K function

K function

We previously looked at the $F(d)$ and $G(d)$ functions, which corresponded to *CDFs, as a function of distance, for open space and distance between points.*

Another interesting feature of a point process is the number of points in a specified area. Consider $E(Num(\boldsymbol{s}, d, \boldsymbol{S}))$, the expected number of points in $\delta_d \boldsymbol{s}$, a circle of radius d centered at $\boldsymbol{s}$.

Ripley's K or just the K function, considers the expected number of points within a distance d of an arbitrary point. Formally this is defined for CSR as

$$K(d) = \frac{E(\text{number of points within d})}{\lambda}$$

. In other words, this is scaled by λ

With CSR, $K(d) = \frac{\lambda \pi d^2}{\lambda} = \lambda d^2$.

To estimate $K(d)$, *we use*

$$\hat{K}(d) = (\hat{\lambda})^{-1} \sum_i \sum_j 1(||\boldsymbol{s}_i - \boldsymbol{s}_j|| \leq d)/n$$

where $\hat{\lambda} = n/|\mathcal{D}|$, note the typo in the book.

The empirical K statistic is compared with πd^2. *For $K > \pi d^2$, the series exhibits clustering, for $K < \pi d^2$ the process exhibits inhibition.*

Recall the dataset with medieval grave site information.

```r
data(grave)
plot(envelope(grave, Kest, verbose = F))
```