1. `gillespie.m` & `plotresult.m` user guide

`gillespie.m` simulates a birth-death process for a population using the Gillespie algorithm. It uses `plotresult.m` to plot the changes in this population with time.
`gillespie.m` plots a guideline for the expected changes in the population size as if the system were deterministic (by plotting an exponential function dependent on the rates of the Poisson processes) rather than random.

Getting started and inputs

- Download `gillespie.m` and `plotresult.m` to the same directory. Open MATLAB and navigate to the directory containing the two files.
- Open both files within MATLAB. To run the programs, press *Run* (under the *EDITOR* tab) or type `gillespie` into the command window and press Enter.
  - `plotresult.m` is executed by `gillespie.m`, with the inputs it requires.
- `gillespie.m` requires no inputs to run, but has its parameters set in the first block of code in the script (lines 1-7). See 'key parameters'.
- `plotresult.m` takes the data generated by `gillespie.m` for plotting. With the command `plotresult(t,k,y)`, `plotresult.m` is given the timestamps (x-axis), replicate number (to differentiate between the trials), and population size data (y-axis). This command can be run in the command window, as long as `t` and `y` are one-dimensional arrays of equal sizes, and `k` is an integer. The dynamics of multiple simulations can be displayed on the same graph if the figure is kept open.

Key parameters and outputs

`init_pop`: initial population size; `replicates`: number of trials of the gillespie algorithm, to obtain a number of possible dynamics; `birth_rate`: average rate at which new members of the population are added per unit time; `death_rate`: average rate at which members of the population are removed per unit time. Different rates can be trialled by editing the parameters with your own rates.
- `t`, `k`, and `y` are the main outputs of gillespie.m
- `plotresult.m` outputs two graphs, plotting the population size against time, and the natural logarithm of the population size against time. The coloured lines show the system as simulated by the gillespie algorithm, while the black dotted line shows the dynamics according to the deterministic model.

`birth_rate > death_rate`: exponential growth-like dynamics will be most likely to be observed (as indicated by the dotted line indicating the deterministic trend).
`death_rate > birth_rate`: population is more likely to reduce to zero earlier.
While the example rates on lines 5-7 produce the same deterministic dynamics as `birth_rate - death_rate` remains constant, a death_rate equal to zero never results in population extinction, while non-zero `death_rate` sometimes does.
When the absolute values of each rate are larger, the system shows larger fluctuations, and it is more likely for the population to go extinct.

2.

a) $\frac{dN}{dt} = rN$ is the ODE which is the deterministic counterpart to the simple birth-death process in `gillespie.m`. $r = r_{birth} - r_{death}$ and the ODE has the solution $N = N_0 e^{rt}$.

b) See gillespie2.m.

c) t_{max} is set to different values for comparing the mean final states with the value $N = N_0 e^{t_{max} r}$. N_0 is set to 1 for all tests. The deterministic variance is always zero, by nature, whereas the variance of the final states of the Gillespie algorithm is inherently quite large the more progressed the system is in time.

t_{max}	Number of trials	Birth rate	Death rate	Mean final state	Variance of final states	Deterministic mean
10	5,000	1.5	0.7	2.9630×10^3	2.3457×10^7	$e^{10(1.5-0.7)} = 2980.9580$
10	5,000	1.0	0.2	2.9902×10^3	1.3918×10^7	2980.9580
10	5,000	1.0	1.0	0.9936	20.2832	$e^0 = 1$
10	5,000	1.1	1.0	2.5648	100.1866	1
5	100,000	1.5	0.7	54.0900	7.7846×10^3	$e^{5(1.5-0.7)} = 54.5982$
2	100,000	1.5	0.7	5.1340	55.8857	$e^{2(1.5-0.7)} = 4.9530$
1	100,000	1.5	0.7	2.4785	8.6780	$e^{(1.5-0.7)} = 2.2255$
0.5	100,000	1.5	0.7	1.7694	2.9457	$e^{(1.5-0.7)/2} = 1.4918$

I increased the number of trials for smaller t_{max} as it became computationally feasible to.

In general, it can be seen that the Gillespie algorithm results in the same mean final state as the deterministic system, with a few exceptions. Notably, what would be a stable system at its equilibrium state in the deterministic system (r=0, N=1), the mean final state is doubled in the randomly varying system. This will be because some of the independent realisations of this system will be able to grow well beyond the mean, by chance, while others are limited in decreasing down to zero, and no further (since negative populations cannot exist). Fluctuations have a lower limit of extinction, but no upper limit, causing an inflated mean. This reasoning is why the mean of the final states is larger than is deterministically expected when the random systems are observed early on in their existence.

The variance of final states increases when the values of birth rate and death rate are greater (as commented on in q1), when the difference between them is greater, and for

larger values of t_{max}. It makes sense that the variance is large for large t_{max} since the final states range from zero to large fluctuations around the expected exponential growth.

d) Setting `birth_rate` and `death_rate` to 1.5 and 0.7, as originally set in q1, results in the **probability of 45.76%** of a population going extinct by t=10 (out of 5,000 trials). There is still a chance of the non-zero populations at t=10 to go extinct, as time passes further. However, as can be seen when running an adapted version of **`gillespie.m`** from q1 (where the axes extend further and simulations run until y>1000), t=10 appears to be close to the asymptotic limit of the deterministic system (see figure 1). Therefore, this probability is likely to be a good estimate of the likelihood of extinction.

Fig 1: `gillespie.m` script being run until t=10, demonstrating the asymptotic nature of the system dynamics at this point in time.

Change in population size with time
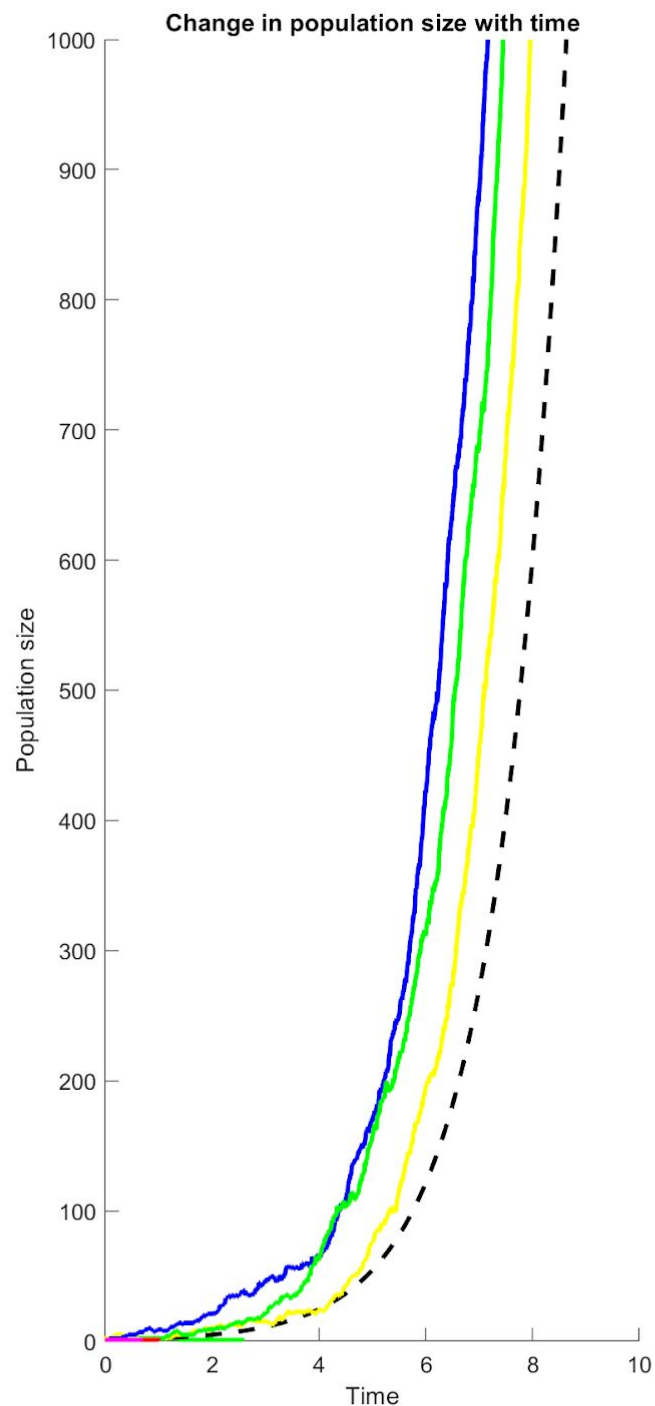
As shown in equation 8 in the work of Allen *et al.* (2016)[1], the probability of eventual extinction is given by $\lim\limits_{t\to\infty} p_0(t) = (\frac{d}{b})^i$, *if* $b > d$, and $\lim\limits_{t\to\infty} p_0(t) = 1$, *if* $b \leq d$, where i is the initial population size. Now, $\lim\limits_{t\to\infty} p_0(t) = (\frac{0.7}{1.5})^1 = 0.466...$ which is similar to the extinction probability found at t=10 in our simulations. The fact that it is slightly larger makes intuitive sense, as extinction continues to be a threat for the survivors after t=10.

I also tested for a starting population of 2, and found a 22.56% chance of extinction across 5,000 simulations, close to the 21.78% chance given by the equation given by Allen *et al.* Again, it makes intuitive sense that starting off with a larger population increases the chances of survival.

3.

The original script has been adapted and saved as `gillespie3.m`. The crucial changes are the additional parameters added at the beginning, and the line `B = birth_rate*(1-(y(j)/C));` where B is then used instead of `birth_rate`, as a population-dependent variable.

The population dynamics here do not display exponential growth anymore. With the carrying capacity set to 25, the simulated populations appear to never reach this value and fluctuate well below it (when studying the dynamics from T=0 until T=10). This can be seen in figure 2, obtained with the `plotresult` function. The rate of extinction roughly appears to be similar, from preliminary observations.

Using the code from q2, the probability of extinction (measured at t=10 across 5,000 replicates) is found to be 0.5138, which is larger than found in q2. This is as expected since the simulations reach smaller population sizes and smaller populations are more likely to go extinct. More importantly, the **mean population size is 5.9616**, with a variance of 44.1065. This mean is far below the ~3,000 average in the non-constrained system, and the variance is roughly 5 orders of magnitude smaller, too.

This long term mean describes the system's "quasi equilibrium state" because, while the system has a tendency to deviate from this value with time, it fluctuates around it - therefore it is almost at equilibrium. If it were a true equilibrium state, then it would not continue to change with time once reaching this mean. It is still quite possible, for example, for a population previously at a size of ~6 (the mean) to go extinct.

[1] Linda J. S. Allen, Sophia R. Jang & Lih-Ing Roeger (2017) Predicting population extinction or disease outbreaks with stochastic models, Letters in Biomathematics, 4:1, 1-22, DOI: 10.1080/23737867.2016.1264870
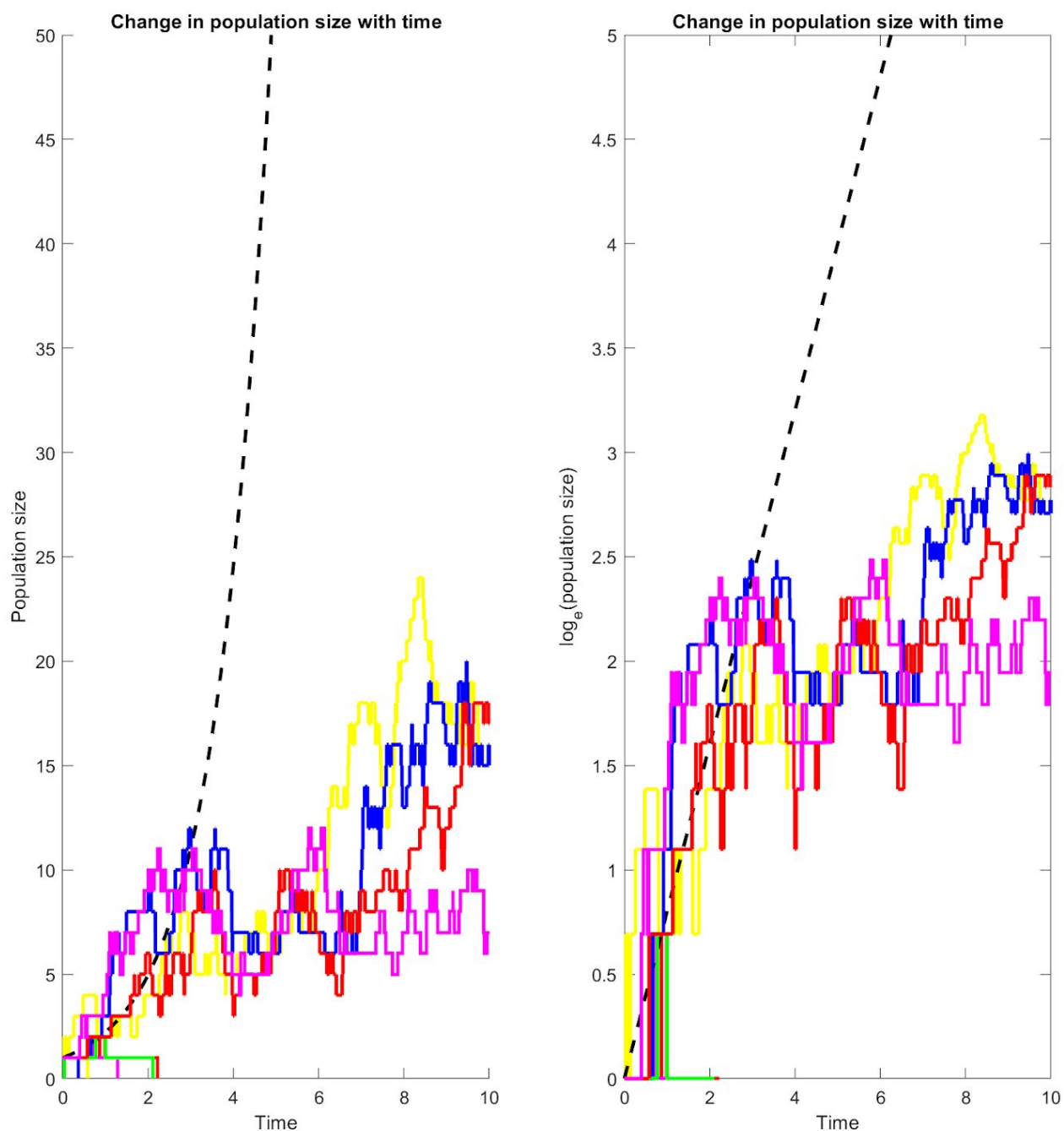
Fig. 2: population dynamics of a constrained stochastic system governed by the Gillespie system with a carrying capacity of 25. r_{birth} = 1.5 and r_{death} = 0.7.