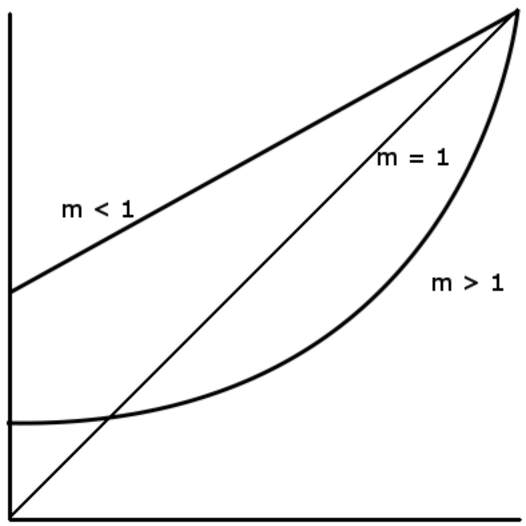
Lecture 14

Scribe: Andrew Owens

February 18, 2009

Previously:

Let the generating function $f(x) = \sum_i^\infty p_i x^i$, where p_i is the probability of having i children. Let m be the slope of f at 1 (i.e. $m = f'(1)$). As we showed in the previous lecture, you can think of $m = 1$ as being the threshold between infinite and non-infinite trees. We also showed in class that $q = f(f(f(f(...))))$ is the probability of extinction.



Lemma: Let $m = f'(1)$. If $m \neq 1$ then the expected size of an extinct family is finite. If $m = 1$ and $p_1 = 1$, then the tree is infinite and has regular degree, so there is no extinct family. If $m = 1$ and $p_1 < 1$, then the expected size is infinite.

Proof: Let z_i be the size of generation i. The expected value of z_1 given extinction may differ from the expected value of z_1 in general. This makes sense, because if you know that a process is extinct, then the first generation is more likely to be small.

By Bayes rule, $P(z_1 = k|extinction) = \frac{P(extinction|z_1=k)P(z_1=k)}{P(extinction)}$. You can verify this statement by multiplying both sides by $P(extinction)$ which yields the joint probability $P(extinction, z_i = k)$ on both sides of the equation. Note that if $z_1 = k$, the probability of extinction is the probability that each of the k children become extinct. Thus $P(extinction|z_1 = k) = q^k$ and $P(z_1 = k|extinction) = \frac{p_k q^k}{q} = p_k q^{k-1}$.

To verify that $\sum_{k=0}^\infty P(z_1 = k|extinction) = 1$, note that $\sum_{k=0}^\infty p_k q^{k-1} = \frac{1}{q}\sum_{k=0}^\infty p_k q^k = \frac{f(q)}{q}$. But $f(q) = q$, which proves the statement.

Now we calculate: $E(z_1|extinction) = \sum_{k=0}^\infty kP(z_i = k|extinction) = \sum_{k=0}^\infty kp_k q^{k-1}$. Note that $f'(x) = \sum_{k=0}^\infty kp_k x^{k-1}$ and hence $E(z_1|extinction) = f'(q)$.

What is the expected size of z_1? $\sum_{k=0}^\infty kp_k = \sum_{k=0}^\infty kp_k x^{k-1}\big]_{x=1} = f'(1) = m$.

Note that if $m < 1$, there are no infinite trees. Thus $E(z_1|extinction) = E(z_1) = m$.

Expected size of generation i (slightly different than what was presented in lecture)

Intuitively, this value seems to be $[E(z_1|extinction)]^i = [f'(q)]^i$, but we need to be more rigorous.

Any member of generation i must have been a child from some vertex in generation $i-1$. So if we number the vertices from generation $i-1$, the number of vertices in generation i is the number of children from vertex 1, plus the number of children from vertex 2, etc. Furthermore, the number of children from a particular vertex in generation $i-1$ is distributed the same as z_1, since each vertex is an independent branching process. Putting these ideas together: since there are z_{i-1} vertices in generation $i-1$, we find $z_i = \sum_{i=1}^{z_{i-1}} z_1$. Note that z_{i-1} is a random variable, so we now need a technique that will allow us to calculate the expectation of this expression.

Let n be a fixed constant. Then by linearity of expectation, $E(\sum_i^n X_i) = \sum_i^n E(X_i)$. If all of the X_i have the same distribution, then $E(\sum_i^n X_i) = nE(X_i)$.

What do you do if the number of variables, n is a random variable? In this case, $E(\sum_{i=1}^n X_i) = P(n = 1)E(X_1|n = 1) + P(n = 2)E(X_1 + X_2|n = 2) + P(n = 3)E(X_1 + X_2 + X_3|n = 3) + ...$ But the X_i are statistically independent, so applying linearity of expectation, $E(\sum_{i=1}^n X_i) = P(n = 1)E(X_1) + 2P(n = 2)E(X_1) + 3P(n = 3)E(X_1) + ... = E(X_1)[1P(n = 1) + 2P(n = 2) + 3P(n = 3) + ... = E(n)E(X_1)$.

It follows that $E(z_i|extinction) = E(\sum_{i=1}^{z_{i-1}} z_1|extinction) = E(z_{i-1}|extinction)E(z_1|extinction)$. Therefore $E(z_i|extinction) = [E(z_1)]^i = [f'(q)]^i$. Thus the expected size of a tree is $E(tree) = \sum_{i=0}^{\infty}(f'(q))^i = \frac{1}{1-f'(q)}$.

If $m < 1$, then $f(x) = x$ when $x = 1$. Therefore $q = 1$ and so $f'(q) < 1$. If $m > 1$, then $f(x) = x$ at $x = q$ and $x = 1$. Therefore the slope at q must have been less than 1. Thus if $m \neq 1$, the expected size of the tree is finite.

Generating functions

Here are some more uses for generating functions. Let $a_0, a_1, a_2, ...$ denote the generating function $g(x) = \sum_{i=0}^{\infty} a_i x^i$.

- $a_1, 2a_2, 3a_3 = x\frac{d}{dx}(g(x)) = \sum_{i=0}^{\infty} i a_i x^{i-1}$.
- $1, 1, 1, 1 = \sum_{i=0}^{\infty} x^i = \frac{1}{1-x}$.
- $0, 1, 2, 3 = \sum_{i=0}^{\infty} i x^{i-1} = \frac{d}{dx}(\sum_{i=0}^{\infty} x^i) = \frac{x}{(1-x)^2}$.

Choosing items Suppose A can be selected 0 or 1 times: generating function $1 + x$. B: 0, 1, or 2 times; $1 + x + x^2$. C: 0, 1, 2, 3 times; $1 + x + x^2 + x^3$. How many ways can you select five items? Multiplying the three generating functions gives us $(1+x)(1+x+x^2)(1+x+x^2+x^3) = 1+5x+5x^2+6x^3+5x^4+3x^5+x^6$. Thus there are 3 ways to choose five items, the coefficient of x^5. These are CCCBB, CCCBA, and CCBBA.

Making change: How many ways are there to make 23 cents with pennies, nickels, and dimes?

Pennies: $1 + x + x^2 + ...$

Nickels: $1 + x^5 + x^{10} + ...$

Dimes: $1 + x^{10} + x^{20} + ...$

Multiplying these three generating functions together, we find that the coefficient of x^{23} is 9. Indeed, there are nine ways: $ddppp$, $dnnppp$, dnp^8, dp^{13}, n^4p^3, n^3p^8, n^2p^{13}, np^{18}, and p^{23}.