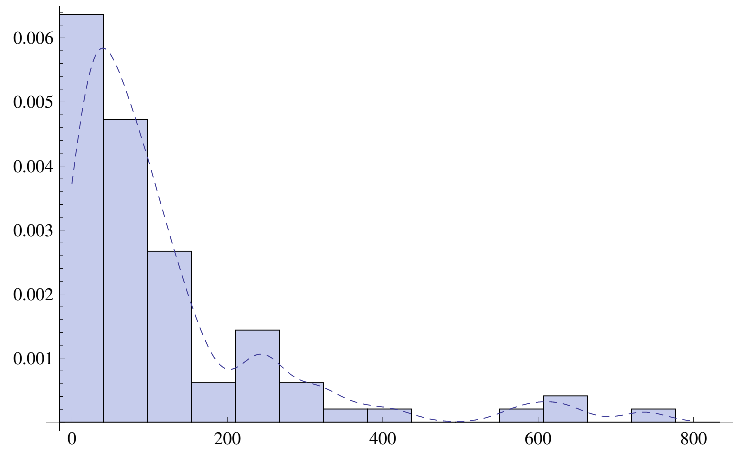

Figure 7: Density estimate $\hat{f}_{207.58}(x, 25)$ and 12-bin histogram (Length of treatment spells)

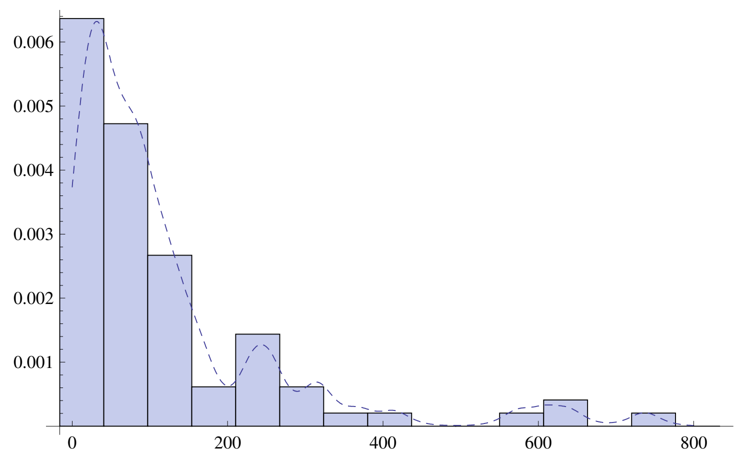

Figure 8: Density estimate $\hat{f}_{41.72}(x, 2.5)$ and 12-bin histogram (Length of treatment spells)

Consider the data set consisting of 86 spells (in days) of psychiatric treatments undergone by a cohort of suicidal patients. This data set has been frequently used to illustrate various density estimation techniques. It was first analyzed by Copas and Fryer (1980) and then discussed by Silverman (1986). It can be seen from the histogram shown in Figure 7 that the distribution has a heavy tail and is strongly right-skewed. Letting $w = 25$ and $w = 2.5$, we used the direct plug-in approach to select the optimal bandwidths which were found to be 207.58 and 41.72, respectively. The resulting density estimates, that is, $\hat{f}_{207.58}(x, 25)$ and

$\hat{f}_{41.72}(x, 2.5)$ are respectively plotted in Figures 7 and 8.

4.3 The Galaxy Velocities Data Set

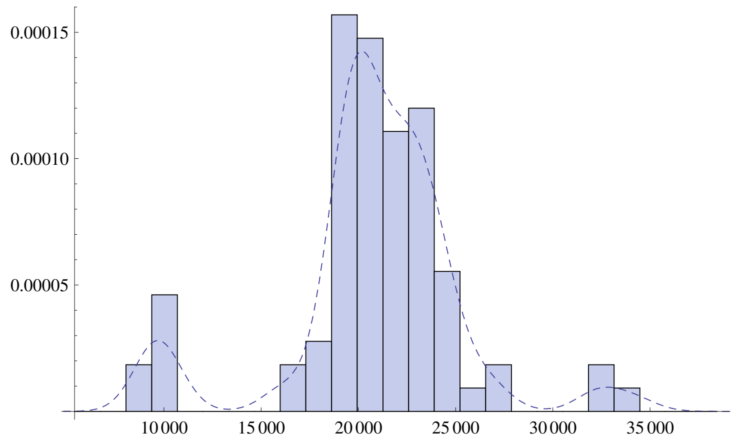

Figure 9: Density estimate $\hat{f}_{7897.87}(x, 25)$ and 19-bin histogram (Galaxy velocities)

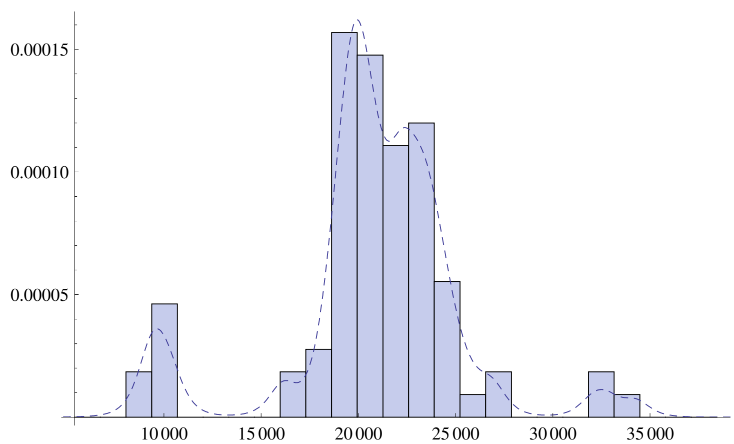

Figure 10: Density estimate $\hat{f}_{1501.47}(x, 2.5)$ and 19-bin histogram (Galaxy velocities)

Consider the data set consisting of 82 galaxies velocities in km/sec from six well-separated conic sections of an unfilled survey of the Corona Borealis region. Multimodality in such surveys is indicative of voids and superclusters in the far universe. This data set which was originally analyzed by Roeder (1990) has served as a benchmark example in mixture analysis.

The histogram clearly indicates that the underlying distribution is multimodal. The density estimates $\hat{f}_{7897.87}(x, 25)$ and $\hat{f}_{1501.47}(x, 2.5)$ whose associated bandwidths were determined by making use the algorithm described in Section 2.2, are respectively plotted in Figures 9 and 10.

Acknowledgements

The financial support of the Natural Sciences and Engineering Research Council of Canada is gratefully acknowledged. We would also like to thank a reviewer for carefully reading this paper and pointing out some typos.

References

[1] Chen, J.L. and Lee, H.J. (2001). Field data extraction for form document processing using a gravitation-based algorithm. *Pattern Recognition*, **34**(9), 1741-1750.

[2] Copas, J.B. and Fryer, M.J. (1980). Density estimation and suicide risks in psychiatric treatment, *Journal of the Royal Statistical Society*, Series A, **143**, 167-176.

[3] Fotheringham, A.S. and O'Kelly, M.E. (1989). *Spatial Interaction Models, Formulations and Applications*, Kluwer Academic Publishers, Boston.

[4] Gradshteyn, I.S. and Ryzhik, I.M. (2000). *Table of Integrals, Series, and Products.* Academic Press, London.

[5] Gregory, R. (2006). *Classical Mechanics*, Cambridge University Press, Cambridge.

[6] Huff, D.L. (1966). A programmed Solution for Approximating an Optimum Retail Location, *Land Economics*, **42**, 293-303.

[7] Izenman, A.J. (1991). Recent development in nonparametric density density estimation. *Journal of the American Statistical Association*, **86**, 205-224.

[8] Ottensmann, J.R. (1995). Using a Gravity Model to Predict Circulation in a Public Library System, *Library and Information Science Research*, **17**, 387-402.

[9] Ravi, T.V. and Chidananda, K.C. (1999). Clustering of symbolic objects using gravitational approach. *IEEE Transactions on Systems, Man, and Cybernetics*, Part B, **29**(6), 888-894.

[10] Roeder, K. (1990). Density estimation with confidence sets exemplified by superclusters and voids in galaxies. *Journal of the American Statistical Association*, **85**, 617-624.

[11] Scheurer, B. and Debrock, G. (1988). *Newton's Scientific and Philosophical Legacy.* Springer, New York.

[12] Sheather, S.J. and Jones, M.C. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society*, Series B, 683-690.

[13] Sheu, J.B. (2003). Locating manufacturing and distribution centers: an integrated supply chain-based spatial interaction approach. *Transportation Research*, Part E, **39**, 381-397.

[14] Silverman, B. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London.

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