Introduction Is	sues Kerne	el Estimator	Bandwidth	Bivariate	Cond. Kernel	Compact Kernel
00 0	00		0	0	0	000000

A.II. Kernel Estimation of Densities

Olivier Scaillet

University of Geneva and Swiss Finance Institute

swiss:finance:institute

GENEVA FINANCE RESEARCH INSTITUTE 1/16

Introduction	Issues O	Kernel Estimator	Bandwidth 0	Bivariate 0	Cond. Kernel 0	Compact Kernel
Outli	ine					



- 2 Issues with Empirical Averages
- 3 Kernel Estimator
- Optimal Bandwidth
- 5 Bivariate Kernels
- 6 Conditional Kernels
- 7 Kernels with Compact Support

swiss:finance:institute

GENEVA FINANCE RESEARCH INSTITUTE 2/16



The moments of a random variable are a summary of its distributional behavior.

Full information is provided by its distribution.

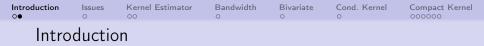
The *cumulative distribution function* for a single asset *i* corresponds to

 $F_i(\zeta_i) = P(Y_{i,t} \leq \zeta_i),$

while for two assets i and j, we have

$$F_{ij}(\zeta_i,\zeta_j) = P(Y_{i,t} \leq \zeta_i, Y_{j,t} \leq \zeta_j).$$

GENEVA FINANCE RESEARCH INSTITUTE 3/16



A cdf may be expressed as an expectation

$$F_i(\zeta_i) = \int_{-\infty}^{\zeta_i} f(y_i) dy_i = \int_{-\infty}^{+\infty} 1_{y_i \leq \zeta_i} f(y_i) dy_i = E[1_{y_i \leq \zeta_i}],$$

where $1_{Y_{i,t} \leq \zeta_i}$ = indicator function of the set $\{Y_{i,t} : Y_{i,t} \leq \zeta_i\}$

$$1_{Y_{i,t} \leq \zeta_i} = \begin{cases} 1 \text{ if } Y_{i,t} \leq \zeta_i \\ 0 \text{ otherwise} \end{cases}$$

GENEVA FINANCE RESEARCH INSTITUTE 4/16

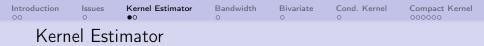
Introduction lssues Kernel Estimator Bandwidth Bivariate Cond. Kernel Compact Kernel

As previously in order to estimate expectations, we need to replace E by an empirical average:

$$\hat{F}_i(\zeta_i) = \frac{1}{T} \sum_{t=1}^T \mathbb{1}_{y_{i,t} \leq \zeta_i},$$

$$\hat{F}_{ij}(\zeta_i,\zeta_j) = \frac{1}{T}\sum_{t=1}^T \mathbf{1}_{y_{i,t} \leq \zeta_i, y_{j,t} \leq \zeta_j},$$

 $\implies \text{We obtain step functions which are$ *not* $differentiable.}$ $\implies \text{We cannot build empirical counterparts of densities, i.e.,}$ $f_i(\zeta_i) = \frac{dF_i(y_i)}{dy_i} \mid_{y_i = \zeta_i}$



To build empirical counterparts of densities, we rely on kernel estimation.

<u>Idea behind:</u> We start from the histogram,

$$\hat{f}_i(\zeta_i) = rac{1}{T}\sum_{t=1}^T \mathbb{1}_{y_{i,t}=\zeta_i}$$

and replace bars by smooth bumps

$$\hat{f}_i(\zeta_i) = rac{1}{Th} \sum_{t=1}^T K\left(rac{y_{i,t}-\zeta_i}{h}
ight)$$

The bump K is called a *Kernel*. It should be positive and integrate $SWISS: fit \Omega \in One Content C$

GENEVA FINANCE RESEARCH INSTITUTE 6/16



Gaussian Kernel = Gaussian density

$$K\left(\frac{y_{i,t}-\zeta_i}{h}\right) = \frac{1}{\sqrt{2\pi}}\exp{-\frac{1}{2}\left(\frac{y_{i,t}-\zeta_i}{h}\right)^2}$$

GENEVA FIN

The smoothing parameter *h* is called the *bandwidth*.

The bandwidth h plays the same role as the class length for histograms.



If h is too large (large class), we get *oversmoothing*.

If h is too small (small class), we get *undersmoothing*.

Rule of thumb to select the bandwidth:

 $h = \hat{\sigma} T^{-1/5}$

DESEADCH

where $\hat{\sigma}$ is empirical standard deviation of the data.

Introduction Issues Kernel Estimator Bandwidth Bivariate Cond. Kernel Compact Kernel Condoc Kernel Compact Kern

It is possible to extend to higher dimensions and to the conditional case.

$$\hat{f}_{ij}(\zeta_i,\zeta_j) = \frac{1}{Th^2} \sum_{t=1}^T K\left(\frac{y_{i,t}-\zeta_i}{h}\right) K\left(\frac{y_{j,t}-\zeta_j}{h}\right)$$

Note that the *curse of dimensionality* appears when we are above five dimensions.

We need a lot of information (data) to get an accurate estimation of the high dimensional object to be estimated.



Recall the definition (Bayes Theorem)
$$f(\zeta_i | y_{j,t} = \zeta_j) = \frac{f_{ij}(\zeta_i, \zeta_j)}{f_j(\zeta_j)}$$

 \Longrightarrow we only need to replace the unknown quantities by their estimates

$$\hat{f}(\zeta_i | y_{j,t} = \zeta_j) = \frac{f_{ij}(\zeta_i, \zeta_j)}{\hat{f}_j(\zeta_j)}$$



Previous estimators have good properties when the data take values in $\Re.$

When data are bounded from below at zero (losses with a positive sign), they exhibit boundary bias (edge effect).

This *boundary bias* is due to weight allocation by the fixed symmetric kernel outside the density support when smoothing is carried out near the boundary.

DESEADCHI



One of the remedy consists in replacing symmetric kernels by asymmetric kernels, which never assigns weight outside the support.

The form of the estimators is the same

$$\hat{f}_i(\zeta_i) = \frac{1}{Th} \sum_{t=1}^T K(y_{i,t};\zeta_i,h)$$

but K is replaced by an asymmetric kernel.



Gamma Kernel:

$$K(y;\zeta,h) = \frac{y^{\varsigma/h}e^{-y/h}}{h^{\varsigma/h+1}\Gamma(\varsigma/h+1)}$$

where
$$\Gamma(x) = \int_0^\infty e^{-u} u^{x-1} du$$

<u>Reciprocal Inverse Gaussian Kernel:</u> $\mathcal{K}(y;\zeta,h) = \frac{1}{\sqrt{2\pi h y}} \exp\left(-\frac{\varsigma - h}{2h}\left(\frac{y}{\varsigma - h} - 2 + \frac{\varsigma - h}{y}\right)\right)$

swiss:finance:institute

GENEVA FINANCE RESEARCH INSTITUTE 13/16



When the data are defined on [0, 1], we face two boundaries.

It is then useful to use a kernel whose support is also [0, 1], for example the Beta kernel:

$$K(y;\zeta,h)=\frac{1}{B(\varsigma/h+1,(1-\varsigma)/h+1)}y^{\varsigma/h}(1-y)^{(1-\varsigma)/h}$$

where $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$.

swiss:finance:institute

GENEVA FINANCE RESEARCH INSTITUTE 14/16



This estimator is useful to analyze the distribution of recovery rates at default.

There is a renewed interest in LGD (loss given default), which is mainly prompted by Basle II and the explosion of the credit derivatives market.

Data are scarce, in particular outside the US. The market standard to model LGD is a parametric assumption of beta distributed recoveries.

There are several measures of LGD

- \rightarrow ultimate recoveries
- \rightarrow trading price recoveries



These measures often give very different results. Which one should be used depends who you are and what you do with your defaulted positions.

The data concern 623 US defaulted bond issues spanning form 1981 to end 1999. These are trading price recoveries which are classified by industry and seniority.

The data comes from the S&P/PMD database.

The market assumption of a beta distribution is often severely wrong. This could lead to underestimation of risk measures.