Generating Functions and Networks

Last updated: 2019/01/14, 23:14:28

Complex Networks | @networksvox
CSYS/MATH 303, Spring, 2019

Prof. Peter Dodds | @peterdodds

Dept. of Mathematics & Statistics | Vermont Complex Systems Center
Vermont Advanced Computing Core | University of Vermont

These slides are brought to you by:

These slides are also brought to you by:

Special Guest Executive Producer

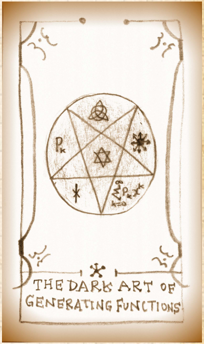
On Instagram at pratchett_the_cat

Outline


THE SUN
THE GOLDEN AGE OF REDUCTIONISM
THE MANIFESTO
THE WELL OF SCALES
THE LAWS OF DECAYING POWER
THE UNEXPECTED TRANSFIGURATION
MINISTRY OF RANDOM WALKS
THE RICH GET RICHER
THE DREAM OF EFFICIENT LANGUAGE
THE LAW OF DIGITS FIRST
THE SURPRISE OF BEING ROBUST YET FRAGILE
THE MISREADING OF THE BOOK OF SAND
THE MANY TALES OF MANY TAILS
THE MEASURING
THE DATA POOR
THE POURING DATA
THE CHARMS OF THE MANY WRONG PATHS
THE KNOWLEDGE OF DEAD MONKS
THE KNOWLEDGE OF DEAD MONKS
THE EMERGENCE OF STRUCTURE
THE EMERGENCE OF DESTRUCTION
THE EMERGENCE OF THINKING
It was the best of times
THE EMERGENCE OF STORIES
THE MECHANICS STATISTICAL
THE THRESHOLDS OF PERCOLATION
THE NETWORKS COMPLEX
THE ELEMENTS OF WEBS
THE BIRDS OF A FEATHER
THE BIRDS OF NOT A FEATHER
THE STRANGENESS OF FRIENDS
THE SMALLNESS OF WORLDS
THE THEORY OF SIX DEGREES
THE NETWORKS FREE OF SCALE
CLASSIFIED
PHOENIX
THE CONTAGION
THE CONFUSION OF CONTAGIONS
THE DISEASE OF RANDOM MIXING
THE UNENDING PANDEMICS
THE THRESHOLDS OF THE MIND
THE SOCIAL WILD
THE SOCIAL CONTAGION
THE RATIO OF GAIN
THE MECHANICS OF FAME
THE TALK OF FAME
THE STORIES CONTAGIOUS
THE JOHN DORY
THE ADJACENT NARRATIVES
THE STORYTELLERS
THE PATHS OF UNIVERSALITY
THE DEPENDENT PATHS
THE TROUSERS OF BIFURCATING REALITIES
THE UNKNOWN MECHANISM
THE LANDSCAPES OF FORKING PATHS
THE NETWORKS OF BLOOD
THE TREES OF REALITY
THE ORDERS OF STREAMS
THE LAWS OF BRANCHING
THE SCALING OF BROKEN HIERARCHIES
THE BRAID OR THE BRANCH
THE LAW OF OPTIMAL FORKS
THE CHURCH OF QUARTEROLOGY
THE TRUTHICIDE
THE FAMILY OF PANCAKE COWS
THE FAMILY OF SPHERICAL COWS
THE IMPROBABLE PACHYDERM
THE ILL-SCALING OF ALLOMETRIC BODIES
THE SUPPLY MOST FAIR
THE SUPPLY TO THE MOST
THE DARK ART OF GENERATING FUNCTIONS
THE RANDOM SUM OF RANDOMNESS
THE TWO-SIDED WORLDS
THE STORIES AND THE TROPES
THE TRUTH OF DIFFUSION
THE UNVEILING OF HIDDEN POWERS
THE SCRYING FOR THE SHAPES OF THINGS
THE LIMITS OF ORGANIZATIONS
THE OVERCOMING OF CATASTROPHE
THE TROUBLES ILL-FORMED
THE NETWORKS OF MANY ILKS
THE CHAOS OF FETTERED CONTAGION
THE END


"I am the Monarch of the Sea..."
N,m
THE ARK OF THE RANDOM NETWORK


THE SUN
THE GOLDEN AGE OF REDUCTIONISM
THE MANIFESTO
THE WELL OF SCALES
THE LAWS OF DECAYING POWER
THE UNEXPECTED TRANSFIGURATION
MINISTRY OF RANDOM WALKS
THE RICH GET RICHER
THE DREAM OF EFFICIENT LANGUAGE
THE LAW OF DIGITS FIRST
THE SURPRISE OF BEING ROBUST YET FRAGILE
THE MISREADING OF THE BOOK OF SAND
THE MANY TALES OF MANY TAILS
THE MEASURING
THE DATA POOR
THE POURING DATA
THE CHARMS OF THE MANY WRONG PATHS
THE KNOWLEDGE OF DEAD MONKS
THE KNOWLEDGE OF DEAD MONKS
THE EMERGENCE OF STRUCTURE
THE EMERGENCE OF DESTRUCTION
THE EMERGENCE OF THINKING
THE EMERGENCE OF STORIES
THE MECHANICS STATISTICAL
THE THRESHOLDS OF PERCOLATION
THE NETWORKS COMPLEX
THE ELEMENTS OF WEBS
THE BIRDS OF A FEATHER
THE BIRDS OF NOT A FEATHER
THE ARKS OF THE RANDOM NETWORKS
THE STRANGENESS OF FRIENDS
THE SMALLNESS OF WORLDS
THE THEORY OF SIX DEGREES
THE NETWORKS FREE OF SCALE
THE CONTAGION
THE CONFUSION OF CONTAGIONS
THE DISEASE OF RANDOM MIXING
THE UNENDING PANDEMICS
THE THRESHOLDS OF THE MIND
THE SOCIAL WILD
THE SOCIAL CONTAGION
THE RATIO OF GAIN
THE MECHANICS OF FAME
THE TALK OF FAME
THE STORIES CONTAGIOUS
THE JOHN DORY
THE ADJACENT NARRATIVES
THE STORYTELLERS
THE PATHS OF UNIVERSALITY
THE DEPENDENT PATHS
THE TROUSERS OF BIFURCATING REALITIES
THE UNKNOWN MECHANISM
THE LANDSCAPES OF FORKING PATHS
THE NETWORKS OF BLOOD
THE TREES OF REALITY
THE ORDERS OF STREAMS
THE LAWS OF BRANCHING
THE SCALING OF BROKEN HIERARCHIES
THE BRAID OR THE BRANCH
THE LAW OF OPTIMAL FORKS
THE CHURCH OF QUARTEROLOGY
THE TRUTHICIDE
THE FAMILY OF PANCAKE COWS
THE FAMILY OF SPHERICAL COWS
THE IMPROBABLE PACHYDERM
THE ILL-SCALING OF ALLOMETRIC BODIES
THE SUPPLY MOST FAIR
THE SUPPLY TO THE MOST
THE RANDOM SUM OF RANDOMNESS
THE TWO-SIDED WORLDS
THE STORIES AND THE TROPES
THE TRUTH OF DIFFUSION
THE UNVEILING OF HIDDEN POWERS
THE SCRYING FOR THE SHAPES OF THINGS
THE LIMITS OF ORGANIZATIONS
THE OVERCOMING OF CATASTROPHE
THE TROUBLES ILL-FORMED
THE NETWORKS OF MANY ILKS
THE CHAOS OF FETTERED CONTAGION
THE END


THE DARK ART OF GENERATING FUNCTIONS

Generatingfunctionology [1]

- Idea: Given a sequence $a_0, a_1, a_2, \ldots$, associate each element with a distinct function or other mathematical object.
- Well-chosen functions allow us to manipulate sequences and retrieve sequence elements.

Definition:

- The generating function (g.f.) for a sequence $\{a_n\}$ is

$$F(x) = \sum_{n=0}^{\infty} a_n x^n.$$

- Roughly: transforms a vector in R^∞ into a function defined on R^1.
- Related to Fourier, Laplace, Mellin, ...

Simple examples:

Rolling dice and flipping coins:

- $p_k^{(\text{⚂})} = \textbf{Pr}(\text{throwing a } k) = 1/6$ where $k = 1, 2, \ldots, 6.$

$$F^{(\text{⚂})}(x) = \sum_{k=1}^{6} p_k^{(\text{⚂})} x^k = \frac{1}{6}(x + x^2 + x^3 + x^4 + x^5 + x^6).$$

- $p_0^{(\text{coin})} = \textbf{Pr}(\text{head}) = 1/2$, $p_1^{(\text{coin})} = \textbf{Pr}(\text{tail}) = 1/2.$

$$F^{(\text{coin})}(x) = p_0^{(\text{coin})} x^0 + p_1^{(\text{coin})} x^1 = \frac{1}{2}(1 + x).$$

- A generating function for a probability distribution is called a Probability Generating Function (p.g.f.).
- We'll come back to these simple examples as we derive various delicious properties of generating functions.

Example

- Take a degree distribution with exponential decay:

$$P_k = ce^{-\lambda k}$$

where geometricsumfully, we have $c = 1 - e^{-\lambda}$

- The generating function for this distribution is

$$F(x) = \sum_{k=0}^{\infty} P_k x^k = \sum_{k=0}^{\infty} ce^{-\lambda k} x^k = \frac{c}{1 - xe^{-\lambda}}.$$

- Notice that $F(1) = c/(1 - e^{-\lambda}) = 1$.
- For probability distributions, we must always have $F(1) = 1$ since

$$F(1) = \sum_{k=0}^{\infty} P_k 1^k = \sum_{k=0}^{\infty} P_k = 1.$$

- Check die and coin p.g.f.'s.

Properties:

- Average degree:

$$\langle k \rangle = \sum_{k=0}^{\infty} k P_k = \sum_{k=0}^{\infty} k P_k x^{k-1} \bigg|_{x=1}$$

$$= \frac{\mathrm{d}}{\mathrm{d}x} F(x) \bigg|_{x=1} = F'(1)$$

- In general, many calculations become simple, if a little abstract.
- For our exponential example:

$$F'(x) = \frac{(1 - e^{-\lambda})e^{-\lambda}}{(1 - xe^{-\lambda})^2}.$$

- So: $\langle k \rangle = F'(1) = \dfrac{e^{-\lambda}}{(1 - e^{-\lambda})}.$
- Check for die and coin p.g.f.'s.

Useful pieces for probability distributions:

- Normalization:

$$F(1) = 1$$

- First moment:

$$\langle k \rangle = F'(1)$$

- Higher moments:

$$\langle k^n \rangle = \left(x \frac{\mathrm{d}}{\mathrm{d}x} \right)^n F(x) \bigg|_{x=1}$$

- kth element of sequence (general):

$$P_k = \frac{1}{k!} \frac{\mathrm{d}^k}{\mathrm{d}x^k} F(x) \bigg|_{x=0}$$

A beautiful, fundamental thing:

- The generating function for the sum of two random variables

$$W = U + V$$

is

$$F_W(x) = F_U(x)F_V(x).$$

- Convolve yourself with Convolutions: Insert question from assignment 5 .
- Try with die and coin p.g.f.'s.
  1. Add two coins (tail=0, head=1).
  2. Add two dice.
  3. Add a coin flip to one die roll.

Edge-degree distribution

- Recall our condition for a giant component:

$$\langle k \rangle_R = \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle} > 1.$$

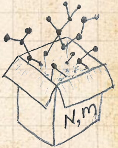
- Let's re-express our condition in terms of generating functions.
- We first need the g.f. for R_k.
- We'll now use this notation:

  $F_P(x)$ is the g.f. for P_k.
  $F_R(x)$ is the g.f. for R_k.

- Giant component condition in terms of g.f. is:

$$\langle k \rangle_R = F_R'(1) > 1.$$

- Now find how F_R is related to F_P ...



Edge-degree distribution

- We have

$$F_R(x) = \sum_{k=0}^{\infty} R_k x^k = \sum_{k=0}^{\infty} \frac{(k+1)P_{k+1}}{\langle k \rangle} x^k.$$

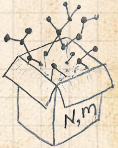
Shift index to $j = k + 1$ and pull out $\frac{1}{\langle k \rangle}$:

$$F_R(x) = \frac{1}{\langle k \rangle} \sum_{j=1}^{\infty} j P_j x^{j-1} = \frac{1}{\langle k \rangle} \sum_{j=1}^{\infty} P_j \frac{\mathrm{d}}{\mathrm{d}x} x^j$$

$$= \frac{1}{\langle k \rangle} \frac{\mathrm{d}}{\mathrm{d}x} \sum_{j=1}^{\infty} P_j x^j = \frac{1}{\langle k \rangle} \frac{\mathrm{d}}{\mathrm{d}x} \left(F_P(x) - P_0 \right) = \frac{1}{\langle k \rangle} F_P'(x).$$

Finally, since $\langle k \rangle = F_P'(1)$,

$$\boxed{F_R(x) = \frac{F_P'(x)}{F_P'(1)}}$$



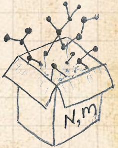
Edge-degree distribution

- Recall giant component condition is $\langle k \rangle_R = F'_R(1) > 1$.
- Since we have $F_R(x) = F'_P(x)/F'_P(1)$,

$$F'_R(x) = \frac{F''_P(x)}{F'_P(1)}.$$

- Setting $x = 1$, our condition becomes

$$\boxed{\frac{F''_P(1)}{F'_P(1)} > 1}$$



Size distributions

To figure out the size of the largest component (S_1), we need more resolution on component sizes.

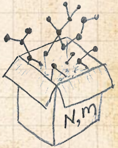
Definitions:

- π_n = probability that a random node belongs to a finite component of size $n < \infty$.
- ρ_n = probability that a random end of a random link leads to a finite subcomponent of size $n < \infty$.

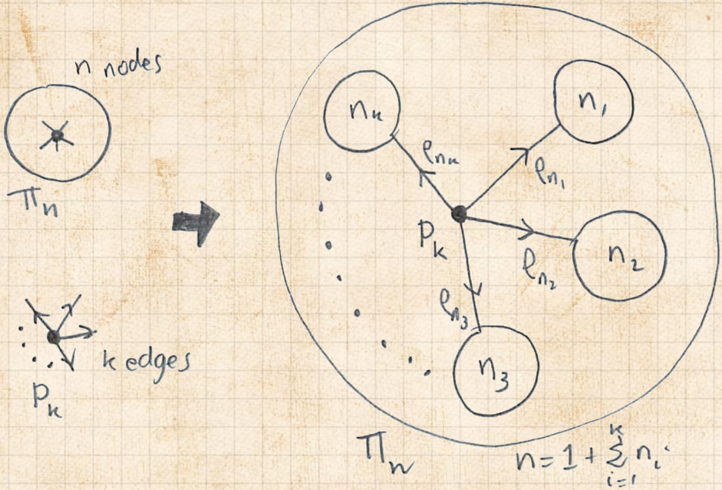
Local-global connection:

$$P_k, R_k \Leftrightarrow \pi_n, \rho_n$$

neighbors ⇔ components

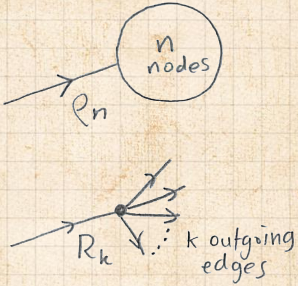


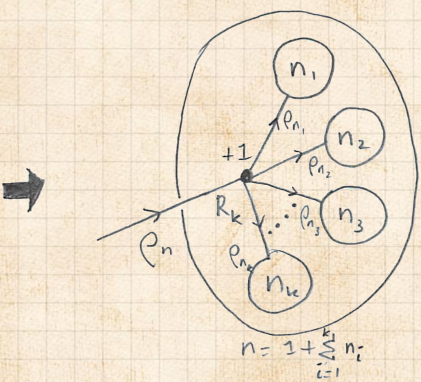
Connecting probabilities:



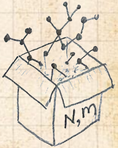
- Markov property of random networks connects π_n, ρ_n, and P_k.

Connecting probabilities:





- Markov property of random networks connects ρ_n and R_k.



G.f.'s for component size distributions:

- $$F_\pi(x) = \sum_{n=0}^{\infty} \pi_n x^n \text{ and } F_\rho(x) = \sum_{n=0}^{\infty} \rho_n x^n$$

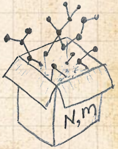
The largest component:

- Subtle key: $F_\pi(1)$ is the probability that a node belongs to a finite component.
- Therefore: $S_1 = 1 - F_\pi(1)$.

Our mission, which we accept:

- Determine and connect the four generating functions

$$F_P, F_R, F_\pi, \text{ and } F_\rho.$$



Useful results we'll need for g.f.'s

Sneaky Result 1:

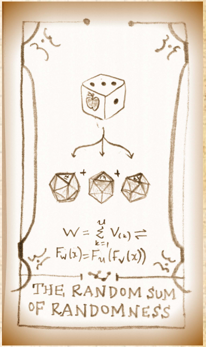
- Consider two random variables U and V whose values may be $0, 1, 2, \ldots$
- Write probability distributions as U_k and V_k and g.f.'s as F_U and F_V.
- SR1: If a third random variable is defined as

$$W = \sum_{i=1}^{U} V^{(i)} \text{ with each } V^{(i)} \stackrel{d}{=} V$$

then

$$\boxed{F_W(x) = F_U\left(F_V(x)\right)}$$


THE SUN
THE GOLDEN AGE OF REDUCTIONISM
THE MANIFESTO
THE WELL OF SCALES
THE LAWS OF DECAYING POWER
THE UNEXPECTED TRANSFIGURATION
MINISTRY OF RANDOM WALKS
THE RICH GET RICHER
THE DREAM OF EFFICIENT LANGUAGE
THE LAW OF DIGITS FIRST
THE SURPRISE OF BEING ROBUST YET FRAGILE
THE MISREADING OF THE BOOK OF SAND
THE MANY TALES OF MANY TAILS
THE MEASURING
THE DATA POOR
THE POURING DATA
THE CHARMS OF THE MANY WRONG PATHS
THE KNOWLEDGE OF DEAD MONKS
THE KNOWLEDGE OF DEAD MONKS
THE EMERGENCE OF STRUCTURE
THE EMERGENCE OF DESTRUCTION
THE EMERGENCE OF THINKING
It was the best of times
THE EMERGENCE OF STORIES
THE MECHANICS STATISTICAL
THE THRESHOLDS OF PERCOLATION
THE NETWORKS COMPLEX
THE ELEMENTS OF WEBS
THE BIRDS OF A FEATHER
THE BIRDS OF NOT A FEATHER
THE ARKS OF THE RANDOM NETWORK
THE STRANGENESS OF FRIENDS
THE SMALLNESS OF WORLDS
THE THEORY OF SIX DEGREES
THE NETWORKS FREE OF SCALE
CLASSIFIED PHOENIX
BIO HAZARD
THE CONTAGION
THE CONFUSION OF CONTAGIONS
THE DISEASE OF RANDOM MIXING
THE UNENDING PANDEMICS
THE THRESHOLDS OF THE MIND
THE SOCIAL WILD
THE SOCIAL CONTAGION
THE RATIO OF GAIN
THE MECHANICS OF FAME
THE TALK OF FAME
THE STORIES CONTAGIOUS
THE JOHN DORY
THE ADJACENT NARRATIVES
THE STORYTELLERS
THE PATHS OF UNIVERSALITY
THE DEPENDENT PATHS
THE TROUSERS OF BIFURCATING REALITIES
THE UNKNOWN MECHANISM
THE LANDSCAPES OF FORKING PATHS
THE NETWORKS OF BLOOD
THE TREES OF REALITY
THE ORDERS OF STREAMS
THE LAWS OF BRANCHING
THE SCALING OF BROKEN HIERARCHIES
THE BRAID OR THE BRANCH
THE LAW OF OPTIMAL FORKS
THE CHURCH OF QUARTEROLOGY
THE TRUTHICIDE
THE FAMILY OF PANCAKE COWS
THE FAMILY OF SPHERICAL COWS
THE IMPROBABLE PACHYDERM
THE ILL-SCALING OF ALLOMETRIC BODIES
THE SUPPLY MOST FAIR
THE SUPPLY TO THE MOST
THE DARK ART OF GENERATING FUNCTIONS
THE TWO-SIDED WORLDS
THE STORIES AND THE TROPES
THE TRUTH OF DIFFUSION
THE UNVEILING OF HIDDEN POWERS
THE SCRYING FOR THE SHAPES OF THINGS
THE LIMITS OF ORGANIZATION
THE OVERCOMING OF CATASTROPHE
THE TROUBLES ILL-FORMED
THE NETWORKS OF MANY ILKS
THE CHAOS OF FETTERED CONTAGION
THE END


$W = \sum_{k=1}^{U} V_{(k)}$
$F_W(x) = F_U(F_V(x))$
THE RANDOM SUM OF RANDOMNESS

Proof of SR1:

Write probability that variable W has value k as W_k.

$$W_k = \sum_{j=0}^{\infty} U_j \times \text{Pr(sum of } j \text{ draws of variable } V = k)$$

$$= \sum_{j=0}^{\infty} U_j \sum_{\substack{\{i_1, i_2, \ldots, i_j\} | \\ i_1 + i_2 + \ldots + i_j = k}} V_{i_1} V_{i_2} \cdots V_{i_j}$$

$$\therefore F_W(x) = \sum_{k=0}^{\infty} W_k x^k = \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} U_j \sum_{\substack{\{i_1, i_2, \ldots, i_j\} | \\ i_1 + i_2 + \ldots + i_j = k}} V_{i_1} V_{i_2} \cdots V_{i_j} x^k$$

$$= \sum_{j=0}^{\infty} U_j \sum_{k=0}^{\infty} \sum_{\substack{\{i_1, i_2, \ldots, i_j\} | \\ i_1 + i_2 + \ldots + i_j = k}} V_{i_1} x^{i_1} V_{i_2} x^{i_2} \cdots V_{i_j} x^{i_j}$$

Proof of SR1:

With some concentration, observe:

$$F_W(x) = \sum_{j=0}^{\infty} U_j \sum_{k=0}^{\infty} \underbrace{\underbrace{\sum_{\substack{\{i_1, i_2, \ldots, i_j\} | \\ i_1+i_2+\ldots+i_j=k}} V_{i_1} x^{i_1} V_{i_2} x^{i_2} \cdots V_{i_j} x^{i_j}}_{x^k \text{ piece of } \left(\sum_{i'=0}^{\infty} V_{i'} x^{i'}\right)^j}}_{\left(\sum_{i'=0}^{\infty} V_{i'} x^{i'}\right)^j = (F_V(x))^j}$$

$$= \sum_{j=0}^{\infty} U_j \left(F_V(x)\right)^j$$

$$= F_U\left(F_V(x)\right)$$

- Alternate, groovier proof in the accompanying assignment.

Useful results we'll need for g.f.'s

Sneaky Result 2:

- Start with a random variable U with distribution U_k $(k = 0, 1, 2, \ldots)$
- SR2: If a second random variable is defined as

$$V = U + 1 \quad \text{then} \quad \boxed{F_V(x) = xF_U(x)}$$

- Reason: $V_k = U_{k-1}$ for $k \geq 1$ and $V_0 = 0$.
-

$$\therefore F_V(x) = \sum_{k=0}^{\infty} V_k x^k = \sum_{k=1}^{\infty} U_{k-1} x^k$$

$$= x \sum_{j=0}^{\infty} U_j x^j = xF_U(x).$$

Useful results we'll need for g.f.'s

Generalization of SR2:

- (1) If $V = U + i$ then

$$F_V(x) = x^i F_U(x).$$

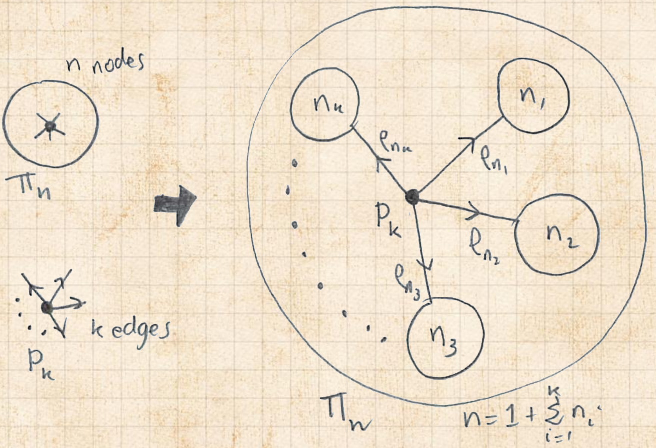
- (2) If $V = U - i$ then

$$F_V(x) = x^{-i} F_U(x)$$

$$= x^{-i} \sum_{k=0}^{\infty} U_k x^k$$

Connecting generating functions:

- Goal: figure out forms of the component generating functions, F_π and F_ρ.



- Relate π_n to P_k and ρ_n through one step of recursion.

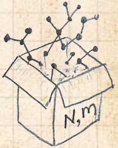
Connecting generating functions:

- π_n = probability that a random node belongs to a finite component of size n

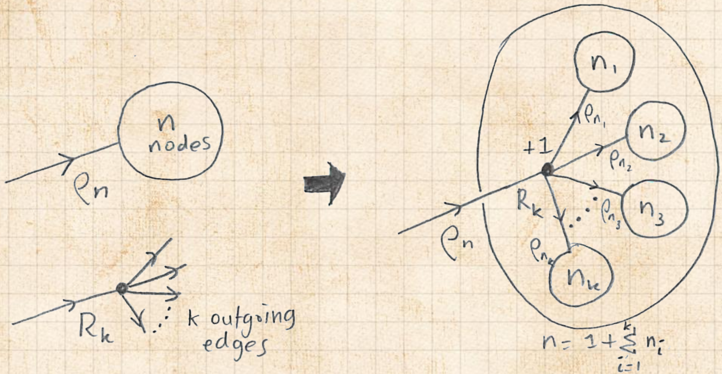
$$= \sum_{k=0}^{\infty} P_k \times \Pr\left(\begin{array}{l}\text{sum of sizes of subcomponents} \\ \text{at end of } k \text{ random links} = n-1\end{array}\right)$$

- Therefore: $$\boxed{F_\pi(x) = \underbrace{x}_{\text{SR2}} \underbrace{F_P\left(F_\rho(x)\right)}_{\text{SR1}}}$$

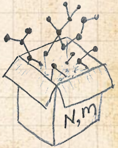
- Extra factor of x accounts for random node itself.



Connecting generating functions:



- Relate ρ_n to R_k and ρ_n through one step of recursion.

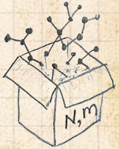


Connecting generating functions:

- ρ_n = probability that a random link leads to a finite subcomponent of size n.
- Invoke one step of recursion:
  ρ_n = probability that in following a random edge, the outgoing edges of the node reached lead to finite subcomponents of combined size $n-1$,

$$= \sum_{k=0}^{\infty} R_k \times \text{Pr}\left(\begin{array}{l} \text{sum of sizes of subcomponents} \\ \text{at end of } k \text{ random links} = n-1 \end{array} \right)$$

- Therefore: $$F_\rho(x) = \underbrace{x}_{\text{SR2}} \underbrace{F_R\left(F_\rho(x)\right)}_{\text{SR1}}$$
- Again, extra factor of x accounts for random node itself.

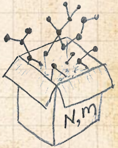


Connecting generating functions:

- We now have two functional equations connecting our generating functions:

$$F_\pi(x) = xF_P\left(F_\rho(x)\right) \text{ and } F_\rho(x) = xF_R\left(F_\rho(x)\right)$$

- Taking stock: We know $F_P(x)$ and $F_R(x) = F'_P(x)/F'_P(1)$.
- We first untangle the second equation to find F_ρ
- We can do this because it only involves F_ρ and F_R.
- The first equation then immediately gives us F_π in terms of F_ρ and F_R.



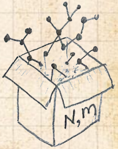
Component sizes

- Remembering vaguely what we are doing:

  Finding F_{π} to obtain the fractional size of the largest component $S_1 = 1 - F_{\pi}(1)$.
- Set $x = 1$ in our two equations:

$$F_{\pi}(1) = F_P\left(F_{\rho}(1)\right) \quad \text{and} \quad F_{\rho}(1) = F_R\left(F_{\rho}(1)\right)$$

- Solve second equation numerically for $F_{\rho}(1)$.
- Plug $F_{\rho}(1)$ into first equation to obtain $F_{\pi}(1)$.



Component sizes

Example: Standard random graphs.

- We can show $F_P(x) = e^{-\langle k \rangle (1-x)}$

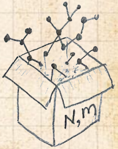
$$\Rightarrow F_R(x) = F_P'(x)/F_P'(1)$$

$$= \langle k \rangle e^{-\langle k \rangle (1-x)} / \langle k \rangle e^{-\langle k \rangle (1-x')}|_{x'=1}$$

$$= e^{-\langle k \rangle (1-x)} = F_P(x)$$

...aha!

- RHS's of our two equations are the same.
- So $F_\pi(x) = F_\rho(x) = xF_R(F_\rho(x)) = xF_R(F_\pi(x))$
- Consistent with how our dirty (but wrong) trick worked earlier ...
- $\pi_n = \rho_n$ just as $P_k = R_k$.

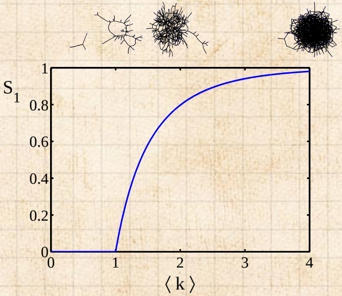


Component sizes

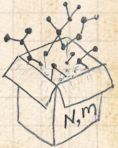
- We are down to $F_\pi(x) = xF_R(F_\pi(x))$ and $F_R(x) = e^{-\langle k \rangle (1-x)}$.
- $$\therefore F_\pi(x) = xe^{-\langle k \rangle (1-F_\pi(x))}$$
- We're first after $S_1 = 1 - F_\pi(1)$ so set $x = 1$ and replace $F_\pi(1)$ by $1 - S_1$:

$$1 - S_1 = e^{-\langle k \rangle S_1}$$

Or: $\langle k \rangle = \frac{1}{S_1} \ln \frac{1}{1 - S_1}$

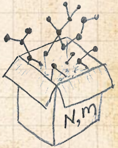


- Just as we found with our dirty trick ...
- Again, we (usually) have to resort to numerics ...



A few simple random networks to contemplate and play around with:

- Notation: The Kronecker delta function $\delta_{ij} = 1$ if $i = j$ and 0 otherwise.
- $P_k = \delta_{k1}$.
- $P_k = \delta_{k2}$.
- $P_k = \delta_{k3}$.
- $P_k = \delta_{kk'}$ for some fixed $k' \geq 0$.
- $P_k = \frac{1}{2}\delta_{k1} + \frac{1}{2}\delta_{k3}$.
- $P_k = a\delta_{k1} + (1-a)\delta_{k3}$, with $0 \leq a \leq 1$.
- $P_k = \frac{1}{2}\delta_{k1} + \frac{1}{2}\delta_{kk'}$ for some fixed $k' \geq 2$.
- $P_k = a\delta_{k1} + (1-a)\delta_{kk'}$ for some fixed $k' \geq 2$ with $0 \leq a \leq 1$.



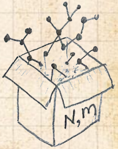
A joyful example:

$$P_k = \frac{1}{2}\delta_{k1} + \frac{1}{2}\delta_{k3}.$$

- We find (two ways): $R_k = \frac{1}{4}\delta_{k0} + \frac{3}{4}\delta_{k2}$.
- A giant component exists because: $\langle k \rangle_R = 0 \times 1/4 + 2 \times 3/4 = 3/2 > 1$.
- Generating functions for P_k and R_k:

$$F_P(x) = \frac{1}{2}x + \frac{1}{2}x^3 \text{ and } F_R(x) = \frac{1}{4}x^0 + \frac{3}{4}x^2$$

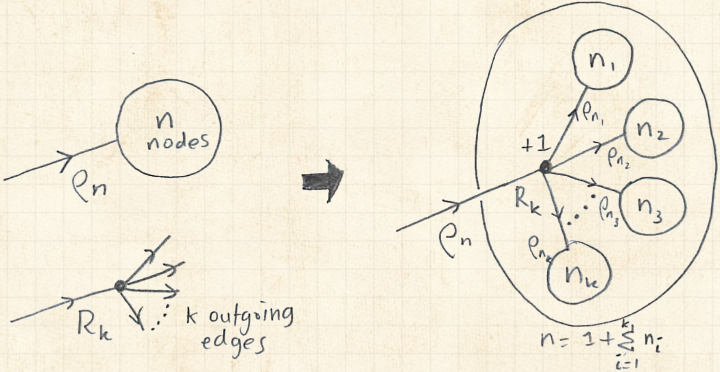
- Check for goodness:
  - $F_R(x) = F_P'(x)/F_P'(1)$ and $F_P(1) = F_R(1) = 1$.
  - $F_P'(1) = \langle k \rangle_P = 2$ and $F_R'(1) = \langle k \rangle_R = \frac{3}{2}$.
- Things to figure out: Component size generating functions for π_n and ρ_n, and the size of the giant component.

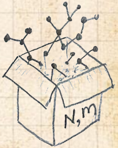


Find $F_\rho(x)$ first:

- We know:

$$F_\rho(x) = xF_R\left(F_\rho(x)\right).$$





- Sticking things in things, we have:

$$F_\rho(x) = x\left(\frac{1}{4} + \frac{3}{4}\left[F_\rho(x)\right]^2\right).$$

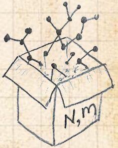
- Rearranging:

$$3x\left[F_\rho(x)\right]^2 - 4F_\rho(x) + x = 0.$$

- Please and thank you:

$$F_\rho(x) = \frac{2}{3x}\left(1 \pm \sqrt{1 - \frac{3}{4}x^2}\right)$$

- Time for a Taylor series expansion.
- The promise: non-negative powers of x with non-negative coefficients.
- First: which sign do we take?



- Because ρ_n is a probability distribution, we know $F_\rho(1) \leq 1$ and $F_\rho(x) \leq 1$ for $0 \leq x \leq 1$.
- Thinking about the limit $x \to 0$ in

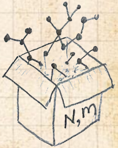
$$F_\rho(x) = \frac{2}{3x}\left(1 \pm \sqrt{1 - \frac{3}{4}x^2}\right),$$

we see that the positive sign solution blows to smithereens, and the negative one is okay.
- So we must have:

$$F_\rho(x) = \frac{2}{3x}\left(1 - \sqrt{1 - \frac{3}{4}x^2}\right),$$

- We can now deploy the Taylor expansion:

$$(1+z)^\theta = \binom{\theta}{0}z^0 + \binom{\theta}{1}z^1 + \binom{\theta}{2}z^2 + \binom{\theta}{3}z^3 + \ldots$$



- Let's define a binomial for arbitrary θ and $k = 0, 1, 2, \ldots$:

$$\binom{\theta}{k} = \frac{\Gamma(\theta+1)}{\Gamma(k+1)\Gamma(\theta-k+1)}$$

- For $\theta = \frac{1}{2}$, we have:

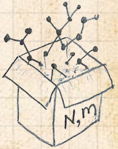
$$(1+z)^{\frac{1}{2}} = \binom{\frac{1}{2}}{0} z^0 + \binom{\frac{1}{2}}{1} z^1 + \binom{\frac{1}{2}}{2} z^2 + \ldots$$

$$= \frac{\Gamma(\frac{3}{2})}{\Gamma(1)\Gamma(\frac{3}{2})} z^0 + \frac{\Gamma(\frac{3}{2})}{\Gamma(2)\Gamma(\frac{1}{2})} z^1 + \frac{\Gamma(\frac{3}{2})}{\Gamma(3)\Gamma(-\frac{1}{2})} z^2 + \ldots$$

$$= 1 + \frac{1}{2} z - \frac{1}{8} z^2 + \frac{1}{16} z^3 - \ldots$$

where we've used $\Gamma(x+1) = x\Gamma(x)$ and noted that $\Gamma(\frac{1}{2}) = \frac{\sqrt{\pi}}{2}$.

- Note: $(1+z)^\theta \sim 1 + \theta z$ always.



- Totally psyched, we go back to here:

$$F_\rho(x) = \frac{2}{3x}\left(1 - \sqrt{1 - \frac{3}{4}x^2}\right).$$

- Setting $z = -\frac{3}{4}x^2$ and expanding, we have:

$$F_\rho(x) =$$

$$\frac{2}{3x}\left(1 - \left[1 + \frac{1}{2}\left(-\frac{3}{4}x^2\right)^1 - \frac{1}{8}\left(-\frac{3}{4}x^2\right)^2 + \frac{1}{16}\left(-\frac{3}{4}x^2\right)^3\right] + \ldots\right)$$

- Giving:

$$F_\rho(x) = \sum_{n=0}^{\infty} \rho_n x^n =$$

$$\frac{1}{4}x + \frac{3}{64}x^3 + \frac{9}{512}x^5 + \ldots + \frac{2}{3}\left(\frac{3}{4}\right)^k \frac{(-1)^{k+1}\Gamma(\frac{3}{2})}{\Gamma(k+1)\Gamma(\frac{3}{2}-k)} x^{2k-1} + \ldots$$

- Do odd powers make sense?

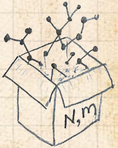
- We can now find $F_\pi(x)$ with:

$$F_\pi(x) = x F_P\left(F_\rho(x)\right)$$

$$= x\frac{1}{2}\left(\left(F_\rho(x)\right)^1 + \left(F_\rho(x)\right)^3\right)$$

$$= x\frac{1}{2}\left[\frac{2}{3x}\left(1-\sqrt{1-\frac{3}{4}x^2}\right) + \frac{2^3}{(3x)^3}\left(1-\sqrt{1-\frac{3}{4}x^2}\right)^3\right].$$

- Delicious.
- In principle, we can now extract all the π_n.
- But let's just find the size of the giant component.



- First, we need $F_\rho(1)$:

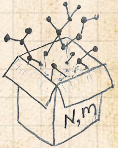
$$F_\rho(x)\big|_{x=1} = \frac{2}{3\cdot 1}\left(1-\sqrt{1-\frac{3}{4}1^2}\right) = \frac{1}{3}.$$

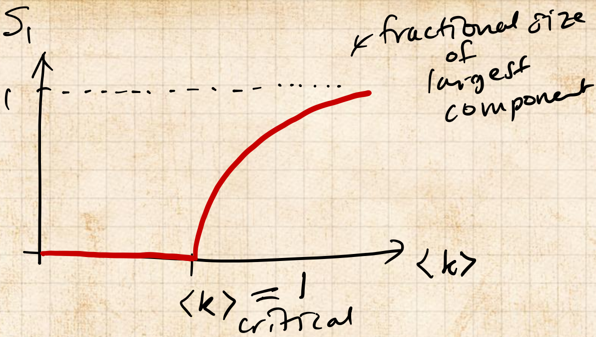
- This is the probability that a random edge leads to a sub-component of finite size.
- Next:

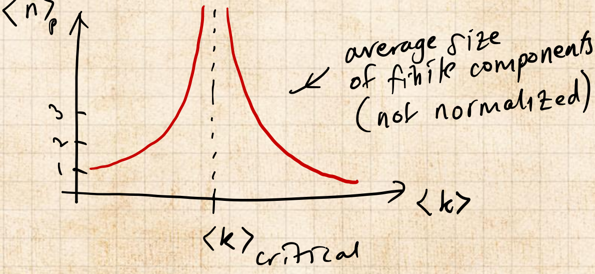
$$F_\pi(1) = 1\cdot F_P\left(F_\rho(1)\right) = F_P\left(\frac{1}{3}\right) = \frac{1}{2}\cdot\frac{1}{3}+\frac{1}{2}\left(\frac{1}{3}\right)^3 = \frac{5}{27}.$$

- This is the probability that a random chosen node belongs to a finite component.
- Finally, we have

$$S_1 = 1 - F_\pi(1) = 1 - \frac{5}{27} = \frac{22}{27}.$$




S_1
1
fractional size of largest component
$\langle k \rangle$
$\langle k \rangle_{\text{critical}} = 1$


$\langle n \rangle_p$
3
2
1
average size of finite components (not normalized)
$\langle k \rangle$
$\langle k \rangle_{\text{critical}}$

Average component size

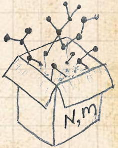
- Next: find average size of finite components $\langle n \rangle$.
- Using standard G.F. result: $\langle n \rangle = F'_{\pi}(1)$.
- Try to avoid finding $F_{\pi}(x)$ ...
- Starting from $F_{\pi}(x) = xF_P\left(F_{\rho}(x)\right)$, we differentiate:

$$F'_{\pi}(x) = F_P\left(F_{\rho}(x)\right) + xF'_{\rho}(x)F'_P\left(F_{\rho}(x)\right)$$

- While $F_{\rho}(x) = xF_R\left(F_{\rho}(x)\right)$ gives

$$F'_{\rho}(x) = F_R\left(F_{\rho}(x)\right) + xF'_{\rho}(x)F'_R\left(F_{\rho}(x)\right)$$

- Now set $x = 1$ in both equations.
- We solve the second equation for $F'_{\rho}(1)$ (we must already have $F_{\rho}(1)$).
- Plug $F'_{\rho}(1)$ and $F_{\rho}(1)$ into first equation to find $F'_{\pi}(1)$.



Average component size

Example: Standard random graphs.

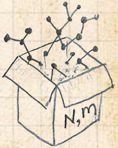
- Use fact that $F_P = F_R$ and $F_\pi = F_\rho$.
- Two differentiated equations reduce to only one:

$$F'_\pi(x) = F_P\left(F_\pi(x)\right) + xF'_\pi(x)F'_P\left(F_\pi(x)\right)$$

Rearrange: $$F'_\pi(x) = \frac{F_P\left(F_\pi(x)\right)}{1 - xF'_P\left(F_\pi(x)\right)}$$

- Simplify denominator using $F'_P(x) = \langle k \rangle F_P(x)$
- Replace $F_P(F_\pi(x))$ using $F_\pi(x) = xF_P(F_\pi(x))$.
- Set $x = 1$ and replace $F_\pi(1)$ with $1 - S_1$.

End result: $$\langle n \rangle = F'_\pi(1) = \frac{(1 - S_1)}{1 - \langle k \rangle (1 - S_1)}$$



Average component size

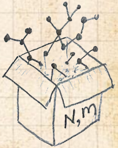
- Our result for standard random networks:

$$\langle n \rangle = F'_\pi(1) = \frac{(1 - S_1)}{1 - \langle k \rangle (1 - S_1)}$$

- Recall that $\langle k \rangle = 1$ is the critical value of average degree for standard random networks.
- Look at what happens when we increase $\langle k \rangle$ to 1 from below.
- We have $S_1 = 0$ for all $\langle k \rangle < 1$ so

$$\langle n \rangle = \frac{1}{1 - \langle k \rangle}$$

- This blows up as $\langle k \rangle \to 1$.
- Reason: we have a power law distribution of component sizes at $\langle k \rangle = 1$.
- Typical critical point behavior ...



Average component size

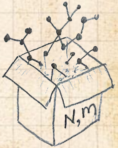
- Limits of $\langle k \rangle = 0$ and ∞ make sense for

$$\langle n \rangle = F'_{\pi}(1) = \frac{(1 - S_1)}{1 - \langle k \rangle (1 - S_1)}$$

- As $\langle k \rangle \to 0$, $S_1 = 0$, and $\langle n \rangle \to 1$.
- All nodes are isolated.
- As $\langle k \rangle \to \infty$, $S_1 \to 1$ and $\langle n \rangle \to 0$.
- No nodes are outside of the giant component.

Extra on largest component size:

- For $\langle k \rangle = 1$, $S_1 \sim N^{2/3}/N$.
- For $\langle k \rangle < 1$, $S_1 \sim (\log N)/N$.



- Let's return to our example: $P_k = \frac{1}{2}\delta_{k1} + \frac{1}{2}\delta_{k3}$.
- We're after:

$$\langle n \rangle = F'_\pi(1) = F_P\left(F_\rho(1)\right) + F'_\rho(1) F'_P\left(F_\rho(1)\right)$$

where we first need to compute

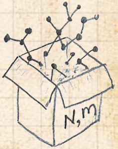
$$F'_\rho(1) = F_R\left(F_\rho(1)\right) + F'_\rho(1) F'_R\left(F_\rho(1)\right).$$

- Place stick between teeth, and recall that we have:

$$F_P(x) = \frac{1}{2}x + \frac{1}{2}x^3 \text{ and } F_R(x) = \frac{1}{4}x^0 + \frac{3}{4}x^2.$$

- Differentiation gives us:

$$F'_P(x) = \frac{1}{2} + \frac{3}{2}x^2 \text{ and } F'_R(x) = \frac{3}{2}x.$$



- We bite harder and use $F_\rho(1) = \frac{1}{3}$ to find:

$$F'_\rho(1) = F_R\left(F_\rho(1)\right) + F'_\rho(1) F'_R\left(F_\rho(1)\right)$$

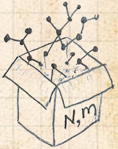
$$= F_R\left(\frac{1}{3}\right) + F'_\rho(1) F'_R\left(\frac{1}{3}\right)$$

$$= \frac{1}{4} + \frac{\cancel{3}}{4}\frac{1}{3^{\cancel{2}}} + F'_\rho(1)\frac{\cancel{3}}{2}\frac{1}{\cancel{3}}.$$

- After some reallocation of objects, we have $F'_\rho(1) = \frac{13}{2}$.
- Finally: $\langle n \rangle = F'_\pi(1) = F_P\left(\frac{1}{3}\right) + \frac{13}{2} F'_P\left(\frac{1}{3}\right)$

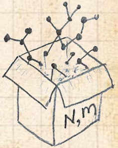
$$= \frac{1}{2}\frac{1}{3} + \frac{1}{2}\frac{1}{3^3} + \frac{13}{2}\left(\frac{1}{2} + \frac{\cancel{3}}{2}\frac{1}{3^{\cancel{2}}}\right) = \frac{5}{27} + \frac{13}{3} = \frac{122}{27}.$$

- So, kinda small.



Nutshell

- Generating functions allow us to strangely calculate features of random networks.
- They're a bit scary and magical.
- We'll find generating functions useful for contagion.
- But we'll also see that more direct, physics-bearing calculations are possible.



References I

[1] H. S. Wilf.
Generatingfunctionology.
A K Peters, Natick, MA, 3rd edition, 2006. pdf