Example 1: Single stochastic growth model

We begin with EXAMPLE1. This is a simple simulation of a birth process. Each individual has a birth rate (births/unit time), r. In a short time, dt, we would expect $r \times dt$ new births for each individual. Thus, in a population of size N there will be $rNdt$ births. So the ODE equation for change in population size is

$$\frac{dN}{dt} = rN$$

where r is the growth rate per individual per unit time. The solution of this equation is exponential growth i.e.

$$N(t) = N(0)e^{rt},$$

where $N(0)$ is the initial population at time $t = 0$.

In this model, how long does it take for the population to double?

How long does it take for the population to reach four times its original size?

For a stochastic simulation, the probability that an individual gives birth in a short time, dt, is given by $r\,dt$. The same probability applies to each of the N individuals. Hence the number of new births will be binomially distributed:

$$\text{births in time dt} \sim Bin(rdt, N\,)$$

Exercise
Check the Berkeley Madonna code in EXAMPLE 1 to see how this is implemented. Note that as the ODE equation deals with mean values, N doesn't have to be a whole number. For the stochastic simulation, N is always a whole number of individuals.

1. Open the Berkeley Madonna model EXAMPLE1.

2. Click *<RUN>* and you will see the model solution. The exact solution lies on the same line as N_ODE, which is a numerical solution of the model and the stochastic solution is shown in green. Click *<RUN>* multiple times.

What do you notice?

What is the difference between the solutions' behaviour at early, low populations and at later times with larger populations? Explain this behaviour.

In the long term, are the stochastic runs getting closer to the exact solution or further away, on the log scale?

What about on the linear scale?

4. Try some different values of r using the slider.

How does the value of r affect your comparisons between the exact and stochastic simulations?

5. Madonna allows the user to perform multiple simulations using Batch Runs. Try this:
   a. Select *<Parameters><Batch Runs>* then *<OK>*.
   b. Set the number of runs to 10 and choose the "Keep Runs Separate" option.
   c. Press *<OK>*. Now repeat this but enter "100" in the # of Runs box.

6. Finally repeat once more but select the "Compute Mean" option. Madonna will now perform 100 simulations and calculate the mean at each time point. This is plotted instead of the individual simulations.

How does the mean compare with the exact solution?

For linear differential equation models, the exact solution is the mean of the stochastic simulations. The more stochastic simulations are run, the closer the mean gets to the exact solution. However, this conclusion breaks down when there is the possibility of extinction, as examined in Example 3, and also for non-linear models.

Example 2: Stochastic growth model

In this example we extend the code developed in Example 1 to perform multiple runs simultaneously using arrays in Berkeley Madonna. This also allows us look at estimates of the mean over many runs and how that depends on the number of runs.

Exercise

1. Open the Madonna model EXAMPLE2.

2. Look at the equations for the stochastic simulations. Notice that there is now a subscript [i] to denote an individual simulation and this ranges from 1 to nsim (written as e.g. N[1..nsim]).

3. Click <RUN> and you will see the model solutions as before. Increase the number of simulations slider, nsim, to perform 10, 100 and 1000 simulations. On Page 3 you will see the mean of these simulations.

How do they compare with the exact solution? How does the number of simulations affect agreement?

Example 4: Stochastic SIS model

The models discussed so far have looked at population size: growth, extinction, or simply "wandering about". In the next example we consider a model of infection transmission within a closed popluation (no births or deaths): the susceptible-infected-susceptible (SIS) model.

In this model we consider two types of event:
*   Infection events, which occur in each time step with probability $\beta I \, dt / N$ per susceptible individual, and
*   Recovery events, which occur in each time step with a probability $\nu \, dt$ per infected individual.

The ODE representation of the model is:

$$
\frac{dS}{dt} = - \beta S \frac{I}{N} + \nu I
$$
$$
\frac{dI}{dt} = \beta S \frac{I}{N} - \nu I
$$

The analytical solution to this ODE can be shown to be:

\begin{equation}
I(t) = I* \frac{1}{ 1 + (I*/I_0 - 1) exp(-(\beta - \nu) t)}
\end{equation}

where I_0 is the initial number of infected individuals and

\begin{equation}
I*=N(1-\nu/\beta) \mbox{.}
\label{eqn:sis_eqm}
\end{equation}

After sufficient time, the infection and recovery processes balance each other out. This occurs when the number of infectious individuals reaches its equilibrium value I*. We can write the equilibrium value as

\begin{equation}
I* = N (1-1/R_0)
\end{equation}

since R_0 is defined as β / ν. At this point, we say that the disease is \emph{endemic} in the population.

The stochastic simulation of this system is constructed in a similar way to the previous examples (see the code). Its behaviour, differs significantly from the deterministic simulation for some parameter values. Weíll investigate these in the rest of this section.

Exercise

1. Open the model EXAMPLE4.

2. First, we look at the endemic situation where there is a continuous population of infected individuals (check in the code that the correct initial condition is in place).

   Using Equation \ref{eqn:sis_eqm} How do you expect the contact rate (β), recovery rate (ν) and total population (N) to affect the equilibrium number of infected individuals?

########################### empty table here#######################

   Now try varying these parameters in the model. Does the model behave as you expect?

3. Now look at the accuracy of the deterministic solution. With default parameter values, vary N from around 400 down to about 50. How does deterministic solution compare to the mean value of I from the stochastic simulation?

   For large N,

   For small N,

   The differences between the simulations are due to the non-linear infection term, $ \beta SI/N$. The number of susceptibles (S) and infected (I) are negatively correlated, since when an individual leaves one class, he enters the other. This correlation term is absent from the ODE model which then overestimates the force of infection, giving a higher population of infecteds.

4. Fade-out can occur when the endemic infected population is small.

   Set $\beta =0.5$, $\nu=0.3$, and gradually reduce N from the default value of 100.

   From what value do fadeouts start happening?

   Can you explain whatís happening to the difference between mean stochastic behaviour and the deterministic?

   Do a parameter plot of the proportion of simulations that have ``faded out'' by the end of the simulation, against N, from 150 to 10. What do you notice? Estimate a critical community size.

5. Finally, we look at a growing epidemic starting from a single individual. In the code, comment out the initial condition for endemic behaviour and comment in the epidemic one. Set N high (300+) and $nsim$ to 1000. What proportion of fade-outs are you getting?

   You should get the same fade-out rate as in example 3. Can you explain why?