Non-Parametric Density Estimation

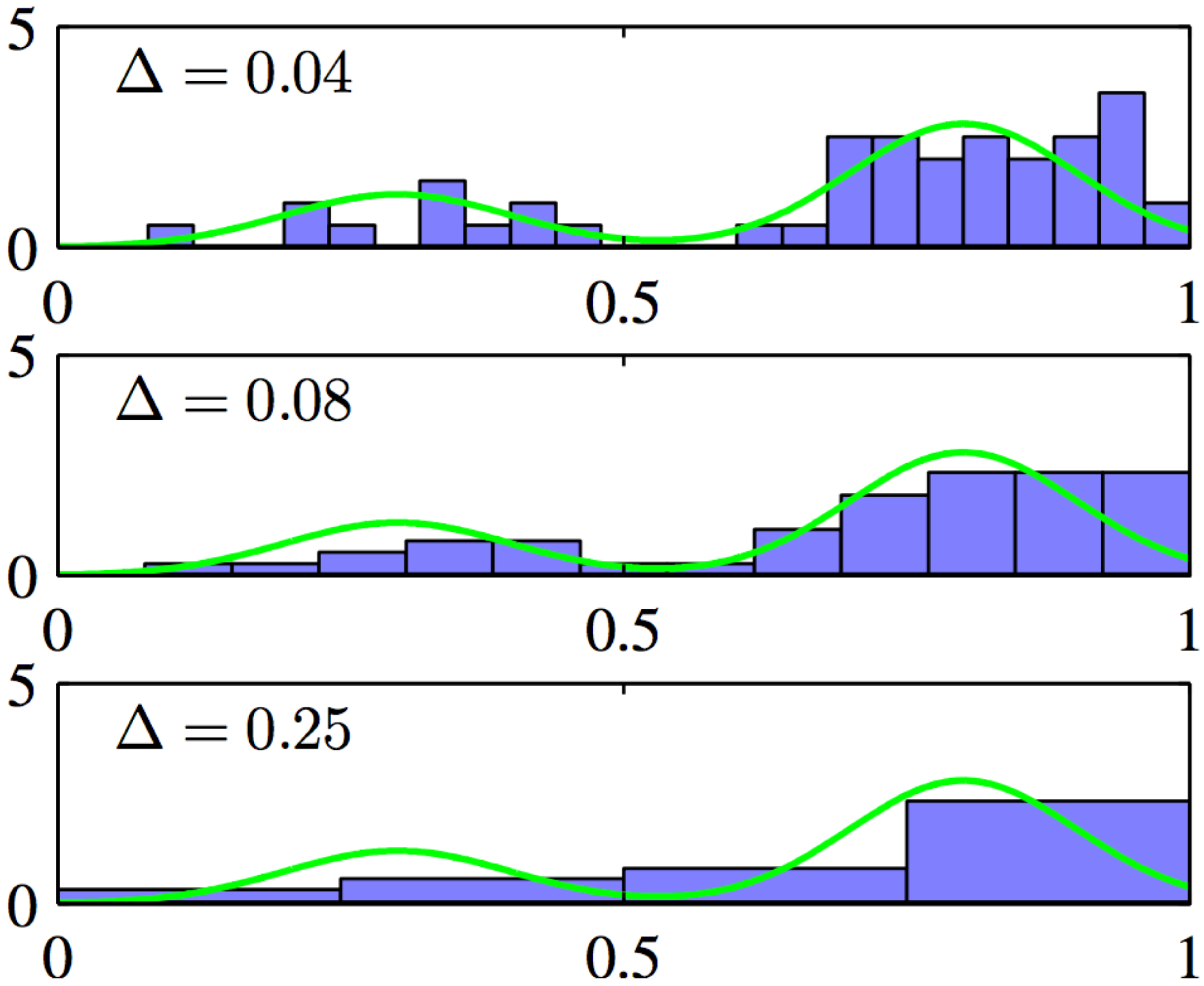
Histogram for Density Estimation

Density of r.v. $x \in \mathbb{R}$:

$$p(x{=}a) = \lim_{\Delta \to 0} \frac{P\big(x{\in}[a, a{+}\Delta]\big)}{\Delta}$$

$$p(x) \approx \frac{K}{NV}$$

V is fixed

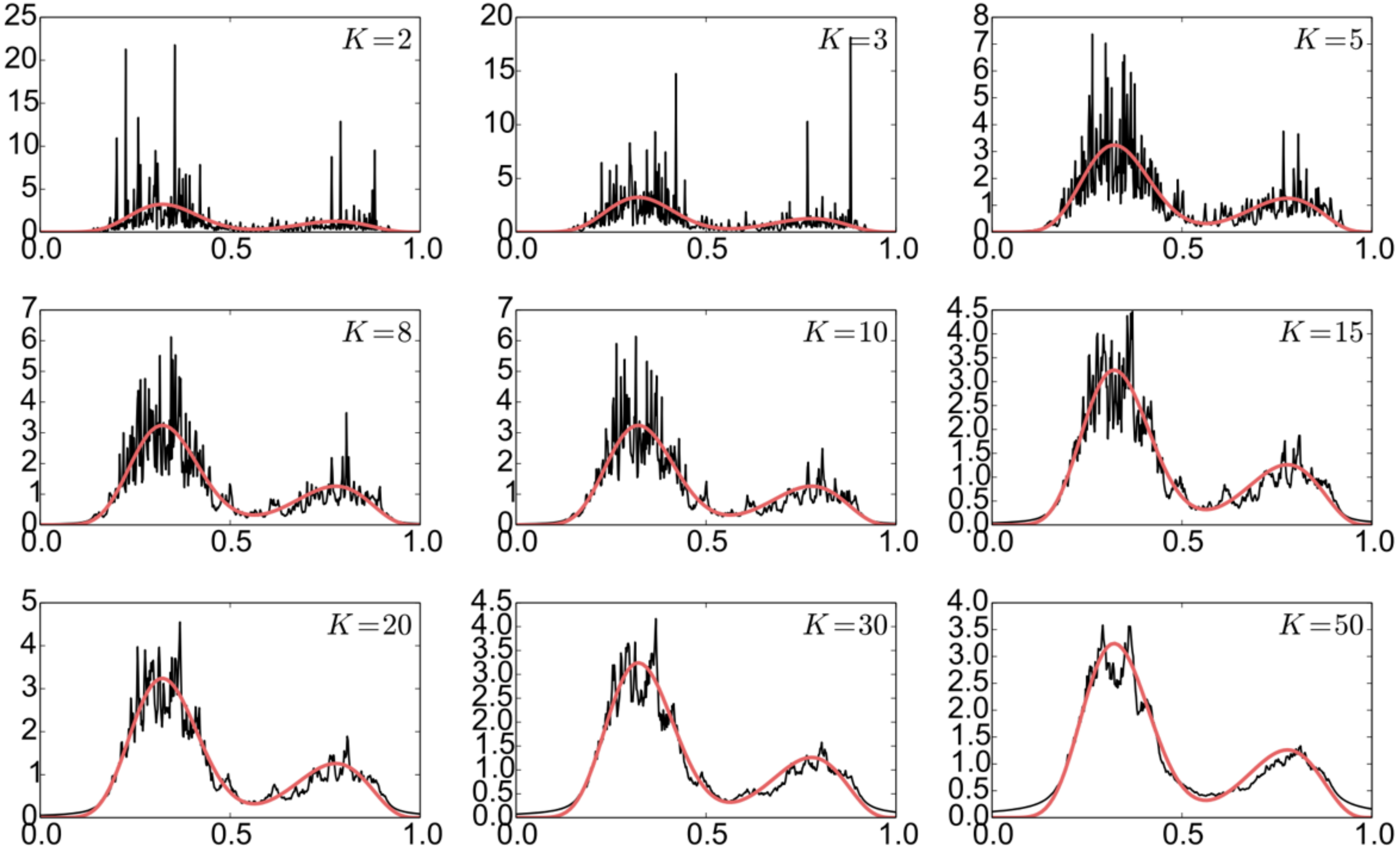


Converges to true density in the limit $N \to \infty$ and shrinking V appropriately

Find K neighbors, the density estimate is then $p \sim 1/V$ where V is the volume of a minimum cell containing K NNs. Example ($p \sim$ inverse distance to K-th NN, same 1000 samples as before):

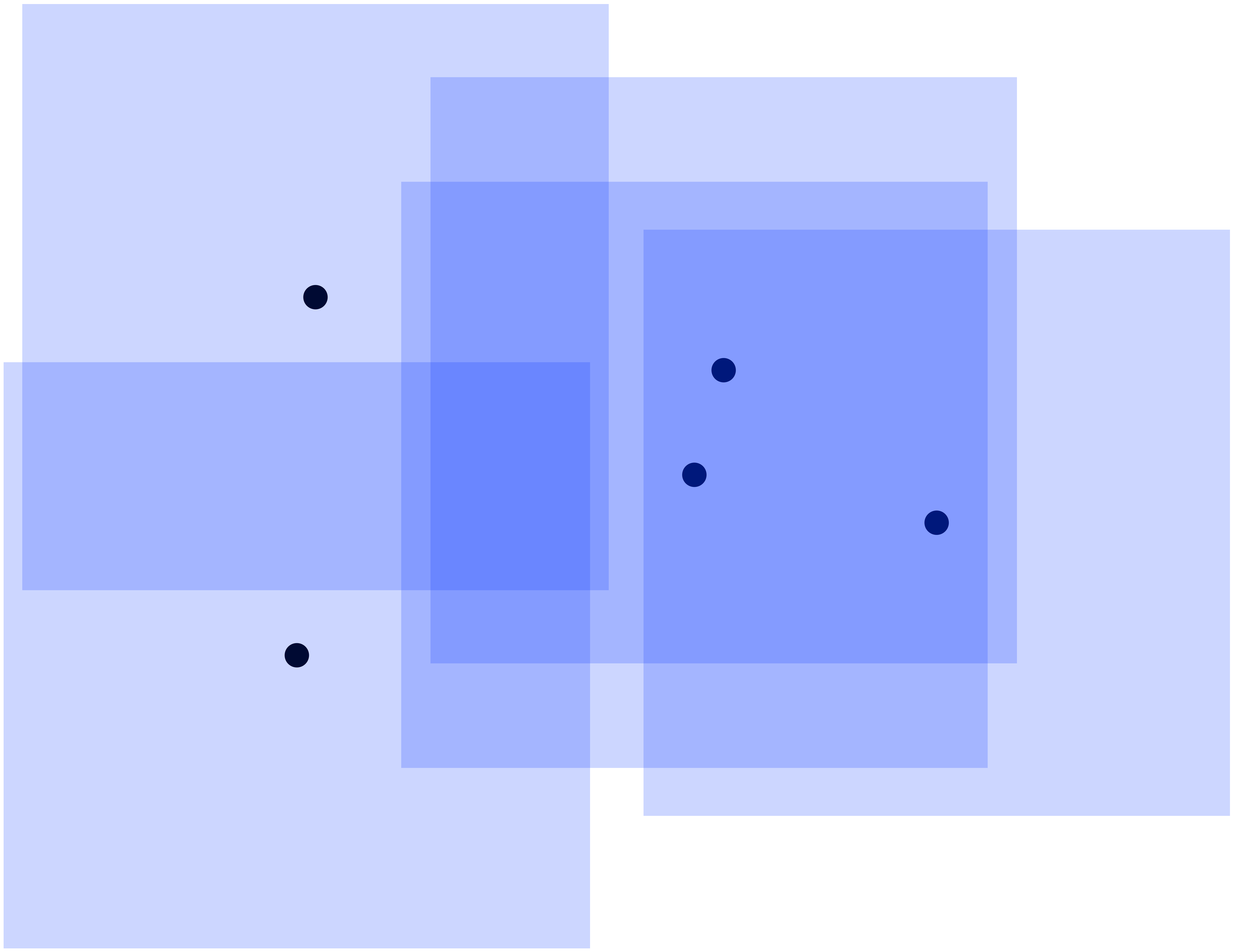
$$p(x) \approx \frac{K}{NV}$$

K is fixed



Converges to true density in the limit $N \to \infty$ provided that K is increased appropriately

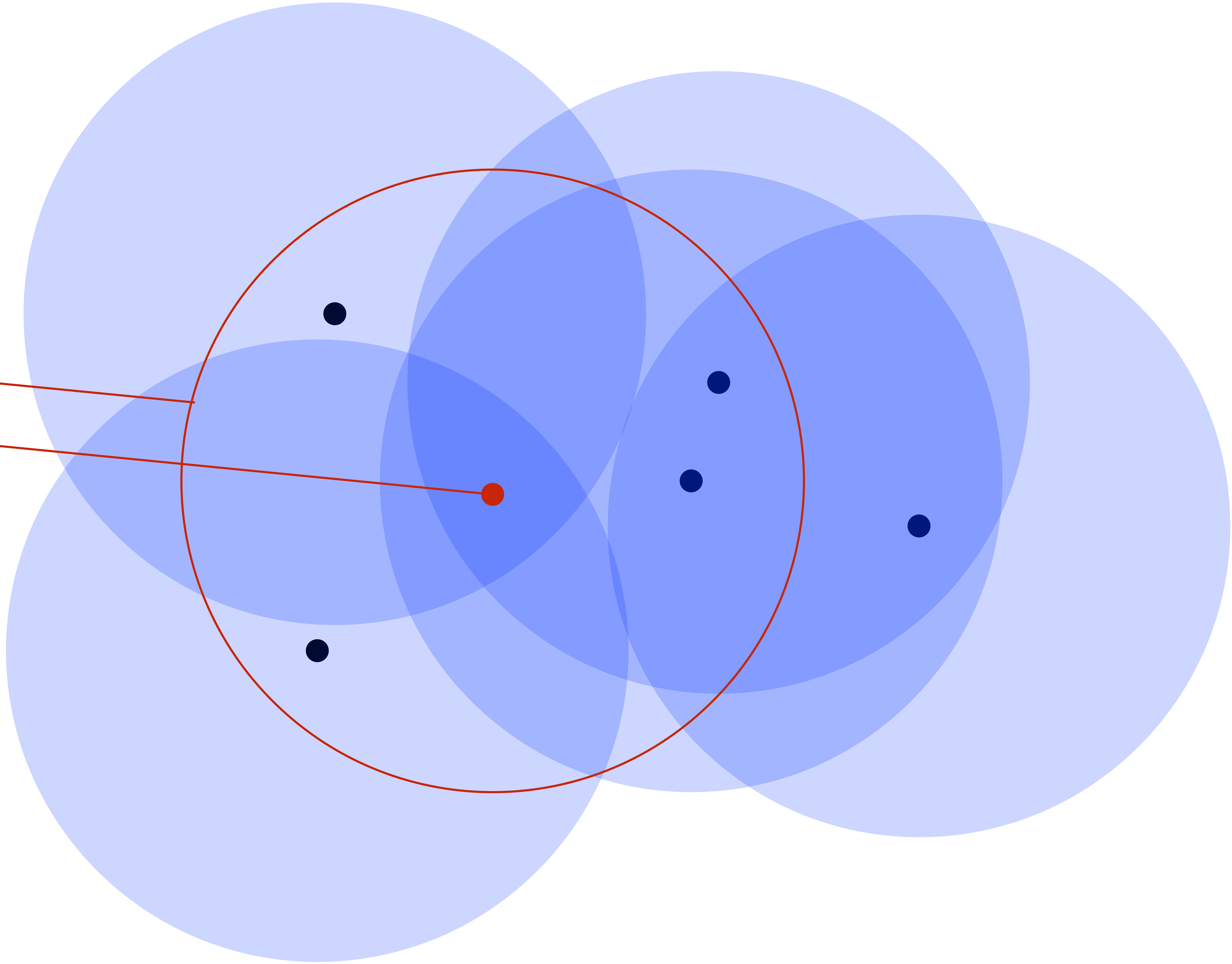
Kernel Density Estimate



Kernel Density Estimate

4 points inside ball

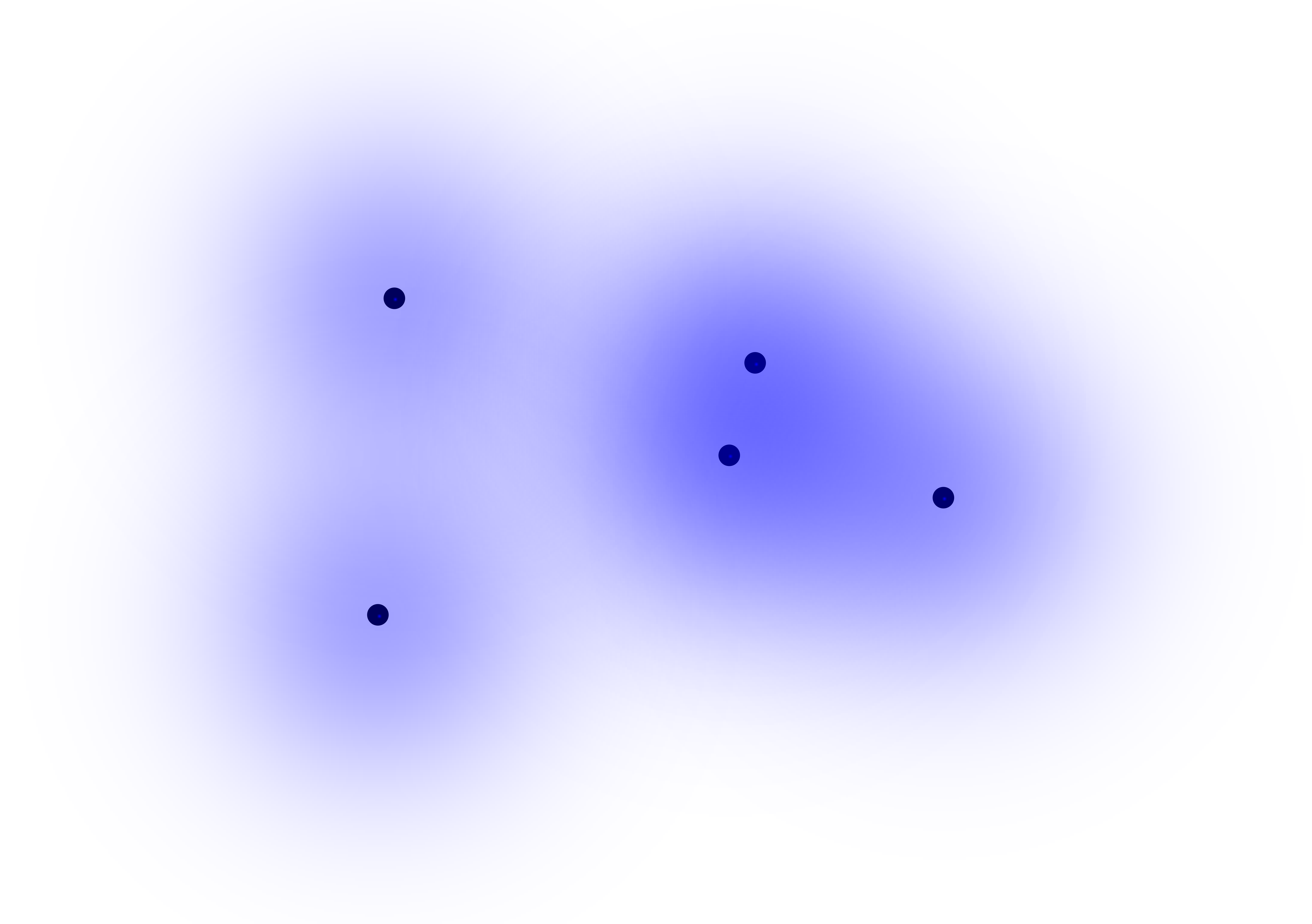
4 copies of kernel overlap



$$p(x) \approx \frac{K}{NV}$$

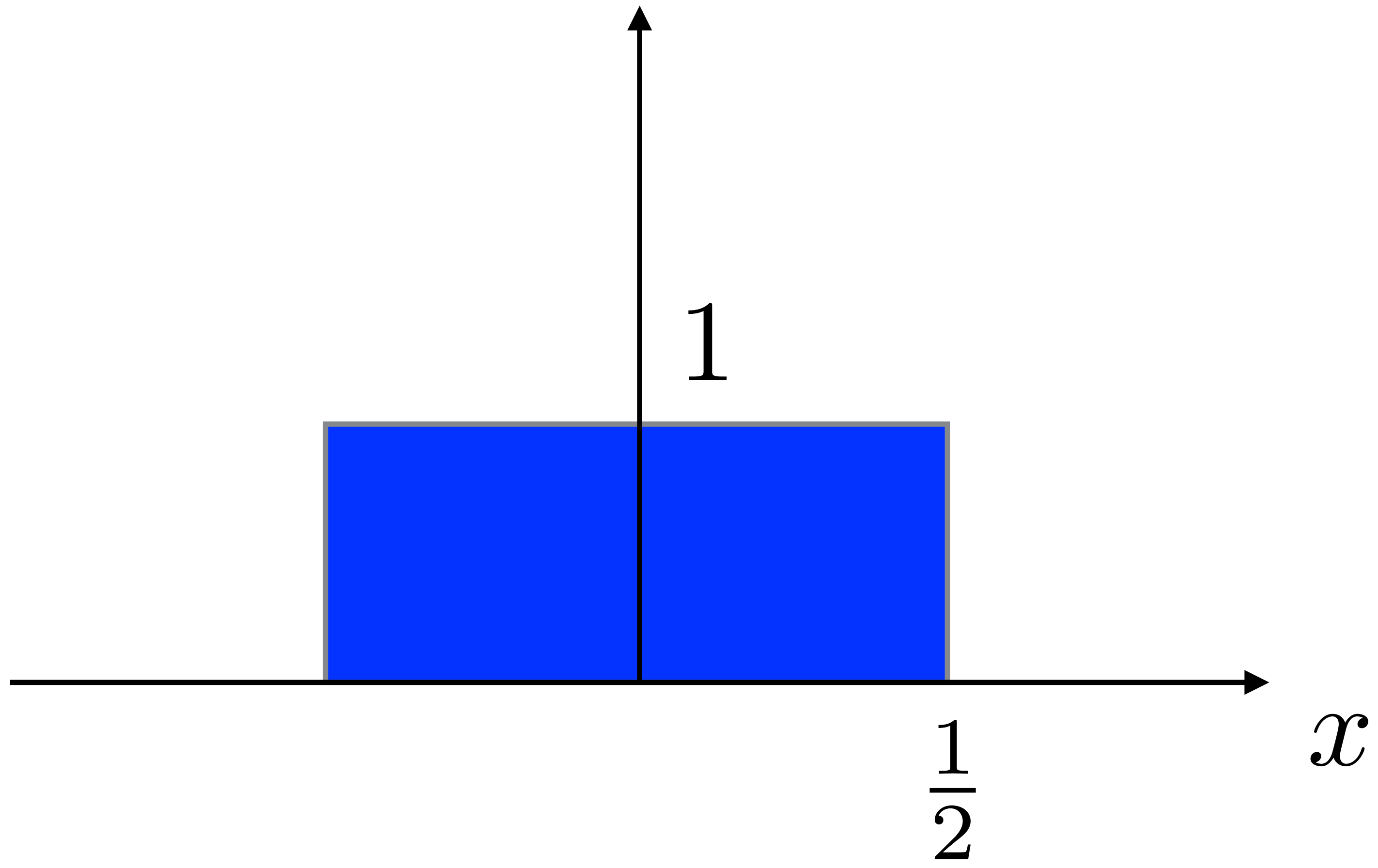
Kernel Density Estimate

Kernel Density Estimate



Kernel Density Estimation

Kernel



In this context "Kernel" = "density":

$K : \mathcal{X} \mapsto \mathbb{R}$

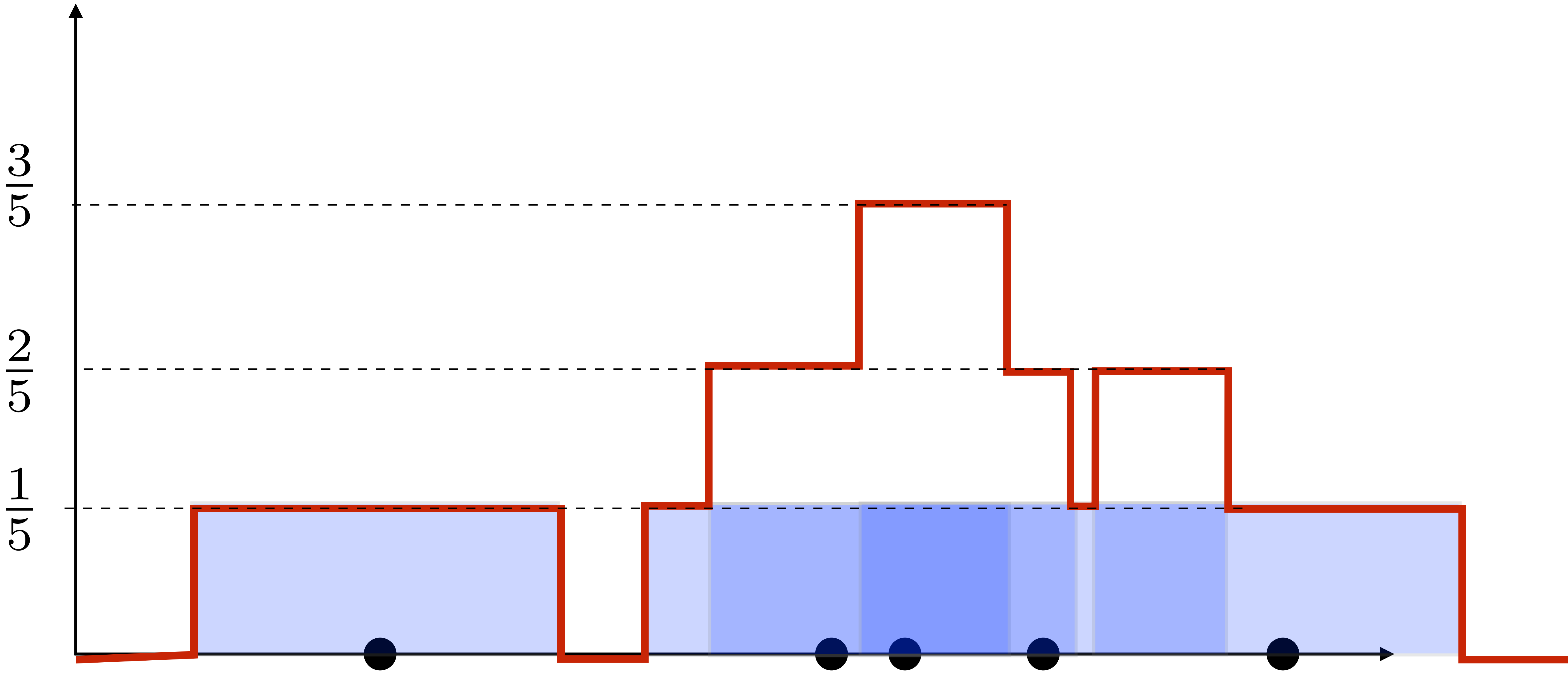
$K(x) \geq 0$

$\int K(x)dx = 1$

Shifted and scaled: $K_h(x, y) = K((x - y)/h)$

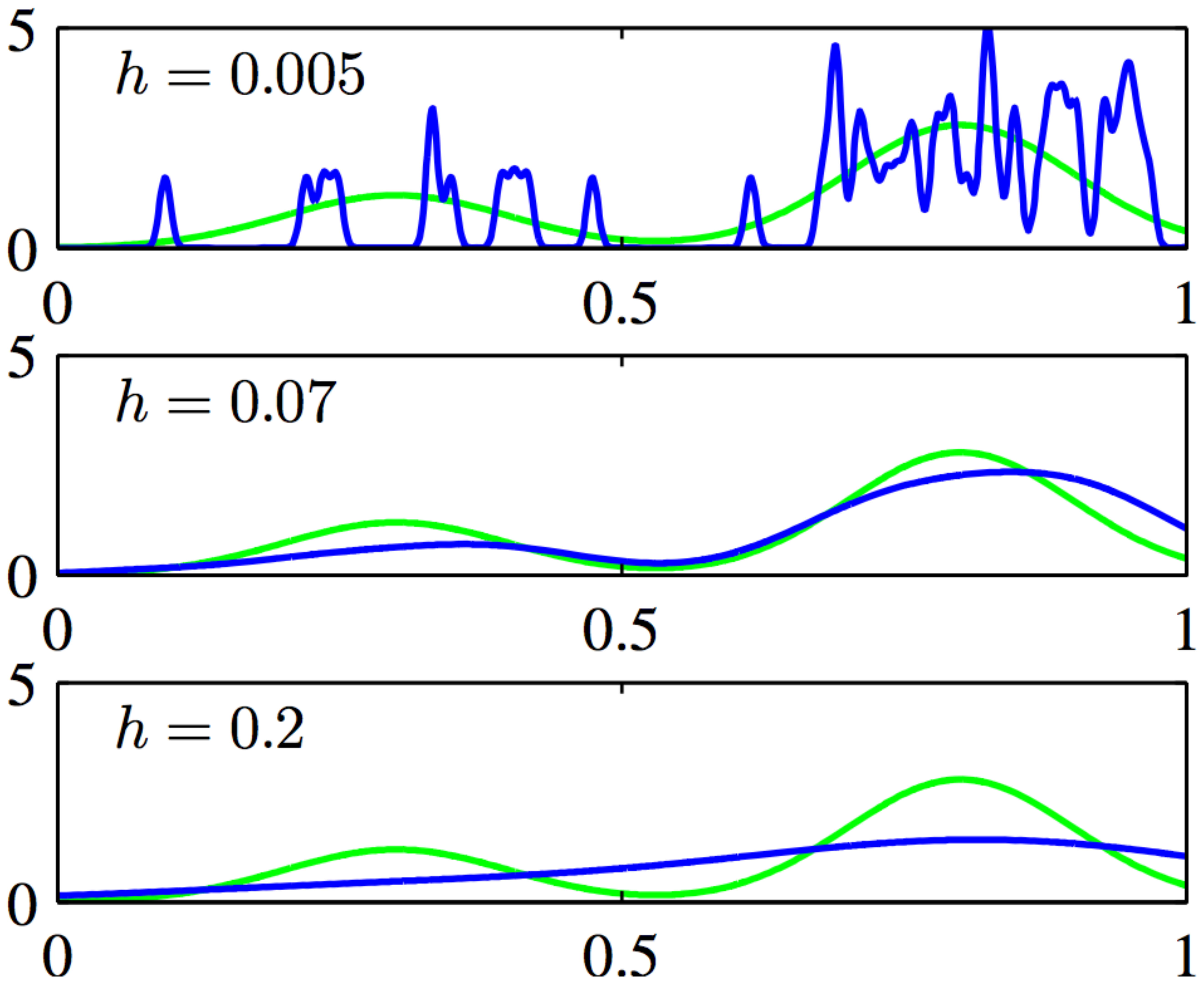
Scale parameter h, e.g. standard deviation

Sum of Kernels at different shifts:



$$p(y) = \sum_{i=1}^{N} \frac{1}{N} K_h(y - x_i)$$

Kernel Density Estimation



Histogram and Kernel Density Estimation in 2D