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. Examples and code.

1. Poisson processes.

Last week we discussed Poisson processes.

If N is a simple point process with conditional intensity λ, where λ 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 $k = 0, 1, 2, ...$,

where $A = \int_B \lambda(t,x,y)\, dtdxdy$,

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.

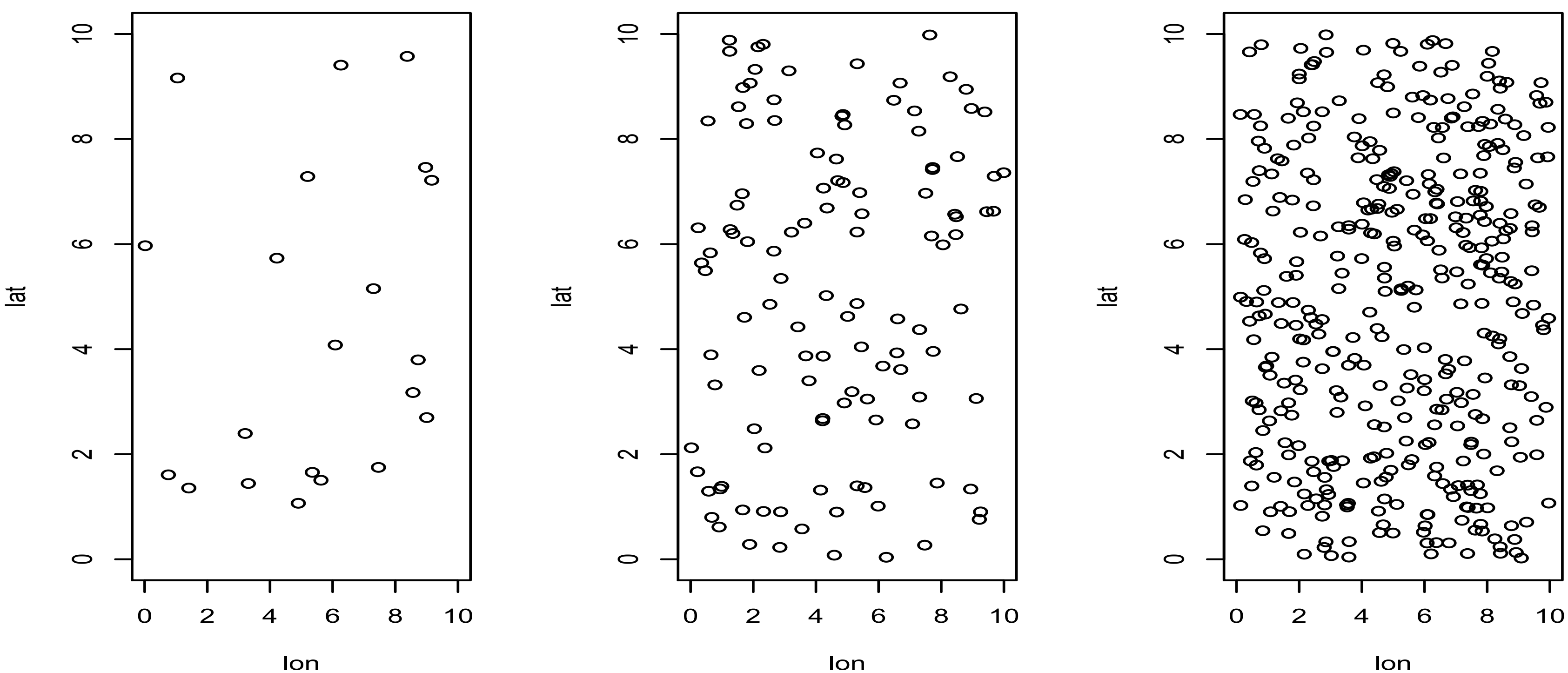
Siméon Poisson

Siméon Denis Poisson (1781–1840)

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*.

$E(N(B) \mid c) = V(N(B)|c) = c|B|$, 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*.

$E(N(B) \mid c) = V(N(B)|c) = c|B|$, 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.

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

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

$= |B|^2\, E(c^2) + |B|E(c)$,

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

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

3. Compound Poisson processes.

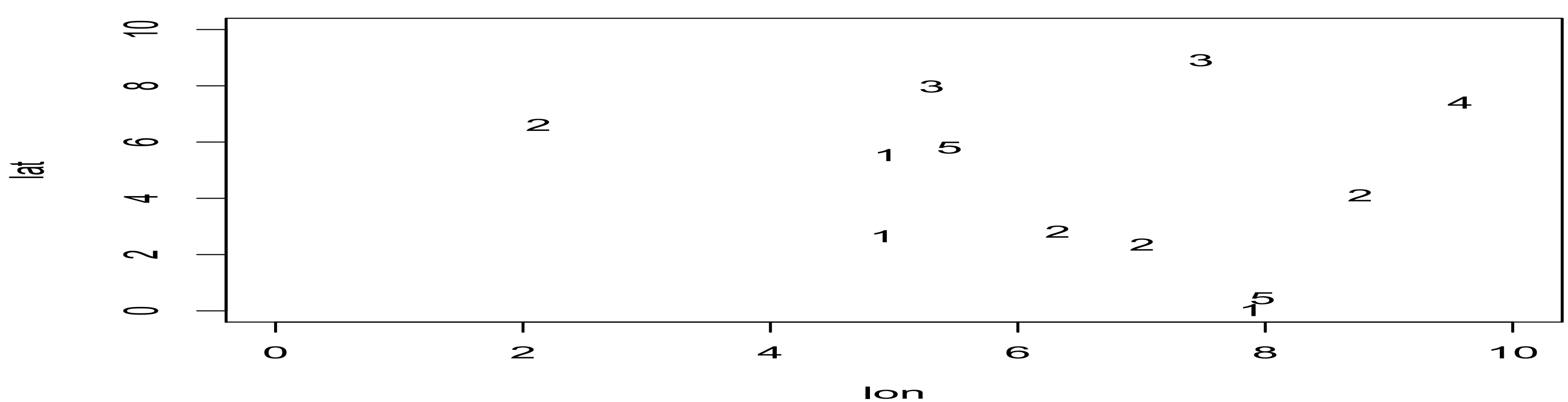
Suppose N is not simple, and instead, it is generated as follows. You first generate a stationary Poisson process M with intensity c, and then for each point τ_i of M, N will have some non-negative number Z_i of points right at τ_i, where Z_i are all iid and independent of M. Then N is a compound Poisson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

For instance, the counts Z_i might themselves have a Poisson distribution.

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

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

$V(N(B)) = c|B|V(Z) + c|B|(E(Z))^2 = c|B|E(Z^2) \geq EN(B)$, 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. Compound Poisson processes.

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

$EN(B) = \sum_{m=0}^{\infty} E(N(B)|m)\ f(m)$

$= \sum_{m=0}^{\infty} (mE(Z))\ f(m)$

$= E(Z) \sum_{m=0}^{\infty} m f(m) = E(Z)\ E(M) = c|B|\ E(Z)$, and

$EN(B)2 = \sum_{m=0}^{\infty} E(N(B)^2|m) f(m)$

$= \sum_{m=0}^{\infty} (m\ E(Z^2) + (m^2 - m)E(Z)^2)\ f(m)$

$= \sum_{m=0}^{\infty} E(Z^2)\ E(M) - E(Z)^2\ E(M) + E(Z)^2\ E(M^2)$

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

So $V(N(B)) = E(N(B)^2) - (E(N(B))^2$

$= V(Z)E(M) + E(Z)^2\ E(M^2) - E(M)^2 E(Z)^2$

$= E(Z)^2\ (E(M^2) - E(M)^2) + V(Z)E(M)$

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

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

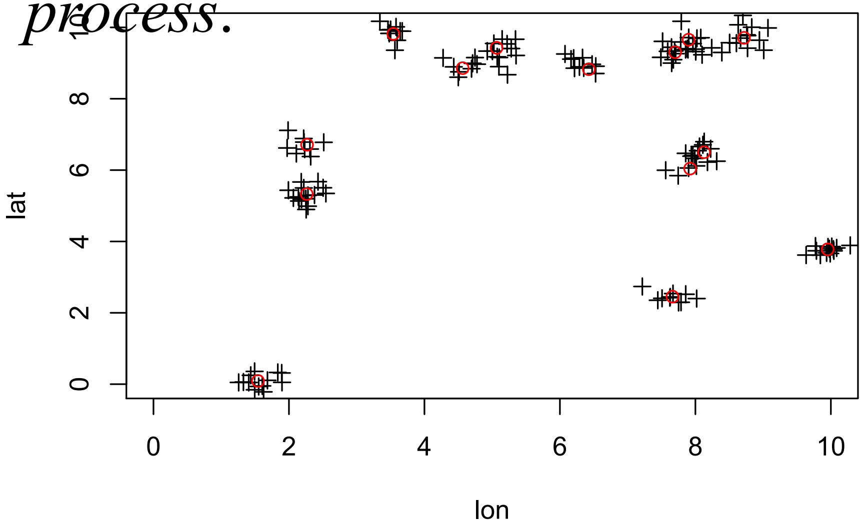
$V(N(B)) = c|B|\ (E(Z)^2 + V(Z)) = c|B|E(Z^2) \geq EN(B)$, because, for a non-negative integer-valued random variable Z, $E(Z^2) \geq E(Z)$ with equality iff. Z can only be 0 or 1.

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 τ_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 τ_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*.



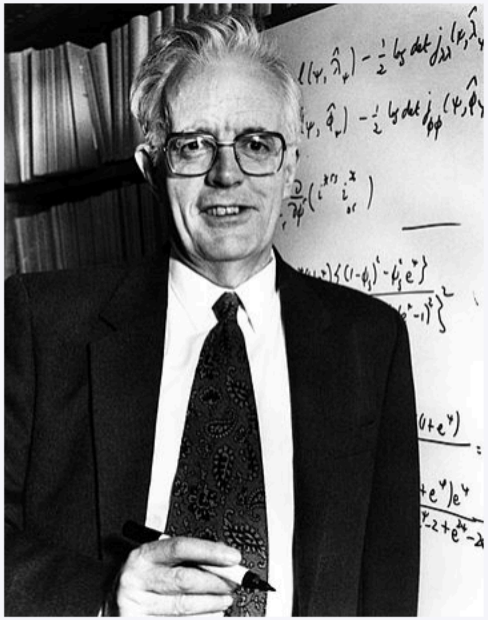
Jerzy Neyman

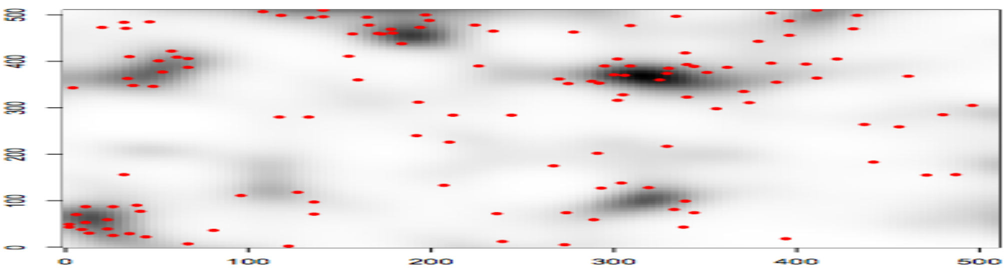
5. Cox process.

Suppose you somehow generate a stochastic process λ(t,x,y) such that λ(t,x,y) ≥ 0 for all t, x, and y. Then you let *N* be a Poisson process with intensity λ(t,x,y). So λ(t,x,y) can be random, but conditional on λ, *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





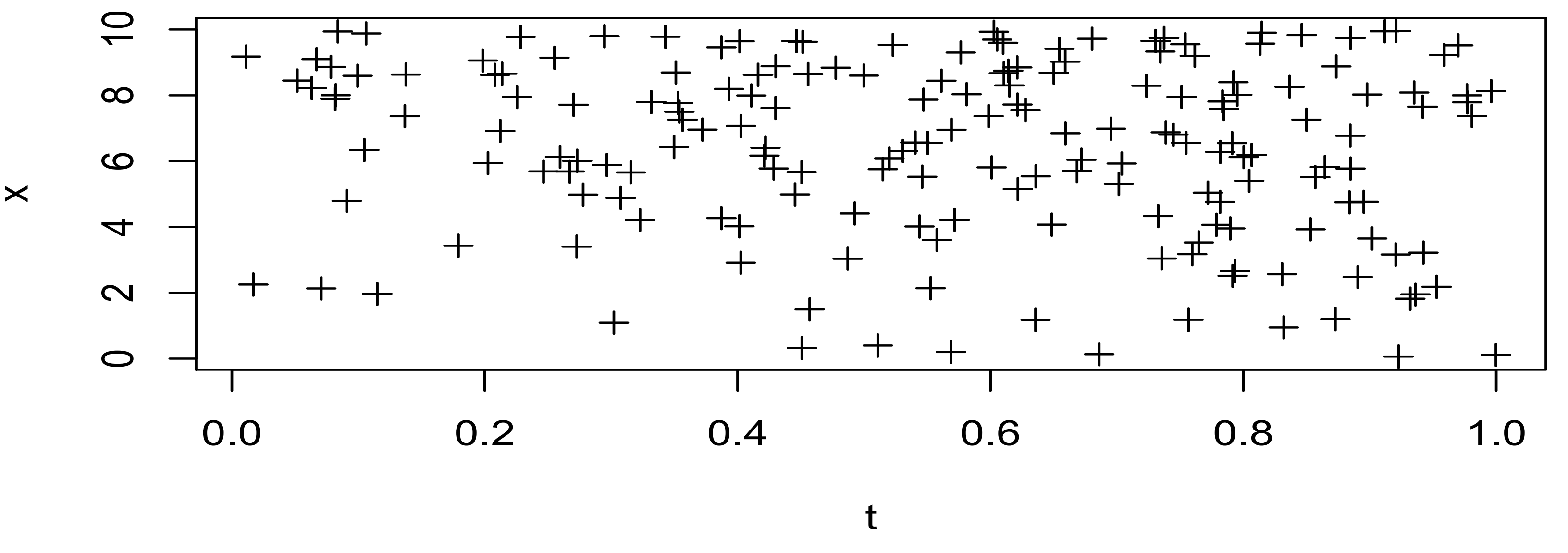
Exercises.

1. Suppose *N* is a Poisson process with $\lambda(t,x,y) = 1.5 + 10t + 2x$

on B = [0,1] x [0,10] x [0,1].

time X Y

What is EN(B)?



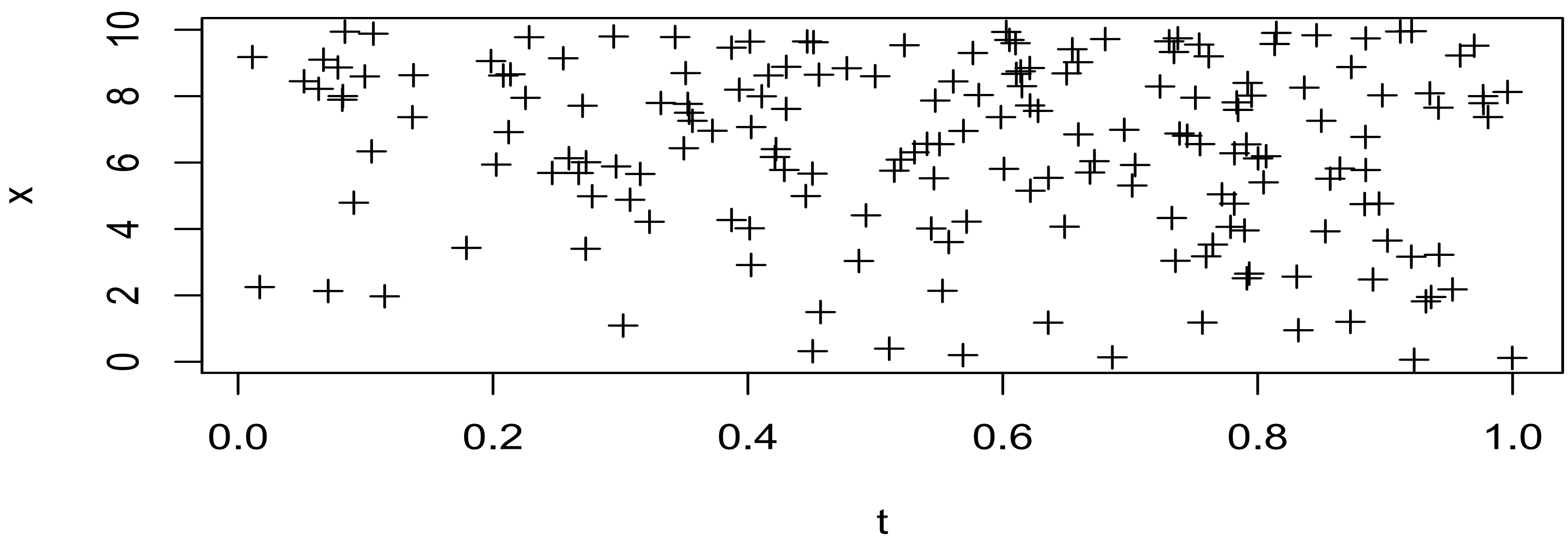
Exercises.

1. Suppose *N* is a Poisson process with $\lambda(t,x,y) = 1.5 + 10t + 2x$

on $B = [0,1] \times [0,10] \times [0,1]$.

What is EN(B)?

$\int_B \lambda(t,x,y)\, dt\, dx\, dy = 1.5(10) + 10(10)(1^2)/2 + 2(1)(10^2)/2 =$
$15+50+100=165$.



Exercises.

2. 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 λ.

d. $\lambda = 1$.

Exercises.

2. 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 λ.

d. $\lambda = 1$.

a. means λ 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")
```