Nonparametric Density Estimation

Fan and Yao 2005, Ch.5

WS 2010, 390027 UK Non-linear Time Series Analysis, Prof. Robert Kunst

Florian Kaulich, 0225548

Introduction

- Nonparametric function estimation: Smoothing and densitiy estimation
- Smoothing techniques: Useful graphic tool for summarizing the marginal distribution of a given time series
- Time domain smoothing and State domain smoothing (actual vs. lagged series)
- Smoothing used to estimate spectal density to examine cyclic patterns, white noise of residuals (of a fitted model),…
- Density estimation is the simplest nonparametric function estimation

Kernel density estimation (1)

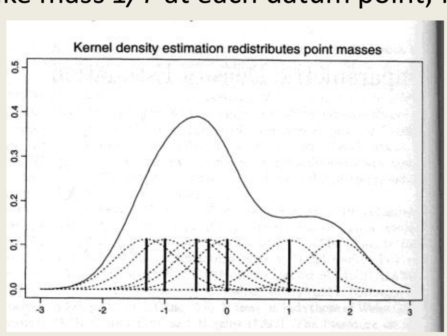
- Used to examine the overall distribution of a data set
- Improvement to histogram method
- Number and locations of peaks/throughs and symmetry of a density
- Empirical distribution function:

$$\hat{F}(x) = \frac{1}{T}\sum_{t=1}^{T} I(X_t \le x)$$

- *T* data points $X_1,\dots,X_T$
- Putting mass *1/T* at each observed datum

Kernel density estimation (2)

- Problem: Nondecreasing function, not very useful in examining the overall structure
- Density function would be better, but density of an empirical function does not exist
- Solution: take mass *1/T* at each datum point, redistribute it, add up

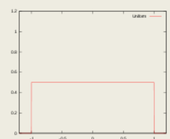


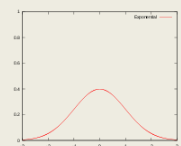
Different kernel functions (1)

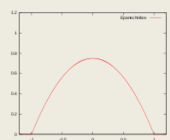
- Kernel density estimate

$$\hat{f}_h(x) = \frac{1}{T}\sum_{t=1}^{T} K_h(X_t - x) = \int K_h(X_t - x) d\hat{F}(u) \qquad K_h(\cdot) = K(\cdot/h)/h$$

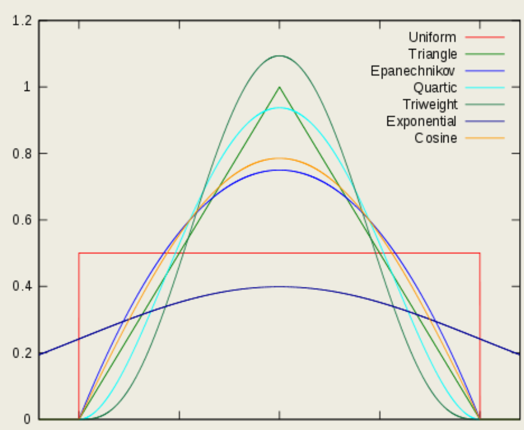
- Uniform $K(u) = \frac{1}{2}\mathbf{1}_{\{|u|\le 1\}}$
- Gaussian $K(u) = \left(\sqrt{2\pi}\right)^{-1} e^{-u^2/2}$
- Epanechnikov $K(u) = \frac{3}{4}(1-u^2)\mathbf{1}_{\{|u|\le 1\}}$
- Biweight $K(u) = \frac{15}{16}(1-u^2)^2\mathbf{1}_{\{|u|\le 1\}}$
- Triweight $K(u) = \frac{35}{32}(1-u^2)^3\mathbf{1}_{\{|u|\le 1\}}$







Different kernel functions (2)



How to do it?

- Practically: Choose Kernel function and bandwidth h
- „It is well-known both empirically and theoretically that the choice ot kernel functions is not very important to the kernel density estimator.“
- … if the kernel is symmetric and unimodal, and if bandwidth *h* is optimally chosen

Optimal bandwidth *h* (1)

- Mean square error (pointwise):

$$MSE(x) \equiv E\left\{\hat{f}_h(x) - f(x)\right\}^2$$

- Mean integrated squared error (global):

$$MISE \equiv E\int_{-\infty}^{\infty}\left\{\hat{f}_h(x) - f(x)\right\}^2 dx$$

- Minimize MISE w.r.t. h:

$$h_{opt} = \alpha(K)\|f''\|_2^{-2/5} T^{-1/5}$$

$$\alpha(K) = \mu_2(K)^{-2/5}\|K\|_2^{2/5}$$

$$\mu_2(K) = \int_{-\infty}^{+\infty} u^2 K(u)\,du \quad \text{is the variance of } K$$

$$\|g\|_2^2 = \int_{-\infty}^{+\infty} g(u)^2\,du \quad \text{is the } L_2-\text{norm}$$

Optimal bandwidth *h* (2)

- Optimal MISE (with optimal bandwidth):

$$\frac{5}{4}\beta(K)\|f''(x)\|_2^{2/5}T^{-4/5} \qquad \beta(K)=\mu_2(K)^{2/5}\|K\|_2^{8/5}$$

TABLE 5.1. Some useful constants related to the kernel functions.

Functional	Gaussian	Uniform	Epanechnikov	Biweight	Triweight
$\mu_2(K)$	1	0.3333	0.2000	0.1429	0.1111
$\|K\|_2^2$	0.2821	0.5000	0.0600	0.7143	0.8159
$\alpha(K)$	0.7764	1.3501	1.7188	2.0362	2.3122
$\beta(K)$	0.3633	0.3701	0.3491	0.3508	0.3529

- Note: β nearly the same
- $h_1 = \dfrac{\alpha(K_1)}{\alpha(K_2)} h_2$

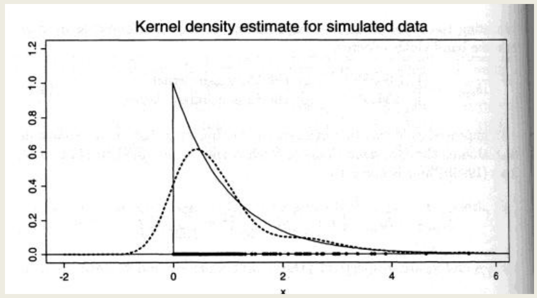
Optimal bandwidth *h* (3)

- Practically, most parameters cannot be observed

$$h_{opt,T} = \left(8\sqrt{\pi}/3\right)^{1/5}\alpha(K)\sigma T^{-1/5}$$

- Replace unknown parameters by observed (sample) measures, solve constants numerically
- $$\hat{h}_{opt,n} = \begin{cases} 1.06sT^{-1/5} & \text{for Gaussian kernel} \\ 2.34sT^{-1/5} & \text{for Epanechnikov kernel} \end{cases}$$
- Good if data is Gaussian, oversmooths if data is asymmetric or multimodal

Boundary Correction



- Reflection: $\hat{f}_h^*(x) = \frac{1}{T}\left\{\sum_{t=1}^{T} K_h(X_t - x) + \sum_{t=1}^{T} K_h(-X_t - x)\right\} \quad x \geq 0$
- Transformation: $\hat{f}_x(x) = g'(x) T^{-1} \sum_{t=1}^{T} K_h(g(x) - g(X_t))$