3.5: Systems: Power laws; 80/20 Principle

“The 80/20 Principle asserts that a minority of causes, inputs, or effort usually lead to a majority of the results, outputs, or rewards.”

—Richard Koch

Main point: General rules/consistencies exist that are common to many systems …

1: Power Laws (80/20 Principle)
- most websites have a few links; a few have a whole lot
- general (counterintuitive) property of networks: vastly different linkages
  o would expect each node to have roughly the same (or gaussian)
  o but just a few have most of the connections – hubs
    ▪ most have few; a few have most
    ▪ highly skewed — far from normal (80/20 distribution)
- also in metabolic networks: just a few molecules participate in most reactions
- also in scientific influence – number of references to a paper
  o original reason for Google algorithm (cf. number of connections to a website)
- original formulation by Pareto: a small percentage of people have most of the wealth
  o order individuals in terms of wealth; top 20% (or so) have 80% of total wealth
    ▪ not always exactly 80/20; could be 80/10, or 90/20
- also true of many other things in life
  o peas in Pareto’s garden; products of companies; hubs; actors; your friends, etc
  o general law of effect in systems
    ▪ general description: power law. Distribution: \( f(a) = a^{-\gamma} \) for some \( \gamma \)
- most of what you do has little return, but a small fraction has enormous impact
  o decisions you make, things you do
  o papers you write (citations – Redner; also, check google scholar)
  o find this part, and focus on it
Why do power-law distributions exist? [“gist”-Barabasi]

- 1: every entity (person, item) the same -> flat (no linkage)
- 2: every cause the same -> gaussian curve (linkage to interior)
- 3: every opportunity the same -> power law (linkage to exterior)
- individual connections to exterior – to *environment*
  o *preferential attachment*: change in \( a = (ka + b) \)
  o exponential curve \( a = 10^\alpha \)
    - exponent \( \alpha \) reflects degree of internal interaction
  o exponent (=linkage) \( \alpha \) varies, with *gaussian distribution*
  o effect on distribution = power law = \( a^\gamma \)
    - exponent \( \gamma \) reflects degree of external interaction

- **Note:** power-law distributions are *scale-free*
  o distribution above is similar to distribution below
  o can’t tell from the distribution what scale you’re at (no characteristic scale)
    - no average, no standard deviation
      - e.g. earthquake size, frequency
Real-World Segment: Interaction with Industry

- Commonly believed that there are two kinds of science:
  - basic science: unimportant in the “real” world
  - applied science: unimportant in the academic world

- the two are connected:
  1. *basic science often gives rise to applications*
     - e.g. computer industry
  2. can also go the other way: *applications can give rise to basic science*
     - e.g., discovery of bacteria by Pasteur motivated by problems in wine industry
     - more generally: *tests done by engineers are a kind of experiment*; can provide interesting data

- standard way of viewing science:
  - basic vs. applied: points along a *single dimension*

- may be better to view this as *two dimensions*:
  - basic: degree to which work affects *theories*
  - applied: degree to which work affects *applications*
  - these dimensions don’t interact
    - can have various combinations—various types of problems

- area of common interest to industry and academia
  - Pasteur’s quadrant