P
o
C
S
What's
The
Story?

Principles of Complex Systems, Vols. 1 and 2
CSYS/MATH 6701, 6713
University of Vermont, Fall 2025
"Cat in the wall, eh?"
Assignment 08

It's Always Sunny in Philadelphia: Mac and Dennis Break Up, S5E08
Episode links: IMDB, Fandom, TV Tropes.

Due: ~~Friday, October 24~~ Sunday, October 26, by 11:59 pm
https://pdodds.w3.uvm.edu/teaching/courses/2025-2026pocsverse/assignments/08/
Some useful reminders:
Deliverator: Prof. Peter Sheridan Dodds (contact through Teams)
Office: The Ether and/or Innovation, fourth floor
Office hours: See Teams calendar
Course website: https://pdodds.w3.uvm.edu/teaching/courses/2025-2026pocsverse
Overleaf: LaTeX templates and settings for all assignments are available at
https://www.overleaf.com/read/tsxfwwmwdgxj.

Some guidelines:

1. Each student should submit their own assignment.
2. All parts are worth 3 points unless marked otherwise.
3. Please show all your work/workings/workingses clearly and list the names of others with whom you ~~conspired~~ collaborated.
4. We recommend that you write up your assignments in LaTeX (using the Overleaf template). However, if you are new to LaTeX or it is all proving too much, you may submit handwritten versions. Whatever you do, please only submit single PDFs.
5. For coding, we recommend you improve your skills with Python. And it's going to be a no for the catachrestic Excel. Please do not use any kind of AI thing unless directed. The (evil) Deliverator uses (evil) Matlab.
6. There is no need to include your code but you can if you are feeling especially proud.

Assignment submission:

Via **Brightspace** (which is not to be confused with the death vortex of the same name, just a weird coincidence). Again: One PDF document per assignment only.

Please submit your project's current draft in pdf format via Brightspace.

---

1. (3 + 3 + 3 points for each plot)

   Code up Simon's rich-gets-richer model.

   Plot Zipf distributions for $\rho =$ 0.10, 0.01, and 0.001. and perform regressions to test $\alpha = 1 - \rho$.

   Run the simulation for long enough to produce decent scaling laws (recall: three orders of magnitude is good).

   Averaging over simulations will produce cleaner results so try 10 and then, if possible, 100.

   Note the first mover advantage.

2. (3 + 3 + 3 points) For Herbert Simon's rich-get-richer model of what we've called Random Competitive Replication, we found in class that the normalized number of groups in the long time limit, n_k, satisfies the following difference equation:

$$\frac{n_k}{n_{k-1}} = \frac{(k-1)(1-\rho)}{1+(1-\rho)k} \qquad (1)$$

   where $k \geq 2$. The model parameter ρ is the probability that a newly arriving node forms a group of its own (or is a novel word, starts a new city, has a unique flavor, etc.). For $k = 1$, we have instead

$$n_1 = \rho - (1-\rho)n_1 \qquad (2)$$

   which directly gives us n_1 in terms of ρ.

   (a) Derive the exact solution for n_k in terms of Gamma functions and ultimately the Beta function.

   (b) From this exact form, determine the large k behavior for n_k ($\sim k^{-\gamma}$) and identify the exponent γ in terms of ρ. You are welcome to use the fact that $B(x, y) \sim x^{-y}$ for large x and fixed y (or use Stirling's approximation directly on the Gamma functions that will appear).

   Note: Simon's own calculation is slightly awry.

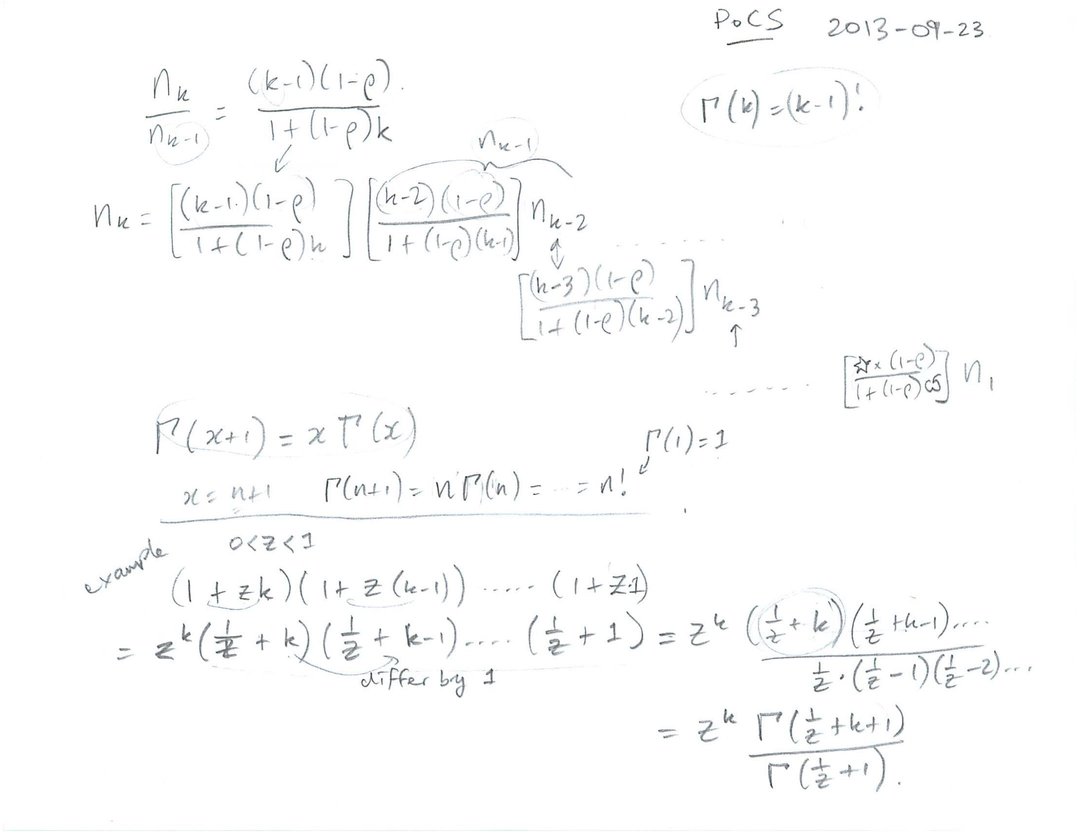
   The end result will be mathematically okay.

   And rich-get-richer mechanisms really are everywhere.

But Simon's simple model is wrong because of the first mover advantage.

Hint—Setting up Simon's model:
http://www.youtube.com/watch?v=OTzI5J5W1K0

The hint's output including the bits not in the video:



PoCS 2013-09-23

$$\Gamma(k) = (k-1)!$$

$$\frac{n_k}{n_{k-1}} = \frac{(k-1)(1-\rho)}{1+(1-\rho)k}$$

$$n_k = \left[\frac{(k-1)(1-\rho)}{1+(1-\rho)k}\right]\overbrace{\left[\frac{(k-2)(1-\rho)}{1+(1-\rho)(k-1)}\right] n_{k-2}}^{n_{k-1}} \cdots$$

$$\left[\frac{(k-3)(1-\rho)}{1+(1-\rho)(k-2)}\right] n_{k-3}$$

$$\cdots\cdots \left[\frac{☆ \times (1-\rho)}{1+(1-\rho)☺}\right] n_1$$

$$\Gamma(x+1) = x\Gamma(x)$$

$$x = n+1 \qquad \Gamma(n+1) = n\Gamma(n) = \cdots = n! \qquad \Gamma(1) = 1$$

example $0 < z < 1$

$$(1+zk)(1+z(k-1))\cdots\cdots(1+z1)$$

$$= z^k\left(\frac{1}{z}+k\right)\left(\frac{1}{z}+k-1\right)\cdots\cdots\left(\frac{1}{z}+1\right) = z^k\frac{\left(\frac{1}{z}+k\right)\left(\frac{1}{z}+k-1\right)\cdots}{\frac{1}{z}\cdot\left(\frac{1}{z}-1\right)\left(\frac{1}{z}-2\right)\cdots}$$

differ by 1

$$= z^k\frac{\Gamma\left(\frac{1}{z}+k+1\right)}{\Gamma\left(\frac{1}{z}+1\right)}.$$

3. (6 + 3 + 3 points)

   In Simon's original model, the expected total number of distinct groups at time t is ρt. Recall that each group is made up of elements of a particular flavor.

   In class, we derived the fraction of groups containing only 1 element, finding

$$n_1^{(g)} = \frac{N_1(t)}{\rho t} = \frac{1}{2-\rho}.$$

   Notation note: We're using a version of Simon's notation. It's not great. Given the connection to the probability distribution view, $P_k \sim k^{-\gamma}$, where $\gamma = 1 + 1/\alpha$, we would now write P_k instead of $n_k^{(g)}$.

   (a) (3 + 3 points)

   Find the form of $n_2^{(g)}$ and $n_3^{(g)}$, the fraction of groups that are of size 2 and size 3. in terms of the innovation rate ρ.

(b) (3 points)

Using data for James Joyce's Ulysses (see below), first show that Simon's estimate for the innovation rate $\rho_{\mathrm{est}} \simeq 0.115$ is reasonably accurate for the version of the text's word counts given below.

We're including all punctuation and we're lowercasing all words.

Your results won't explicitly match Simon's because the parsing is different.

Hint: Do not compute ρ_{est} from an estimate of γ. Instead, $\rho_{\mathrm{est}} =$ number of distinct types divided by the number of total tokens.

(c) (3 points)

Now compare the theoretical estimates for $n_1^{(g)}$, $n_2^{(g)}$, and $n_3^{(g)}$, with empirical values you obtain for Ulysses.

Data:

- Tab-separated text file (first column = 1-gram, second column = 1-gram count k):
https://pdodds.w3.uvm.edu/teaching/courses/2025-2026pocsverse/data/ulysses_1grams_lowercase.tsv