Kernel density estimates

- Useful literature:
  - A good introductory discussion of the estimation of the [a] bandwidth parameter, [b] kernel densities and [c] relative risk ratio analysis can be found in Waller & Gotway (2004). *Applied Spatial Statistics for Public Health Data*. Wiley on pp 130-136 and pp 164-171.
    It is available as eBook in UTD's library.
  - An extended resource is Baddeley et al. (2016). *Spatial Point Patterns. Methodology and Applications with R*. CRC Press
    It centers around the powerful library **spatstat**, which we will be using in the first part of this course.
    The associated web-site is www.spatstat.org .
    An earlier draft manuscript of the **spatstat** book can be found in the document **Rspatialcourse_CMIS_PDF Standard.pdf**.
- The intensity at any pivot location $\mathbf{s}$ in the study area $\Re$, i.e., $\mathbf{s} = (x, y)^T \in \Re$, can be estimated by

$$\hat{\lambda}(\mathbf{s}) = \frac{1}{\delta_\tau(\mathbf{s})} \cdot \sum_{i=1}^{n} \frac{1}{\tau^2} \cdot k\left(\frac{\mathbf{s} - \mathbf{s}_i}{\tau}\right)$$

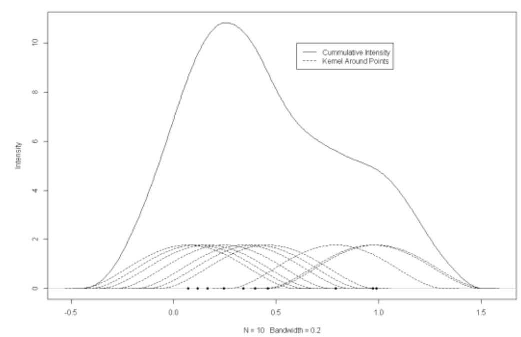
  where the $\mathbf{s}_i$'s are the observed point locations and $\mathbf{s}$ are the roaming locations at which the densities are calculated.

  - In essence, it measures the contribution of the points $\mathbf{s}_i$ in a ***disk*** around the wandering point $\mathbf{s}$. Other shapes such as oriented ellipsoids are also possible.
- The function $k(\cdot)$ is a kernel density function such as the ***quartic*** bivariate density with

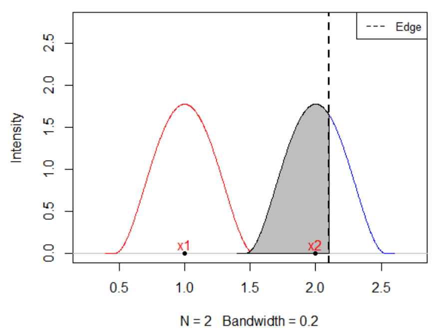
$$\mathbf{u} \equiv \frac{\mathbf{s} - \mathbf{s}_i}{\tau}$$

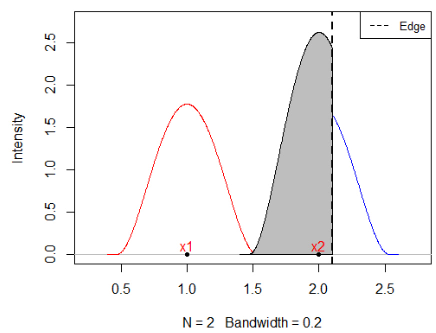
$$k(\mathbf{u}) = \begin{cases} \frac{3}{\pi} \cdot (1 - \mathbf{u}^T \cdot \mathbf{u}) & \text{for } \mathbf{u}^T \cdot \mathbf{u} \le 1 \\ 0 & \text{otherwise} \end{cases}$$

- Note: `spatstat` allows of a range of kernel function.
- The kernels around the reference locations $\mathbf{s}_i$ are combined into the kernel density (see also the script **`KernelDensityExplained.R`**):



- The normalizing edge correction factor is $1/\delta_\tau(\mathbf{s})$ ensures that the intensity integrates within the study area to a proper distribution function (the area underneath the curve is one) and edge cells are not leaking probability mass outside the study area.





- The edge correction factor is $1/\delta_\tau(\mathbf{s})$ increases the intensity for locations $\mathbf{s}$ at the edge because there are potentially less locations $\mathbf{s}_i$ in its vicinity.
  In the **`spatstat::density( )`** function it is implemented as the Diggle adjustment.
- The ***bandwidth parameter*** τ determines the degree of smoothing

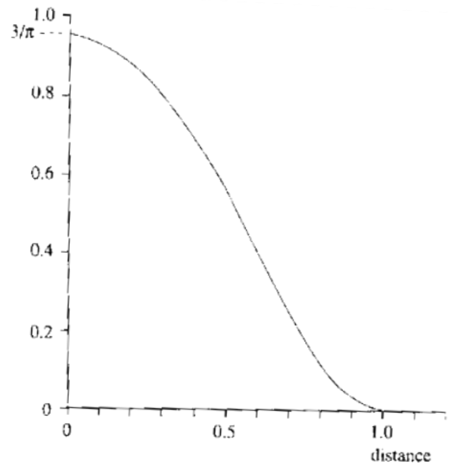


Fig. 3.3 Slice through a quartic kernel

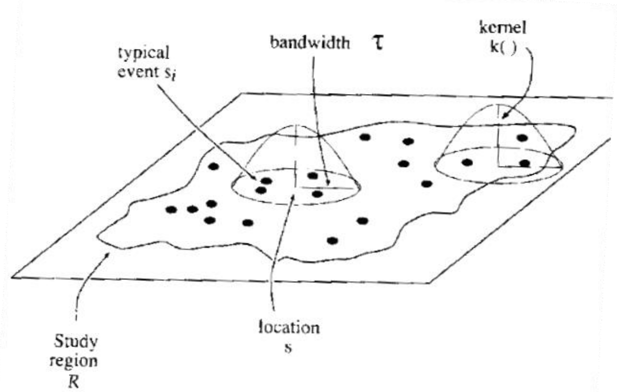


Fig. 3.4 Kernel estimation of a point pattern

For Gaussian kernels approximate $2/3$ of the probability mass falls with a radius τ around the reference point $\mathbf{s}_i$.

- Experiment with the bandwidth and kernel density functions in **`KernelDensityExplained.R`**
- In a GIS setting, the kernel density is evaluated over a fine grid in order to give a smooth surface representation.
- Several algorithms based on varying assumptions are can be used to estimate an "optimal" bandwidth.
  Ultimately, however, the smoothing effect should be decided by logical arguments, visual inspection and experimentation after using the estimated bandwidth as initial guess:
  - Over-smoothing may lead to masking of any relevant point concentrations in the spatial configuration.
  - Under-smoothing will exhibit a granular surface with too much detail and perhaps subregions with zero density.
- Adaptive kernel density estimates can be found in the library **`sparr`**. Adaptive kernel estimates adjust the bandwidths $\tau(\mathbf{s}_i)$ of the underlying intensity surface at location $\mathbf{s}_i$.
  - At locations of high point density the bandwidth should be small in order to preserve regional details.
  - At locations of low density, a large bandwidth is required to pull a sufficient number of points into the kernel.

$\tau(s_i) = \tau_0 \cdot \left(\dfrac{\tilde{\lambda}_g}{\tilde{\lambda}(\mathbf{s}_i)} \right)^{\alpha}$ where $\tilde{\lambda}(\mathbf{s}_i)$ is some initial estimate at τ_0 and $\tilde{\lambda}_g = \sqrt[n]{\tilde{\lambda}(\mathbf{s}_1) \cdot \tilde{\lambda}(\mathbf{s}_2) \cdots \tilde{\lambda}(\mathbf{s}_n)}$ is the geometric mean of the initial estimates at the *n* event locations.

- The parameter α controls the degree of local adaptiveness, i.e., $\alpha = 0$ leads to a global bandwidth estimate.