Week 6.

Stat 222, Spatial Statistics. Lecture MWF 9am, Math-Sci 5203.
Professor: Rick Paik Schoenberg, frederic@ucla.edu, www.stat.ucla.edu/∼frederic

DAY TWELVE. Monday, 4/30/01.

1) Definitions of point processes.

a) $0 - 1$ valued stochastic process on S.
1 where there's a point, 0 everywhere else.
b) Random measure on S, taking values in $\mathbf{Z}^+ \cup \infty$ (the nonnegative integers or infinity).
For any subset B of S, $N(B)$ is the number of points in B.

Sometimes one of the dimensions is time, and sometimes this can be treated like another spatial dimension.

Usually we require in the definition of a point process that it be locally finite, i.e. that $N(B) < \infty$ for any compact set B.

Often there is additional information associated with each point. This additional info is called a *mark*, and the process is then called a marked point process.

2) Examples. see p 580-581.

Occurrances of events, such as plane crashes, outbreaks of disease, environmental disturbances like earthquakes, hurricanes, etc.

In such cases the marks might be the amount of destruction caused by the event.

Note that in the analysis of spatial point processes, the idea is that the points occur at random locations, and one is interested in the spatial configuration of these locations. This differentiates point processes from the coal-ash data, where the locations of the observations are not random, but are selected by the researchers. For the coal-ash data, we would not be interested, for example, in the probability that there's an observation at one location given that there's an observation at a nearby location. In point process analysis, such probabilities are exactly what we're after.

3) Basic point process concepts.

Simple: a point process is simple if all the points τ_i are distinct, i.e. $\tau_i \neq \tau_j$ if $i \neq j$.

Orderly: a point process is orderly if $\underset{|B_s| \downarrow 0}{lim} P\{N(B_s) > 1\}/|B_s| = 0$, for balls B_s around any location s.

Orderliness means the points aren't bunching up anywhere in a probabilistic sense; that is, the probabilities of there being lots of points near any particular location aren't too high.

Note that a point process is allowed to have a location s (or several such locations) where it is *sure* to have a point. But it cannot have infinitely many in any compact set.

4) Integration.

$\int_B dN = N(B) =$ the number of points in B.

Similarly, $\int_B f(x,y)dN(x,y) = \sum_{i:\tau_i in B} f(\tau_i) =$ the sum of the function $f(x,y)$ only over all locations (x,y) in B where there happen to be points.

For instance, suppose there are points at locations $(1,1)$, $(.5,.7)$ and $(2,3)$. Then $\int_B (xy)dN(x,y) = (1 \times 1) + (.5 \times .7) + (2 \times 3) = 1 + .35 + 6 = 7.35$.

5) Intensity and conditional intensity.

Intensity = rate.

$\lambda_1(s) = \lim_{|B_s| \downarrow 0} \frac{E[N(B_s)]}{|B_s|} =$ rate at which we expect points to accumulate around location s.

λ_1 is the overall intensity. There is also the conditional intensity λ (sometimes called $\lambda(s|H_s)$, which is especially relevant when one of the coordinates is *time*. In such cases it is of interest to look at the expected number of points around a location and time s, *given* the whole history H_s consisting of all information on what points have occurred previously.

$\lambda(s) = \lim_{|B_s| \downarrow 0} \frac{E[N(B_s)|H_s]}{|B_s|} =$ rate at which we expect points to accumulate around location (and time) s, given info on what points have occurred prior to s.

If, for disjoint sets B_1, B_2, ..., B_j, the number of points in these sets are *independent* of each other, then $\lambda_1 = \lambda$. If in addition N is simple and orderly, then N is called a *Poisson process*.

6) Poisson process.

Suppose N is simple and orderly, and $N(B_1), N(B_2), \ldots, N(B_j)$ are independent random variables, provided $B_1, \ldots, B_j$ are disjoint sets. Then N is a Poisson process. The name comes from the fact that it follows that for any set B, $N(B)$ has a Poisson distribution with some mean $\mu(B)$. That is:

$$P[N(B) = k] = \frac{\mu(B)^k}{k!} e^{-\mu(B)}, \tag{1}$$

for $k = 0, 1, 2, \ldots$.

DAY THIRTEEN. Wednesday, 5/9/01.

1) More about point processes.

a) How does the intensity λ_1 relate to the mean $\mu(B)$?

When λ_1 exists and is finite (note that this does not necessarily have to be the case), the mean $\mu(B)$ is just the integral of the intensity:

$\mu(B) = \int_B \lambda_1(s)ds.$

λ_1 is the expected number of points per unit area.

b) Does a Poisson process have to be stationary?

No. In general, the intensity $\lambda(s)$ for a Poisson process can vary with s. In other words, the expected number of points in one set B_1 can be different from the expected number in another set B_2, and this implies that the process is not stationary.

A Poisson process is stationary if and only if $\lambda_1(s)$ is the same for (almost) all s. That is, $\lambda_1(s) = \alpha$. In this case $\mu(B) = \alpha|B|$, for any set B.

For some reason the stationary Poisson process is often called a *homogeneous* Poisson process. The two terms are the same.

c) What does it mean for a point process to be a homogeneous Poisson process?

If a point process is homogeneous Poisson, that means it is simple, orderly, and, for any set B, given that there are exactly k points in B, those k points are distributed **uniformly** within B.

In fact, a point process obeys these conditions if and only if it is homeogeneous Poisson.

So, that means that a homogeneous Poisson process has its points uniformly scattered throughout the space S.

d) What's an example of an *inhomogeneous* Poisson process?

To specify a Poisson process, all you need to do is specify the function $\lambda_1(s)$. For instance, suppose the space S is the square $[0, 10] \times [0, 10]$. You could have $\lambda_1(s) = xy$. Or $\lambda_1(s) = x + y$. In both of these cases we'd expect few points near the origin, and lots of points in the upper-right corner. If instead $\lambda_1(s) = 2.5$, then it's a homogeneous Poisson process, and we expect 2.5 points per unit area.

e) What's an example of a *non-Poisson* point process?

May be self-exciting, or self-correcting. That is, the occurrance of points in some set B_1 may influence the number of points that will occur in a nearby set B_2. Remember, in order for the point process to be Poisson, $N(B_1)$ must be independent of $N(B_2)$, for disjoint B_1

and B_2.

Self-exciting = clustered.

Self-correcting = inhibitory.

f) What's an example of a self-exciting point process?

Poisson cluster process: start with a Poisson process, and for each point in it, consider that point a "parent". For each parent, generate a random number of "children", so that the parents' numbers of children are iid from some probability distribution P. Let the spatial locations of the children be dispersed around the parent, independently of each other, according to some spatial density f. Then look at the final map of all the children.

For instance, you might take P to be a Poisson distribution with mean 13.2, and you might let f be a uniform distribution on the unit circle. This is a self-exciting process: if you know there are points in B_1, then it's more likely that there will be points in a nearby region B_2.

g) A simple example of a self-correcting process?

Hard-core process: start with a Poisson process, and pick some number δ. Find all pairs of points that are within δ of one another, and delete **all** such points. The remaining points must be well-dispersed. If you know there's a point in B_1, then that tells you it is less likely that there's a point in a nearby region B_2.

2) Review:

White noise.

Variograms: definition, properties, models (p63?). What they say about the data.

Nugget effect and sill.

Variogram of coal-ash data and of residuals (Figs. on p51).

Pocket plot (p44).

Point process basics.

(extra-credit) Ergodicity.

All necessary figures will be supplied for you on the midterm.

DAY FOURTEEN. Friday, 5/11/01.

Midterm Exam.