

## **Statistics 222, Spatial Statistics.**

### Outline for the day:

1. Poisson process, continued.
2. Mixed Poisson process.
3. Compound Poisson process
4. Poisson cluster process.
5. Cox process.
6. Gibbs processes.
7. Matern processes.
8. Examples and code.

# 1. Poisson processes.

Last week we discussed Poisson processes.

If  $N$  is a simple point process with conditional intensity  $\lambda$ , where  $\lambda$  does not depend on what points have occurred previously, then  $N$  is a *Poisson process*.

For such a process, for any set  $B$ ,  $N(B)$  has a Poisson distribution.

$$P(N(B) = k) = e^{-A} A^k / k! ,$$

for  $k = 0, 1, 2, \dots$ ,

where  $A = \int_B \lambda(t, x, y) dt dx dy$ ,

and with the convention  $0! = 1$ .

The mean of  $N(B)$  is  $A$  and the variance is also  $A$ .  $E(N(B)^2) = A^2 + A$ .

We will now discuss a few extensions of Poisson processes.

**Siméon Poisson**



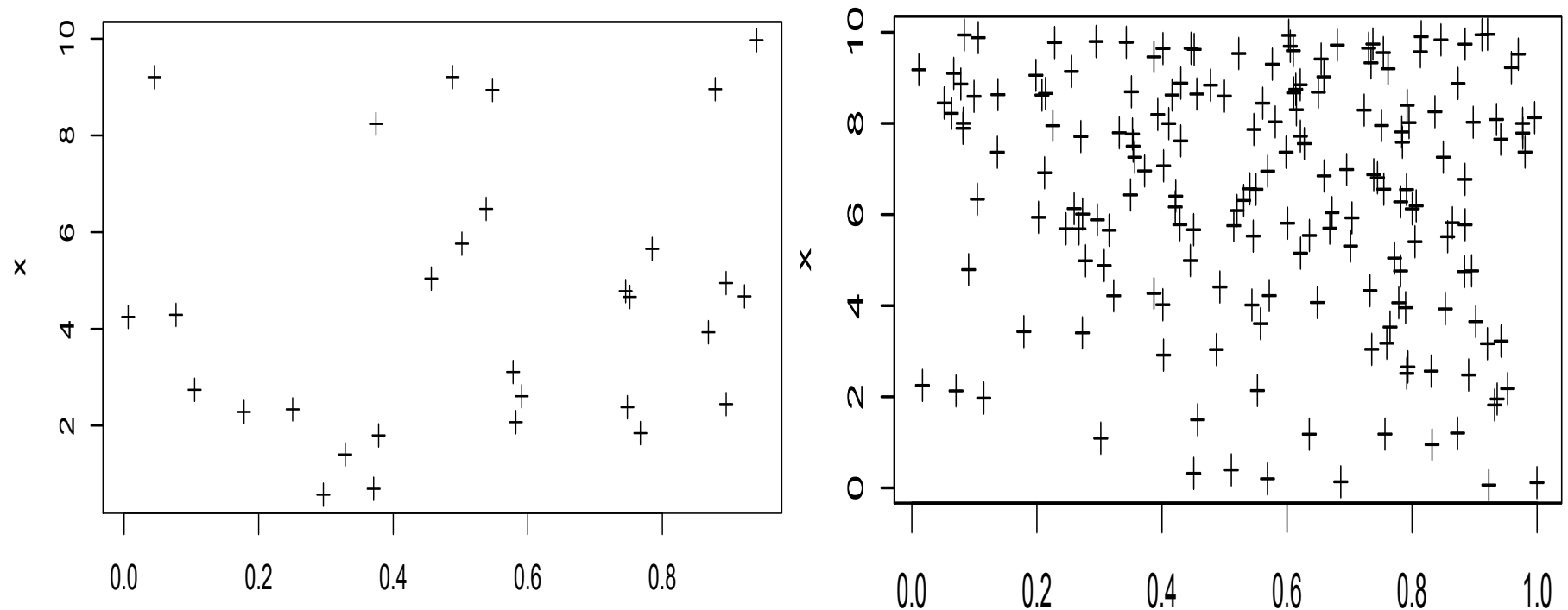
Siméon Denis Poisson (1781–1840)

## Poisson processes, continued.

On the left is a stat. Poisson process with  $\lambda(t,x) = 2.5$  on  $[0,1] \times [0,10]$ , and on the right is a Poisson process with  $\lambda(t,x) = 1.5 + 10t + 2x$ .

The key thing about Poisson processes is their complete independence.

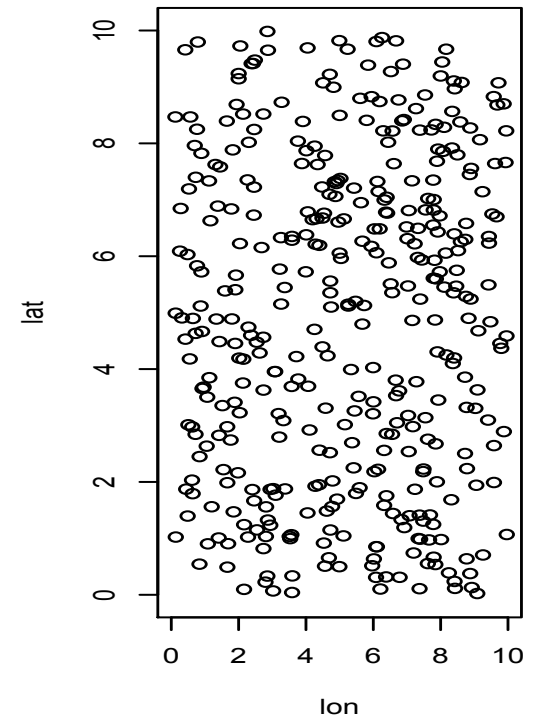
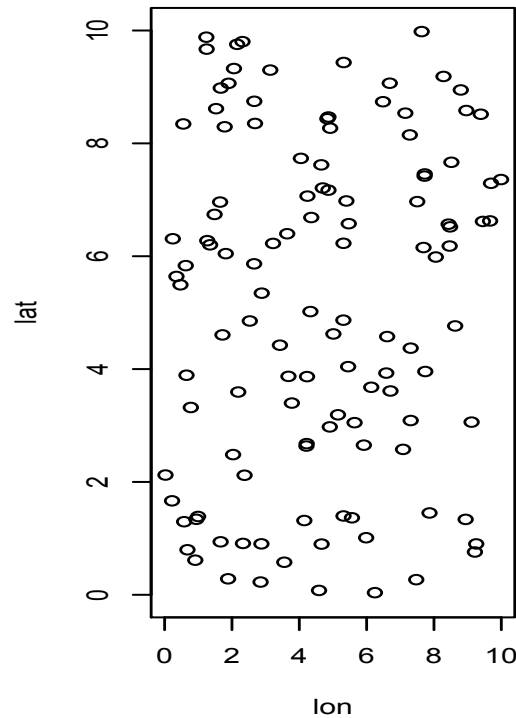
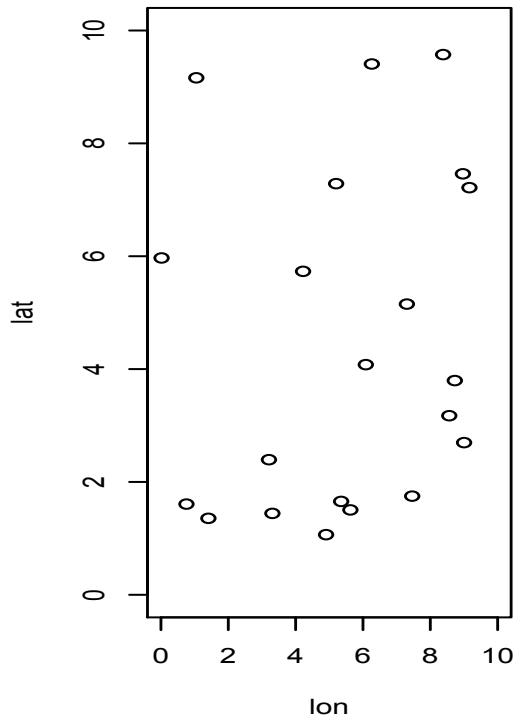
For a Poisson process  $N$ ,  $N(B_1)$  and  $N(B_2)$  are independent for any disjoint sets  $B_1$  and  $B_2$ .



## 2. Mixed Poisson processes.

Suppose  $\lambda(t,x,y) = c$ , where  $c$  is a random variable. For example,  $c$  might be Poisson or exponential, or half normal, or something constrained to be positive. Then conditional on  $c$ ,  $N(B)$  is Poisson distributed. Then  $N$  is a *mixed Poisson process*.

$E(N(B) | c) = V(N(B)|c) = c|B|$ , but unconditionally,  $N(B)$  is not Poisson distributed now.



## 2. Mixed Poisson processes.

Suppose  $\lambda(t,x,y) = c$ , where  $c$  is a random variable. For example,  $c$  might be Poisson or exponential, or half normal, or something constrained to be positive. Then conditional on  $c$ ,  $N(B)$  is Poisson distributed. Then  $N$  is a *mixed Poisson process*.

$E(N(B) | c) = V(N(B)|c) = c|B|$ , but unconditionally,  $N(B)$  is not Poisson distributed now. If we imagine simulating the process repeatedly, each time with a different draw of  $c$ , then the distribution of  $N(B)$  will not be Poisson.  $N(B)$  will typically be overdispersed relative to the Poisson process, i.e. will have higher variance.

$$E(N(B)) = \int E(N(B)|c) f(c)dc = \int c|B| f(c)dc = |B|E(c).$$

$$E(N(B)^2) = \int E(N(B)|c)^2 f(c) dc = \int [c^2|B|^2 + c|B|] f(c) dc \\ = |B|^2 E(c^2) + |B|E(c),$$

$$\text{so } V(N(B)) = |B|^2 E(c^2) + |B|E(c) - |B|^2 [E(c)]^2 = E(N(B)) + |B|^2 V(c).$$

$$\text{So, } V(N(B)) \geq E(N(B)).$$

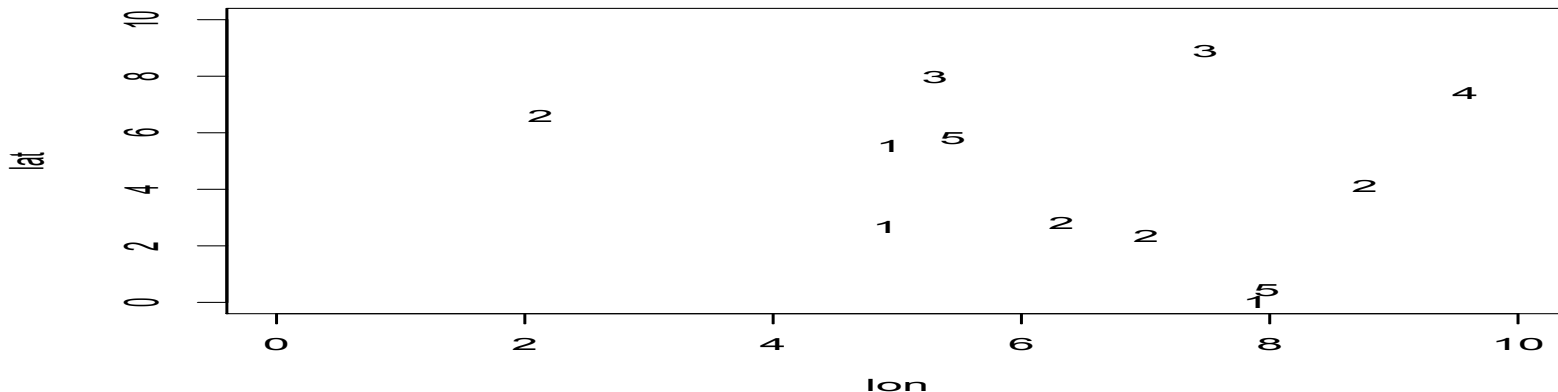
### 3. Compound Poisson process.

Suppose  $N$  is not simple, and instead, it is generated as follows. You first generate a stationary Poisson process  $M$  with intensity  $c$ , and then for each point  $\tau_i$  of  $M$ ,  $N$  will have some non-negative number  $Z_i$  of points right at  $\tau_i$ , where  $Z_i$  are all iid and independent of  $M$ . Then  $N$  is a compound Poisson process.

For a compound Poisson process, again the variance  $\geq$  the mean.

$E(N(B)) = c|B|E(Z)$ , and

$V(N(B)) = c|B|V(Z) + c|B|(E(Z))^2 = c|B|E(Z^2) \geq E(N(B))$ , because, for a non-negative integer-valued random variable  $Z$ ,  $E(Z^2) \geq E(Z)$  with equality iff.  $Z$  can only be 0 or 1.



### 3. Variance of the compound Poisson process.

Fix  $B$ . Let  $M$  denote  $M(B)$ . For a compound Poisson process,

$$E(N(B)) = c|B| E(Z).$$

$$E(N(B)^2) = V(Z) E(M) + (E(Z))^2 E(M^2).$$

$$\text{So } V(N(B)) = V(Z) E(M) + E(Z)^2 V(M).$$

$M$  is Poisson, so  $E(M) = V(M) = c|B|$ , so

$$V(N(B)) = c|B| (V(Z) + E(Z)^2) = c|B| E(Z^2) \geq E(N(B)),$$

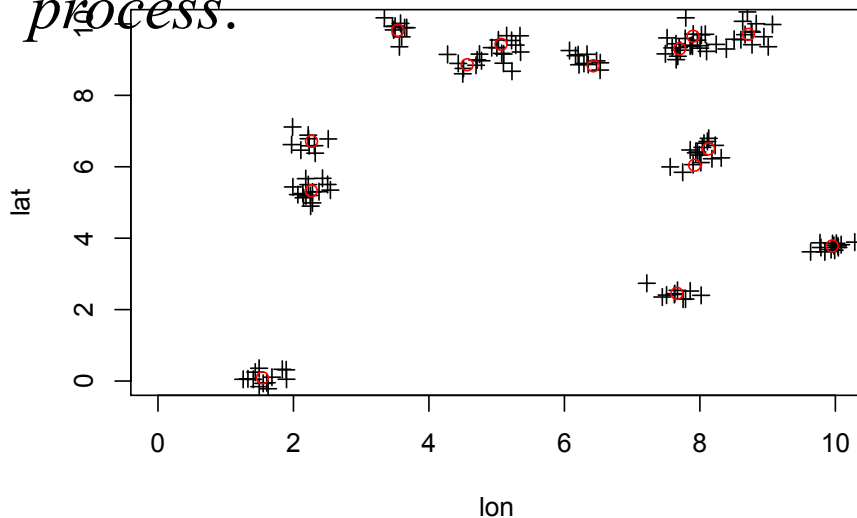
since  $E(Z^2) \geq E(Z)$  because  $Z$  is nonnegative integer valued.

#### 4. Poisson cluster processes.

Another extension of the Poisson process is the Poisson cluster process. Imagine first generating *parent* points  $M$  according to a Poisson process. Then for each parent point  $\tau_i$ , you generate some random number  $Z_i$  of offspring points, and these offspring points are scattered spatially and temporally, independently of each other, with some distribution centered at  $\tau_i$ . Let  $N$  be the collection of just the offspring, not the parents.  $N$  is called *Poisson cluster process*.

Usually  $M$  is assumed *stationary* Poisson.

In the particular case where the  $Z_i$  are iid Poisson random variables independent of  $M$ , the process is called a *Neyman-Scott cluster process*.



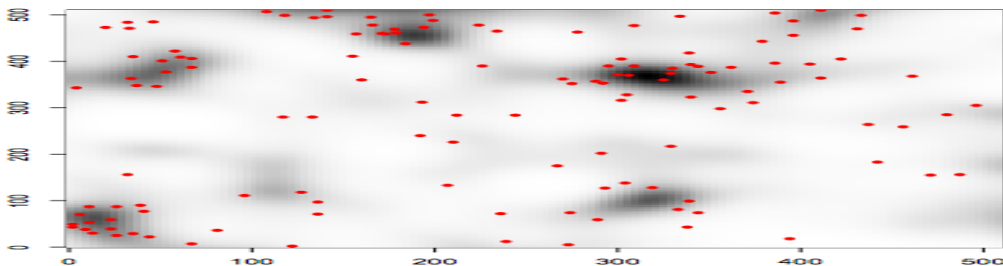
Jerry Neyman



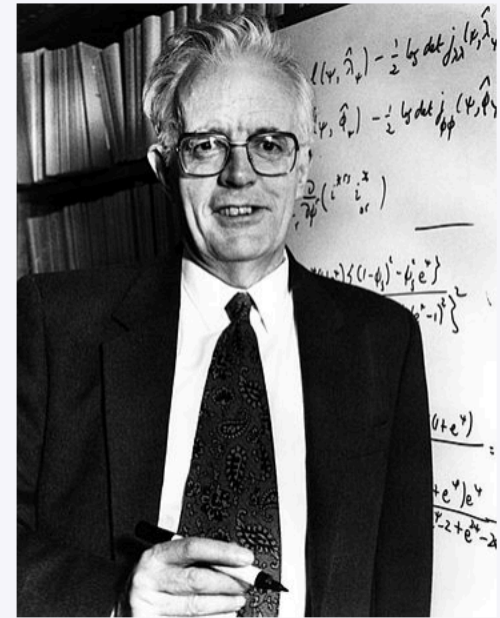
## 5. Cox process.

Suppose you somehow generate a stochastic process  $\lambda(t,x,y)$  such that  $\lambda(t,x,y) \geq 0$  for all  $t$ ,  $x$ , and  $y$ . Then you let  $N$  be a Poisson process with intensity  $\lambda(t,x,y)$ . So  $\lambda(t,x,y)$  can be random, but conditional on  $\lambda$ ,  $N$  is a Poisson process. In this case we say  $N$  is a *Cox process* or equivalently a *doubly stochastic Poisson process*.

Cox processes arise in practice when modeling events depending on some other random phenomenon. For instance, the points of  $N$  might be the times and locations of flu epidemics, which might depend on the temperature and this might in turn be modeled as evolving stochastically.



**David Cox**



## 6. Gibbs process.

For any finite collection  $(\tau_1, \tau_2, \dots, \tau_n)$  of points in space-time, if the joint density is  $C(\theta) \exp[-\theta \{ \sum_i \psi_1(\tau_i) + \sum_{i,j} \psi_2(\tau_i, \tau_j) \}]$ , then  $N$  is a Gibbs process.

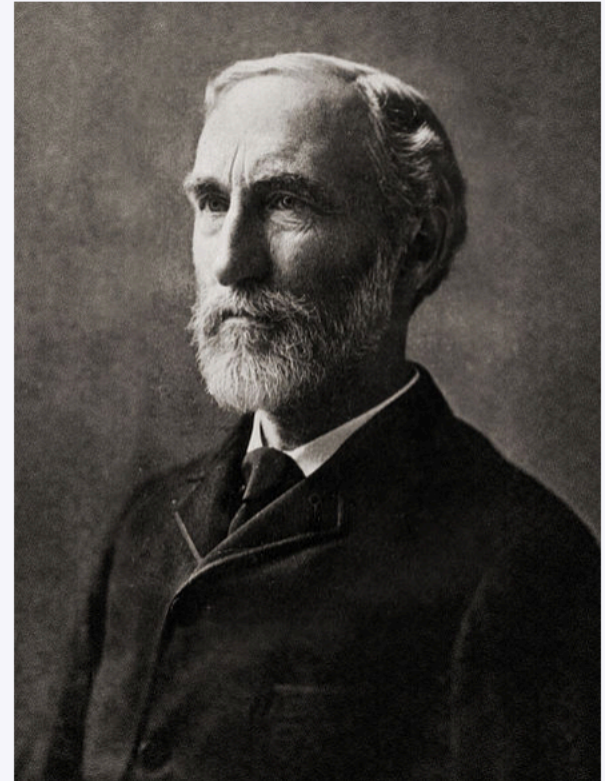
Often  $\psi_2(x_i, x_j)$  can be written  $\psi(r)$ , where  $r = |x_i - x_j|$ .

Some special cases are important.

a. When  $\psi(r) = 0$ , there are no interactions and the process is an inhomogeneous Poisson process with intensity  $\psi_1(x)$ .

b.  $\psi(r) = -\log[1 - e^{-(r/\sigma)^2}]$  defines a *soft-core* model. Weak repulsion.

**Josiah Willard Gibbs**



Josiah Willard Gibbs

<b>Born</b>	February 11, 1839 New Haven, Connecticut, U.S.
<b>Died</b>	April 28, 1903 (aged 64) New Haven, Connecticut, U.S.

## 6. Gibbs process, continued.

$\psi_2(\mathbf{r})$  is called the *interaction potential*.

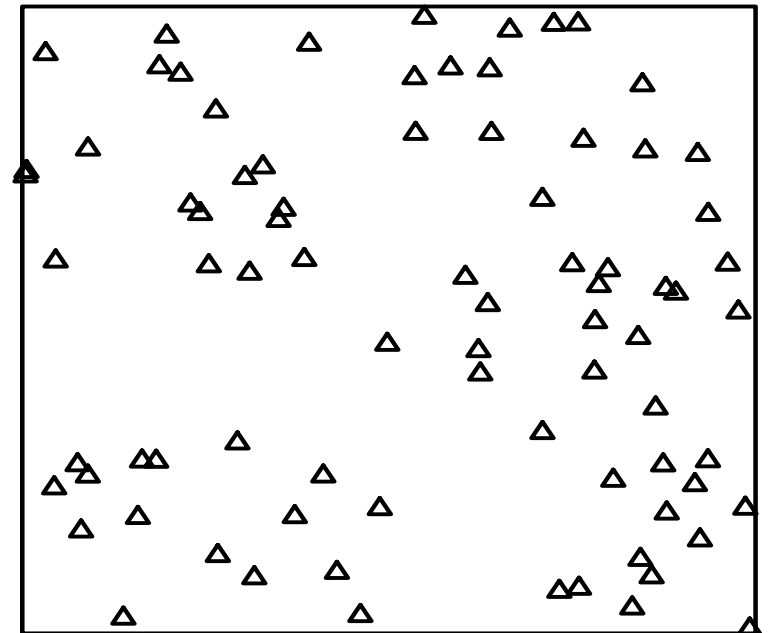
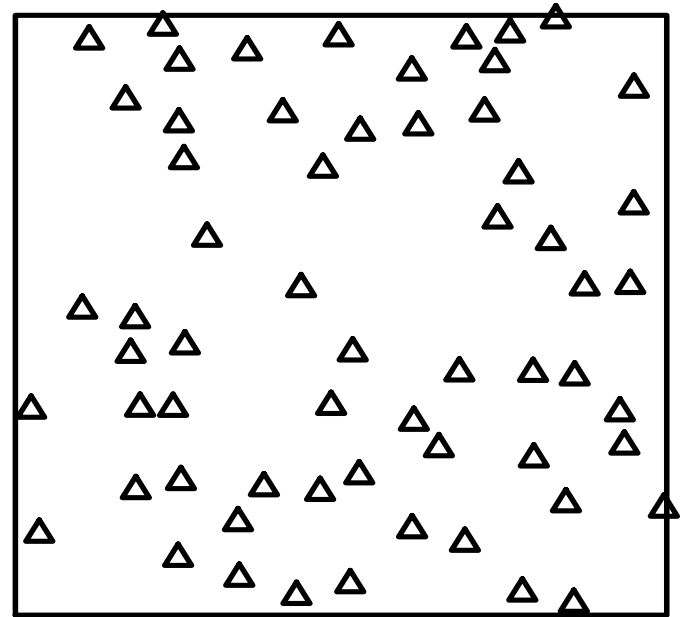
$$\begin{aligned} \text{c. } \psi(\mathbf{r}) &= \infty \text{ for } r \leq \sigma \\ &= 0 \text{ for } r > \sigma \end{aligned}$$

defines a *hard-core* process.

d.  $\psi(\mathbf{r}) = (\sigma/r)^n$  is an intermediate choice between the soft-core and hard-core models.

e. Strauss process.

$$\begin{aligned} \psi_1(\mathbf{x}) &= \alpha, \text{ and} \\ \psi_2(\mathbf{r}) &= \beta, \text{ for } r \leq R, \\ \psi_2(\mathbf{r}) &= 0, \text{ for } r > R. \end{aligned}$$



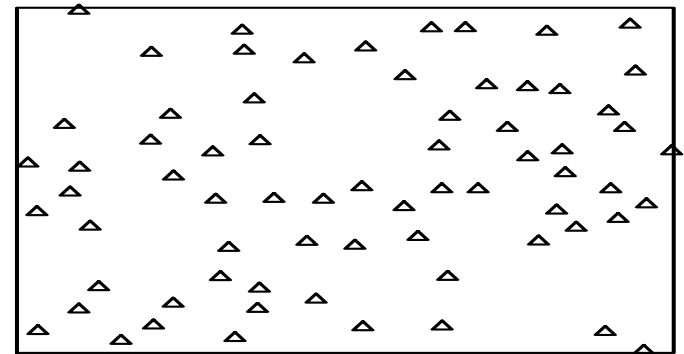
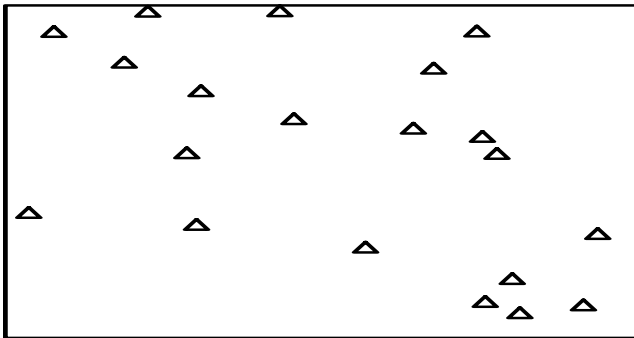
## 7. Matern process.

The Matern(I) process is generated as follows.

- Generate  $M$  according to a stationary Poisson process.
- Let  $N$  be all points of  $M$  that are not within some fixed distance  $r$  of any other point of  $M$ .

The Matern(II) process is generated a bit differently.

- Generate points  $\tau_1, \tau_2, \dots$  according to a stationary Poisson process.
- For  $i = 1, 2, \dots$ , keep point  $i$  if there is no *previous* kept point  $\tau_j$  with  $|\tau_i - \tau_j| \leq r$ .



## Exercises.

1. A mixed Poisson process is a Cox process where
  - a.  $\lambda = E(\lambda)$  in every realization.
  - b.  $\lambda(t,x,y) = \lambda(t',x',y')$ , for any locations  $(t,x,y)$  and  $(t',x',y')$ .
  - c. The cluster sizes are Poisson distributed with mean  $\lambda$ .
  - d.  $\lambda = 1$ .

## Exercises.

1. A mixed Poisson process is a Cox process where
  - a.  $\lambda = E(\lambda)$  in every realization.
  - b.  $\lambda(t, \mathbf{x}, y) = \lambda(t', \mathbf{x}', y')$ , for any locations  $(t, \mathbf{x}, y)$  and  $(t', \mathbf{x}', y')$ .**
  - c. The cluster sizes are Poisson distributed with mean  $\lambda$ .
  - d.  $\lambda = 1$ .
- a. means  $\lambda$  is a constant, so  $N$  is a stationary Poisson process.
- d. Also defines a stationary Poisson process, with rate 1.

## Code from Day 2.

```
## nonsimple point process
```

```
n = 20
```

```
x = runif(n)
```

```
y = runif(n)
```

```
plot(x,y,xlab="t",ylab="lat",pch=2)
```

```
points(x[20],y[20],pch=3)
```

```
## nonsimple ground process
```

```
plot(x,y,xlab="t",ylab="lat",pch=2)
```

```
points(x[20],y[20]+.05,pch=3)
```

```
## nonorderly process
```

```
plot(c(0,1),c(0,1),type="n",xlab="t",ylab="lat")
```

```
n = 100
```

```
for(i in 1:n) points(1/i,runif(1),pch=3,cex=.5)
```

## Code from Day 2.

```
## points at (i,i) with prob. 1/i.
```

```
plot(c(0,100),c(0,100),type="n",xlab="t",ylab="lat")
```

```
for(i in 1:100) if(runif(1) < 1/i) points(i,i,pch=3)
```

```
## stationary Poisson process with intensity 2.5 on B=[0,1]x[0,10].
```

```
n = rpois(1,2.5*1*10)
```

```
t = runif(n)
```

```
x = runif(n)*10
```

```
plot(t,x,pch=3)
```



Code from Day 2.

```
## nonstationary Poisson process with intensity  $1.5+10t+2x$  on B.
```

```
n = rpois(1,15+50+100)
```

```
n1 = 0
```

```
t = c()
```

```
x = c()
```

```
while(n1<n){
```

```
  t2 = runif(1) ## candidate point
```

```
  x2 = runif(1)*10
```

```
  if(runif(1) < (1.5+10*t2+2*x2)/(1.5+10+20)){ ## keep it
```

```
    t = c(t,t2)
```

```
    x = c(x,x2)
```

```
    n1 = n1 + 1
```

```
    cat(n1," ")
```

```
  }
```

```
}
```

```
plot(t,x,pch=3)
```

Code from Today.

```
## mixed Poisson process
```

```
par(mfrow=c(1,3))
```

```
m = rexp(1,rate=.5)
```

```
n1 = rpois(1,m*10*10)
```

```
x1 = runif(n1)*10
```

```
y1 = runif(n1)*10
```

```
plot(c(0,10),c(0,10),xlab="lon",ylab="lat",type="n")
```

```
points(x1,y1)
```

```
## I ran the previous 5 lines 3 times.
```

Code.

```
## compound Poisson.  
par(mfrow=c(1,1))  
n1 = rpois(1,.12*10*10)  
x1 = runif(n1)*10  
y1 = runif(n1)*10  
a = as.character(rpois(n1,3))  
plot(c(0,10),c(0,10),xlab="lon",ylab="lat",type="n")  
text(x1,y1,a)
```

Code.

```
## Neyman-Scott.  
n1 = rpois(1,.12*10*10)  
x1 = runif(n1)*10  
y1 = runif(n1)*10  
x2 = c()  
y2 = c()  
## parents are (x1,y1).  
for(i in 1:n1){  
  c = rpois(1,8) ## number of offspring  
  if(c>0) for(j in 1:c){  
    x2 = c(x2,rnorm(1,sd=.2)+x1[i])  
    y2 = c(y2,rnorm(1,sd=.2)+y1[i])  
  }  
}  
plot(c(0,10),c(0,10),xlab="lon",ylab="lat",type="n")  
points(x2,y2,pch=3)  
points(x1,y1,col="red")
```