Statistics 222, Spatial Statistics.

Outline for the day:

1. MLE.
2. Simulation.
3. Purely spatial processes, Papangelou intensity,
   and the Georgii-Zessin-Nguyen formula.
4. Kernel smoothing.
5. Exercises and code.

1. Maximum likelihood estimation.

Find $\hat{\theta}$ (= θˆ) maximizing $l(\theta) = \sum \log(\lambda(\tau_i)) - \int \lambda(t,x,y)\, dt\, dx\, dy$.

Ogata (1978) showed that the resulting estimate, θˆ, is, under standard conditions,
asymptotically unbiased, $E(\hat\theta) \rightarrow \theta$,
consistent, $P(|\hat\theta - \theta| > \varepsilon) \rightarrow 0$ as $T \rightarrow \infty$, for any $\varepsilon > 0$,
asymptotically normal, $\hat\theta \rightarrow_D$ Normal as $T \rightarrow \infty$,
and asymptotically efficient, min. variance anong asymptotically unbiased estimators.

Further, he showed standard errors for θˆ can be constructed using the diagonal elements of the inverse of the Hessian of L evaluated at θˆ .

```
sqrt(diag(solve(loglikelihood$hess)))
```

Ogata, Y. (1978). The asymptotic behaviour of maximum likelihood estimators for stationary point processes. Ann. Inst. Statist. Math. 30, 243-261.

Maximum likelihood estimation continued.

The conditions of Ogata (1978) can be relaxed a bit for Poisson processes [1], and for certain spatial-temporal process in general [2].

Even if the process is not Poisson, under some circumstances [3] the parameters governing the unconditional intensity, $E\lambda$, can be consistently estimated by maximizing $L_P(\theta) = \sum \log(E\lambda(\tau_i)) - \int E\lambda(t,x,y)\, dt\, dx\, dy$. Basically pretend the process is Poisson.

Suppose you are missing some covariate that might affect λ. Under general conditions, the MLE will nevertheless be consistent, provided the effect of the missing covariate is small [4].

[1] Rathbun, S.L., and Cressie, N. (1994). Asymptotic properties of estimators for the parameters of spatial inhomogeneous Poisson point processes. Adv. Appl. Probab. 26, 122–154.

[2] Rathbun, S.L., (1996). Asymptotic properties of the maximum likelihood estimator for spatio-temporal point processes. *JSPI* 51, 55–74.

[3] Schoenberg, F.P. (2004). Consistent parametric estimation of the intensity of a spatial-temporal point process. *JSPI* 128(1), 79--93.

[4] Schoenberg, F.P. (2016). A note on the consistent estimation of spatial-temporal point process parameters. *Statistica Sinica,* 26, 861-879.

Maximum likelihood estimation continued.

λ is completely separable if $\lambda(t,x,y;\theta) = \theta_3\, \lambda_0(t;\theta_0)\, \lambda_1(t, x; \theta_1)\, \lambda_2(t, y; \theta_2)$.
Suppose N has marks too. λ is separable in mark (or coordinate) *i* if
$\lambda(t, x, y, m_1, m_2, ..., m_k; \theta) = \theta_2\, \lambda_i(t, m_i; \theta_i)\, \lambda_{-i}(t, x, y, m_{-i}; \theta_{-i})$.

Suppose you are neglecting some *mark* or coordinate of the process. Under some conditions, the MLE of the other parameters will nevertheless be consistent [1].

In maximizing $L(\theta) = \sum \log(\lambda(\tau_i)) - \int \lambda(t,x,y)\, dt\, dx\, dy$,
it is typically straightforward to compute the sum, but the integral can be tricky esp. when the conditional intensity is very volatile. One trick noted in [2] is that, for a Hawkes process where $\lambda(t,x,y) = \mu(x,y) + \kappa \sum_{\{t',x',y':\, t' < t\}} g(t-t',x-x',y-y')$, where g is a density, and $\int \mu(x,y)dxdy = \mu$,

$$\begin{aligned} \int \lambda(t,x,y)\, dt\, dx\, dy &= \mu T + \kappa \int \sum g(t-t',x-x',y-y')\, dt\, dx\, dy \\ &= \mu T + \kappa \sum \int g(t-t',x-x',y-y')\, dt\, dx\, dy \\ &\sim \mu T + \kappa N. \end{aligned}$$

[1] Schoenberg, F.P. (2016). A note on the consistent estimation of spatial-temporal point process parameters. *Statistica Sinica,* 26, 861-879.
[2] Schoenberg, F.P. (2013). Facilitated estimation of ETAS. *Bulletin of the Seismological Society of America,* 103(1), 601-605.

2. Simulation.

Jesper Møller

One can simulate spatial-temporal point processes by *thinning*.
Lewis, P. and Shedler, G. (1979). Simulation of nonhomogeneous poisson processes by thinning. *Naval Research Logistics Quarterly*, 26:403–413, 1979.

Suppose λ has some upper bound, B. $\lambda(t,x,y) \leq B$ everywhere.
First, simulate a stationary Poisson process *N* with intensity B.
For i = 1,2, ... keep point τ_i with probability $\lambda(\tau_i)/B$. We saw this in the code for Day 3 for simulating inhomogeneous Poisson processes and it works for other processes too.

Boundary issues can be important in simulation. For Gibbs processes, for instance, the simulation can be biased because of missing points outside the observation region. For Hawkes processes, the simulation will tend to be biased by having too few points at the beginning of the simulation. One can have burn-in, by simulating points outside the observation region or before time 0, or in some cases some fancy weighting schemes can be done to achieve *perfect* simulation without burn-in.
Møller, J. and Waagepetersen, R. (2003). *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall, Boca Raton.

3. Purely spatial processes, Papangelou intensity and the Georgii-Zessin Nguyen formula.

For point processes in R^2, there is no natural ordering as there is in time. One could just use the x-coordinate in place of time and define a conditional intensity, but most models for spatial processes would be very awkward to define this way.

Instead, a more natural and useful tool is the Papangelou intensity, $\lambda(x,y)$, which is the conditional rate of points around location (x,y), given information on everywhere else. Letting

$l(\theta) = \sum \log(\lambda(\tau_i)) - \int \lambda(x,y)\, dx\, dy,$

where $\lambda(x,y)$ is the Papangelou intensity,

$l(\theta)$ is called the *pseudo-loglikelihood*.

A key formula for space-time point processes is called the *martingale formula*:

for any predictable function f(t,x,y),

$E \int f(t,x,y)\, dN = E \int f(t,x,y)\, \lambda(t,x,y)\, d\mu.$

$= E \sum_i f(t_i,x_i,y_i) = E \int f(t,x,y)\, \lambda(t,x,y)\, dt\, dx\, dy$

For spatial point processes the corresponding formula,

$E \int f(x,y)\, dN = E \int f(x,y)\, \lambda(x,y)\, dx\, dy$

is called the Georgii-Zessin-Nguyen formula.

When f = 1, this means $EN(B) = E \int \lambda\, d\mu$.

4. Kernel smoothing.

A simple way to start summarizing a spatial point process is by kernel smoothing.

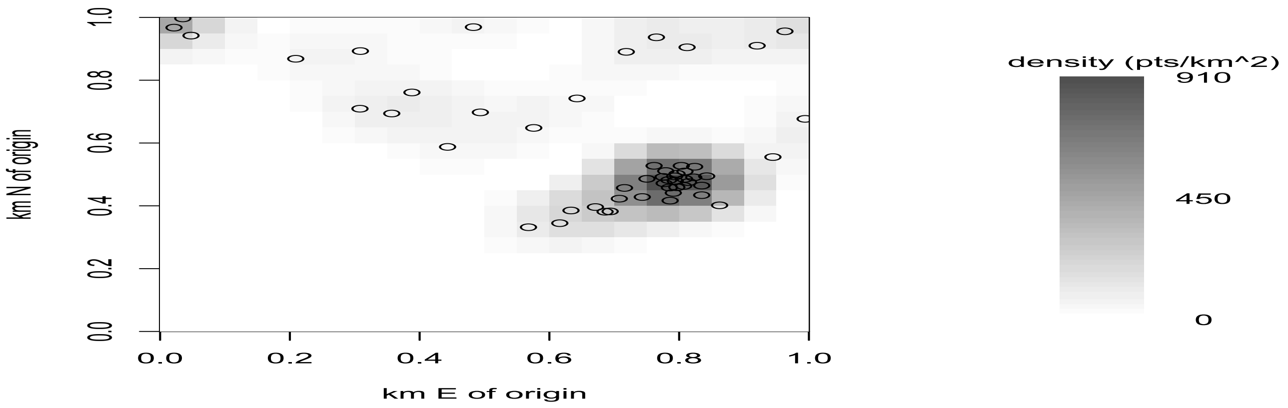
Suppose your observation region is B.

Let k(x,y) be a spatial density function, called a kernel, and construct, for each location (x,y),

$\hat{\lambda}(x,y) = \int_B k((x',y') - (x,y))\, dN(x',y') / \rho(x,y)$,

where $\rho(x,y) = \int_B k((x',y') - (x,y))\, dx'\, dy'$ is an edge correction term.

The resulting function $\hat{\lambda}(x,y)$ is a natural estimator of $\lambda(x,y)$ and, when N is a Poisson process, can be an asymptotically unbiased estimator of $\lambda(x,y)$.



5. Exercises.

a. Suppose N is a Poisson process with intensity $\lambda(t,x,y) = \exp(3t)$ over t in [0,10], x in [0,1], y in [0,1].

N happens to have points at (1.5, .4, .2)
(2, .52, .31)
(4, .1, .33)
(5, .71, .29).

What is the log-likelihood of this realization?

5. exercises.

a. Suppose N is a Poisson process with intensity $\lambda(t,x,y) = \exp(-3t)$ over t in [0,10], x in [0,1], y in [0,1].

N happens to have points at (1.5, .4, .2)
(2, .52, .31)
(4, .1, .33)
(5, .71, .29).

What is the log-likelihood of this realization?

$-4.5-6-12-15 - \iiint \exp(-3t)\, dt\, dx\, dy$

$= -37.5 - \int_0^{10} \exp(-3t)\, dt$, because x and y go from 0 to 1,

$= -37.5 - \exp(-3t) / (-3)]_0^{10}$

$= -37.5 + \exp(-30)/3 - \exp(0)/3$

$= -37.5 + \exp(-30)/3 - 1/3$

~ -37.83.

5. exercises.

Which of the following is not typically true of the MLE of a spatial-temporal point process?

a. It is unbiased.
b. It is consistent.
c. It is asymptotically normal.
d. It is asymptotically efficient.

5. exercises.

Which of the following is not typically true of the MLE of a spatial-temporal point process?

a. It is unbiased.
b. It is consistent.
c. It is asymptotically normal.
d. It is asymptotically efficient.

Entering data example.

```
## First, input 54 points using the mouse.
n = 54
plot(c(0,1),c(0,1),type="n",xlab="longitude",ylab="latitude",
                    main="locations")
 x1 = rep(0,n)
 y1 = rep(0,n)
 for(i in 1:n){
 z1 = locator(1)
 x1[i] = z1$x
 y1[i] = z1$y
 points(x1[i],y1[i])
 }
```

```
##### PLOT THE POINTS WITH A 2D KERNEL SMOOTHING IN
    GREYSCALE PLUS A LEGEND
library(splancs)
bdw = sqrt(bw.nrd0(x1)^2+bw.nrd0(y1)^2)  ## possible default bandwid
b1 = as.points(x1,y1)
bdry = matrix(c(0,0,1,0,1,1,0,1,0,0),ncol=2,byrow=T)
z = kernel2d(b1,bdry,bdw)
attributes(z)
par(mfrow=c(1,2))
image(z,col=gray((64:20)/64),xlab="km E of origin",ylab="km N of
    origin")
points(b1)
x4 = seq(min(z$z),max(z$z),length=100)
```

```
plot(c(0,10),c(.8*min(x4),1.2*max(x4)),type="n",
        axes=F,xlab="",ylab="")
image(c(1,1.5),x4,matrix(rep(x4,2),ncol=100,byrow=T),
        add=T,col=gray((64:20)/64))
text(2,min(x4),as.character(signif(min(x4),2)),cex=1)
text(2,(max(x4)+min(x4))/2,
        as.character(signif((max(x4)+min(x4))/2,2)),cex=1)
text(2,max(x4),as.character(signif(max(x4),2)),cex=1)
mtext(s=3,l=-3,at=1,"density (pts/km^2)")
```

We will continue this next time, fitting models to this by MLE.