Introduction to modelling

Stéphanie Portet

Department of Mathematics
University of Manitoba
Winnipeg, MB, Canada
Stephanie.Portet@umanitoba.ca

March 2023

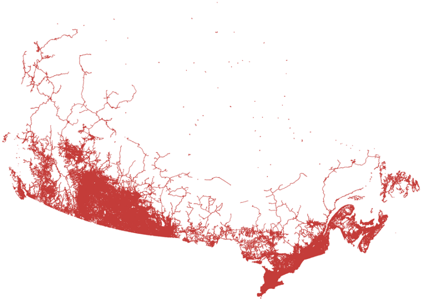
Outline

Question

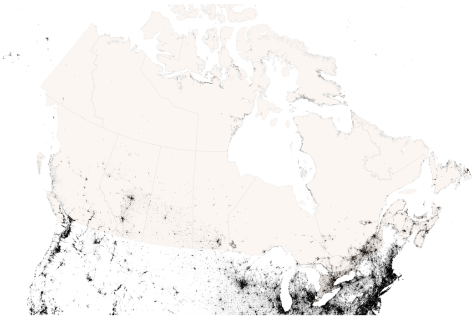
What do you need to cross Canada by bike ?

- a bike
-

Maps



Canada represented by its population density



Canada represented by its road system

Mathematical model

- A model is a map
- A model is a simplification of the reality (a map is not the territory it describes)

A model is an idealization of a real-world problem

> *It is a common fallacy to confuse scientists' models of reality with reality itself.* **A model is a map**. *A map is not the territory it describes. (Richard Casement)*

Modelling approaches: Phenomenological vs Mechanistic

Phenomenological approach: "A phenomenologically motivated approach generally constitutes a sketch or a summary of observations, and although it could be of high predictive utility, it is not directly connected to the underlying generative mechanism presumed to have produced available observations." $\Rightarrow$ **Descriptive models**

Mechanistic approach: ".. a mechanistically motivated approach is meant to constitute an explanation for observations, aimed at incorporating basic knowledge, and can typically be cast as a generalization or revision of a phenomenological approach." $\Rightarrow$ **Explanatory models**

N. Rodrigue, H. Philippe (2010) Trends in Genetics, 26: 248-252.

Outline

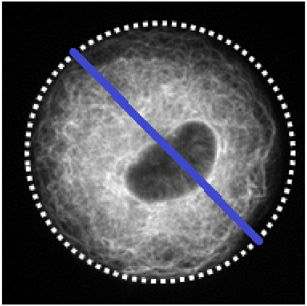
1. Introduction
   - Phenomenological approach
   - Mechanistic approach

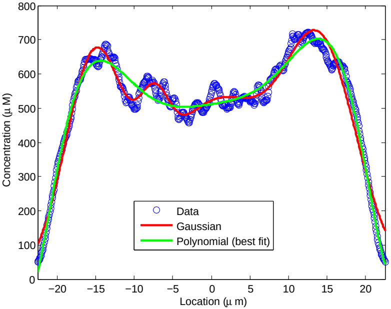
2. Model formulation
   - Single population models
   - Structured populations
   - Interacting populations
   - Compartmental models
   - Spatial models
   - Stochastic approaches

Phenomenological approach:

- **Data:** Limited number of data points representing values of a function for a limited number of values of the independent variable (e.g. time, space, . . .)
- **Hypothesize the form of the function**
- **Interpolation:** estimate values of data for intermediate values of the independent variable
  $\Rightarrow$ **Curve fitting** (find the curve that has the best fit to data points)
  The best fit minimizes the difference between the actual value (data) and the predicted value (curve)

Example: Spatial profile of protein concentration in a cell





- **Gaussian function**: $\text{Concentration}(x) = \sum_{i=1}^{n} a_i e^{-((x-b_i)/c_i)^2}$
- **Polynomial function**: $\text{Concentration}(x) = \sum_{i=0}^{n} p_i x^i$

Phenomenological approach:

No model is formulated, but trend or main feature in data can be extracted

Descriptive aspect only

Predictive aspect

- Extrapolation (to find data points outside of the range of known data points)

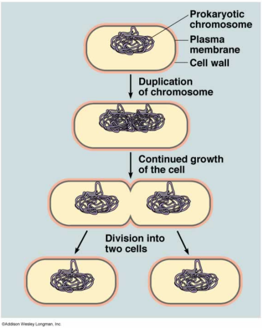
Outline

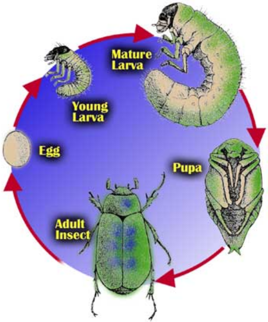
To write a mechanistic model

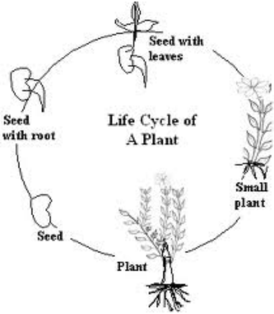
1. Identify the problem $\Rightarrow$ **Define the question to be answered**
2. **Define model variables and parameters** $\Leftarrow$ Experimental data
3. Identify the important processes governing the problem $\Rightarrow$ **Make assumptions**
4. Identify the **basic principles** that govern the quantities studied (physical laws, interactions..)
5. Express mathematically these principles $\Rightarrow$ **Choice of the formalism**
6. Verify that units are consistent
7. Verify that model is well-posed (existence of solutions, positivity of solutions..)

Idealization of real-world problems (never a completely accurate representation)

How does a population grow?







Outline

1. Introduction
   - Phenomenological approach
   - Mechanistic approach

2. Model formulation
   - Single population models
   - Structured populations
   - Interacting populations
   - Compartmental models
   - Spatial models
   - Stochastic approaches

Outline

How does a population grow? - Modelling change

Future value = Present value + Change

or

Change= Future value - Present value

$\Rightarrow$ Dynamical systems

How does a population grow?

- $N(t)$ population observed at time t
- Δt time interval
- r rate of growth per unit time

$$
\begin{aligned}
N(t+\Delta t) &= N(t) + r\Delta t N(t)\\
N(t+\Delta t) - N(t) &= r\Delta t N(t)\\
\frac{N(t+\Delta t) - N(t)}{\Delta t} &= rN(t)
\end{aligned}
$$

Assume that $\Delta t \to 0$

$$
\begin{aligned}
\lim_{\Delta t \to 0} \frac{N(t+\Delta t) - N(t)}{\Delta t} &= rN(t)\\
\frac{dN}{dt} &= rN
\end{aligned}
$$

How to represent time?

Time can be described as a

- discrete variable (time interval, generation..) $\Rightarrow$ **Difference equations**

$$N(t+\Delta t) = N(t) + r\Delta t N(t) = (1 + r\Delta t)N(t)$$

- continuous variable (instantaneous) $\Rightarrow$ **Differential equations**

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

$\frac{dN}{dt}$ = instantaneous rate of change of the state variable N with respect to time t

Exponential decay or growth model

Let N be the amount at time t.

Assumption about the variation of the amount

the rate of change of N is proportional to the current value of N

$$\frac{dN}{dt} = rN, \qquad N(0) = N_0$$

r is called the **rate of growth** $(r > 0)$ **or decay** $(r < 0)$

Solution to the IVP

$$N(t) = N_0 e^{rt}$$

- $r > 0$, $\lim_{t \to \infty} N(t) = \infty$
- $r < 0$, $\lim_{t \to \infty} N(t) = 0$

Difference equations (discrete time): First-order linear homogeneous equations with constant coefficients

Consider the first-order linear homogeneous difference equation with constant coefficients

$$N_{t+1} = N_t + r\Delta_t N_t = (1 + r\Delta_t)N_t = aN_t$$

If an initial value N_0 is known, the solution is unique and is given by

$$N_t = a^t N_0.$$

The **asymptotic behavior** of the solution depends on the value of a:

- if $|a| < 1$, then $\lim_{t\to\infty} N_t = 0$, i.e., N_t converges to 0,
- if $a = 1$, then for all $t \geq 0$, $N_t = N_0$, i.e., N_t remains constant,
- if $a = -1$, then for all $t \geq 0$, $N_t = (-1)^t N_0$, i.e., N_t alternates,
- if $|a| > 1$ then N_t diverges (either approaches infinity if $a > 1$ or diverges with alternating signs if $a < -1$).

Choice of formalism

How to represent a problem

- Static vs Dynamic
- Stochastic vs Deterministic
- Continuous vs Discrete
- Homogeneous vs Detailed

Formalism

Differential equations (ODE, PDE, DDE, SDE), difference equations, integral equations, integro-differential equations, Markov chains, game theory, graph theory, agent-based model, cellular automata, L-systems . . .

Malthus' population growth model

Consider a population with $N(t)$ individuals at time t. Suppose that

- individuals are born with rate constant b,
- individuals die with rate constant d.

Then the evolution of $N(t)$ over time is governed by an ODE, Malthus' equation:

$$\frac{dN}{dt} = bN - dN = (b - d)N \qquad (1)$$

The solution is

$$N(t) = N(0)e^{(b-d)t}$$

- $b > d$, $\lim_{t\to\infty} N(t) = \infty$, the population grows,
- $b < d$, $\lim_{t\to\infty} N(t) = 0$, the population extincts.

Logistic equation: a refinement of Malthus' model

Assumptions

Additionally to birth and death, suppose that

- individuals are subject to *intraspecific competition* with other members of their species:
    - competition for food, competition for nesting space..
- intraspecific competition occurs with rate constant κ

Model: logistic equation

$$\frac{dN}{dt} = (b-d)N - \kappa N^2$$

by setting $r = b - d > 0$, $K = (b-d)/\kappa$,

$$\frac{dN}{dt} = r\left(1 - \frac{N}{K}\right)N \qquad (2)$$

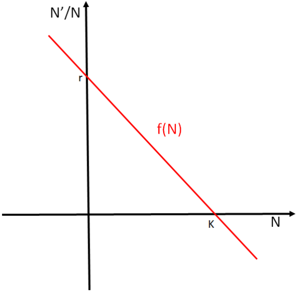
r: intrinsic rate of growth; K: environmental carrying capacity

Per capita growth rate depends on the population density

N'/N is the *per capita* instantaneous rate of growth:

$$\frac{N'}{N} = \underbrace{r}_{\text{Growth}} \underbrace{\left(1 - \frac{N}{K}\right)}_{\text{Regulation by crowding}} = r\left(1 - \frac{N}{K}\right) = f(N)$$

- $N < K$: population below carrying capacity, no regulation
- $N = K$: population at carrying capacity
- $N > K$: population over carrying capacity, crowding effects, population diminishes



Environment is capable of sustaining no more than a fixed number K of cells

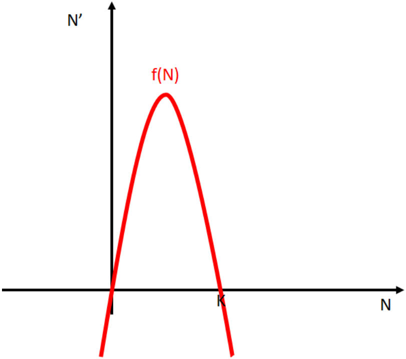
Solution of the logistic equation

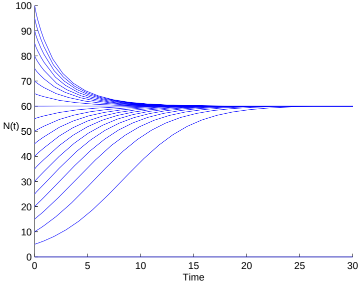
- Logistic equation is separable (or a Bernoulli equation with $n = 2$)

$$\frac{dN}{dt} = rN - \frac{r}{K}N^2 = f(N), \quad N_0 = N(0)$$

- Explicit solution

$$N(t) = \frac{KN_0}{N_0 + (K - N_0)e^{-rt}} \quad t \geq 0; \quad \lim_{t\to\infty} N(t) = K$$





Effect of formalism – discrete vs continuous in time

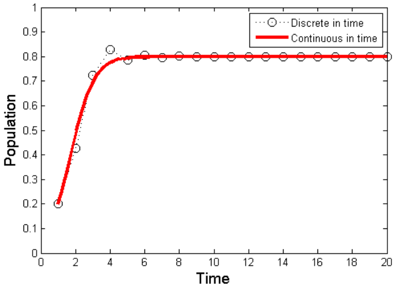
Logistic equation

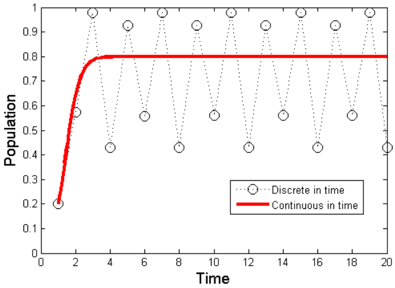
- discrete in time

$$N(t+1) = (1+r)N(t) - (r/K)N(t)^2$$

- continuous in time

$$\frac{dN}{dt} = rN - (r/K)N^2$$



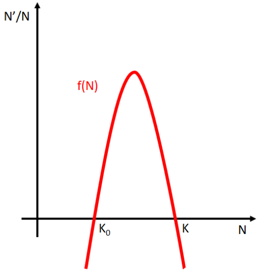


(Left) $r = 1.5,\ K = 0.8$ (Right) $r = 2.5,\ K = 0.8$

Other examples of non-constant per capita rate of growth

Allee effect

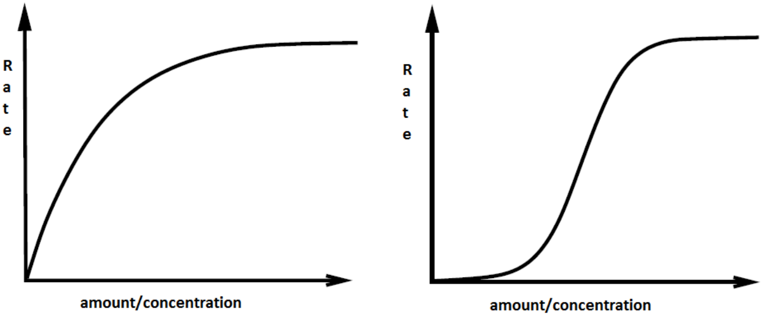
$$\frac{dN}{dt} = r\left(1 - \frac{N}{K}\right)\left(\frac{N}{K_0} - 1\right)N = f(N)N$$



Gompertz equation

$$\frac{dN}{dt} = re^{-\alpha t}N = f(t)N$$

Saturating rates : different translations

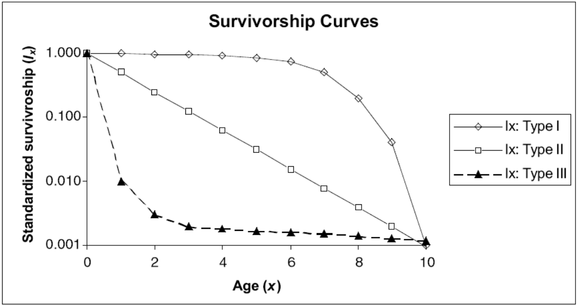


- Hyperbolic saturation: as the amount increases the rate increases but by slowing down (Michaelis-Menten dynamics), $f(N) = \frac{\alpha N}{\kappa + N}$
- Sigmoidal saturation: from a slow to rapid rate, “switch-like” rise toward to the limiting value (Cooperativity), $f(N) = \frac{\alpha N^m}{\kappa + N^m}$

Outline

Structured populations

In some species, the amount of reproduction varies greatly with the age of individuals.



Type I: high survivorship throughout life until old age sets in, and then survivorship declines dramatically to 0. Humans are type I organisms. Type III: In contrast, very low survivorship early in life, and few individuals live to old age.

Age structure or developmental stage of population matters

McKendrick–von Foerster equation

Model with age-structure

$$\frac{\partial N}{\partial t} + \frac{\partial N}{\partial a} + \mu(a)N(t,a) = 0$$

with boundary condition (birth function)

$$N(t,0) = \int_0^\infty b(a)N(t,a)da$$

and initial condition (initial age distribution)

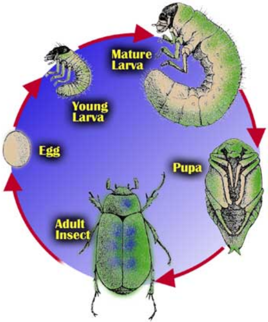
$$N(0,a) = f(a)$$

where

- $N(t,a)$ population density at time t and age a
- $b(a)$ birth rate of individuals of age a
- $\mu(a)$ death rate of individuals of age a
- $f(a)$ initial age distribution of population

Structured population models

used when the population can be organized or divided into various subclasses following traits such as age, life-stage or size. The variable that describes this trait is called the structuring variable.



The dynamic interactions among the stages, ages or sizes determine how the population structure changes over the time.

Structured population dynamics: discrete models

- Population categorized into a finite number of classes $i = 1, 2, \cdots, m$
- $x_i(t)$ number or density of individuals in the i^{th} class at time $t = 0, \Delta t, 2\Delta t, \ldots$
- If only birth and death processes (no migration):

$$x(t + \Delta t) = Px(t)$$

where $P = T + F$ is the projection matrix

$T = [t_{ij}]$ transition matrix, $0 \leq t_{ij} \leq 1$ and $\sum_{i=1}^{m} t_{ij} \leq 1$ for all j

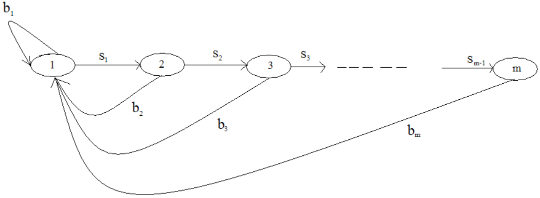
- t_{ij} fraction of $j-$class individual expected to survive and move to class i per unit of time
- t_{ii} fraction of individuals in class i that survive and remain in class i after one unit of time
- No individual can shrink or grow more than one class in one unit of time

$F = [f_{ij}]$ fertility matrix, $f_{ij} \geq 0$

- f_{ij} the expected number of (surviving) $i-$class offspring per $j-$class individual per unit of time

A particular case : Leslie model

(the time interval coincides with the structure interval)



$$
\begin{aligned}
x_1(t+1) =& b_1 x_1(t) + b_2 x_2(t) + b_3 x_3(t) + \ldots b_m x_m(t) \\
x_2(t+1) =& s_1 x_1(t) \\
& \vdots \\
x_m(t+1) =& s_{m-1} x_{m-1}(t)
\end{aligned}
$$

$$
X(t+1) = \begin{pmatrix} x_1(t+1) \\ x_2(t+1) \\ \vdots \\ x_m(t+1) \end{pmatrix} = \begin{pmatrix} b_1 & b_2 & \ldots & b_{m-1} & b_m \\ s_1 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & s_{m-1} & 0 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_m(t) \end{pmatrix} = LX(t)
$$

Outline

Example: Rabbits and foxes in Australia

Question: What are the interactions between rabbit and fox populations?

Assumptions:

- Rabbit = prey
- Fox = specialist predator (has a limited diet and stricter habitat requirements.. opposed to generalist)

Example: Rabbits and foxes in Australia

Propose a model to describe the interactions of the population of foxes (F) and the population rabbits (R) using ODEs formalim

To do so, the following mechanisms have to be described:

- Dynamics of predators in absence of preys
- Dynamics of preys in absence of predators
- Interactions between preys and predators

Mathematical modelling = Translation to mathematics

Propose a model to describe the interactions of the population of foxes (F) and the population rabbits (R) using ODEs formalim

To do so, the following mechanisms have to be described:

- Dynamics of predators in absence of preys
-
-

$$\frac{dR}{dt} =$$
$$\frac{dF}{dt} = -\,cF$$

Mathematical modelling = Translation to mathematics

Propose a model to describe the interactions of the population of foxes (F) and the population rabbits (R) using ODEs formalim

To do so, the following mechanisms have to be described:

- Dynamics of predators in absence of preys
- Dynamics of preys in absence of predators
-

$$\frac{dR}{dt} = aR \qquad \qquad \frac{dR}{dt} = a\left(1 - \frac{R}{k}\right)R$$

Interacting populations

Mass Action law

The rate of an elementary reaction (defined by reduction of reactant or formation of product) is proportional to the concentration of each individual species involved in the elementary reaction.

Mathematical modelling = Translation to mathematics

Propose a model to describe the interactions of the population of foxes (F) and the population rabbits (R) using ODEs formalim

To do so, the following mechanisms have to be described:

- Dynamics of predators in absence of preys
- Dynamics of preys in absence of predators
- Interactions between preys and predators

$$\begin{aligned} \frac{dR}{dt} &= aR - bRF \\ \frac{dF}{dt} &= -cF + eRF \end{aligned}$$

$$\begin{aligned} \frac{dR}{dt} &= a\left(1 - \frac{R}{k}\right)R - bRF \\ \frac{dF}{dt} &= -cF + eRF \end{aligned}$$

Mathematical modelling = Translation to mathematics

Propose a model to describe the interactions of the population of foxes (F) and the population rabbits (R) using ODEs formalim

To do so, the following mechanisms have to be described:

- Dynamics of predators in absence of preys
- Dynamics of preys in absence of predators
- Interactions between preys and predators

$$\frac{dR}{dt} = aR - bRF \qquad \frac{dR}{dt} = a\left(1 - \frac{R}{k}\right)R - bRF \qquad \frac{dR}{dt} = aR - \frac{bRF}{1+fR}$$

$$\frac{dF}{dt} = -cF + eRF \qquad \frac{dF}{dt} = -cF + eRF \qquad \frac{dF}{dt} = -cF + \frac{eRF}{1+fR}$$

Different assumptions ⇒ Different terms ⇒ Different models (a collection of models)

Other types of interactions

$$\frac{dN_1}{dt} = r_1 N_1 \left(1 \underbrace{- \frac{N_1}{K_1}}_{\text{Intra-population}} + \underbrace{b_{12} \frac{N_2}{K_1}}_{\text{Inter-population}} \right)$$

$$\frac{dN_2}{dt} = r_2 N_2 \left(1 \underbrace{- \frac{N_2}{K_2}}_{\text{Intra-population}} + \underbrace{b_{21} \frac{N_1}{K_2}}_{\text{Inter-population}} \right)$$

b_{ij} strength of interaction exerted by an individual of species j on an individual of species i

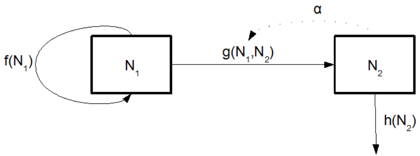
b_{12}	b_{21}	Interactions
-	-	Competitive
+	+	Mutualistic
+	-	Parasitic
-	+	Parasitic

Consumer-Resource models

- N_1 resource
- N_2 consumer

$$\frac{dN_1}{dt} = \underbrace{f(N_1)}_{\text{resource-renewal}} - \underbrace{g(N_1, N_2)}_{\text{consumption of resource by consumers}}$$

$$\frac{dN_2}{dt} = \underbrace{\alpha}_{\text{conversion factor}} g(N_1, N_2) - \underbrace{h(N_2)}_{\text{change of consumers in absence of resource}}$$



Possible forms

Resource-renewal term

$f(N_1) = \pi$	Inflow of resources at a constant rate
$f(N_1) = -\pi$	Outflow of resources at a constant rate
$f(N_1) = rN_1$	Constant per capita growth of resource
$f(N_1) = rN_1(1 - \frac{N_1}{K})$	Per capita growth of resources declines linearly with resource level (Logistic)
$f(N_1) = rN_1 e^{-\beta N_1}$	Per capita growth of resources declines exponentially with resource level

Resource consumption term

$g(N_1, N_2) = acN_1 N_2$	Linear rate of resource consumption
$g(N_1, N_2) = \frac{acN_1}{b+N_1} N_2$	Saturating rate of resource consumption
$g(N_1, N_2) = \frac{acN_1^k}{b+N_1^k} N_2$	Saturating rate of resource consumption

Change of consumers in absence of resource

$h(N_2) = dN_2$	Constant per capita death rate of consumers
$h(N_2) = (dN_2)N_2$	Per capita death rate of consumers increases linearly with consumer population size

Outline

Compartmental models

Definition

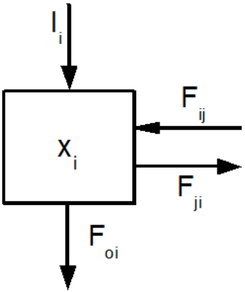
Systems in which there are flows of material between units called **compartments**

Compartments

A compartment is an amount of some material : the material of a compartment is at all times homogeneous; any material entering it is instantaneously mixed with the material of the compartment.

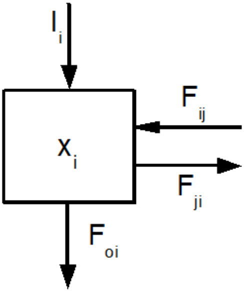
Dynamic models that depends on local mass balance conditions

The i^{th} compartment of a differential compartmental model



- x_i mass of compartment i
- I_i flows into the compartment i from the environment (**inflows**)
- F_{oi} **outflows** from compartment i to the environment (out of the system)
- F_{ji} transfers from compartment i to compartment j
- F_{ij} transfers from compartment j to compartment i

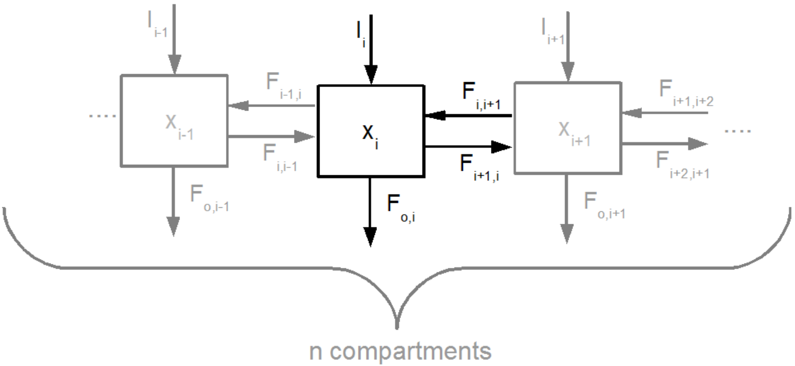
The i^{th} compartment of a differential compartmental model



$$\frac{dx_i}{dt} = \sum_{j \neq i} (-F_{j,i} + F_{i,j}) + I_i - F_{o,i}$$

- x_i nonnegative state variable
- all flows are nonnegative ($F_{i,j} \geq 0,\ I_i \geq 0,\ F_{o,i} \geq 0,\ \forall i, j$)
- sign in the equation translates the directions of flow
- if $x_i = 0$, then $F_{o,i} = 0$ and $F_{j,i} = 0\ \forall j$

An example of model of n compartments



n state variables x_i for $i \in \{1, \cdots, n\}$

$$\frac{dx_i}{dt} = -F_{i-1,i} + F_{i,i-1} - F_{i+1,i} + F_{i,i+1} + I_i - F_{oi}$$

If $\mathbf{F}$ is C^k, then

$$F_{j,i}(\mathbf{x}) = f_{j,i}(\mathbf{x}) \cdot x_i$$

for some function $f_{j,i}(\mathbf{x})$ which is at least C^{k-1}.

System

$$\frac{dx_i}{dt} = \sum_{j \neq i} (-F_{j,i} + F_{i,j}) + I_i - F_{o,i}$$

can then be rewritten as

$$\frac{dx_i}{dt} = -\left(f_{o,i} + \sum_{j \neq i} f_{j,i} \right) x_i + I_i + \sum_{j \neq i} f_{i,j} x_j$$

- $f_{i,j}$ constants or functions only of time $\Rightarrow$ Linear system
- $f_{i,j}$ functions of $\mathbf{x}$ (not constant function) $\Rightarrow$ Nonlinear system

Compartmental models

$$\frac{dx}{dt} = \text{input rate} - \text{output rate}$$

Applications

- pharmacokinetics
- physiology
- immunology
- epidemiology
- ecology
- ..

Epidemiological Models (1/2)

Classification of individuals according to their disease status

- susceptible (S): individuals not infective but who are capable of contracting the disease
- latent or exposed (E): infected by the disease, but not yet infectious
- infective (I): infectious individual; an individual can be infectious before symptoms appear.
- removed (R): no longer infectious, whether by acquiring immunity or death

Types of models

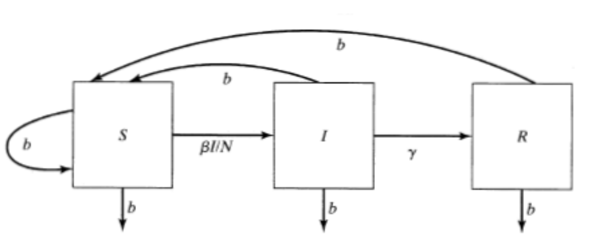
- SI model: no recovery.
- SIS model: recovery but no immunity.
- SIR model: recovery with permanent immunity.
- SIRS model: recovery with temporary immunity.

Epidemiological Models (2/2)

Parameters

- β transmission rate
- b rate of a birth
- d rate of death
- γ rate of recovery; $1/\gamma$ is the average length of the infectious period when there are no death.
- ν rate of loss of immunity; $1/\nu$ average length of immunity.

SIR model



Outline

Spatial models

If the population/amount is not homogeneous in space $\Rightarrow$
Space-dependent processes + Movement/Motion

- **Discrete in space:** compartmental models, metapopulations, network models, cellular automata, lattice gas models, Potts model...
- **Continuous in space:** integro-differential equations, partial differential equations...

Different types of motions

- **Diffusion**: random motion of objects in a fluid
- **Advection**: objects are carried along by a current in the fluid
- **Taxis**: motion in response to a stimulus (environment sensing $\rightarrow$ respond to environment)
  - Chemotaxis: response to a chemical gradient
  - Phototaxis: response to a light source
  - Geotaxis: response to a gravitational field
  - Galvanotaxis: response to an electrical field (human skin cells migrate toward the negative pole in direct current electric fields of physiological strength (wound healing))
  - Haptotaxis: response to an adhesive gradient

Macroscopic theory – Conservation law

- $C(x,t)$ concentration of particles at location x at time t
- $J(x,t)$ (flux rate) rate at which C moves across the boundary at position x from left to right at time t (amount/area/time)
- $f(x,t,C)$ source function

The conservation law is

$$\frac{\partial C}{\partial t} + \frac{\partial J}{\partial x} = f(x,t,C)$$

(evolution equation for C)

(in $n-$dimension)

$$\frac{\partial C}{\partial t} + \nabla \cdot J = f(\mathbf{x},t,C)$$

Diffusive flux – Fick's law

"C moves from regions of high concentration to regions of low concentration", at a rate proprotional to the gradient concentration

$$J(x,t) = -D\frac{\partial C}{\partial x}$$

D is the diffusion coefficient.

Under Fick's law, C evolves as follows

$$\frac{\partial C}{\partial t} + \frac{\partial}{\partial x}\left(-D\frac{\partial C}{\partial x}\right) = f(x,t,C)$$

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} + f(x,t,C)$$

(in $n-$dimension)

$$\frac{\partial C}{\partial t} = \nabla\cdot(D\nabla C) + f(\mathbf{x},t,C) = D\nabla^2 C + f(\mathbf{x},t,C) = D\Delta C + f(\mathbf{x},t,C)$$

Advective flux

There is a uniform macroscopic flow of the solvent, with a speed u along the $x-$axis, which carries solutes along with it.

$$J(x,t) = uC(x,t)$$

So C evolves as follows

$$\frac{\partial C}{\partial t} + \frac{\partial uC}{\partial x} = f(x,t,C)$$

$$\frac{\partial C}{\partial t} = -u\frac{\partial C}{\partial x} + f(x,t,C)$$

(in $n-$dimension)

$$\frac{\partial C}{\partial t} = -\nabla \cdot (\mathbf{u}C) + f(\mathbf{x},t,C)$$

(Advective and diffusive flux)

$$J(x,t) = uC(x,t) - D\frac{\partial C}{\partial x}$$

Attraction-Repulsion

Φ represents a source of attraction/repulsion for solutes/particles/cells. An attractive/repulsive force would pull/push particles towards/forwards the sites of greatest attraction/repulsion:

- direction and magnitude of motion is determined by the gradient of Φ,
- α a scalar to characterize the sensitivity of solutes/particles/cells to the attraction($\alpha > 0$)/repulsion($\alpha < 0$)

$$J(x,t) = \alpha C(x,t)\frac{\partial \Phi}{\partial x}$$

So C evolves as follows

$$\frac{\partial C}{\partial t} = -\frac{\partial}{\partial x}\left[\alpha C(x,t)\frac{\partial \Phi}{\partial x}\right] + f(x,t,C)$$

(in $n-$dimension)

$$\frac{\partial C}{\partial t} = -\nabla \cdot (\alpha C \nabla \Phi) + f(\mathbf{x},t,C)$$

Boundary conditions

Boundary conditions reflect certain physical conditions of the experiment.

One-dimensional case:

- Dirichlet boundary condition: C(L,t)=f(t)
- Neumann boundary condition: $J(L, t) = g(t)$
- Robin condition: $J(L, t) = h(t) - aC(L, t)$

(L is a boundary point)

Diffusion equation

$$\begin{aligned}\frac{\partial C}{\partial t} =& D\frac{\partial^2 C}{\partial x^2}, \quad -\infty < x < \infty, \quad t > 0,\\ C(x,0) =& C_0\delta(x),\\ \lim_{x\to\pm\infty} C(x,t) =& 0,\end{aligned}$$

where C_0 is the total amount of material, and $\delta(x)$ is the Dirac delta function.
The solution is

$$C(x,t) = \frac{C_0}{\sqrt{4\pi Dt}} exp^{-\frac{x^2}{4Dt}}.$$

(normal density function with mean 0 and variance $2Dt$)

Drift-diffusion equation

$$\begin{aligned}\frac{\partial C}{\partial t} =& - u\frac{\partial C}{\partial x} + D\frac{\partial^2 C}{\partial x^2}, \quad -\infty < x < \infty, \quad t > 0,\\ C(x,0) =& C_0\delta(x - x_0),\\ \lim_{x\to\pm\infty} C(x,t) =& 0,\end{aligned}$$

The solution is

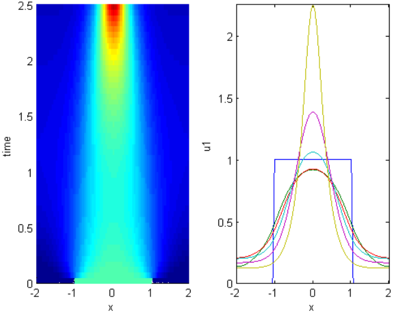
$$C(x,t) = \frac{C_0}{\sqrt{4\pi Dt}} exp^{-\frac{(x-x_0-ut)^2}{4Dt}}.$$

(normal density function with mean $x_0 + ut$ and variance $2Dt$)

Example: Chemotaxis

Keller-Segel model describes directed motion of cells stimulated by the chemical which they produce themselves.

$$\frac{\partial u}{\partial t} = -\nabla\left(-\mu\nabla u + \chi u \nabla v\right)$$
$$\frac{\partial v}{\partial t} = \nabla\left(D\nabla v\right) + fu - kv$$



Keller, E., Segel, L. (1970) J. Theor. Biol. 26:399-415.

Outline

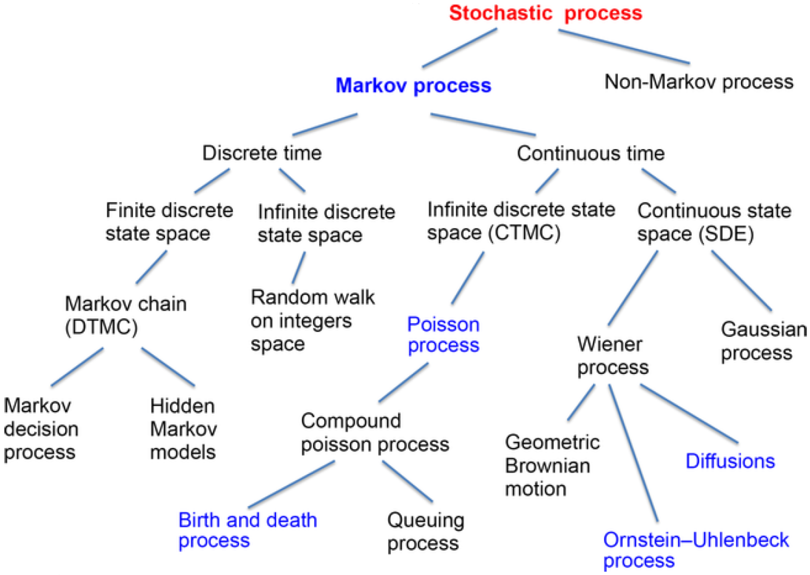
"The dynamics of real systems are often influenced by internal and external factors which are not completely understood and, therefore, cannot be described precisely. In order to understand such systems, **deterministic models** *(which explain broad overall behavior and growth patterns) are modified to incorporate complex variations in the mechanisms underlying the system. These variations, defined in terms of probabilities which evolve over time alongside the populations, result in new structures referred to as* **stochastic models***."*

Mubayi *et al. Handbook of Statistics*

Stochastic effects (variations)

- from differences among individuals (demographic stochasticity)
- from fluctuations in the environment (due to forces/mechanisms not described in the model), (environmental stochasticity)

Different types of stochastic processes



Mubayi *et al.* *Handbook of Statistics*

Stochastic differential equations

A SDE can be defined as a stochastic process $X_t = X(t)$ satisfying

$$dX_t = f(X_t, t)dt + g(X_t, t)dW_t$$

or

$$X_t = X_{t_0} + \int_{t_0}^{t} f(X_s, s)ds + \int_{t_0}^{t} g(X_s, s)dW_s$$

where

- $W_t = W(t)$ is the Wiener process (a Gaussian process with zero mean $E(W(t)) = 0$ and variance proportional to the elapsed time $Var(W(t)) = t$)
- $f(\cdot)$ the deterministic component
- $g(\cdot)$ the stochastic component
  - when $g(\dots)$ does not depend on $X_t \Rightarrow$ additive noise
  - when $g(\dots)$ depends on $X_t \Rightarrow$ multiplicative noise

SDE - Example

Stochastic logistic equation with multiplicative noise

$$dx_t = rx_t(1 - \frac{x_t}{k})dt + cx_t dw_t$$

Since a Wiener process is nondifferentiable, Ito's formula is often used to find the explicit solution of simple SDEs with a Wiener process.

When no explicit solution is available different characteristics of the process can be approximated by simulation, such as sample paths, moments, qualitative behavior..

Stochastic approach: Chemical Master Equation

Assume a system composed of N different types of molecules, and there exist M reactions.

- $\mathbf{X}(t) = [X_1(t), X_2(t), \ldots, X_N(t)]^T$ random variable
- $X_i(t)$ ($i \in \{1, \ldots, N\}$) the number of molecules of type i at time t
- $\mathbf{x} = [x_1, \ldots, x_N]^T$ state of system ($\mathbf{x} \in \mathbb{N}^N$)
- $p(\mathbf{x}, t) = Prob[\mathbf{X}(t) = \mathbf{x}]$ probability of the state $\mathbf{x}$ at time t

$$\frac{dp(\mathbf{x}, t)}{dt} = \sum_{j=1}^{M} [a_j(\mathbf{x} - \mathbf{v}_j)p(\mathbf{x} - \mathbf{v}_j, t) - a_j(\mathbf{x})p(\mathbf{x}, t)]$$

- $a_j(\mathbf{x})$ propensity function of j^{th} reaction at the state $\mathbf{x}$
- $\mathbf{v}_j$ j^{th} column of the stoichiometric matrix $\mathbf{v}$ (the reaction j will lead the system from the state $\mathbf{x}^i$ to the state $\mathbf{x}^i + \mathbf{v}_j$)

Propensity functions

- $a_j(\mathbf{x})\Delta t$ is the probability that reaction j will occur in $(t, t + \Delta t)$ when the system is at state $\mathbf{x}$
- $a_j(\mathbf{x}) =$ rate of reaction j X number of reactant combinations available in the state $\mathbf{x}$ to allow reaction j

Chemical Master Equation $\rightarrow$ Forward Chapman Kolmogorov Equation

$P(t) = [p(\mathbf{x}^1, t), p(\mathbf{x}^2, t), \ldots]^T$ vector whose the i^{th} entry is the probability of the i^{th} state $\mathbf{x}^i$ at time t

$$\frac{dP}{dt} = QP(t)$$

The solution is (if the number of states is finite)

$$P(t) = \exp(Qt)P(0)$$

A very simple chemical reaction: degradation

$$X \xrightarrow{k} \emptyset$$

Deterministic approach: $X(t)$ concentration of molecule X at time t

$$\frac{dX}{dt} = -kX, \quad X(0) = x_0, \Rightarrow X(t) = x_0 \exp^{-kt}.$$

Stochastic approach: $X(t)$ number of molecule X at time t (random variable), $p(x, t|x_0, t_0) := Prob\{X(t) = x, given\, X(t_0) = x_0\}$

$$\frac{dp(x, t|x_0, t_0)}{dt} = k(x+1)p(x+1, t|x_0, t_0) - k(x)p(x, t|x_0, t_0).$$

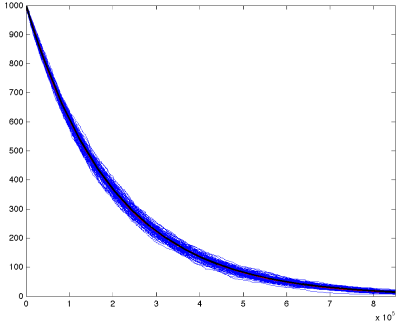
The solution is (Binomial probability density function)

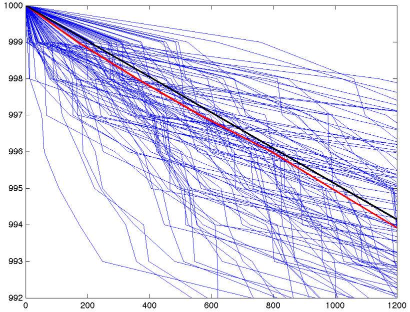
$$p(x, t|x_0, 0) = \frac{x_0!}{x!(x_0 - x)!} \exp^{-kxt}(1 - \exp^{-kt})^{x_0 - x}, \quad (x = 0, \dots, x_0)$$

$$E(X) = x_0 \exp^{-kt}, \quad Var(X) = x_0 \exp^{-kt}(1 - \exp^{-kt})$$

$$X \xrightarrow{k} \emptyset$$

Simulations done with Gillespie's algorithm

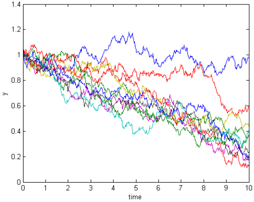


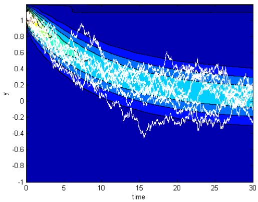


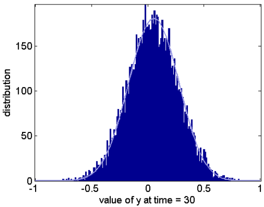
blue = realization, red = mean of realizations, black = ODE solution

Individual vs Population

Langevin equation: (stochastic equation of motion for the time evolution of y) motion of a particle in a viscous medium subject to friction and noise (white Gaussian noise).
Fokker-Planck equation: equation of motion for the time dependent probability distribution $p(y,t)$.







Deterministic vs Stochastic

- Deterministic model output is fully determined by the parameter values and the initial conditions
  - deterministic models capture the mean behavior of a system
- Stochastic models possess some inherent randomness. The same set of parameter values and initial conditions will lead to an ensemble of different outputs
  - stochastic models capture the ways that a system's behavior may deviate (variability) from the mean
  - small populations
  - when addressing questions related to variations in population

Once the model is written

Mathematical analysis: identify the type of mathematical techniques and theories required for the analysis of the model.. and characterize the behavior of the model

Numerical experiment: conduct numerical simulations of the model..

Model calibration: identify and estimate the values of parameters..

Sensitivity analysis: understand the effect of model inputs (parameters or initial conditions) on model outputs.. which parameter is the key driver of the model responses

Validation: model must represent accurately the real process, it must reproduce known states of the real process.. if several models are considered, model selection has to be used.