Random Graphs G22.3033-007
Assignment 9. Due Monday, Apr 10,2006

`Note:` NO CLASS Monday, April 24. LAST CLASS Monday, May 1

1. Let D denote the unit disk in the plane with center $\vec{0} = (0, 0)$. Let $\vec{P}$ be at distance $1 - s$ from the $\vec{0}$, $s \in (0, 1)$. Let L be the line through $\vec{P}$ perpendicular to the line from $\vec{0}$ to $\vec{P}$. Then L splits D into two parts. Let $f(s)$ denote the area of the smaller part.

   (a) Find $f(s)$ precisely.

   (b) Give an asymptotic formula $f(s) \sim As^B$ as $s \to 0^+$.

   (c) Let $\vec{P_1}, \ldots, \vec{P_n}$ be chosen uniformly and independently from D. Call $\vec{P_i}$ *extremal* if, splitting D by a line through $\vec{P_i}$ as defined above, all the $\vec{P_j}$, $j \neq i$, lie in the larger part. Find an exact formula (as an integral) for the expected number of extremal points.

   (d) Find an asymptotic formula (as $n \to \infty$) for the above formula. (Note: The main term will be when s is "small" but finding the right parametrization for s in terms of n is the key to the problem.) (Remark: The convex hull of n randomly chosen points has been the object of much study. Extremal points are necessarily on the convex hull, though the converse is not true.)

2. By a *dumbbell* we mean two cycles joined by a path. Let r, s denote the cycle lengths and t the number of interior points of the path so that the dumbbell has $k = r + s + t$ points.

   (a) Find the number of dumbbells with parameters r, s, t on n vertices.

   (b) Find an exact expression $A = A(n, p)$ for the expected number of dumbbells in $G(n, p)$. The expression should be a sum over k as above.

   (c) Let $c < 1$ be fixed and let $G \sim G(n, p)$ with $p = \frac{c}{n}$. Prove (by showing $A(n, p) \to 0$) that almost surely G does not contain a dumbbell.

   (d) (*) Let $p = \frac{1}{n} + \lambda n^{-4/3}$ where $\lambda = \lambda(n) \to -\infty$. (This is known as the *subcritical* range.) Prove (by showing $A(n, p) \to 0$) that almost surely G does not contain a dumbbell. [Note: Connected components of G which are neither trees nor unicyclic are called *complex*. Complex components either contain dumbbells or something similar. Extending this analysis one can show that complex components do not appear in the subcritical range.]

3. Consider a branching process beginning with root Eve in which each node independently has number of children given by a Poisson distribution with mean one. List all possible outcomes (trees) for which the family (including Eve) has precisely five nodes. For each give the probability of obtaining that outcome. What is the total probability of getting a family of size precisely five?

4. Consider a branching process beginning with root Eve in which each node independently has number of children given by a Poisson distribution with mean c.

   (a) What is the probability Eve has precisely two children?

   (b) What is the probability Eve has no children with precisely two children?

   (c) Draw the family tree (including males) with root your maternal grandmother. [If you'd rather not, just make one up or use your officemate's grandmother.]

   (d) True or false: she had no children that had no children that had no children.

   (e) With the branching process defined for Eve what is the probability that she has no children that have no children that have no children?

5. For $1 \leq i < j \leq n$ let X_{ij} be independent and uniform in $[0, 1]$. Let T be the size of the minimal spanning tree (MST) on K_n with X_{ij} being the length of the edge ij. Here we derive (minus some technical details) a remarkable formula for the asymptotic expectation of T.

   (a) Argue $E[T] = \binom{n}{2} \int_0^1 t f^-(t) dt$ where $f^-(t)$ is the probability that edge $\{1, 2\}$ is in the MST conditional on $X_{12} = t$.

   (b) Assume $X_{12} = t$. Argue that $\{1, 2\}$ is in the MST if and only if there is no path from 1 to 2 consisting of edges all of weight less than t. (Hint: Consider Kruskal's algorithm.)

(c) Deduce $f^-(t)$ is the probability that $1, 2$ don't lie in the same component of $G(n, t)$, conditional on $1, 2$ not being adjacent.

(d) Now set $t = \frac{c}{n}$ with $c > 0$ arbitrary and asymptotics as $n \to \infty$. Let $f(t)$ be the probability that $1, 2$ don't lie in the same component of $G(n, t)$ (with no conditioning). Argue that $f^-(t) = f(t) + o(1)$. (This is part of a very general principle that conditioning on an almost sure event changes a probability by $o(1)$.)

(e) (The meat of the problem.) Find the limiting value of $f(t)$ as a function $g(c)$. (There will be two cases and $g(c)$ may be given implicitly. The key is to look at the almost sure picture of $G(t)$ given in class.)

(f) For $K > 0$ let $E_K[T]$ denote the sum of the lengths of the minimal spanning tree, counting only lengths that are at most K/n. Express $\lim_{n\to\infty} E_K[T]$ as an integral.

(g) It can be shown that $\lim_{n\to\infty} E[T] = \lim_{K\to\infty} \lim_{n\to\infty} E_K[T]$ and lets assume that. (But I hope you see that this is *not* obvious. If, for example, most of the weight from the minimal spanning tree came from edges of weight between $n^{-1/2}$ and $2n^{-1/2}$ then it wouldn't be true.) Given this assumption find $\lim_{n\to\infty} E[T]$ as an integral. Evaluate the integral numerically.

(h) (*) Express $\lim_{n\to\infty} E[T]$ in a nice form which overlaps the name of a movie star.

I think that it is a relatively good approximation to truth - which is much too complicated to allow anything but approximations - that mathematical ideas originate in empirics, although the genealogy is sometime long and obscure. But, once they are so conceived, the subject begins to live a particular life of its own and it is better compared to a creative one, governed by almost entirely aesthetical motivations, than to anything else and, in particular, to an empirical science.
– John von Neumann