Toy models

David Aldous

October 6, 2014
Some probability models of real-world phenomena are “quantitative” in the sense that we believe the numerical values output by the model will be approximately correct. At the other extreme, a toy model is a consciously over-simplified model of some real-world phenomenon that typically attempts to study the effect of only one or two of the factors involved while ignoring many complicating real-world factors. It is thus “qualitative” in the sense that we do not believe that numerical outputs will be accurate.

An example is the Galton-Watson branching process model I will describe. This is a textbook topic in STAT 150 (a first course in stochastic processes), but I want to emphasize the “just supercritical” formula (2).
This **Galton-Watson branching process** model is used as a toy model in many different settings. To have a concrete language, in this section we talk about “individuals” and “offspring”. To visualize individuals and offspring, you can either imagine asexual reproduction or look only at males or only at females in a two-sex species like humans.

The model is that there is a probability distribution \( \mathbf{p} := (p_i; i = 0, 1, 2, \ldots) \) and that each individual in a generation has a random number of offspring in the next generation, this number being picked from \( \mathbf{p} \) independently for different parents.

By default we assume the process starts with 1 individual in generation 0; so there is some random number \( Z_n \geq 0 \) of individuals in each generation \( n = 0, 1, 2, 3, \ldots \). There are two logical possibilities for what might happen in the long run:

- either “extinction” meaning \( Z_n = 0 \) for all large \( n \)
- or “survival”, meaning \( Z_n \geq 1 \) for all \( n \).
One of the highlights of an undergraduate course in stochastic processes is the following theorem.

Write $\mu$ and $\sigma$ for the mean and s.d. of the number of offspring.

**Theorem.** (a) If $\mu < 1$ then $P(\text{extinction}) = 1$.
(b) If $\mu > 1$ then $\rho = P(\text{extinction}) < 1$ and is the solution of the equation

$$\rho = \Phi(\rho)$$

(1)

where $\Phi$ is the probability generating function defined by

$$\Phi(z) = \sum_{i=0}^{\infty} p_i z^i.$$

Keep in mind that the “independence” assumptions are tantamount to assuming there is no “interaction” between individuals and that there are no external constraints on population size – both assumptions are unrealistic in almost all imaginable real-world contexts.
I won’t repeat the textbook derivation of the Theorem, but I will derive an interesting approximate formula for a particular setting. The cases \( \mu < 1, \mu = 1, \mu > 1 \) are called \textit{subcritical}, \textit{critical}, \textit{supercritical}. I want to consider the “just supercritical” case where \( \mu > 1 \) but \( \mu - 1 \) is small.

For a just supercritical Galton-Watson process, \( \mathbb{P}(\text{survival}) \approx \frac{2(\mu-1)}{\sigma^2} \). 

(2)
This is often not mentioned in textbooks, so let me give

**Derivation of formula (2).** Textbook facts about the probability generating function for the random number $X$ of offspring are

$$
\Phi(1) = 1, \quad \Phi'(1) = \mu, \quad \Phi''(1) = \mathbb{E}[X(X - 1)] = \sigma^2 + \mu^2 - \mu \approx \sigma^2
$$

the approximation holding because $\mu \approx 1$.

We want the survival probability $\bar{\rho} = 1 - \rho$. The equation in the
Theorem, $\rho = \Phi(\rho)$, can be rewritten in terms of $\bar{\rho}$ as $h(1 - \bar{\rho}) = 0$,
where $h(x) = \Phi(x) - x$. Consider the series expansion: for small $x$,

$$
h(1 - x) \approx h(1) - xh'(1) + \frac{1}{2}x^2 h''(1).
$$

Since $h(1) = 0$, $h'(1) = \mu - 1$, $h''(1) \approx \sigma^2$ the rewritten equation becomes

$$
0 \approx -\bar{\rho}(\mu - 1) + \frac{1}{2}\bar{\rho}^2 \sigma^2
$$

and solving for $\bar{\rho}$ gives the stated formula (2).
Keep in mind that the “independence” assumptions are tantamount to assuming there is no “interaction” between individuals and that there are no external constraints on population size – both assumptions are unrealistic in almost all imaginable real-world contexts. This is why I call it a “toy model”.

Let’s think of a toy model for the spread of epidemics such as influenza. Each infected person will infect some random number of other people; the mean such number is called the reproduction number $\mu$. We can use the previous Galton-Watson process to model the number of cases in the initial phase; if $\mu < 1$ the epidemic will not occur; if $\mu > 1$ and there are (at least) several initial cases then there will be an epidemic. Once the epidemic grows it is natural to work with

$$g(t) = \text{proportion of population infected}.$$ 

If we ignore the fact that people recover, then the rate of growth of the epidemic is roughly proportion to the number of infected-infected contacts, and this is most simply modeled by the
**logistic equation**

\[ g'(t) = cg(t)(1 - g(t)) \]

The solution, up to an arbitrary time-shift, is

\[ g(t) = \frac{1}{1 + e^{-ct}} \]
**Reality check.** In that toy model, 100% of population is eventually infected. In fact, in the annual “seasonal influenza” epidemic in the U.S., typically the percentage of population infected is in the range 5% - 20%.

So what’s wrong with the toy model? Many things, in particular

- people recover (infective for about 7 days)

and the population is not homogeneous:

- different levels of partial immunity
- different people have different numbers of inter-personal contacts
- spatial locations matter.

These all affect the eventual proportion of population infected, but the S-shaped curve remains typical.

[show flu-rates]