Epidemics and zombies

This exercise is based on Alemi and Bierbaum's class project, published in "You Can Run, You Can Hide: The Epidemiology and Statistical Mechanics of Zombies", Alexander A. Alemi, Matthew Bierbaum, Christopher R. Myers, and James P. Sethna, Phys. Rev. E 92, 022146 (2015). See also the Zombietown site http://hey.runat.me/pages/programming/zombietown_usa.html and simulator http://mattbierbaum.github.io/zombies-usa. The computational aspects of the exercise start in part (e).

Epidemics are studied by disease control specialists using statistical methods, modeling the propagation of the disease as susceptible people are infected, infect others, and recover. The SIR model is the simplest commonly studied model, with three coupled differential equations: $dS/dt = -\beta SI$ reflects the rate β at which each infected person I infects each susceptible person S, and $dR/dt = \kappa I$ reflects the rate κ that each infected person joins the recovered population R.

(a) *What is the equation for dI/dt implied by these first two equations, assuming no infected people die or shift groups other than by new infections or recoveries?*

We shall use a less common, but even simpler SZR model, designed to predict the evolution of a zombie outbreak.
  $dS/dt = -\beta S Z$
  $dZ/dt = (\beta-\kappa) S Z$
  $dR/dt = \kappa S Z.$
Here the zombie population Z never recovers, but if it is destroyed by a member of the surviving population S, it joins the removed population R. The bite parameter β describes the rate at which a zombie bites a human it encounters, and the kill parameter κ gives the rate that a human may destroy a zombie it finds.

The SZR model is even simpler than the SIR model, in that one can write an explicit solution for the evolution as a function of time. We do so in two steps.

(b) *Argue that the only stationary states have all zombies or all humans. Both dS/dt and dZ/dt are linear in SZ, so there must be a linear combination of the two that has no time dependence. Show that $P=Z+(1-\kappa/\beta)S$ satisfies $\dot{P}=0$. Argue from these two facts that for P<0 the zombies must lose.*

(c) *Show that $\chi=S/Z$ satisfies $d\chi/dt=\gamma\chi$, and so $\chi(t)=\chi_0 \exp(\gamma t)$. Show that $\gamma = -\beta P$. Check that this answer concurs with your criterion for human survival in part (b).*

So the fraction of the doomed species exponentially decays, and the population of the surviving species is determined by P. If desired, one could use the added equation S(t)+Z(t)+R(t)=N with your answers to parts (b) and (c) to solve analytically for the explicit time evolution of S, Z, and R. We shall solve these

equations numerically instead (below).

Suppose now that we start with a single zombie Z0=1, and the number of humans S0 is large. It would seem from our equation for our invariant P from part (b) that if the bite parameter β is greater than the kill parameter κ that the humans are doomed. But surely there is some chance that we will be lucky, and kill the zombie before it bites any of us? This will happen with probability $\kappa/(\beta+\kappa)$. If we fail the first time, we can hope to destroy two zombies before either bites again\dots

Here is where the statistical fluctuations become important. The SZR differential equations above are a continuum approximation to the discrete transitions between integer numbers of the three species. These three equations are similar to reaction rate equations in chemistry (as in eqn 6.50) with molecules replaced by people:

$S + Z \rightarrow 2\,Z$ (rate $\beta\,S\,Z$)

$S + Z \rightarrow S + R$ (rate $\kappa\,S\,Z$).

Just as for disease outbreaks, if the number of molecules is small then chemical reactions exhibit important statistical fluctuations. These fluctuations are important, for example, for the biology inside cells, where the numbers of a given species of RNA or protein can be small, and the number of DNA sites engaging in creating RNA is usually either zero or one (see Exercises 8.10 and 8.11).

We can simulate each individual bite and kill event for a population of S humans and Z zombies. (Indeed, this can be done rather efficiently for the entire population of the USA; see the simulator above.) The time to the next event is an exponential random variable given by the total event rate. Which event happens next is then weighted by the individual rates for the different events.

It is well known that decay rates add. Let us nonetheless derive this. Let the probability density of event type #n be $\rho_n(t)$, with survival probability $S_n(t) = 1 - \int_0^t \rho_n(\tau)d\tau$. Let there be two types of events.

(d) Write the probability density of the first event, $\rho_{tot}(t)$, in terms of ρ_1, ρ_2, S_1, and S_2. (The final answer should involve no integrals or derivatives.) Specialize to systems with constant rates, which have an exponentially decaying survival time, Subscript[S, n] $(t) = \exp(-\gamma_n t)$. Show that the total event rate γ_{tot} is the sum $\gamma_1 + \gamma_2$. Show that the probability of the next event being γ_1 is γ_1/γ_{tot}.

To simulate one step of our discrete SZR model, we

(i) find the total rate of events γ_{tot},

(ii) increment t by Δt, a random number pulled from an exponential distribution with decay rate γ_{tot},

(iii) choose to bite or to kill by choosing a random number uniform in $(0, \gamma_{tot})$, and checking if it is less than γ_1,

(iv) change S, Z, and R appropriately for the event, and

(v) perform any observations needed.

This is a simple example of the Gillespie algorithm, discussed in more detail in Exercises 8.10 and 8.11.

(e) *Write a routine to use the Gillespie algorithm to solve the discrete SZR model for the two reaction-rate equations above, keeping track of t, S, Z, and R for each event, ending when S=0 or Z=0, and adding an extra point at t_{max} if the system terminates early. Write a routine numerically solve the three continuum equations above. Use β=0.001, κ=0.0008, and t_{max}=5. (Should the zombies win?) Compare the two for initial conditions Z_0=100, S_0=9900, and R_0=0 (a simultaneous outbreak of a hundred zombies). Is the continuum limit faithfully describing the behavior for large numbers of zombies and humans?*

```
β = 0.001; κ = 0.0008;
tMax = 5; S0 = ...; Z0 = ...; R0 = 0;
SZReqns = {S'[t] == ..., ..., ...}
{Ssol, Zsol, Rsol} =
  NDSolveValue[{SZReqns, S[0] == S0, ...}, {S, Z, R}, {t, 0, tMax}];

gContinuous = Plot[{Ssol[t], ..., ...}, {t, 0, tMax},
  PlotLegends → {"Survivors", "Zombies", "Removed"}, AxesLabel → {"Time t",}]


Gillespie[S0_, Z0_, R0_, tMax_] :=
 Module[{t = 0., S = S0, Z = Z0, R = R0, biteRate, killRate,
   totalRate, whichEvent, Soft, Zoft, Roft, St, Zt, Rt},
  (* Reap and Sow are much faster than appending to a list. Sow the points S(t),
  Z(t), and R(t) during the simulation, and reap them at the end. *)
  {{Soft}, {Zoft}, {Roft}} =
   Reap[While[t < ... && ...,
      biteRate = ...;
       ...;
      totalRate = ...; t += RandomVariate[ExponentialDistribution[...]];
      whichEvent = ...;
      If[whichEvent < ...,
        (* Zombie bites human *)
        S = ...; ...,
        (* Human kills zombie *)
        ...];
      Sow[{t, S}, St];
      Sow[{t, Z}, Zt];
      Sow[{t, R}, Rt]];
     If[t < tMax,
      Sow[{tMax, S}, St];
       ...];
     , {St, Zt, Rt}][[2]]; {Soft, Zoft, Roft}]

tMax = 5; S0 = ...; Z0 = ...; R0 = 0;
{Soft, Zoft, Roft} = Gillespie[...];
gGillespie = ListPlot[{Soft, Zoft, Roft}, Joined → True]
Show[gContinuous, gGillespie]
```

Now let us examine the likelihood of zombies being stamped out despite their advantage in biting, if we start with only one zombie.

(f) *Compare the continuum solution with the zombie simulation for twenty initial conditions using $Z_0=1$, $S_0=9999$, and $R_0=0$. Are the simulations suffering from more fluctuations than they did for larger number of zombies? Do you see evidence for zombie extinction early in the outbreak? What fraction of the initial outbreaks appear to have killed off all the humans? Zoom in to early times ($Z≤5$, $t<0.5$) and note a few*

trajectories where the first zombie is killed before biting, and trajectories where the zombie population goes extinct after reaching a peak population of two or three.

```
tMax = 5.; S0 = ...; Z0 = ...; R0 = 0;
{Ssol, Zsol, Rsol} = NDSolveValue[..., ..., ...];
gContinuousExtinct = Plot[...];
nRuns = 20;
For[n = 1, n ≤ nRuns, n = n + 1,
  {{SG[n], ZG[n], RG[n]} = Gillespie[...];
    gG[n] = ListPlot[..., Joined → True];}];
Show[gContinuousExtinct, Table[gG[n], {n, 1, nRuns}]]

Show[Table[gG[n], {n, 1, nRuns}], PlotRange → {{0, 0.5}, {0, 5}}]
```

We can estimate the probability P^∞_{ext} that a single zombie with bite rate larger than kill rate will go extinct before taking over a large initial human population $S_0 \to \infty$. As described in Alemi and Bierbaum's zombie paper in PRE, the probability P_{ext} that the zombies go extinct, in the limit of many humans, is equal to the probability that the first one is destroyed, plus the probability that it bites first times the probability that both zombie lines go extinct:

$$P^\infty_{ext} = \kappa/(\beta+\kappa) + \beta/(\beta+\kappa) \, (P^\infty_{ext})^2.$$

This can be solved to show $P^\infty_{ext} = \kappa/\beta$.

Exactly this same argument holds for regular disease outbreaks. Similar arguments can be used to determine the likelihood that an advantageous gene mutation will take over a large population.

(g) *Was the fraction of extinctions you observed in part (f) roughly given by the calculation above? Write a simpler routine that simulates the Gillespie algorithm over n epidemics, reporting the fraction in which zombies go extinct (observing only whether Z=0 happens before S=0, ignoring the time and the trajectory). For S_0=9999 and 1000 epidemics, how good is the prediction? Draw a plot of P^∞_{ext} versus log S_0, for S_0=1, 2, 4, ..., 512, using 1000 epidemics each. How large must the human population be before our estimate P^∞_{ext} is within about 10% of the correct answer?*

```
Extinction[S0_, Z0_, R0_, nEpidemics_] :=
 Module[{S, Z, R, biteRate, killRate, totalRate, whichEvent},
  nExtinct = ...;
  For[epidemic = 1, ..., ...,
   S = S0; Z = Z0; R = R0;
   While[S Z > 0,
    {biteRate = ...;
      ...;
     If[whichEvent < ...,
       (* Zombie bites human *)
       ...,
       (* Human kills zombie *)
       ...]}];
   If[S > 0, ...]];
  N[nExtinct/nEpidemics]]
```

```
Extinction[..., ..., 0, ...]

extinctionVsPop = Table[{2^n, Extinction[2^n, ...]}, {n, 0, 10}]

ListLogLinearPlot[extinctionVsPop, Joined → True,
 AxesLabel → {"Initial Population", "Zombie Extinction Probability P_ext"}]
```