10 October 2014

INTRODUCTION TO MATHEMATICAL MODELING. A PROBABILISTIC POPULATION MODEL

David A. Meyer

Department of Mathematics
University of California/San Diego
La Jolla, CA 92093-0112
http://math.ucsd.edu/~dmeyer/; dmeyer@math.ucsd.edu

A microscale probabilistic model

Suppose we want to model population changes by considering individuals, which for the purposes of simplicity let's take to be *E. coli* bacteria (single-celled). Roughly speaking, the *E. coli* in your gut reproduce with a timescale of 1 day, so in our discrete model, the length of a time step can be thought of as 1 day. There are many factors which go into a cell fissioning; we summarize them by making the process random: we assume each cell splits with probability b during each time step. Furthermore, since individual bacteria certainly die, we call the probability of that happening in each time step, d.

Now let N_t = the number of bacteria at time t; it is a *random variable*, with some probability distribution. Before seeing what we can compute about N_t analytically, we can simply simulate this process. The figures below show the results of doing so a couple of times, with $b = 0.3$ and $d = 0.2$. Notice that although the population is growing, on the average, in both runs, the time series of values, and the final populations, are completely different.

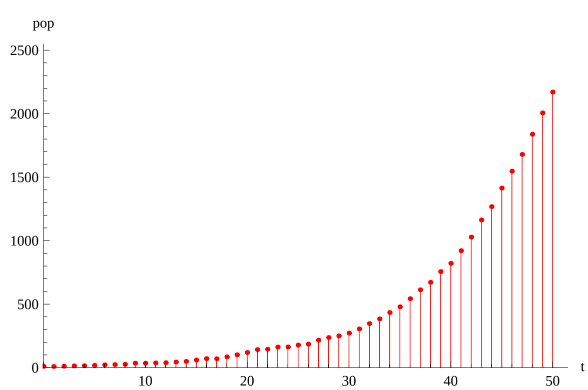

Figure 1a. Time series of values of N_t, starting at $N_0 = 10$.

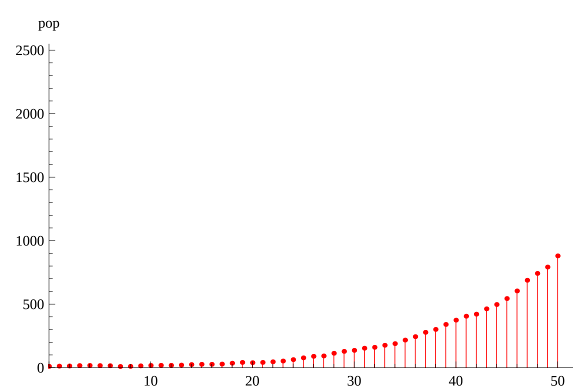

Figure 1b. A very different time series of values of N_t.

Analyzing the model

Suppose there are N_{t-1} cells at the beginning of time step t. Then we can write:

$$N_t = X_1 + \cdots + X_{N_{t-1}} + Y_1 + \cdots + Y_{N_{t-1}}, \tag{1}$$

where $X_i = 1$ if cell i does not die in the t^{th} time step, which happens with probability $1 - d$, and is 0 otherwise; and $Y_i = 1$ if cell i splits in the t^{th} time step, which happens with probability b, and is 0 otherwise. From (1) we can compute the expectation value of N_t, conditional on N_{t-1}, to be:

$$\begin{aligned} \mathsf{E}[N_t \mid N_{t-1}] &= \sum_{i=1}^{N_{t-1}} \mathsf{E}[X_i] + \sum_{i=1}^{N_{t-1}} \mathsf{E}[Y_i] \\ &= (1-d)N_{t-1} + bN_{t-1} =: rN_{t-1}, \end{aligned} \tag{2}$$

where we've used the linearity of expectation value in the first line, and $\mathsf{E}[X_i] = 1 \cdot (1 - d) + 0 \cdot d = 1 - d$ (and similarly for $\mathsf{E}[Y_i] = b$) in the second. Notice that $r = 1 - d + b$ is greater than 1 when $b > d$ and less than 1 when $b < d$, so the expected value of N_t is bigger or smaller than N_{t-1}, respectively.

Now recall that the "law of total probability" tells us that

$$\Pr(N = n) = \sum_{n'} \Pr(N = n \mid N' = n') \Pr(N' = n'),$$

so we can compute the *unconditional* expectation value of N_t as:

$$\begin{aligned} \mathsf{E}[N_t] &= \sum_n n \Pr(N_t = n) \\ &= \sum_n n \sum_{n'} \Pr(N_t = n \mid N_{t-1} = n') \Pr(N_{t-1} = n') \\ &= \sum_{n'} \Big(\sum_n n \Pr(N_t = n \mid N_{t-1} = n') \Big) \Pr(N_{t-1} = n') \\ &= \sum_{n'} \mathsf{E}[N_t \mid N_{t-1} = n'] \Pr(N_{t-1} = n') \\ &= \sum_{n'} rn' \Pr(N_{t-1} = n') = r\mathsf{E}[N_{t-1}], \end{aligned}$$

which implies

$$\mathsf{E}[N_t] = r^t N_0. \tag{3}$$

Figure 2 shows these expected values plotted on the same graphs as the simulation runs shown in Figure 1. In the first run the simulated values are much larger, in the second,

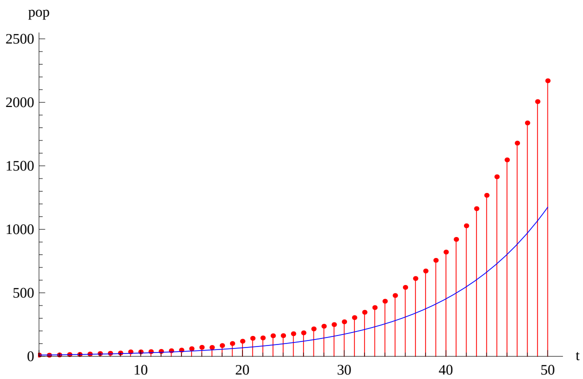

Figure 2a. Time series of values of N_t, starting at $N_0 = 10$, with the exponential growth of the expectation value.

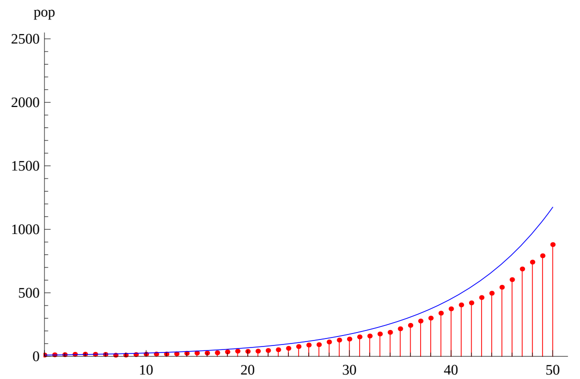

Figure 2b. A very different time series of values of N_t, with the exponential growth of the expectation value.

somewhat smaller; unsurprisingly the population does not grow exactly the same way as the expected value in every run.

Notice that except in the uninteresting case $d = 1$ and $b = 0$, we can write $0 < r = e^k$ for $k \in \mathbb{R}$ in (3), and get $\mathsf{E}[N_t] = e^{kt}N_0$, which is the same as the solution to the ODE:

$$\frac{\mathrm{d}N}{\mathrm{d}t} = kN. \tag{4}$$

Thus we should understand the macroscale ODE model (4) as describing the *expectation value* of the microscale probabilistic model. The former gives us no information, for example, about the random fluctuations away from the expectation value that we see in the simulation results shown in the figures.

