Stochastic individual-based models

Modern Techniques in Modelling

Introduction

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Previously

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

– We have explored discrete time-step compartmental models, and continuous ODE-based compartmental models, both of which are deterministic and track *groups* of individuals.

– Now we'll be exploring another kind of model, **stochastic individual-based models** (also called: agent-based models, microsimulations).

– To understand this type of model, we'll discuss what **individual-based** means and what **stochastic** means in this context.

Stochastic **individual-based** models

Individual-based logic:

- Different kind of logic to a compartmental model.
- Instead of thinking about and tracking groups of individuals:
  - There are 634 infected individuals and 366 uninfected individuals.
  - What happens next to each group?
- We are thinking about and tracking individuals:
  - Individual 1 is infected.
  - Individual 2 is not infected.
  - Individual 3 is infected.
  - …
  - What happens next to each individual?

Stochastic individual-based models

Probabilistic rules:

- Models are specified in terms of probabilistic "rules" for each individual.
- Usually, these rules are evaluated in discrete time steps Δt, often 1 day.
- For example, for an SIS model (Susceptible-Infectious-Susceptible), each individual might follow these rules once each day:

If susceptible,
then with probability X become infectious
If infectious,
then with probability Y become susceptible

How do we choose these probabilities?

- Typically, X would be higher or lower depending on the force of infection, and Y would be chosen to reflect the duration of infectiousness (we'll see more on the next few slides).
- Later we'll discuss more complex kinds of rules.

Turning event rates into probabilities

In an ODE model, a state transition is an event that happens at a certain rate.

$$\begin{aligned} dS/dt &= -\lambda S + \gamma I \\ dI/dt &= \lambda S - \gamma I \\ \lambda &= \beta \frac{I}{N} \end{aligned}$$

Above, the S -> I transition happens at rate $\lambda = \beta \frac{I}{N}$ for each S individual, and the I -> S transition happens at rate γ for each I individual.

Both λ and γ are "per-capita" rates, in units of "events per unit time, per person" (events $\cdot$ unit time^{-1} $\cdot$ person^{-1}). The overall rate of "events per unit time" is obtained by multiplying these by *S* and *I* respectively.

Turning event rates into probabilities

Suppose we wanted to set up our stochastic individual-based model in the same way, with an S -> I transition at rate λ and an I -> S transition at rate γ.

– For an SIS model (Susceptible-Infectious-Susceptible), each individual might follow these rules once each day:

If susceptible,

then with probability X become infectious

If infectious,

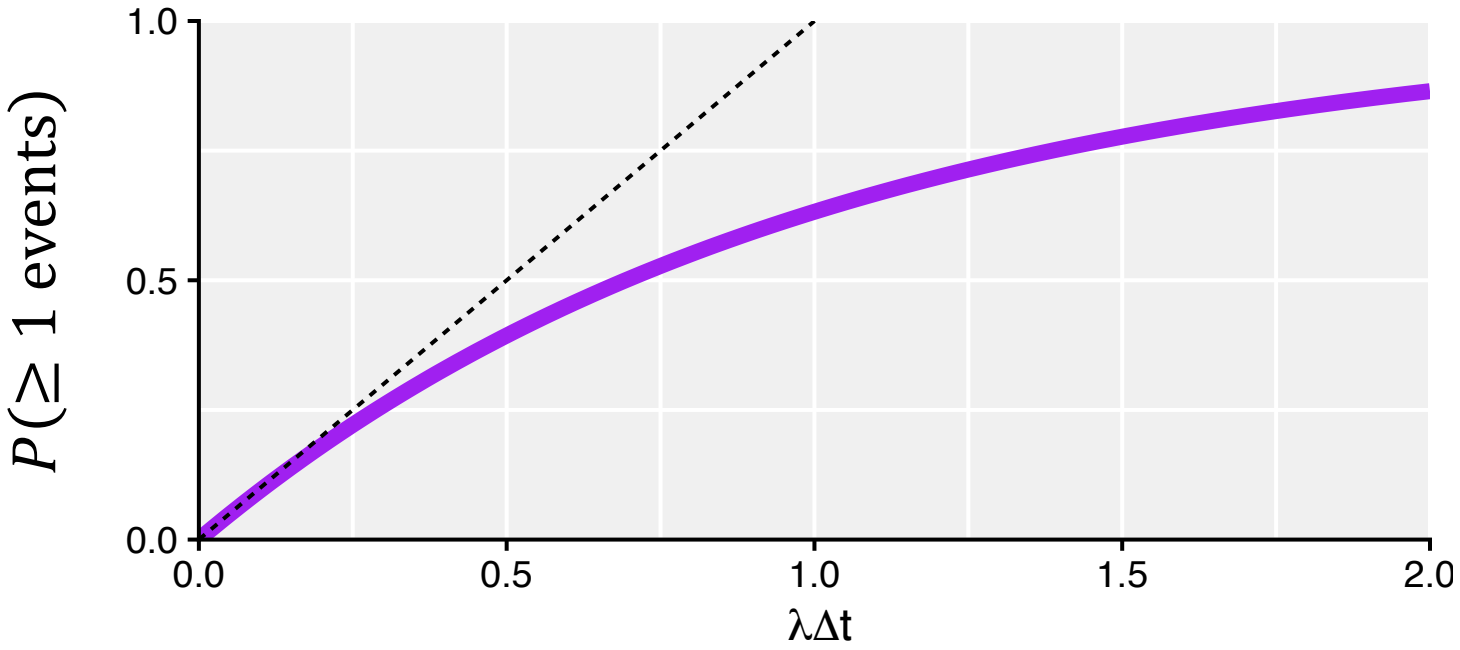
then with probability Y become susceptible

How do we turn a rate (e.g. λ) into a probability (e.g. X) that, during a given time step, a given individual experiences the event?

Turning event rates into probabilities

– If some event happens at per-individual rate λ, the probability of it happening at least once to a given individual in a given time step of duration Δt is

$$P(\geq 1 \text{ events}) \quad = \quad 1 - e^{-\lambda \Delta t}$$



Turning event rates into probabilities

We might think:

probability = rate × Δt

e.g.

rate = 0.1 events/day

Δt = 1 day

probability = 0.1 × 1 = 0.1

However, this starts to break down for higher rates:

e.g.

rate = 1.2 events/day

Δt = 1 day

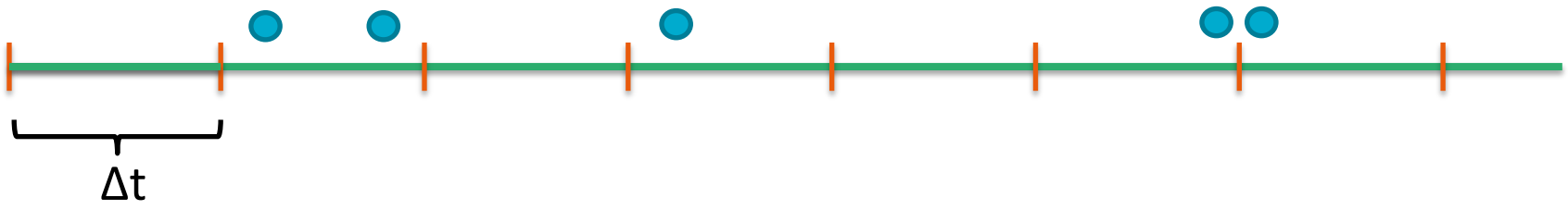
probability = 1.2 × 1 = 1.2???

Problems:

- Not clear what to do with this probability that is > 1.
- Even when an event happens, on average, once per day, there can be days where the event doesn't happen at all and days when it happens multiple times. We want to preserve this property of random events.

Turning event rates into probabilities

Consider a length of time, which is divided into time steps Δt, “scattered” randomly with events.

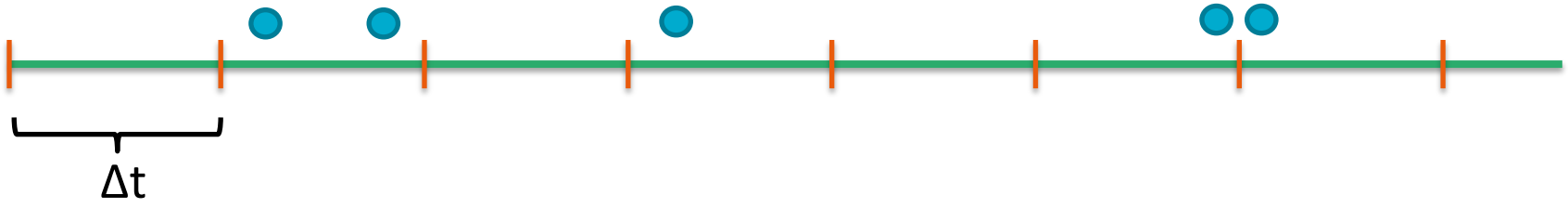


Some time steps have 0 events, some have 1, some have >1. How do we account for this mathematically?

The difficulty here is that some time steps contain >1 events. So let’s take our time step and break it down even further into pieces that are so small, it’s safe to assume that no more than one event can happen in each tiny piece.

Turning event rates into probabilities

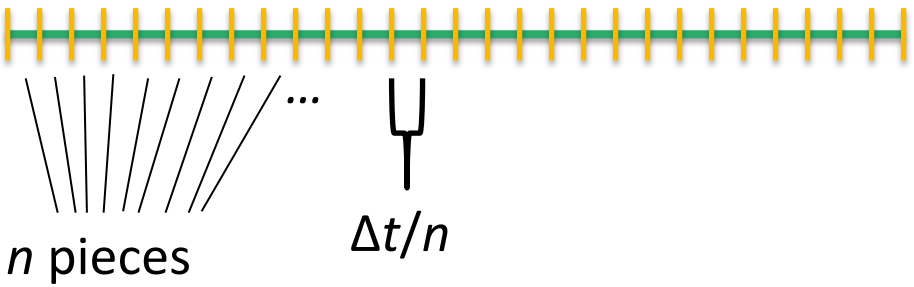
Consider a length of time, which is divided into time steps Δt, “scattered” randomly with events.



Some time steps have 0 events, some have 1, some have >1. How do we account for this mathematically?

The difficulty here is that some time steps contain >1 events. So let’s take our time step and break it down even further into pieces that are so small, it’s safe to assume that no more than one event can happen in each tiny piece.

Turning event rates into probabilities



There are n pieces, each of length $\Delta t/n$.
The probability of the event, whose rate is λ, occurring during any given time-step piece is $\lambda \times \Delta t/n$.

Now we can calculate the probability that 0, 1, 2, 3, … events happen in Δt.

$$P(0 \text{ events}) = (1 - \lambda\,\Delta t/n)^n$$

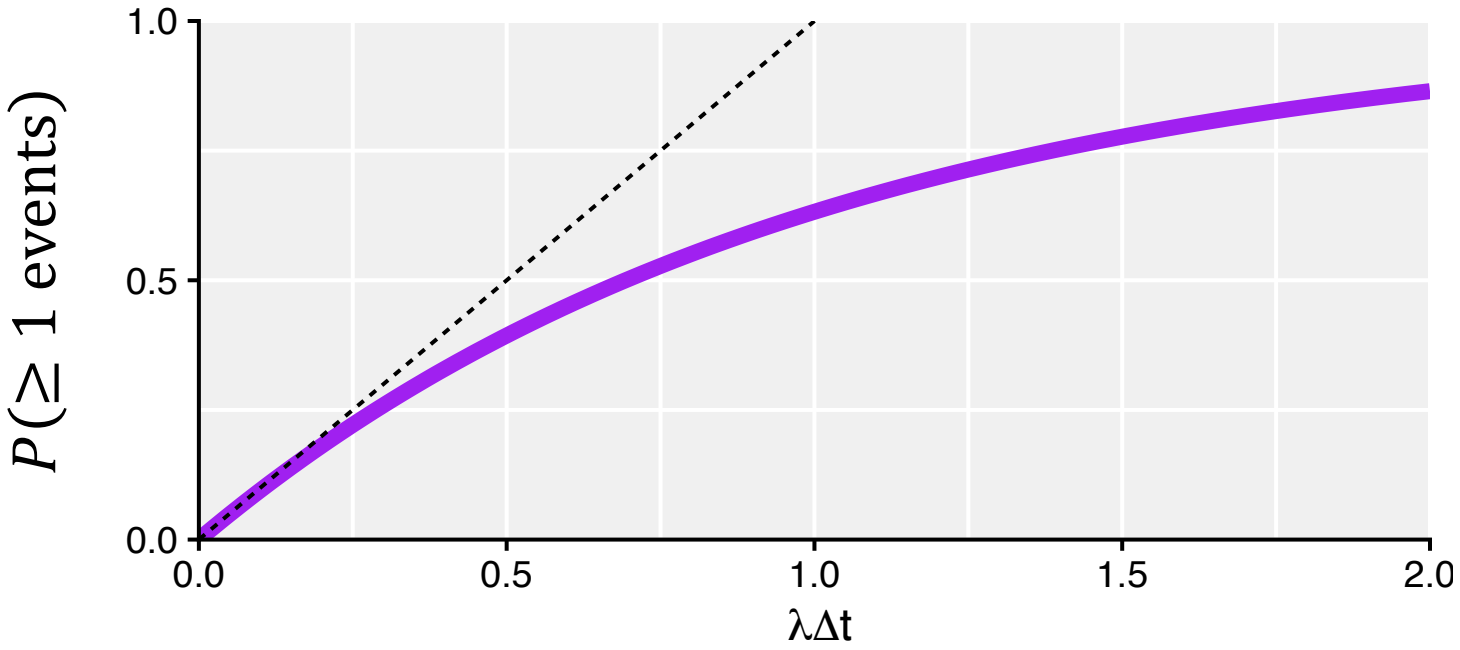
$$\lim_{n\to\infty} (1 - \lambda\,\Delta t/n)^n = e^{-\lambda\Delta t}$$

$$P(0 \text{ events}) = e^{-\lambda\Delta t}\ ;\ \ P(\geq 1 \text{ events}) = 1 - e^{-\lambda\Delta t}$$

Turning event rates into probabilities

– If some event happens at per-individual rate λ, the probability of it happening at least once to a given individual in a given time step of duration Δt is

$$P(\geq 1 \text{ events}) \quad = \quad 1 - e^{-\lambda \Delta t}$$



An example stochastic individual-based model

Let's say we have N individuals who can be either susceptible or infected.

The force of infection is $\lambda = \beta I/N$ and the recovery rate is γ.

We loop over time steps 1 to T and our time step size is Δt.

```
For each ts from 1 to T {
  lambda <- beta * I/N

  For each i from 1 to N {
    If individual i is susceptible:
      with prob 1-exp(-lambda·Δt) make infected.
    Else-if individual i is infected:
      with prob 1-exp(-gamma·Δt) make susceptible.
  }

  Record population state
}
```

An example stochastic individual-based model

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Pseudocode

```
Initialize state


For each ts from 1 to T {
  lambda <- beta * I/N

  For each i from 1 to N {
    If individual i is susceptible:
      with prob 1-exp(-lambda·Δt),
        make infected.

    Else-if individual i is infected:
      with prob 1-exp(-gamma·Δt),
        make susceptible.
  }


  Record population state
}
```

R code

```
state <- rep("S", n)
state[1:10] <- "I"

for (ts in 1:steps) {
  lambda <- beta * sum(state == "I") / n

  for (i in 1:n) {
    if (state[i] == "S") {
      if (runif(1) < 1 - exp(-lambda * dt)) {
        state[i] <- "I"
      }
    } else if (state[i] == "I") {
      if (runif(1) < 1 - exp(-gamma * dt)) {
        state[i] <- "S"
      }
    }
  }

  # Record population state . . .
}
```

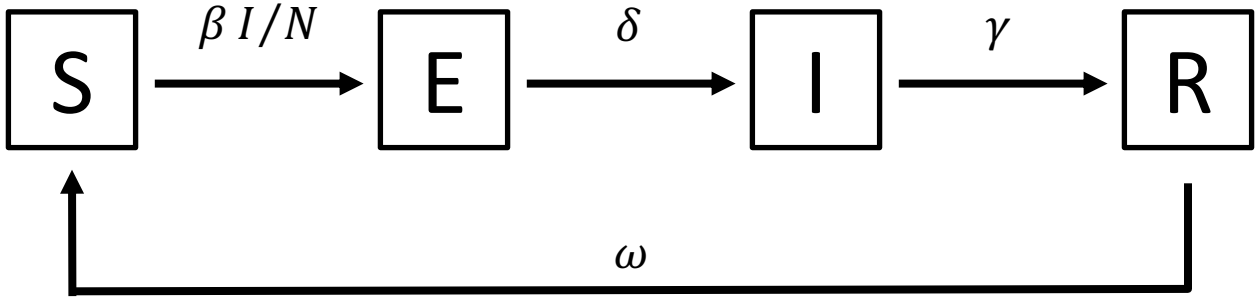
Outline of the practical

1. An individual-based SEIR model of SARS-CoV-2 transmission.
2. Adding more complex dynamics to the SEIR model.
3. Optimizing the model to run faster.

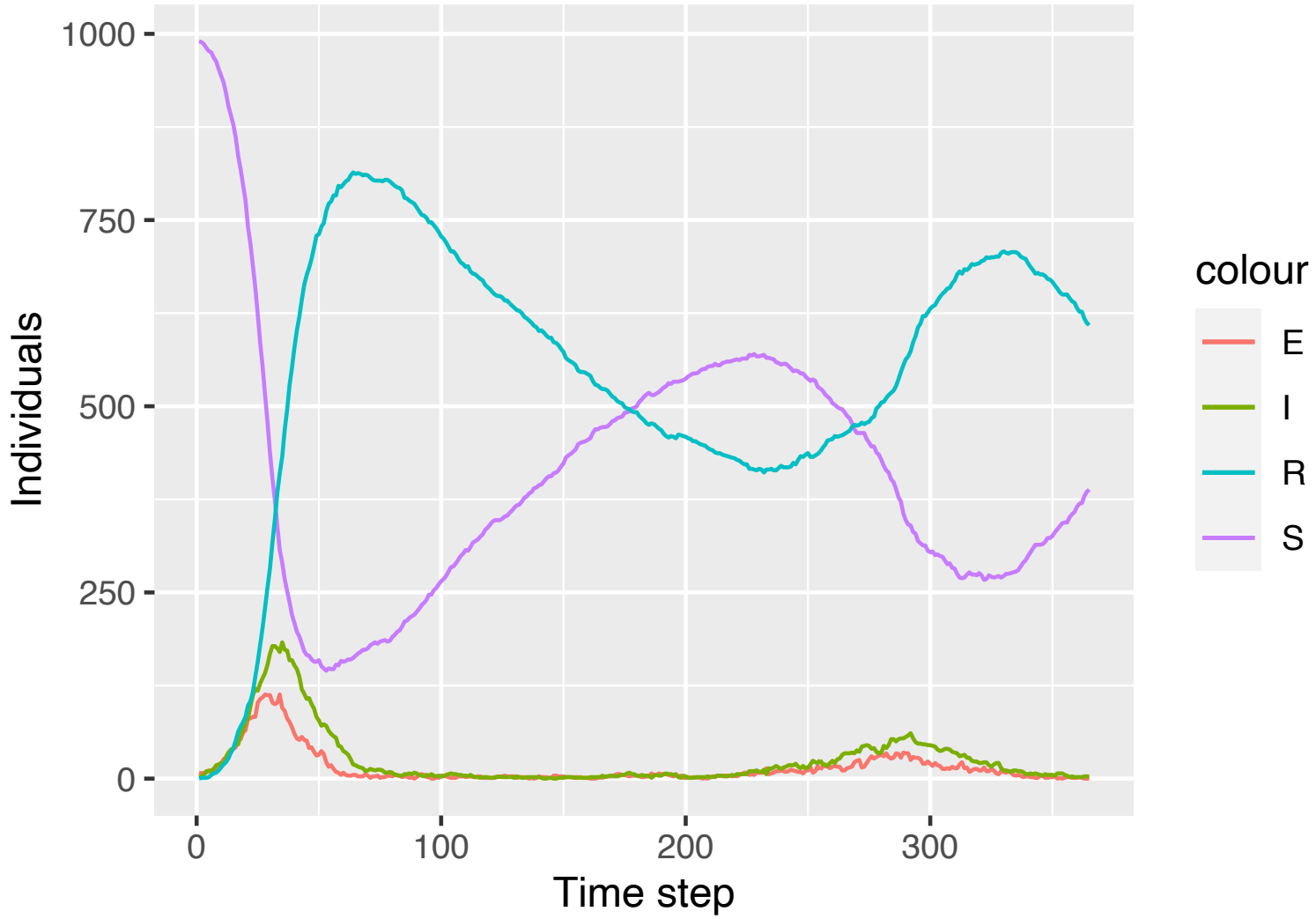
Practical 1. An individual-based SEIR model of SARS-CoV-2 transmission

Practical 1: Individual-based SEIR model

In the first practical, you will be implementing an SEIR model with waning immunity as a stochastic individual-based model.

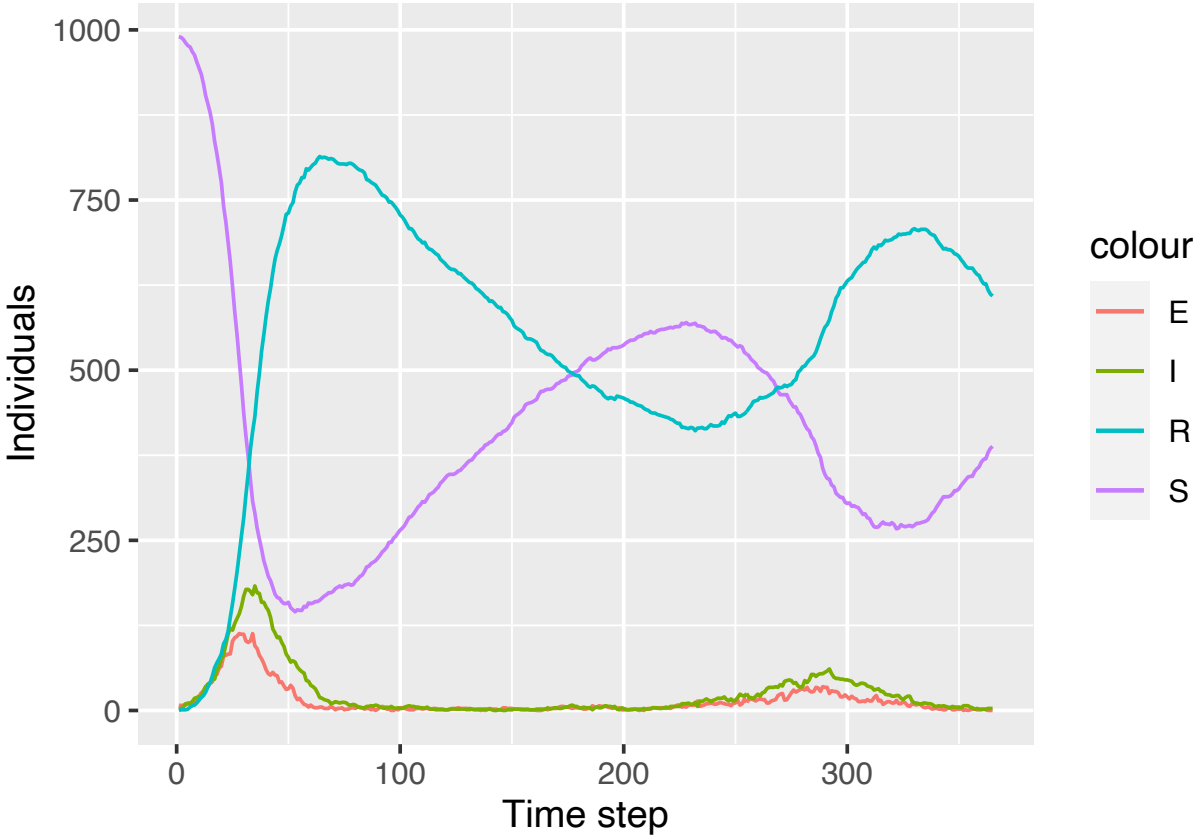


Practical 1: Individual-based SEIR model



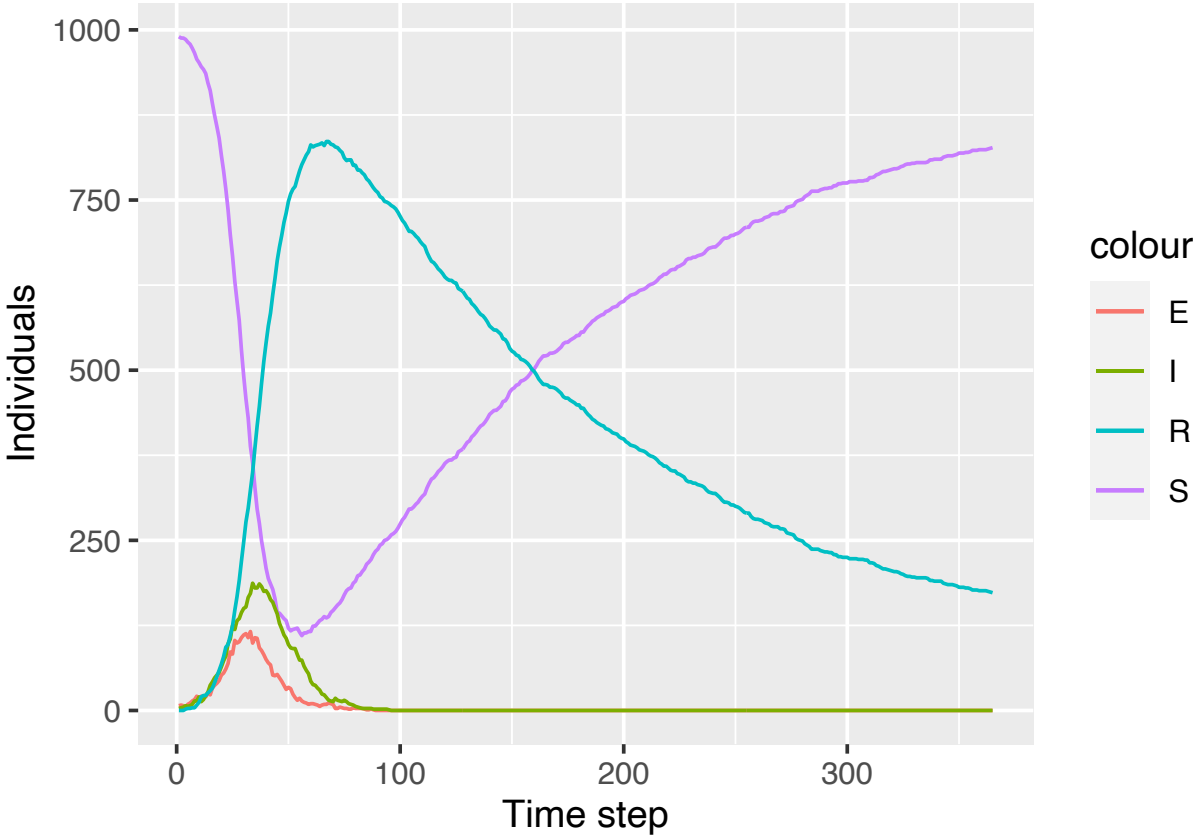
Practical 1: Individual-based SEIR model

6. What happens if you use `set.seed(123456)` instead of `12345`?



Practical 1: Individual-based SEIR model

6. What happens if you use `set.seed(123456)` instead of `12345`?



Advantages of individual-based models

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

- Each individual can have properties (age, sex, physical location, network of contacts…) that do not need to be "grouped" as in a compartmental model, and which can change in precisely specified ways over time.

- Can be easier to think about.

- Can allow for more complex dynamics.

- Because rules are probabilistic, allows consideration of stochastic effects.

Disadvantages of individual-based models

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

- Usually hard to "prove" results. (*i.e.* not analytically tractable)
- Can be more complex to implement than compartmental models.
- Slower to run than compartmental models, especially in R, especially if not coded carefully.
- Stochastic effects introduce complications:
  - Results may depend upon population size, time step size
  - You have to average over multiple "runs" to have confidence in results

Practical 2. Adding more complex dynamics to the model

More complex dynamics

Up until now, in our model we have only kept track of one “piece of information” for each individual, their state (S, E, I, or R).

In an individual-based model, we can track as many different properties for each individual as we can think of – such as their age, viral load, antibody level, geographical location, household, etc…

These properties are not limited to discrete states (such as S/E/I/R, old/young, etc) but can be values such as real numbers (1.234) or character strings (AGTAGCTAGGC…)

We will now explore a few such examples.

Example 1. Age

We can give each individual an age (in years) in addition to their state:

```
state <- rep("S", n)
state[1:10] <- "E"
age <- runif(n, 0, 80)
```

Using this, we can make other properties of the simulation depend upon a person's age, such as making infectiousness depend on age:

```
# before
lambda <- beta * sum(state == "I") / n
# after
lambda <- beta * sum(infectiousness(state, age)) / n

infectiousness <- function(state, age) {
    ifelse(state == "I", 1.25 - age/160, 0)
}
```

Example 1. Age

We can also simulate aging in the model:

```
for (i in 1:n) {
    age[i] <- age[i] + dt / 365
    # . . .
}
```

But this may not make a difference if we are only simulating short outbreaks.

Example 2. Arbitrary delays

In previous practicals, individuals had a constant rate of transitioning out of, for example, the latent period (the "Exposed" state).

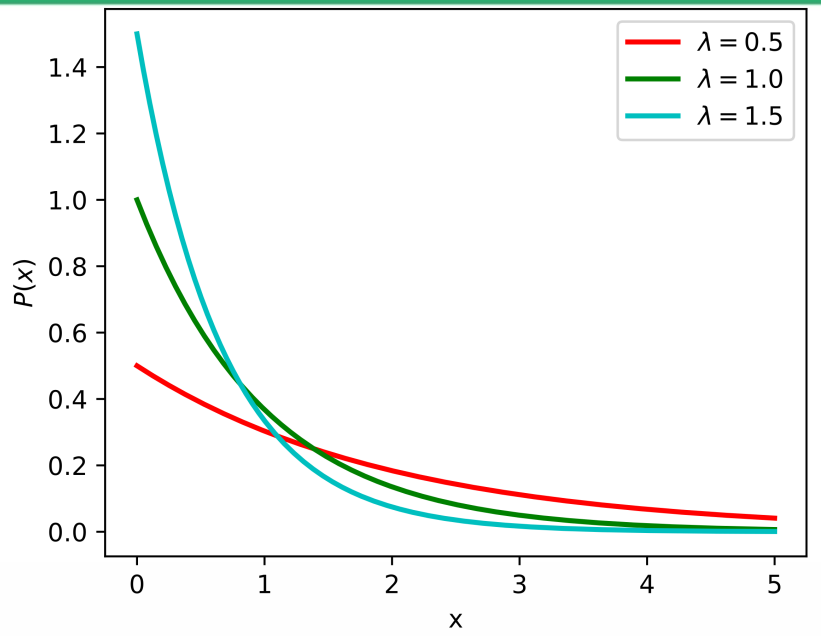
$$dE/dt = (\beta I/N)S - \delta E$$
$$dI/dt = \delta E - \gamma I$$

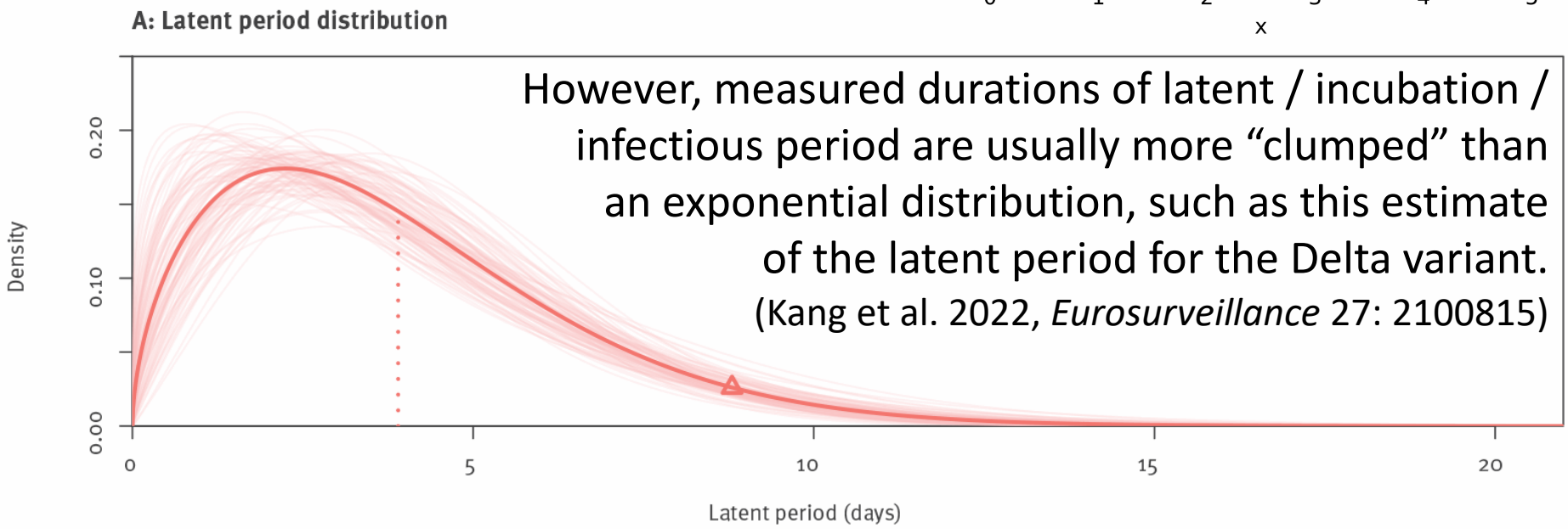
```
if (state[i] == "E") {
    if (runif(1) < 1 - exp(-delta * dt)) {
        state[i] <- "I"
    }
}
```

Example 2. Arbitrary delays

When individuals leave a compartment at a constant per-capita rate, the amount of time each individual spends in the compartment is **exponentially distributed**.

(image: Wikipedia)

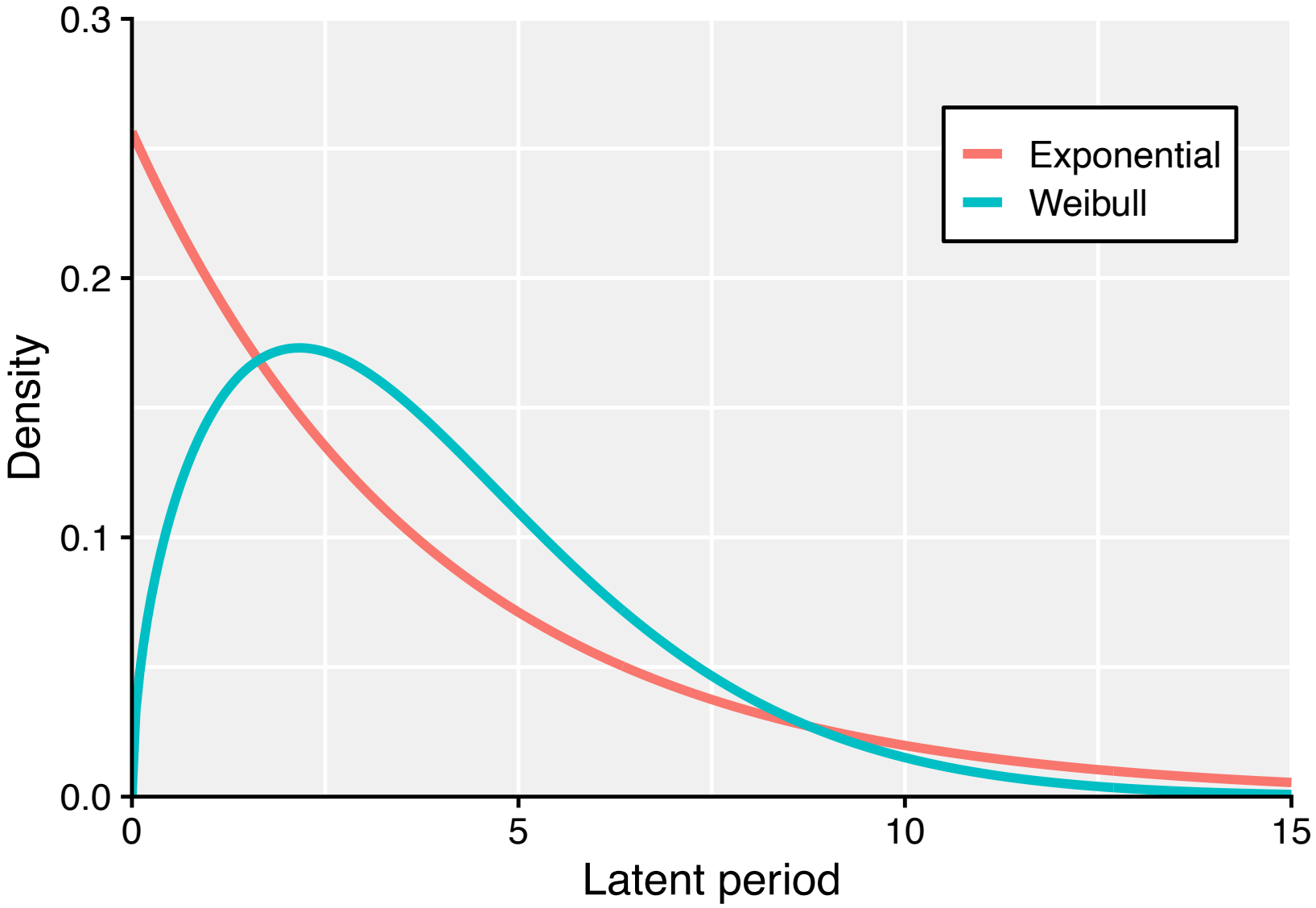




However, measured durations of latent / incubation / infectious period are usually more "clumped" than an exponential distribution, such as this estimate of the latent period for the Delta variant.

(Kang et al. 2022, *Eurosurveillance* 27: 2100815)

Example 2. Arbitrary delays



Example 2. Arbitrary delays

We can give each individual a “delay” property:

```
delay <- rep(0, n)
```

When individual *i* enters the Exposed state, we can assign a value to `delay[i]` representing how much time they have left:

```
if (state[i] == "S") {
    # Transition S -> E (infection) at rate lambda
    if (runif(1) < 1 - exp(-lambda * dt)) {
        state[i] <- "E"
        delay[i] <- rlnorm(1, meanlog = 0.5, sdlog = 0.6)
    }
}
```

Example 2. Arbitrary delays

Each time step, we make sure to subtract `dt` from each individual's time remaining:

```
delay[i] <- delay[i] - dt
```

Finally, we make the transition E -> I dependent upon this delay instead of the old rate δ:

```
else if (state[i] == "E") {
    # Transition E -> I (latent to infectious)
    if (delay[i] < 0) {
        state[i] <- "I"
    }
}
```

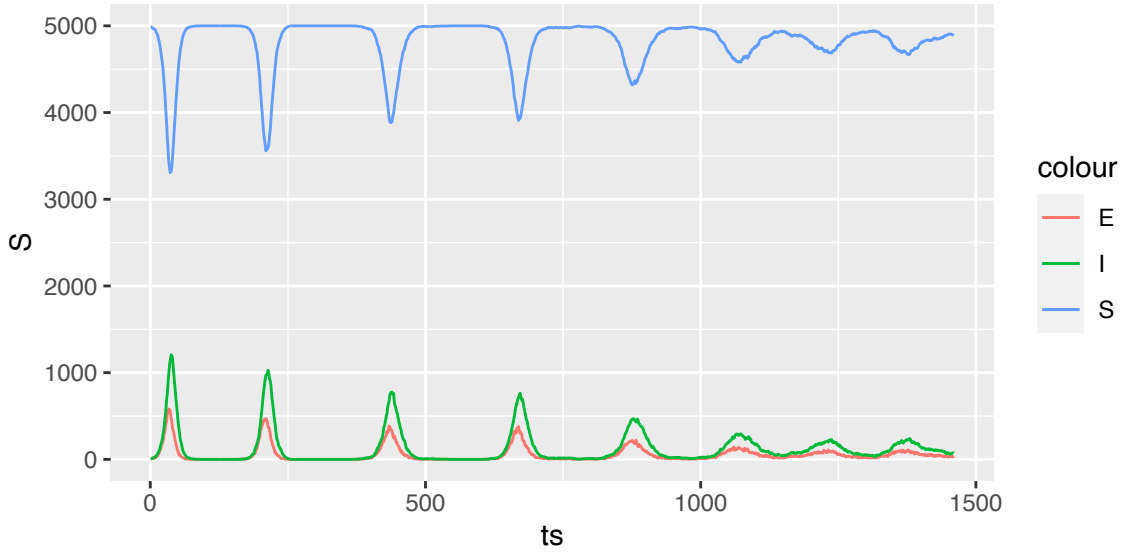
Practical 2. More complex dynamics

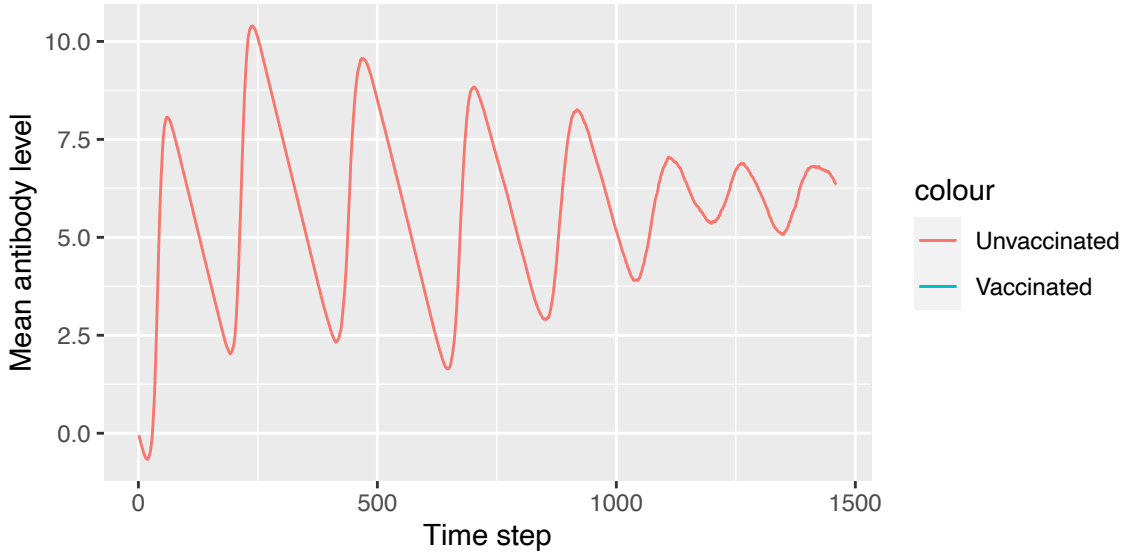
The instructions online will guide you through adding:

- delay distributions for the latent and infectious period,
- age and age-dependent infectiousness,
- immunity based upon antibody levels and
- vaccination

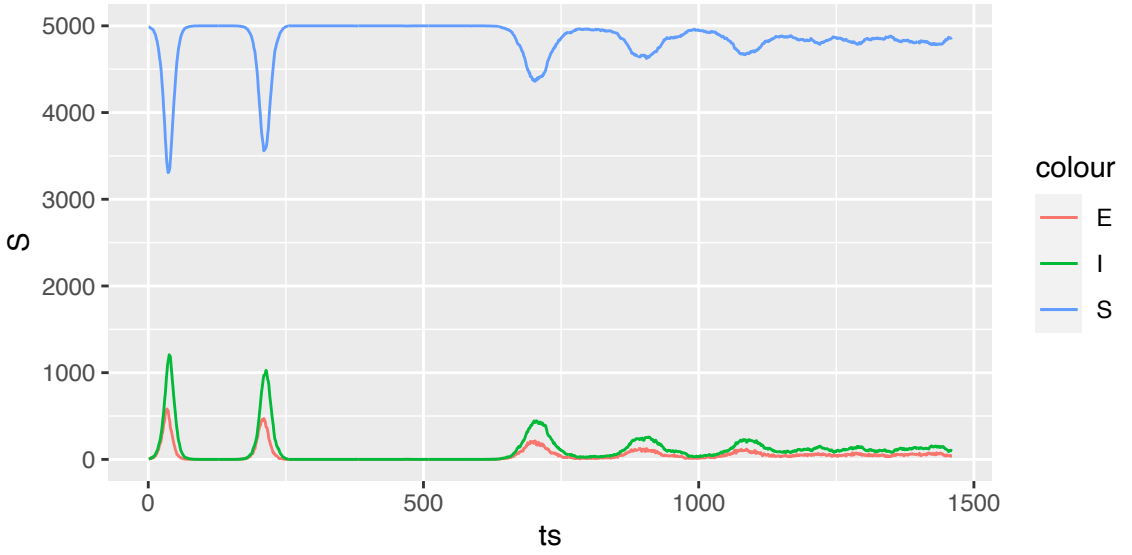
to the model from Practical 1.

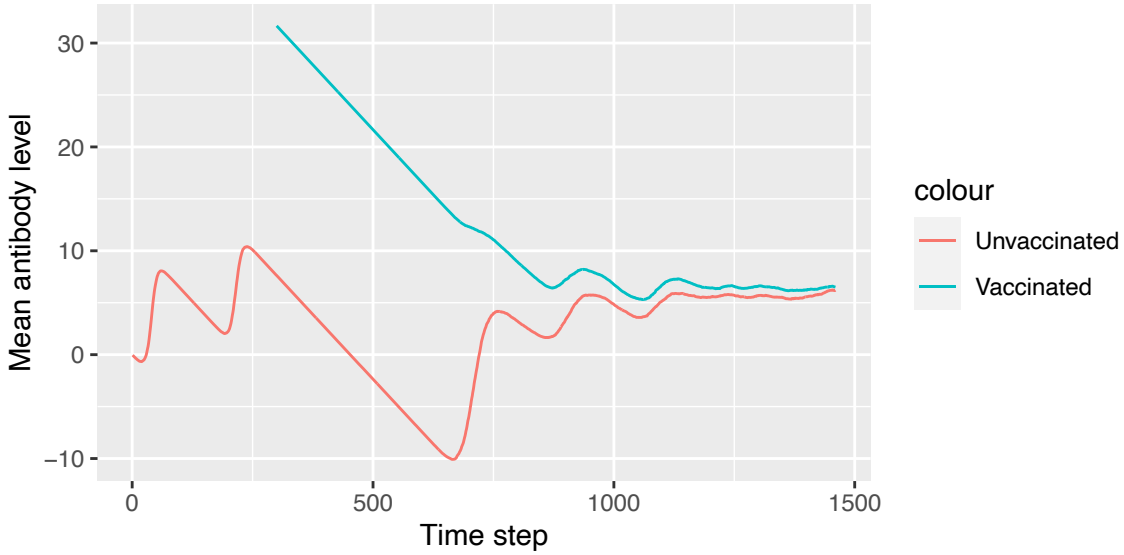
Practical 2. More complex dynamics





Practical 2. More complex dynamics





Practical 3. Optimizing the model to run faster

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Practical 3. Optimizing the model

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

- In R, code that loops over each individual in an individual-based model can be quite slow to run.

- This code can be optimized significantly by vectorizing it—that is, by applying events all at once rather than to each individual in turn.

Vectorizing R code

Compare this code:

```
for (i in 1:n) {
  if (state[i] == "S") {
    if (runif(1) < 1 - exp(-lambda * dt)) {
        state[i] <- "E"
    }
  }
}
```

With this code:

```
trE <- (state == "S") & (runif(n) < 1 - exp(-lambda * dt))
state[trE] <- "E"
```

Summary

- **Stochastic individual-based models** track individual members of a population and assume that events happen probabilistically, which means that they incorporate random chance.

- Individual-based models allow you to do things that compartmental models don't.
  - But they are more complex to implement and slower to run, so only use them if you need that added complexity.
  - Stochasticity also means we will need to run the model multiple times to get a good summary of what the model is doing.

- The run time of an individual-based model in R can be improved by "vectorizing" the code (but an individual-based model will usually still be slower than a deterministic, ODE model equivalent, if one exists).