STAT206 From Random Graphs To Complex Networks **Spring 2003**

Lecture 22: Calculations with Poisson Point Processes

Lecturer: David Aldous *Scribe: Shelley Zhuang*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. They may be distributed outside this class only with the permission of the Instructor.*

22.1 Calculations with Poisson Point Processes: Bounds on TSP Length, and Maxima of Empty Squares

Model:

Porbability density function $f(x)$ on R^d, real $r > 0$
Take random points $X_0, X_1, ..., X_{n-1}$ independent, density $f(\cdot)$
These are n vertices, put edge (i,j) if $\|X_i - X_j\| \leq r$.
Random graph $\mathcal{G}_n = \mathcal{G}_n(f, r)$.
Study asymptotics as $n \to \infty, r_n \downarrow 0$
What does $\mathcal{G}_n$ look like, when viewed from the typical vertex X_0?

Involves random variable $Q = f(X_0) \in (0, \infty)$
Small set A around X_0. E(# of points $X_1, ..., X_{n-1}$ in $A|X_0) \approx nf(X_0)$vol(A) = nQvol(A)

Multiply lengths by $1/r_n$, set $Y_i = \frac{X_i - X_0}{r_n}$, put edge (i,j) if $\|Y_i - Y_j\| \leq 1$.
E(# of points $Y_1, ..., Y_{n-1}$ in $B|X_0) \approx nQr_n^d$vol(B), $\forall B \subset R^d$
$\Rightarrow$ point process $(Y_1, ..., Y_{n-1})$ becomes like Poisson point process on R^d of rate $nQr_n^d \approx \alpha Q$ if $nr_n^d \to \alpha$ (nice case).

N(A) = random # of points in A, has Poisson distribution with mean = λ vol(A).
P(N(A) = 0) = $\exp(-\lambda$ vol(A))

degree(Y_0) = # of other Y's in ball $B(0,1)$, has Poisson distribution with mean = αQ vol(B(0,1))
Y's like Poisson process with rate αQ
V_d = volume of unit ball in R^d
degree($X_0|Q$) = Poisson($\alpha V_d Q$)
E degree(X_0) = αV_d EQ

Poisson process, rate 1 in R^2
Idea: relative to typical point, the other points are like the Poisson process relative to **0**.
Let D_1 be the distance from **0** to its nearest point, and D_2 be the distance from **0** to its second nearest point.
P($D_1 > x$) = P(N(Ball(0,x))=0) = exp($-\pi x^2$)
ED_1 = $\int_0^\infty$ P($D_1 > x$) dx = 1/2
P($D_2 > x$) - P($D_1 > x$) = P(N(Ball(0,x)) = 1) = $\pi x^2 e^{-\pi x^2}$
ED_2 - ED_1 = 1/4
ED_2 = 1/2 + 1/4 = 3/4

L_n = length of shortest tour through n uniform random points in $[0,1]^2$. Consider the Poisson process of rate 1 on $[0, n^{1/2}]^2$, and let L_n^* = length of tour in $[0, n^{1/2}]^2$.

$\frac{EL_n^*}{n}$ = $\frac{1}{2}$ E(sum of length of 2 edges at X_0 in optimal tour) $\geq \frac{1}{2}$ E$(D_1 + D_2)$
$lim_{n\to\infty}\frac{EL_n^*}{n} \geq 5/8$
$8/\pi \geq lim_{n\to\infty}\frac{EL_n}{n^{1/2}} \geq 5/8$

$L(X_1, ..., X_n)$ = length of shortest tour through $X_1, ..., X_n$

Lemma:
$\exists C$ such that $L(X_1, ..., X_n) \leq Cn^{1/2}$, all $(X_1, ..., X_n) \in [0,1]^2$
Given $\mathbf{X}$ = $(X_1, ..., X_n)$, set $2a = min_{i\neq j}|X_i - X_j|$
The discs $B(X_i, a)$ don't overlap
$\Rightarrow n\pi a^2 \leq (1+2a)^2$
$\Rightarrow a \leq \frac{1}{\sqrt{n\pi}}$ for n large

Given $\mathbf{X}$ = $(X_1, ..., X_n)$, $2a = |X_i - X_j|$ for some i
$L(\mathbf{X}) \leq L(\mathbf{X}\backslash\{X_i\}) + 4a \leq L(\mathbf{X}\backslash\{X_i\}) + \frac{4}{\pi}n^{-1/2}$
Inductively $L(\mathbf{X}) \leq \frac{4}{\pi}\sum_{i=1}^n i^{-1/2} \approx \frac{8}{\pi}n^{1/2}$

Probability Approximations via Poisson Clumping Heuristic Example:
L_n = side-length of largest empty square in n uniform random points in $[0,1]^2$
Fix l
$\mathcal{S}$ = random set of all $x \in R^2$ such that square $[x \pm l/2]^2$ has no points of the Poisson process
$P(L_n^* < l) \approx P(\mathcal{S} \cap [0, n^{1/2}]^2$ is empty)
Fix $x \in R^2$
P($x \in \mathcal{S}$) = P(square $[x \pm l/2]^2$ is empty) = P(N(square) = 0) = $\exp(-l^2)$
E area$(\mathcal{S} \cap [0, n^{1/2}]^2) = \int_{x\in[0,n^{1/2}]^2} P(x \in \mathcal{S})dx$ = ne^{-l^2}

Qualitatively, $\mathcal{S}$ must consist of clumps of mean size EC, spatially distributed as Poisson process of some rate λ.

$P(L_n^* < l) \approx P(\mathcal{S} \cap [0, n^{1/2}]^2$ is empty)
$\approx$ P(Poisson, rate λ, process of clumps has 0 clumps in $[0, n^{1/2}]^2$) = $\exp(-\lambda n)$

E area$(\mathcal{S} \cap [0, n^{1/2}]^2)$ = E(X E(# of clumps in $[0, n^{1/2}]^2$)) = E(X λn), where X is the clump size. Solve for $\lambda = \frac{e^{-l^2}}{EC}$.
So $P(L_n^* < l) \approx \exp(-\frac{n}{EC}e^{-l^2})$

Clump size $\widetilde{C}$ here is $\widetilde{C} = (L_1 + L_2)(L_3 + L_4)$
$P(L_1 > x) = P(N(A) = 0) = \exp$(-area A) = $\exp^{-lx}$
L_1 has exp(λ) distribution
This $\widetilde{C}$ is actually an area-biased random pick of clump area
We want C = area of uniform random clump
$Y > 0$ density $h(y) = \lambda e^{-\lambda y}$
$\widetilde{Y}$ size biased Y-density $yh(y)/E(y)$
unsize bias gives $C = L^{(1)}L^{(2)}$, $L^{(i)}$ independent, exp(l)
$EC = \frac{1}{l} \times \frac{1}{l}$
So $P(L_n^* < l) \approx \exp(-nl^2e^{-l^2})$.