13 October 2014

INTRODUCTION TO MATHEMATICAL MODELING. A PROBABILISTIC POPULATION MODEL. II

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

Analyzing the model: population variance

We can run many simulations to understand the range of fluctuations around the expected population. Figure 3 shows a histogram of the resulting populations at each time, *i.e.*, each constant time slice is a sample of size 1000 from the probability distribution of possible populations at that time. Figure 4 shows this distribution for $t = 50$. The median value is 1116, which compares well with the expected value of approximately 1174. Also, the central 95% of these runs lie between 133 and 2786. Our next goal is to try to understand this range analytically, which we can do in this case, although in more complicated models we might have to rely on a large number of simulation runs.

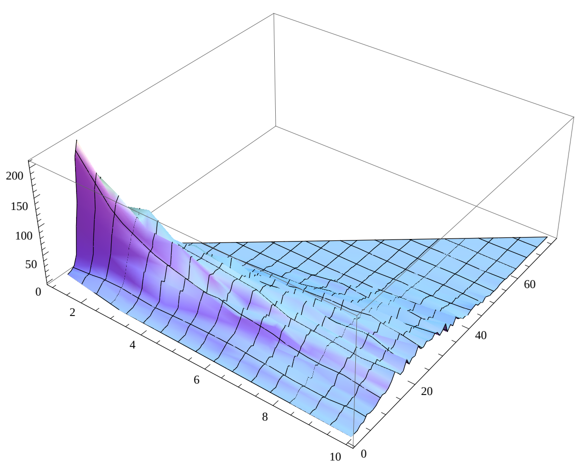

Figure 3. Histogram of population for 1000 runs, with $b = 0.3$, $d = 0.2$.

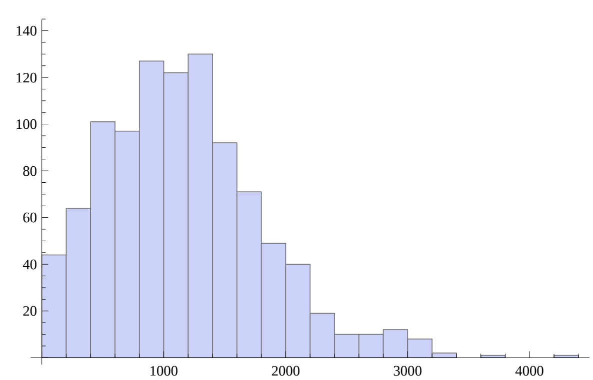

Figure 4. The histogram of 1000 population values at time $t = 50$, also with $b = 0.3$, $d = 0.2$.

To do this we calculate the variance of N_t:

$$\mathsf{Var}[N_t] = \mathsf{E}[(N_t - \mathsf{E}[N_t])^2] = \mathsf{E}[N_t^2] - \big(\mathsf{E}[N_t]\big)^2. \tag{5}$$

Equation (3) gives us the last term in this expression. For the first, we begin by calculating

$$\begin{aligned}
\mathsf{E}[N_t \mid N_{t-1}] &= \mathsf{E}[(X_1 + \cdots + X_{N_{t-1}} + Y_1 + \cdots + Y_{N_{t-1}})^2] \\
&= N_{t-1}\big(\mathsf{E}[X_1^2] + \mathsf{E}[Y_1^2]\big) + 2N_{t-1}^2\mathsf{E}[X_1Y_2] \\
&\qquad + N_{t-1}(N_{t-1}-1)\big(\mathsf{E}[X_1X_2] + \mathsf{E}[Y_1Y_2]\big) \\
&= N_{t-1}\big(1-d+b\big) + 2N_{t-1}^2(1-d)b + N_{t-1}(N_{t-1}-1)\big((1-d)^2+b^2\big) \\
&= N_{t-1}^2(1-d+b)^2 + N_{t-1}\big(1-d+b-(1-d)^2-b^2\big) \\
&=: \alpha N_{t-1}^2 + \beta N_{t-1},
\end{aligned}$$

where $\alpha = r^2$ as defined in (2). Now we use the "law of total probability" again:

$$\begin{aligned}
\mathsf{E}[N_t^2] &= \sum_n n^2 \Pr(N_t = n) \\
&= \sum_n n^2 \sum_{n'} \Pr(N_t = n \mid N_{t-1} = n') \Pr(N_{t-1} = n') \\
&= \sum_{n'} \Big(\sum_n n^2 \Pr(N_t = n \mid N_{t-1} = n')\Big) \Pr(N_{t-1} = n') \\
&= \sum_{n'} (\alpha n'^2 + \beta n') \Pr(N_{t-1} = n') \\
&= \alpha\mathsf{E}[N_{t-1}^2] + \beta\mathsf{E}[N_{t-1}] \\
&= \alpha\mathsf{E}[N_{t-1}^2] + \beta r^{t-1}N_0 \\
&= \alpha\big(\alpha\mathsf{E}[N_{t-2}^2] + \beta r^{t-2}N_0\big) + \beta r^{t-1}N_0 \\
&= \alpha^2\mathsf{E}[N_{t-1}^2] + \big(\frac{\alpha}{r} + 1\big)\beta r^{t-1}N_0 \\
&\ \vdots \\
&= \alpha^t N_0^2 + \frac{1-(\alpha/r)^t}{1-(\alpha/r)}\beta r^{t-1}N_0 \\
&= r^{2t}N_0^2 + \frac{1-r^t}{1-r}\beta r^{t-1}N_0.
\end{aligned}$$

Using this result in (5) we obtain

$$\mathsf{Var}[N_t] = \frac{1-r^t}{1-r}\beta r^{t-1}N_0.$$

Evaluating this for same case as the simulations we find a standard deviation (the square root of the variance) of about 955 at time $t = 50$, which compares well with the histogram in Figure 4. (The 95% range would be about ± 2 standard deviations around the expectation value *if this were a normal distribution*, which it clearly is not, since subtracting 2 standard deviations from the expectation value gives a negative result, which does not occur.)