Edge Corrections for Spatial Point Processes

September 22, 2004

Review: $N_B,\ B \subseteq I\!R^p$ denotes a stationary point process with intensity λ. Thus,

$$E[N_B] = \lambda \nu_p(B),$$

where $\nu_p(B)$ is the p-dimensional volume of B. Some quantities of interest are

$$p(t) = P[N_{B(x,t)} > 0],$$
$$G(t) = P[N_{B(x,t)} > 1 | x \in \mathcal{N}],$$

and

$$K(t) = \frac{1}{\lambda} E[N_{B(x,t)} - 1 | x \in \mathcal{N}],$$

where

$$\mathcal{N} = \{x \in I\!R^p : N_{\{x\}} > 0],$$
$$B(x,t) = \{y \in I\!R^p : d(x,y) \leq t\},$$

and $d(x,y)$ is the Euclidean distance from x to y.

Example: For a Poisson process, N_B has a Poisson distribution for every B, and $N_{B_1}, \cdots, N_{B_m}$ are independent when $B_1, \cdots, B_m$ are mutually exclusive. For a Poisson process in $I\!R^2$,

$$p(t) = 1 - e^{-\lambda \pi t^2},$$
$$G(t) = p(t),$$
$$K(t) = \pi t^2.$$

Data: Suppose that the process is observed throughout a region $E \subseteq \mathbb{R}^p$.

Estimating K. Let

$$\tilde{K}_0(t) = \frac{1}{\lambda n} \sum_{x \in \mathcal{N} \cap E} \#\{y \in \mathcal{N} : 0 < d(x,y) \leq t\},$$

where

$$n = N_E,$$

and

$$\hat{K}_0(t) = \frac{1}{\lambda n} \sum_{x \in \mathcal{N} \cap E} \#\{y \in \mathcal{N} \cap E : 0 < d(x,y) \leq t\}.$$

Then (Ripley says)

$$E[\tilde{K}_0(t)] = K(t),$$

$$\hat{K}_0(t) \leq \tilde{K}_0(t),$$

and, therefore,

$$E[\hat{K}_0(t)] \leq E[\tilde{K}_(t)].$$

That is, $\hat{K}_0$ has a negative bias.

Fact: For a Poisson process: Given $N_E = n$, the n points are independently and uniformly distributed over E.

The Magnitude of the Bias: For a Poisson Process. Observe that for $n \geq 2$,

$$\frac{\lambda \hat{K}_0(t)}{n-1} = \frac{\#\{(x,y) : x, y \in \mathcal{N} \cap E,\ 0 < d(x,y) \leq t\}}{n(n-1)}.$$

So, for $n \geq 2$,

$$E\Big[\frac{\lambda \hat{K}_0(t)}{n-1}|n\Big] = P\big[d(X,Y) \leq t\big] = F(t), \text{ say},$$

where

$$X, Y \sim^{ind} \text{Uniform}(E).$$

$F(t)$ can be computed for some shapes and approximated more generally.

Example. If E is a rectangle in $\mathbb{R}^2$, then

$$F(t) = \frac{\pi t^2}{a} - \frac{2ut^3}{3a^2} + \frac{t^4}{2a^2},$$

where

$$a = \text{area},$$
$$u = \text{perimeter},$$

provided that t does not exceed the length of the shorter sides.

Approximation: Based on this and other examples, Ripley suggests the approximation

$$F(t) \approx \frac{\pi t^2}{a} - \frac{2ut^3}{3a^2}$$

for small t for convex E.

Back to $\hat{K}_0$. Writing

$$\hat{K}_0(t) = \frac{n-1}{\lambda} \times \frac{\lambda \hat{K}_0(t)}{n-1}$$

leads to

$$\begin{aligned} E[\hat{K}_0(t)] &= E\{\frac{n-1}{\lambda} E[\frac{\lambda \hat{K}_0(t)}{n-1}|n]\} \\ &= E[\frac{n-1}{\lambda} F(t)] \\ &= \frac{\lambda a - 1}{\lambda} F(t); \end{aligned}$$

or

$$E[\hat{K}_0(t)] \approx \frac{\lambda a - 1}{\lambda}[\frac{\pi t^2}{a} - \frac{2ut^3}{3a^2}].$$

There are two sources of bias here

$$\frac{\lambda a - 1}{\lambda} < a$$

and

$$a[\frac{\pi t^2}{a} - \frac{2ut^3}{3a^2}] < \pi t^2 = K(t).$$

The second is typically the more serious, at least of $a\lambda = E(n)$ is large.

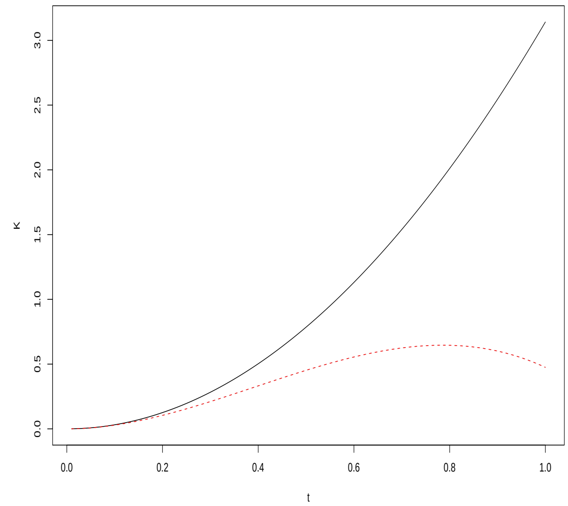

Figure 1: Approximate bias for a square and large λ

Some Corrected Estimators. *Border Corrections.* Find sets $E_t \subseteq E$ for which $E_t + B(0,t) \subseteq E$. Then let

$$\hat{K}_1(t) = \frac{1}{\lambda N_{E_t}} \#\{(x,y) : x \in E_t, y \in E, y \neq x, \ \& \ d(x,y) \leq t\}.$$

Here

$$E[\lambda N_{E_t}] = \lambda^2 \nu_p(E_t)$$

$$\begin{aligned} E\big[\#\{(x,y) : x \in E_t,\ y \in E,\ 0 < d(x,y) \leq t\}\big] \\ &= E\Big\{ \sum_{x \in E_t} E\big[\#\{(x,y) : x \in E_t,\ y \in E,\ 0 < d(x,y) \leq t\} | x\big] \Big\} \\ &= E\big[N_{E_t} \lambda K(t)\big] \\ &= \lambda^2 \nu_p(E_t) K(t), \end{aligned}$$

and $K(t)$ is the ratio of the expections.

$K_1(t)$ need not be a monotone function, however.

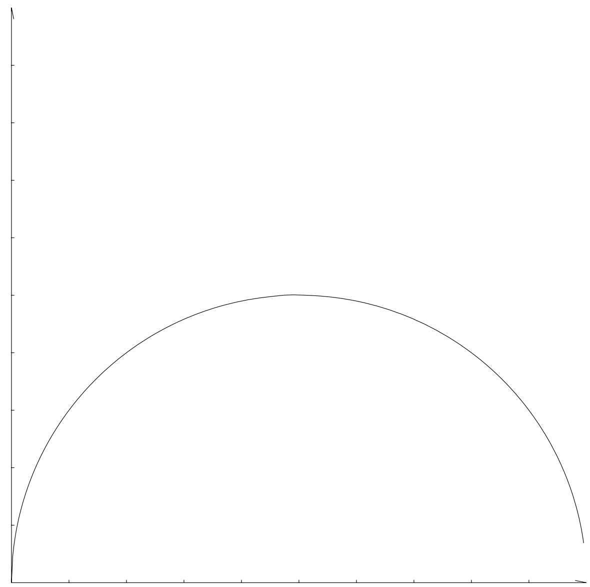

Figure 2: Boundary Length

Reweighting. Define k by

$$\frac{1}{k(x,y)} = \frac{|\partial B[x, d(x,y)] \cap E|}{|\partial B[x, d(x,y)]|},$$

where ∂B denotes the boundard of B and $|\partial B|$ denotes its length, and

$$\hat{K}_2(t) = \frac{1}{\lambda a^2} \sum_{x,y \in \mathcal{N} \cap E} k(x,y) \mathbf{1}_{\{0 < d(x,y) \leq t\}}.$$

Then

$$E[\hat{K}_2(t)] = \cdots = K(t),$$

under some general symmetry assumptions. For example, if E is the unit square and $x = (.5, 0)$ and $y = (0, 0)$, then

$$|\partial B[x, d(x,y)] \cap E| = \frac{1}{2}\pi + 1,$$
$$|\partial B[x, d(x,y)]| = \pi,$$

and

$$k(x,y) = \frac{2\pi}{\pi + 2}.$$

More On Reweighting. There are more elaborate versions of $\hat{K}_2$.