

MA621L-Mathematical Modelling

Model-2 : The SIR Epidemic Model
Dym's Ten Principles and Velten's SQM Framework

Panchatcharam Mariappan¹

¹Associate Professor
Department of Mathematics and Statistics
IIT Tirupati, Tirupati

September 2, 2026





Where We Are

The Contract, Restated

- One model per week. This week the model is **SIR**.
- You are not being handed an equation. You are being handed a *town with a disease in it*.
- The only prerequisite today is calculus: the definition of a derivative, and one limit.

Grading reminder

An **unstated assumption** costs more marks than an arithmetic slip. A correct number with no assumption list cannot exceed 50%.



Plan for These Two Lectures



The same model, built twice, through two different disciplines.

- **Lecture A (Dym).** Ten questions: *Why? Find? Given? Assume? How? Predict? Valid? Verified? Improve? Use?*
This tells you *how to build* the model.
- **Lecture B (Velten).** The triple (S, Q, M) and the classification of the model in SQM space.
This tells you *what you have built*, and where it sits among all other models.

Why both?

Dym gives you a procedure. Velten gives you a coordinate system. A modeller needs both: one to produce the model, one to know its class before solving it.



Lecture A: Dym's Ten Principles

The Ten Questions

Why?	What are we looking for? Identify the need for the model.
Find?	What do we want to know? List the data we are seeking.
Given?	What do we know? Identify the available relevant data.
Assume?	What can we assume? Identify the circumstances that apply.
How?	How should we look at this model? Identify the governing principles.
Predict?	What will our model predict? Equations, calculations, answers.
Valid?	Are the predictions valid? Consistent with principles and assumptions?
Verified?	Are the predictions good? Useful for the reason it was done?
Improve?	Can we improve the model? Which assumption is lifted first?
Use?	How will we exercise the model?

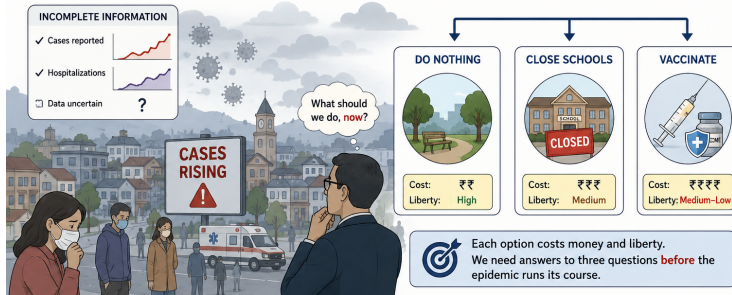


1. Why?



1. Why? — The Need for the Model

A disease is spreading in a town. Officials must decide now, on incomplete information, between



What we are NOT modelling



Not which individuals fall ill.



Not the immune system.



Not the hospital system.



The purpose fixes the level of description.

This is the first modelling decision, made **before** any mathematics.

2. Find?

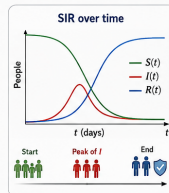
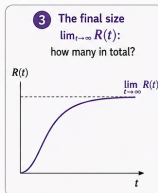
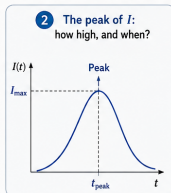
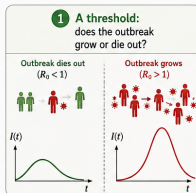


2. Find? – What We Want to Know

Three functions of time t (days):

Symbol	Meaning	Units
$S(t)$	Susceptible: not yet infected, can be infected 	people
$I(t)$	Infective: infected <i>and</i> able to transmit 	people
$R(t)$	Removed: recovered/immune, isolated, or dead 	people

And from them, three derived quantities:



We are not tracking individuals. We are tracking counts in three groups.
Our goal: threshold, peak, final size.

$S(t)$

$I(t)$

$R(t)$

3. Given?

3. Given? — What We Know

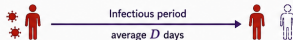
- Total population N of the town.



- Initial data $S(0) = S_0$, $I(0) = I_0$, $R(0) = R_0$,
with $S_0 + I_0 + R_0 = N$.



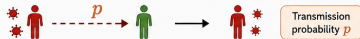
- An infected person stays infectious for an average of D days.



- People make on average c close contacts per day.



- A contact between an infective and a susceptible transmits with probability p .



Flag this now

c and p are very hard to measure *separately*.
Remember this. The derivation in Step 5 will tell us
why it **does not matter** — and why it **can never be fixed**.

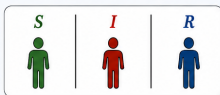


4. Assume?

4. Assume – The Block That Carries the Model

★ Every assumption is **numbered**, so that it can be **lifted** later. An unnumbered assumption is a defect.

A1



A1 (Exhaustive, exclusive compartments.) Everyone is in exactly one of S , I , R at every instant, so $S + I + R = N$.

A2



Births



Deaths



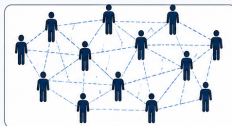
Immigration



Emigration

A2 (Closed population.) No births, deaths, immigration or emigration during the outbreak. Justification is about **timescales**, not about the town: an epidemic lasts weeks, demography takes decades.

A3



A3 (Homogeneous mixing.) Everyone is equally likely to contact everyone else. No households, no classrooms, no geography, no age structure. This is the strongest and least realistic assumption in the model.

4. Assume?

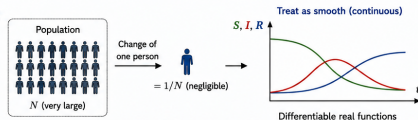


A4

A4 (Large N , continuum approximation.)

One person, a change of $1/N$, is negligible;
so integer-valued S, I, R may be treated
as differentiable real functions.

This is the assumption that licenses calculus.

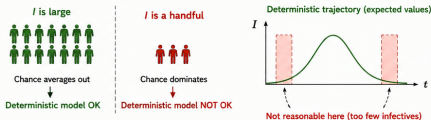


A5

A5 (Deterministic dynamics.)

We track expected numbers and ignore chance.
Reasonable while I is large.

Not reasonable when I is a handful —
i.e. at the start and at the end.



A6

A6 (No latent period.)

Infected at time $t \Rightarrow$ infectious at time t .
(Lifting gives SEIR.)



A7

A7 (Permanent immunity.)

No return from R to S .
(Lifting gives SIRS.)



4. Assume?

A8

A8 (Constant parameters.) c , p , D do not change: no behaviour change in response to news, no seasonality, no pathogen evolution.



No behaviour change in response to news



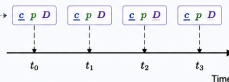
No seasonality



No pathogen evolution

Parameters are constant in time

c p D

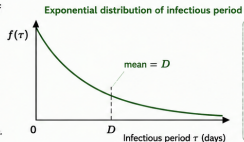


A9

A9 (Constant per-capita recovery rate.) A fixed fraction of infectives recovers per unit time.



Consequence, stated openly: this is equivalent to an **exponentially distributed** infectious period of mean D .
Some recover almost at once, a few stay infectious for a very long time.
The model does not know that a real infectious period is roughly fixed at 5–7 days.



What this means

- Many recover quickly
- Few remain infectious for a very long time



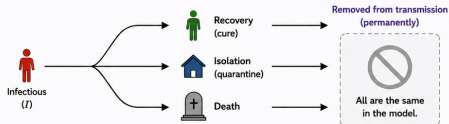
Real infectious period is roughly fixed at 5–7 days.

A10

A10 (Removal is complete.)

Recovery, isolation and death are *mathematically identical*: all remove a person from transmission permanently.

So the model cannot tell a cured patient from a dead one.



5. How?



5. How? — The Governing Principle

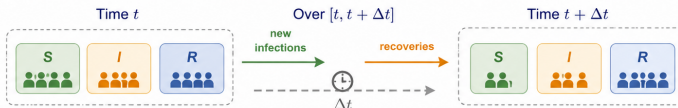


There is no law of physics for epidemics.

The governing principle is **bookkeeping** — conservation of people across compartments.

Over any time interval:

change in a compartment = number who entered – number who left



Apply this on a short interval $[t, t + \Delta t]$, divide by Δt , pass to the limit.

That is the whole model.



Same recipe, different compartments

We derived the heat equation from energy conservation on $[x, x + \Delta x]$. Identical logic. Only *the conserved stuff* changes.

5. How?



5.1 Counting the New Infections

Fix $[t, t + \Delta t]$, short enough that nobody's status changes twice inside it. Take **one** infective.

- 1 By A8, in time Δt this person makes $c \Delta t$ contacts.



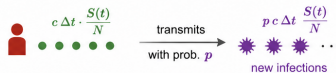
- 2 By A3, contacts are drawn at random from the whole town, so a fraction $S(t)/N$ of them land on susceptibles:

$$c \Delta t \cdot \frac{S(t)}{N} \text{ contacts with susceptibles.}$$



- 3 Each transmits with probability p , so this one infective causes

$$p c \Delta t \frac{S(t)}{N} \text{ new infections.}$$



- 4 There are $I(t)$ infectives, each doing this:

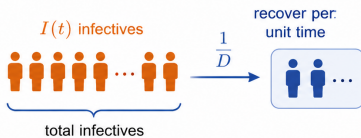


$$\text{new infections in } [t, t + \Delta t] = p c \frac{S(t)I(t)}{N} \Delta t$$

5. How?



By **A9**, a fraction $\frac{1}{D}$ of infectives recovers per unit time.



Define the **recovery rate**

$$\alpha := \frac{1}{D}$$

$$[\alpha] = \frac{1}{\text{time}} .$$

5.2–5.3 Recoveries, and the Balance Equations

By A9, a fraction $1/D$ of infectives recovers per unit time. Define the **recovery rate**

$$\alpha := \frac{1}{D}, \quad [\alpha] = \frac{1}{\text{time}}.$$

Recoveries in $[t, t + \Delta t]$: $\alpha I(t) \Delta t$. Now do the bookkeeping.

$$S(t + \Delta t) - S(t) = -\beta \frac{S(t)I(t)}{N} \Delta t$$

$$I(t + \Delta t) - I(t) = +\beta \frac{S(t)I(t)}{N} \Delta t - \alpha I(t) \Delta t$$

$$R(t + \Delta t) - R(t) = +\alpha I(t) \Delta t$$

Susceptibles only leave; removed only arrive; infectives do both.



5. How?



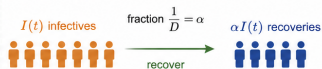
5.2–5.3 Recoveries, and the Balance Equations

By **A9**, a fraction $\frac{1}{D}$ of infectives recovers per unit time.

Define the
recovery rate

$$\alpha := \frac{1}{D}, \quad [\alpha] = \frac{1}{\text{time}}.$$

Per unit time:



Recoveries in $[t, t + \Delta t]$: $\alpha I(t) \Delta t$.

Now do the bookkeeping.

Bookkeeping over $[t, t + \Delta t]$



Compartment	Change = Inflow – Outflow	Differential equation (divide by Δt , let $\Delta t \rightarrow 0$)
S	$-\beta \frac{SI}{N} \Delta t$	$\frac{dS}{dt} = -\beta \frac{SI}{N}$
I	$\beta \frac{SI}{N} \Delta t - \alpha I \Delta t$	$\frac{dI}{dt} = \beta \frac{SI}{N} - \alpha I$
R	$\alpha I \Delta t$	$\frac{dR}{dt} = \alpha I$



Total population
is conserved:

$$S(t) + I(t) + R(t) = N$$

for all t .

Parameter Ranges – What Is Possible?

What are the possible values of each parameter?

- What is the range of the contact rate c ?



Parameter Ranges – What Is Possible?

What are the possible values of each parameter?

- What is the range of the contact rate c ?

$$\boxed{c \geq 0} \quad (\text{usually } c > 0)$$



Parameter Ranges – What Is Possible?



What are the possible values of each parameter?

- What is the range of the contact rate c ?

$$\boxed{c \geq 0} \quad (\text{usually } c > 0)$$

- What is the range of the infectious period D ?

Parameter Ranges – What Is Possible?



What are the possible values of each parameter?

- What is the range of the contact rate c ?

$$\boxed{c \geq 0} \quad (\text{usually } c > 0)$$

- What is the range of the infectious period D ?

$$\boxed{D > 0}$$

Parameter Ranges – What Is Possible?



What are the possible values of each parameter?

- What is the range of the contact rate c ?

$$\boxed{c \geq 0} \quad (\text{usually } c > 0)$$

- What is the range of the infectious period D ?

$$\boxed{D > 0}$$

- What is the range of the transmission probability p ?

Parameter Ranges – What Is Possible?



What are the possible values of each parameter?

- What is the range of the contact rate c ?

$$\boxed{c \geq 0} \quad (\text{usually } c > 0)$$

- What is the range of the infectious period D ?

$$\boxed{D > 0}$$

- What is the range of the transmission probability p ?

$$\boxed{0 \leq p \leq 1}$$

Parameter Ranges – What Is Possible?



What are the possible values of each parameter?

- What is the range of the contact rate c ?

$$\boxed{c \geq 0} \quad (\text{usually } c > 0)$$

- What is the range of the infectious period D ?

$$\boxed{D > 0}$$

- What is the range of the transmission probability p ?

$$\boxed{0 \leq p \leq 1}$$

- What is the range of the transmission rate $\beta = pc$?

Parameter Ranges – What Is Possible?



What are the possible values of each parameter?

- What is the range of the contact rate c ?

$$c \geq 0 \quad (\text{usually } c > 0)$$

- What is the range of the infectious period D ?

$$D > 0$$

- What is the range of the transmission probability p ?

$$0 \leq p \leq 1$$

- What is the range of the transmission rate $\beta = pc$?

$$\beta \geq 0$$

Parameter Ranges – Continued

- What is the range of the recovery rate $\alpha = 1/D$?



Parameter Ranges – Continued

- What is the range of the recovery rate $\alpha = 1/D$?

$$\alpha > 0$$



6. Predict? – Divide by Δt and Let $\Delta t \rightarrow 0$

The left sides become derivatives *by the definition of the derivative*.
Assumption A4 is what licenses this step.

$$\begin{array}{l} \frac{dS}{dt} = -\frac{\beta}{N}SI \\ \frac{dI}{dt} = \frac{\beta}{N}SI - \alpha I \\ \frac{dR}{dt} = \alpha I \end{array} \quad \begin{array}{l} S(0) = S_0 \\ I(0) = I_0 \\ R(0) = R_0 \end{array}$$

Three coupled first-order ODEs, **nonlinear** because of the product SI .

Brace yourselves

There is **no closed-form solution** for $S(t), I(t), R(t)$ in elementary functions.
Everything useful that follows is obtained *without solving the system*.



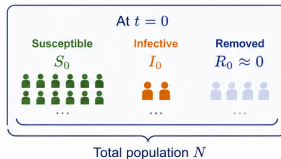
6. Predict?

6.1 The Threshold



The epidemic **grows** at $t = 0$ iff $I'(0) > 0$:

$$\frac{\beta}{N} S_0 I_0 - \alpha I_0 > 0 \iff \frac{\beta S_0}{\alpha N} > 1.$$



Define the **basic reproduction number**

$$\mathcal{R}_0 := \frac{\beta}{\alpha} = \beta D = p c D,$$

the number of people one infective would infect in an **entirely susceptible** population.

With $s_0 = \frac{S_0}{N}$:



epidemic takes off

$\iff$

$$\mathcal{R}_0 s_0 > 1$$

6.1 Why $\mathcal{R}_0 = pcD$ Is the Whole Policy Debate

Read the formula in words:

$$\mathcal{R}_0 = \underbrace{p}_{\text{transmissibility}} \times \underbrace{c}_{\text{sociability}} \times \underbrace{D}_{\text{duration}}$$

- Masks and hygiene attack p .
- School and workplace closure attacks c .
- Testing and isolation attack D .
- Vaccination attacks s_0 .

Herd immunity, for free

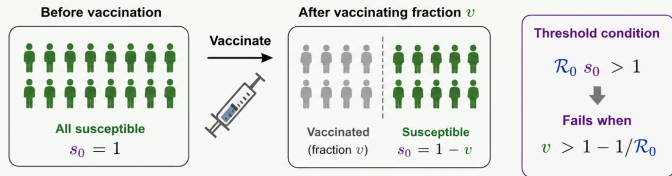
Vaccinating a fraction v leaves $s_0 = 1 - v$, so the threshold fails when $v > 1 - 1/\mathcal{R}_0$.
A policy criterion – and we still have not solved the ODEs.



6. Predict?

Herd immunity, for free

Vaccinating a fraction v leaves $s_0 = 1 - v$, so the threshold fails when $v > 1 - 1/\mathcal{R}_0$.



Herd immunity is achieved when

$$v > 1 - 1/\mathcal{R}_0.$$



A policy criterion –
and we still have not
solved the ODEs.

Parameter Ranges – Continued

- What is the range of the recovery rate $\alpha = 1/D$?



Parameter Ranges – Continued

- What is the range of the recovery rate $\alpha = 1/D$?

$$\alpha > 0$$



Parameter Ranges – Continued

- What is the range of the recovery rate $\alpha = 1/D$?

$$\alpha > 0$$

- What is the range of the dimensionless time $\tau = \alpha t = t/D$?



Parameter Ranges – Continued

- What is the range of the recovery rate $\alpha = 1/D$?

$$\alpha > 0$$

- What is the range of the dimensionless time $\tau = \alpha t = t/D$?

$$\tau \geq 0$$



Parameter Ranges – Continued

- What is the range of the recovery rate $\alpha = 1/D$?

$$\alpha > 0$$

- What is the range of the dimensionless time $\tau = \alpha t = t/D$?

$$\tau \geq 0$$

- What is the range of the basic reproduction number $\mathcal{R}_0 = \beta/\alpha$?



Parameter Ranges – Continued



- What is the range of the recovery rate $\alpha = 1/D$?

$$\alpha > 0$$

- What is the range of the dimensionless time $\tau = \alpha t = t/D$?

$$\tau \geq 0$$

- What is the range of the basic reproduction number $\mathcal{R}_0 = \beta/\alpha$?

$$\mathcal{R}_0 \geq 0$$

Parameter Ranges – Continued



- What is the range of the recovery rate $\alpha = 1/D$?

$$\alpha > 0$$

- What is the range of the dimensionless time $\tau = \alpha t = t/D$?

$$\tau \geq 0$$

- What is the range of the basic reproduction number $\mathcal{R}_0 = \beta/\alpha$?

$$\mathcal{R}_0 \geq 0$$

- What is the range of the initial susceptible fraction $s_0 = S_0/N$?

Parameter Ranges – Continued



- What is the range of the recovery rate $\alpha = 1/D$?

$$\alpha > 0$$

- What is the range of the dimensionless time $\tau = \alpha t = t/D$?

$$\tau \geq 0$$

- What is the range of the basic reproduction number $\mathcal{R}_0 = \beta/\alpha$?

$$\mathcal{R}_0 \geq 0$$

- What is the range of the initial susceptible fraction $s_0 = S_0/N$?

$$0 \leq s_0 \leq 1$$

Parameter Ranges – Continued



- What is the range of the recovery rate $\alpha = 1/D$?

$$\alpha > 0$$

- What is the range of the dimensionless time $\tau = \alpha t = t/D$?

$$\tau \geq 0$$

- What is the range of the basic reproduction number $\mathcal{R}_0 = \beta/\alpha$?

$$\mathcal{R}_0 \geq 0$$

- What is the range of the initial susceptible fraction $s_0 = S_0/N$?

$$0 \leq s_0 \leq 1$$

- What is the range of the vaccinated fraction v ?

Parameter Ranges – Continued



- What is the range of the recovery rate $\alpha = 1/D$?

$$\alpha > 0$$

- What is the range of the dimensionless time $\tau = \alpha t = t/D$?

$$\tau \geq 0$$

- What is the range of the basic reproduction number $\mathcal{R}_0 = \beta/\alpha$?

$$\mathcal{R}_0 \geq 0$$

- What is the range of the initial susceptible fraction $s_0 = S_0/N$?

$$0 \leq s_0 \leq 1$$

- What is the range of the vaccinated fraction v ?

$$0 \leq v \leq 1$$

6.2 Dimensionless Form

Put $s = S/N$, $i = I/N$, $r = R/N$ and rescale time by $\tau = \alpha t$:

$$\frac{ds}{d\tau} = -\mathcal{R}_0 s i, \quad \frac{di}{d\tau} = \mathcal{R}_0 s i - i, \quad \frac{dr}{d\tau} = i.$$

What nondimensionalization bought

Four quantities (β , α , N , and the time unit) have collapsed to **one**: $\mathcal{R}_0$.

Every SIR outbreak in the world is, up to rescaling of time and population, a member of a **one-parameter family**.



Parameter Ranges



Quantity	Meaning	Range	Units
c	contact rate	$c \geq 0$	contacts/time
D	mean infectious period	$D > 0$	time
p	probability of transmission per contact	$0 \leq p \leq 1$	dimensionless
$\beta = pc$	transmission rate	$\beta \geq 0$	1/time
$\alpha = 1/D$	recovery/removal rate	$\alpha > 0$	1/time
$\tau = \alpha t = t/D$	dimensionless time	$\tau \geq 0$	dimensionless
$\mathcal{R}_0 = \beta/\alpha = pcD$	basic reproduction number	$\mathcal{R}_0 \geq 0$	dimensionless
$s_0 = S_0/N$	initial susceptible fraction	$0 \leq s_0 \leq 1$	dimensionless
v	vaccinated fraction	$0 \leq v \leq 1$	dimensionless

6.3 A First Integral – Without Solving Anything

Divide the i equation by the s equation to eliminate τ :

$$\frac{di}{ds} = \frac{\mathcal{R}_0 si - i}{-\mathcal{R}_0 si} = -1 + \frac{1}{\mathcal{R}_0 s}.$$

Integrate:

$$i(s) = i_0 + (s_0 - s) + \frac{1}{\mathcal{R}_0} \ln \frac{s}{s_0}$$

Note what just happened

A system with *no* closed-form solution has a closed-form **orbit** in the (s, i) plane. Give up on $s(t)$ and $i(t)$; keep the geometry.



6.3 Peak Prevalence

Set $di/ds = 0$: the epidemic peaks *exactly* when

$$s = \frac{1}{\mathcal{R}_0}.$$

The same threshold, now read as a turning point: the peak occurs when the susceptible fraction falls through $1/\mathcal{R}_0$, **not** when infectives run out.

Substituting into the orbit:

$$i_{\max} = i_0 + s_0 - \frac{1}{\mathcal{R}_0} - \frac{1}{\mathcal{R}_0} \ln(\mathcal{R}_0 s_0).$$

Interpretation

The peak is reached while the epidemic still has plenty of fuel. Infections continue *after* the peak – roughly as many again. This is why “cases are falling” is not “it is over”.



6. Predict?



6.3 Peak Prevalence

Set $\frac{di}{ds} = 0$: the epidemic peaks exactly when

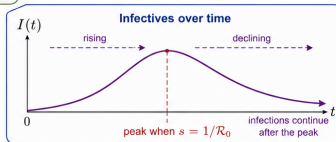
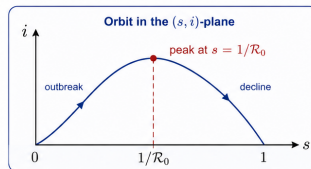
$$s = \frac{1}{\mathcal{R}_0}.$$



The same threshold, now read as a **turning point**: the peak occurs when the susceptible fraction falls through $1/\mathcal{R}_0$, **not** when infectives run out.

Substituting into the orbit:

$$i_{\max} = i_0 + s_0 - \frac{1}{\mathcal{R}_0} - \frac{1}{\mathcal{R}_0} \ln(\mathcal{R}_0 s_0).$$



Interpretation

- The peak is reached while the epidemic still has **plenty of fuel**.
- Infections continue **after** the peak — roughly **as many** again.



This is why
“cases are falling”
is **not** “it is over”.

6.3 Final Size – The Model's Best Prediction

As $\tau \rightarrow \infty$: s decreases and is bounded below, so $s \rightarrow s_\infty$; then $ds/d\tau \rightarrow 0$ forces $i \rightarrow 0$. Setting $i = 0$ with $s_0 + i_0 = 1$:

$$s_\infty - \frac{1}{\mathcal{R}_0} \ln \frac{s_\infty}{s_0} = 1$$

an implicit equation for the fraction *never* infected.

The consequence nobody guesses

$s_\infty > 0$ **always**, for every $\mathcal{R}_0$.

The epidemic ends because infectives run out, not because susceptibles do.

Some people escape entirely without ever being immune.

No amount of intuition supplies this. The model gives it for free.



What Are the Real-World Limits?



Quantity	Real-world maximum?	Why?
c – contacts/time	Finite	A person has only 24 hours/day and a finite number of interactions.
D – infectious period	Finite	Infection eventually clears, changes phase, or the person dies; no universal maximum across diseases.
p – transmission probability/contact	1	Probability cannot exceed 100%.
$\beta = pc$	Finite for a given population/context	Since $p \leq 1$ and real contact rates are finite.
$\alpha = 1/D$	Finite for a given disease	An infectious period cannot realistically be arbitrarily close to zero.
τ – time	No finite maximum	Time can continue indefinitely.
$\mathcal{R}_0 = pcD$	Finite for a particular disease/population	Determined by actual contact, transmission and duration.
$s_0 = S_0/N$	1	At most everyone is susceptible.
v – vaccinated fraction	1	At most 100% of the population can be vaccinated.

7. Valid? – Internal Tests, Before Any Data



Tests against the model's *own* principles and assumptions.

1. **Dimensions.** $[\beta SI/N] = \text{time}^{-1} \cdot \text{people} = \text{people/time} = [dS/dt]$. ✓
And $\mathcal{R}_0 = \beta/\alpha$ is dimensionless, as “people infected per infective” must be. ✓
2. **Conservation.** Adding the three equations: $S' + I' + R' = 0$, so $S + I + R \equiv N$.
Assumption A1 is *recovered as a theorem*, not imposed. ✓
3. **Monotonicity.** $S' \leq 0$ and $R' \geq 0$ always. Consistent with A7. ✓
4. **Positivity.** $S' = 0$ when $S = 0$; $I' = 0$ when $I = 0$. The boundaries cannot be crossed, so $S, I, R \in [0, N]$. A negative number of people is impossible. ✓

What Does $\mathcal{R}_0$ Look Like in Real Epidemics?



Disease	D (days)	$\alpha = 1/D$ (/day)	β (/day)	$\mathcal{R}_0$
Seasonal influenza	~ 2	0.50	0.65	~ 1.3
Ebola	~ 7	0.14	0.36	~ 2.5
COVID-19 (ancestral)	~ 14	0.071	0.21	~ 3.0
SARS-CoV-1	~ 9	0.11	0.39	~ 3.5
Smallpox	~ 7	0.14	0.64	~ 4.5
Rubella	~ 7	0.14	0.93	~ 6.5
Polio	—	—	—	~ 6
Chickenpox	~ 5	0.20	2.2	~ 11
Measles	~ 5	0.20	3.4	~ 17

Values are representative estimates, not universal constants.

Example 1 – Given β and D

A disease has transmission rate

$$\beta = 0.30 \text{ day}^{-1}$$

and mean infectious period

$$D = 5 \text{ days.}$$



Example 1 – Given β and D

A disease has transmission rate

$$\beta = 0.30 \text{ day}^{-1}$$

and mean infectious period

$$D = 5 \text{ days.}$$

Calculate $\mathcal{R}_0$

Recall

$$\mathcal{R}_0 = \beta D.$$

Interpretation

One infective produces, on average, **1.5 secondary infections** in a fully susceptible population.



Example 1 – Given β and D

A disease has transmission rate

$$\beta = 0.30 \text{ day}^{-1}$$

and mean infectious period

$$D = 5 \text{ days.}$$

Calculate $\mathcal{R}_0$

Recall

$$\mathcal{R}_0 = \beta D.$$

Therefore,

$$\mathcal{R}_0 = (0.30)(5) = \boxed{1.5}.$$



Example 1 – Given β and D

A disease has transmission rate

$$\beta = 0.30 \text{ day}^{-1}$$

and mean infectious period

$$D = 5 \text{ days.}$$

Calculate $\mathcal{R}_0$

Recall

$$\mathcal{R}_0 = \beta D.$$

Therefore,

$$\mathcal{R}_0 = (0.30)(5) = \boxed{1.5}.$$

Interpretation

One infective produces, on average, **1.5 secondary infections** in a fully susceptible population.

Example 2 – Given p , c , and D

An infective makes

$$c = 10 \quad \text{contacts/day}$$

and the probability of transmission per contact is

$$p = 0.04.$$



Example 2 – Given p , c , and D

An infective makes

$$c = 10 \text{ contacts/day}$$

and the probability of transmission per contact is

$$p = 0.04.$$

The mean infectious period is

$$D = 7 \text{ days.}$$



Example 2 – Given p , c , and D

An infective makes

$$c = 10 \text{ contacts/day}$$

and the probability of transmission per contact is

$$p = 0.04.$$

The mean infectious period is

$$D = 7 \text{ days.}$$

Step 1: Find β

$$\beta = pc = (0.04)(10) = \boxed{0.4 \text{ day}^{-1}}.$$



Example 2 – Given p , c , and D

An infective makes

$$c = 10 \text{ contacts/day}$$

and the probability of transmission per contact is

$$p = 0.04.$$

The mean infectious period is

$$D = 7 \text{ days.}$$

Step 1: Find β

$$\beta = pc = (0.04)(10) = \boxed{0.4 \text{ day}^{-1}}.$$

Step 2: Find $\mathcal{R}_0$

$$\mathcal{R}_0 = pcD = (0.04)(10)(7) = \boxed{2.8}.$$



Example 3 – Vaccination Changes the Threshold

Suppose a disease has

$$\mathcal{R}_0 = 4$$

and 30% of the population is vaccinated:

$$v = 0.30.$$



Example 3 – Vaccination Changes the Threshold

Suppose a disease has

$$\mathcal{R}_0 = 4$$

and 30% of the population is vaccinated:

$$v = 0.30.$$

Step 1: Find the susceptible fraction

$$s_0 = 1 - v = 1 - 0.30 = \boxed{0.70}.$$



Example 3 – Vaccination Changes the Threshold

Suppose a disease has

$$\mathcal{R}_0 = 4$$

and 30% of the population is vaccinated:

$$v = 0.30.$$

Step 1: Find the susceptible fraction

$$s_0 = 1 - v = 1 - 0.30 = \boxed{0.70}.$$

Step 2: Check the epidemic threshold

$$\mathcal{R}_0 s_0 = 4(0.70) = \boxed{2.8 > 1}.$$

Therefore, the epidemic **takes off**.



Example 3 – Vaccination Changes the Threshold

Suppose a disease has

$$\mathcal{R}_0 = 4$$

and 30% of the population is vaccinated:

$$v = 0.30.$$

Step 1: Find the susceptible fraction

$$s_0 = 1 - v = 1 - 0.30 = \boxed{0.70}.$$

Step 2: Check the epidemic threshold

$$\mathcal{R}_0 s_0 = 4(0.70) = \boxed{2.8 > 1}.$$

Therefore, the epidemic **takes off**.



Example 3 – Vaccination Changes the Threshold



How much vaccination is enough?

$$v_c = 1 - \frac{1}{\mathcal{R}_0} = 1 - \frac{1}{4} = \boxed{0.75}.$$

Thus, under this SIR model, **75% vaccination** is required to prevent initial epidemic growth.

7. Valid? – The Two Best Checks



Degenerate limit $\alpha = 0$ (nobody ever recovers)

Then $R \equiv R_0$ and $I = N - S$, so

$$\frac{dI}{dt} = \frac{\beta}{N} I(N - I)$$

– exactly **logistic growth**. The SIR model *contains* the logistic model as a special case. This is the strongest single sign that the infection term was built correctly. ✓

(And $\beta = 0$ gives $I' = -\alpha I$: pure exponential decay. ✓)

Well-posedness

The right-hand side is polynomial, hence Lipschitz on the bounded set $[0, N]^3$.

Picard–Lindelöf: a unique solution, depending *continuously* on S_0, I_0, β, α .

Continuous dependence is what makes the model **usable**, not merely solvable.

8. Verified? — Is It Good, For the Purpose of Step 1?



Classic test: influenza in an English boys' boarding school, 1978

One boy returned infected to a school of $N = 763$. Daily counts of boys confined to bed were recorded for a fortnight. Least-squares fitting gives roughly

$$\beta \approx 1.7 \text{ day}^{-1}, \quad \alpha \approx 0.45 \text{ day}^{-1}, \quad \mathcal{R}_0 \approx 3.8.$$

Timing, peak height and final size are all reproduced well.

Why here? The setting is nearly ideal: a *closed, well-mixed*, single-age-group population over a *short* interval. A2 and A3 hold about as well as they ever will.

8. Verified? – Where It Fails, and Which Assumption Is To Blame



Applied to	Failure	Culprit
a whole city	over-predicts the peak; real networks are clustered, so local susceptibles are exhausted first	A3
the first ten cases	mis-predicts; chance extinction is common even when $\mathcal{R}_0 > 1$	A5
a full year	curve turns over earlier than predicted; people change behaviour	A8
mortality questions	cannot answer at all	A10

Verdict against the stated purpose

Officials asked: epidemic or not, how big, how many. The model answers all three with **two** parameters, and $\mathcal{R}_0 = pcD$ tells them which lever to pull.

9. Improve? – Which Assumption To Lift First



Lift	Gives	Cost
A6 no latent period	SEIR: corrects timing of the peak	one extra compartment
A5 determinism	stochastic SIR: extinction probability at low I	needs probability
A3 homogeneous mixing	network / age-structured SIR: corrects peak height	many parameters, identifiability problems
A7 permanent immunity	SIRS: endemic state, not a single wave	changes long-run behaviour qualitatively
A9 exponential period	fixed or gamma-distributed period	delay or integro-differential
A2 closed population	recurring epidemics, damped oscillation to endemicity	a new timescale enters
A10 removal complete	separate D compartment, case fatality ratio	trivial, and often what is actually wanted
A8 constant parameters	$\beta(t)$: interventions, seasonality	nonautonomous; <i>destroys</i> the first integral

9. Improve? – And a Non-Uniqueness Warning

Some authors write the infection term as

$$\tilde{\beta} S I \quad (\text{mass action})$$

instead of

$$\frac{\beta}{N} S I \quad (\text{frequency dependent}).$$

These are not the same model

For a *fixed* N they coincide with $\tilde{\beta} = \beta/N$. But they make **different** predictions about how epidemics scale with population size – and no epidemic curve from a single town can distinguish them.

This is the **non-uniqueness** of models, in the wild. The data underdetermine the model. You must choose, and then declare the choice.



10. Use? – How To Exercise the Model



- Solve numerically (RK4, adaptive step) for (s, i, r) vs τ at $\mathcal{R}_0 = 0.9, 1.5, 3, 6$. The threshold appears as a *qualitative* change, not a quantitative one.
- Plot the orbits $i(s)$ with the vertical line $s = 1/\mathcal{R}_0$ marked. Every trajectory peaks on that line.
- **Design curve:** s_∞ vs $\mathcal{R}_0$ from the final-size relation.
- **Policy curve:** $v = 1 - 1/\mathcal{R}_0$.
- **What-if:** impose $\beta \rightarrow \beta/2$ on a chosen day. Show that a *late* intervention barely changes the final size, while an *early* one halves it.

Note the loop: the last item sends you back to *Why*?

How Do Epidemiologists Really Estimate $\mathcal{R}_0$?

$\mathcal{R}_0$ is usually inferred, not directly measured.

Our SIR model

$$\mathcal{R}_0 = \frac{\beta}{\alpha} = \beta D = pcD.$$



How Do Epidemiologists Really Estimate $\mathcal{R}_0$?

$\mathcal{R}_0$ is usually inferred, not directly measured.

Our SIR model

$$\mathcal{R}_0 = \frac{\beta}{\alpha} = \beta D = pcD.$$

But in a real epidemic, we usually do **not** know p , c , and D precisely.



How Do Epidemiologists Really Estimate $\mathcal{R}_0$?

$\mathcal{R}_0$ is usually inferred, not directly measured.

Our SIR model

$$\mathcal{R}_0 = \frac{\beta}{\alpha} = \beta D = pcD.$$

But in a real epidemic, we usually do **not** know p , c , and D precisely.

Observed epidemic data	→	statistical inference	→	$\mathcal{R}_0$
------------------------	---	-----------------------	---	-----------------



How Do Epidemiologists Really Estimate $\mathcal{R}_0$?

$\mathcal{R}_0$ is usually inferred, not directly measured.



Our SIR model

$$\mathcal{R}_0 = \frac{\beta}{\alpha} = \beta D = pcD.$$

But in a real epidemic, we usually do **not** know p , c , and D precisely.

Observed epidemic data $\longrightarrow$ statistical inference $\longrightarrow \mathcal{R}_0$

Key idea

We observe the **consequences of transmission** and infer the parameters that could have produced them.

Step 1 – Measure How Fast the Epidemic Grows

At the beginning of an epidemic, almost everyone is susceptible:

$$S(t) \approx N.$$



Step 1 – Measure How Fast the Epidemic Grows

At the beginning of an epidemic, almost everyone is susceptible:

$$S(t) \approx N.$$

The number of infectives therefore grows approximately exponentially:

$$I(t) \approx I_0 e^{rt}.$$



Step 1 – Measure How Fast the Epidemic Grows

At the beginning of an epidemic, almost everyone is susceptible:

$$S(t) \approx N.$$

The number of infectives therefore grows approximately exponentially:

$$I(t) \approx I_0 e^{rt}.$$

From the early epidemic curve, epidemiologists estimate the **exponential growth rate** r .



Step 1 – Measure How Fast the Epidemic Grows

At the beginning of an epidemic, almost everyone is susceptible:

$$S(t) \approx N.$$

The number of infectives therefore grows approximately exponentially:

$$I(t) \approx I_0 e^{rt}.$$

From the early epidemic curve, epidemiologists estimate the **exponential growth rate** r .

For the simple SIR model

$$I' = (\beta - \alpha) I$$

so that

$$r = \beta - \alpha.$$



Step 1 – Measure How Fast the Epidemic Grows

At the beginning of an epidemic, almost everyone is susceptible:

$$S(t) \approx N.$$

The number of infectives therefore grows approximately exponentially:

$$I(t) \approx I_0 e^{rt}.$$

From the early epidemic curve, epidemiologists estimate the **exponential growth rate** r .

For the simple SIR model

$$I' = (\beta - \alpha) I$$

so that

$$r = \beta - \alpha.$$



Step 1 – Measure How Fast the Epidemic Grows

Since $\alpha = 1/D$,

$$\mathcal{R}_0 = \frac{\beta}{\alpha} = 1 + rD$$



Step 2 – The Generation Interval Matters

Real epidemics are more complicated than the simple SIR model.



Step 2 – The Generation Interval Matters

Real epidemics are more complicated than the simple SIR model.

Let $g(\tau)$ be the **generation-interval distribution**: the probability that transmission occurs τ time units after infection.



Step 2 – The Generation Interval Matters

Real epidemics are more complicated than the simple SIR model.
Let $g(\tau)$ be the **generation-interval distribution**: the probability that transmission occurs τ time units after infection.
If the epidemic grows at rate r , then

$$1 = \mathcal{R}_0 \int_0^{\infty} e^{-r\tau} g(\tau) d\tau \implies \mathcal{R}_0 = \frac{1}{\int_0^{\infty} e^{-r\tau} g(\tau) d\tau}$$



Step 2 – The Generation Interval Matters

Real epidemics are more complicated than the simple SIR model. Let $g(\tau)$ be the **generation-interval distribution**: the probability that transmission occurs τ time units after infection. If the epidemic grows at rate r , then

$$1 = \mathcal{R}_0 \int_0^{\infty} e^{-r\tau} g(\tau) d\tau \implies \mathcal{R}_0 = \frac{1}{\int_0^{\infty} e^{-r\tau} g(\tau) d\tau}$$

Important

Knowing only the **average infectious period** is not always enough. The **timing distribution of transmission** matters.



Step 3 – Fit a Transmission Model

Epidemiologists can combine several sources of information:



Step 3 – Fit a Transmission Model

Epidemiologists can combine several sources of information:



Models may estimate

β , D , reporting rate, latent period, age effects, ...



Step 3 – Fit a Transmission Model

Epidemiologists can combine several sources of information:



Models may estimate

β , D , reporting rate, latent period, age effects, ...

How?

Parameters are estimated using methods such as **maximum likelihood** or **Bayesian inference**.



Step 3 – Fit a Transmission Model

Epidemiologists can combine several sources of information:



Models may estimate

β , D , reporting rate, latent period, age effects, ...

How?

Parameters are estimated using methods such as **maximum likelihood** or **Bayesian inference**.



Step 3 – Fit a Transmission Model

The important distinction

p and c are usually not separately identifiable

$$\beta = pc$$

Independent contact data are needed to separate them.



Example — Estimating $\mathcal{R}_0$ by Bayesian Inference

Suppose an epidemic is growing rapidly.

From the first few weeks of case data, we estimate the exponential growth rate:

$$r \approx 0.20 \text{ day}^{-1}.$$



Example — Estimating $\mathcal{R}_0$ by Bayesian Inference

Suppose an epidemic is growing rapidly.

From the first few weeks of case data, we estimate the exponential growth rate:

$$r \approx 0.20 \text{ day}^{-1}.$$

But the generation interval is uncertain. Suppose previous epidemiological studies suggest

$$\text{generation interval} \sim \mathcal{N}(5, 1^2) \text{ days.}$$



Example — Estimating $\mathcal{R}_0$ by Bayesian Inference

Suppose an epidemic is growing rapidly.

From the first few weeks of case data, we estimate the exponential growth rate:

$$r \approx 0.20 \text{ day}^{-1}.$$

But the generation interval is uncertain. Suppose previous epidemiological studies suggest

$$\text{generation interval} \sim \mathcal{N}(5, 1^2) \text{ days.}$$

Instead of choosing exactly 5 days, we treat it as an **uncertain parameter**.



Example – Estimating $\mathcal{R}_0$ by Bayesian Inference

Bayesian idea

Prior information + observed epidemic data $\longrightarrow$ Posterior distribution



Example – Estimating $\mathcal{R}_0$ by Bayesian Inference



Bayesian idea

Prior information + observed epidemic data $\longrightarrow$ Posterior distribution

We therefore obtain not just one value of $\mathcal{R}_0$, but a **probability distribution**:

$$P(\mathcal{R}_0 \mid \text{data})$$

What Does the Bayesian Result Look Like?

Bayesian inference might produce a posterior such as

$$\mathcal{R}_0 = 2.8 \quad (95\% \text{ credible interval : } 2.2 - 3.5)$$



What Does the Bayesian Result Look Like?

Bayesian inference might produce a posterior such as

$$\mathcal{R}_0 = 2.8 \quad (95\% \text{ credible interval : } 2.2 - 3.5)$$

Interpretation:

- 2.8 is the posterior central estimate.



What Does the Bayesian Result Look Like?

Bayesian inference might produce a posterior such as

$$\mathcal{R}_0 = 2.8 \quad (95\% \text{ credible interval : } 2.2 - 3.5)$$

Interpretation:

- 2.8 is the posterior central estimate.
- There is uncertainty because the epidemic data are limited and the generation interval is uncertain.



What Does the Bayesian Result Look Like?

Bayesian inference might produce a posterior such as

$$\mathcal{R}_0 = 2.8 \quad (95\% \text{ credible interval : } 2.2 - 3.5)$$

Interpretation:

- 2.8 is the posterior central estimate.
- There is uncertainty because the epidemic data are limited and the generation interval is uncertain.
- The interval $[2.2, 3.5]$ represents the Bayesian **95% credible interval**.



What Does the Bayesian Result Look Like?

Bayesian inference might produce a posterior such as

$$\mathcal{R}_0 = 2.8 \quad (95\% \text{ credible interval : } 2.2 - 3.5)$$

Interpretation:

- 2.8 is the posterior central estimate.
- There is uncertainty because the epidemic data are limited and the generation interval is uncertain.
- The interval $[2.2, 3.5]$ represents the Bayesian **95% credible interval**.

This is the key difference

We did not say

$$\mathcal{R}_0 = 2.8 \quad \text{exactly.}$$

Instead, Bayesian inference says:

Given what we know, some values of $\mathcal{R}_0$ are more plausible than others.



The Bayesian Calculation

Let θ denote all unknown parameters, including $\mathcal{R}_0$.
Bayes' theorem gives

$$P(\theta \mid \text{data}) \propto P(\text{data} \mid \theta) P(\theta)$$



The Bayesian Calculation

Let θ denote all unknown parameters, including $\mathcal{R}_0$.
Bayes' theorem gives

$$P(\theta \mid \text{data}) \propto P(\text{data} \mid \theta) P(\theta)$$

$$\underbrace{P(\theta)}_{\text{Prior}} \times \underbrace{P(\text{data} \mid \theta)}_{\text{Likelihood}} \longrightarrow \underbrace{P(\theta \mid \text{data})}_{\text{Posterior}}$$



The Bayesian Calculation

Let θ denote all unknown parameters, including $\mathcal{R}_0$.
Bayes' theorem gives

$$P(\theta \mid \text{data}) \propto P(\text{data} \mid \theta) P(\theta)$$

$$\underbrace{P(\theta)}_{\text{Prior}} \times \underbrace{P(\text{data} \mid \theta)}_{\text{Likelihood}} \longrightarrow \underbrace{P(\theta \mid \text{data})}_{\text{Posterior}}$$

For an epidemic:

Prior knowledge + case data + generation interval + model



The Bayesian Calculation

Let θ denote all unknown parameters, including $\mathcal{R}_0$.
Bayes' theorem gives

$$P(\theta \mid \text{data}) \propto P(\text{data} \mid \theta) P(\theta)$$

$$\underbrace{P(\theta)}_{\text{Prior}} \times \underbrace{P(\text{data} \mid \theta)}_{\text{Likelihood}} \longrightarrow \underbrace{P(\theta \mid \text{data})}_{\text{Posterior}}$$

For an epidemic:

Prior knowledge + case data + generation interval + model



Posterior distribution for $\mathcal{R}_0$





Lecture B: Velten's SQM Framework

From Procedure to Coordinates

Lecture A produced a model. Lecture B asks: *what kind of object is it?*



Definition 1 (Mathematical model, Velten Def. 1.4.1)

A **mathematical model** is a triple (S, Q, M) consisting of

- a **system** S ,
- a **question** Q relating to S ,
- a set of **mathematical statements** $M = \{\Sigma_1, \dots, \Sigma_n\}$ which can be used to answer Q .

The point most students miss

Q is a *constituent* of the model. Change the question and you have a different model — even if you have not touched a single equation.

The System S



S for our model

The population of N individuals of the town during this outbreak, together with the pathogen circulating in it.

- **Boundary:** in the town limit. A2 declares it *closed* — nothing crosses.
- **Explicitly outside in:** individual immune systems, the health-care system, neighbouring towns, the economy.

Drawing the boundary *is* the modelling

Every item on the “outside” list is a decision you must be able to defend. This is exactly the reduction of an open problem to a closed problem.

The Question Q



Q for our model

Given the initial numbers of susceptible and infective people, and given the transmission and recovery rates, how do the aggregate compartment sizes evolve in time – and in particular, does the outbreak grow, how large is its peak, and what fraction of the population is ultimately infected?

Note what Q does **not** ask: *who* gets infected, *where*, or *when* any named individual does.

Same S , different Q , different M

Had Q asked those things, S would be unchanged – but M would have to be an agent-based or network model. Not a refinement. A different model.

The Mathematical Statements M



$$M = \left\{ \begin{array}{l} \Sigma_1 : S' = -\frac{\beta}{N}SI, \quad \Sigma_2 : I' = \frac{\beta}{N}SI - \alpha I, \quad \Sigma_3 : R' = \alpha I, \\ \Sigma_4 : S(0) = S_0, I(0) = I_0, R(0) = R_0, \\ \Sigma_5 : S + I + R = N, \quad \Sigma_6 : \mathcal{R}_0 = \beta/\alpha \end{array} \right\}$$

M is a *set of statements*, not “the equation”. The initial conditions and the invariant are members of M on equal footing with the ODEs – drop Σ_4 and M no longer answers Q .

State Variables and System Parameters (Velten §1.5.1)



State variables	S, I, R – what Q asks about
System parameters	β, α, N – constant with respect to Q

The split is a function of Q , not of S

Change the question to “*what was the transmission rate during this outbreak?*” and:

- β becomes the **unknown**,
- the case counts become **given data**.

The same equations then constitute a **different model** – and, as we will see on the Q axis, one of a different *class*.

SQM Space: the S Axis (Velten §1.7.2)



Criterion	Our model	Justification
Physical – conceptual	Physical	real people, part of the external world
Natural – technical	Natural	biological; nothing engineered
Stochastic – deterministic	system stochastic ; model deterministic	A5. The mismatch is the model's chief weakness at small I
Continuous – discrete	system discrete ; model continuous	A4. Velten: a discrete system may legitimately be described by continuous statements, and vice versa
Dimension	0D	no space variables. A spatial epidemic model would be 2D
Field	biological / medical	epidemiology

Two rows are *mismatches* between S and M . Say so out loud; do not hide them.

SQM Space: the Q Axis (Velten §1.7.3) – 1/2



Criterion	Our model	Justification
Phenomenological – mechanistic	Mechanistic	the infection term was <i>derived</i> from a contact hypothesis (Step 5.1), not fitted. Hence β, α are <i>interpretable</i>
Stationary – instationary	Instationary	Q asks for time evolution
Lumped – distributed	Lumped	no state variable or parameter depends on a space variable
Scale	Macroscopic	population level. The microscopic alternative is agent-based. The scale follows from Q

Consequence of “mechanistic”

The model is a **hypothesis** about internal mechanics. Data can *refute* it; data can never *prove* it.

SQM Space: the Q Axis – 2/2: Direct, Inverse, Control

One system, one set of equations, three different models:



If Q asks for...	the problem is...	our instance
the output, given input and parameters	direct	Lecture A, Step 6
parameters of S	inverse (parameter identification)	fitting β, α to the 1978 record
the input	control	choose v so that $\mathcal{R}_0(1 - v) < 1$

Research – management

Used to understand the threshold structure: a **research** model.

Used to decide a school closure: a **management** model.

The equations are identical; the class is not. Compare Velten on Lotka–Volterra: sound as research, too crude to manage with. Here the honest position is the same – SIR is excellent for orders of magnitude, unreliable for operational forecasts.

SQM Space: the M Axis (Velten §1.7.4)



Criterion	Our model	Justification
Linear – nonlinear	Nonlinear	the product SI multiplies two unknowns. <i>The entire mathematical interest lives here:</i> threshold, peak and final size all come from that one nonlinearity
Analytical – numerical	Mixed	no closed form for $S(t), I(t), R(t)$, so trajectories are numerical; but $i(s)$, the threshold, the peak and the final-size relation are analytical – parameter effects studied <i>without</i> choosing numbers
Autonomous – nonautonomous	Autonomous	t does not appear explicitly (A8). $\beta(t)$ would make it nonautonomous and destroy the first integral
Continuous – discrete	Continuous	t ranges over an interval

M Axis: Which Mathematical Structure?



Our *M*: a system of three coupled first-order ODEs, as an initial value problem

Not an algebraic equation. Not a difference equation. Not a PDE. Not an integral equation.

Change one assumption and you move to a different point on the *M* axis:

Lift	New structure on the <i>M</i> axis
fixed infectious period (A9)	delay differential equation
gamma-distributed period (A9)	integro-differential equation
spatial spread (0D $\rightarrow$ 2D)	reaction-diffusion PDE
discrete generations (A4)	difference equation
lattice with local rules (A3, A4)	cellular automaton

The same epidemic, four different points in SQM space.

M Axis: Well-Posedness



Theorem 2 (Consequence of Picard–Lindelöf)

The right-hand side of the SIR system is polynomial in (S, I, R) , hence locally Lipschitz on the bounded region $[0, N]^3$. Therefore the initial value problem has a **unique** solution, and that solution depends **continuously** on $S_0, I_0, R_0, \beta, \alpha$.

Why a modeller cares

Existence says an answer is there. Uniqueness says it is *the* answer.

Continuous dependence says a small error in the measured β produces only a small error in the prediction – and that is the property that makes the model *usable* rather than merely solvable.

Contrast Week 14: an inverse problem generally loses continuous dependence, and must be regularized.

Where the Model Sits: Black Box to White Box (Velten §1.7.1)



	social (black)	biological (grey)	chemical	mechanical (white)
M	regression	ODE, effective quantities	ODE/PDE	PDE from first principles

SIR sits toward the **white** end *for a biological system* – but well short of mechanics.

- Not a black box: the transmission mechanism is modelled explicitly from contacts.
- Not white either: β aggregates sociology, virology and hygiene into one number that *cannot* be predicted from first principles and must be fitted.

Velten's own term applies

A **grey-box model using effective quantities**. β is an effective transmission rate exactly as an effective permeability is in a porous-medium model.

The Two Frameworks, Side by Side



Dym: ten questions

A procedure for building

Sequential, and looping

Answers “what do I do next?”

Assume? is the load-bearing step

Valid? and *Verified?* kept distinct

Velten: (S, Q, M) + SQM space

A coordinate system for locating

Static classification

Answers “what have I got?”

The $S-M$ mismatches are the
load-bearing observation

Model class predicts behaviour
before solving

Use them together

Dym’s *Assume?* list determines your coordinates on all three Velten axes: A4 puts you at “continuous”, A5 at “deterministic”, A3 at “lumped, macroscopic”, A8 at “autonomous”.

Every assumption is a coordinate.

Closing Warning

Three separate things

The SIR model *fits* the 1978 school data well.

That establishes neither that the model is *true*, nor that it will fit the next outbreak.

Fitting, truth, and usefulness are three different things.



The assumption list — not the fit — is where a model's honesty lives.

And that is why, in this course, an unstated assumption costs you more than a wrong number.

Exercise 1: SIR — Build It Then Break It

1. Derive the SEIR model from scratch by lifting A6. State the new compartment's balance equation and recompute $\mathcal{R}_0$. Does the threshold change? Does the *final size* change?
2. Show that the final-size relation $s_\infty - \mathcal{R}_0^{-1} \ln(s_\infty/s_0) = 1$ has exactly one root in $(0, 1/\mathcal{R}_0)$ when $\mathcal{R}_0 s_0 > 1$. Deduce $s_\infty > 0$.
3. Two towns have populations N and $2N$ with identical α . Compare the predicted final sizes under the mass-action and frequency-dependent transmission forms. What data would settle which is right?
4. Classify the *stochastic* SIR model on all three SQM axes. Which entries change from today's table, and which do not?
5. A colleague fits SIR to city data, obtains $R^2 = 0.98$, and concludes the mixing is homogeneous. In three sentences, explain why that does not follow.

References for This Week

Next Model

Lotka–Volterra Prey-Predator Model, and give the picture behind today's algebra.

