MATH 3610 – 03
Time of residence in states
Introduction to compartmental models

Julien Arino
University of Manitoba
julien.arino@umanitoba.ca

Fall 2024

The University of Manitoba campuses are located on original lands of Anishinaabeg, Ininew, Anisininew, Dakota and Dene peoples, and on the National Homeland of the Red River Métis.
We respect the Treaties that were made on these territories, we acknowledge the harms and mistakes of the past, and we dedicate ourselves to move forward in partnership with Indigenous communities in a spirit of Reconciliation and collaboration.

Outline

Time spent in a state

Exponential distribution

A cohort model

Sojourn times in an SIS disease transmission model

Time spent in a state

Exponential distribution

A cohort model

Sojourn times in an SIS disease transmission model

Some probability theory

We suppose that a system can be in two states, S_1 and S_2.

- At time $t = 0$, the system is in state S_1.
- An event happens at some time $t = \tau$, which triggers the switch from state S_1 to state S_2.

A **random variable** is a variable that takes random values, that is, a mapping from random experiments to numbers.

Let us call T the random variable
"time spent in state S_1 before switching into state S_2"

These states can be anything:

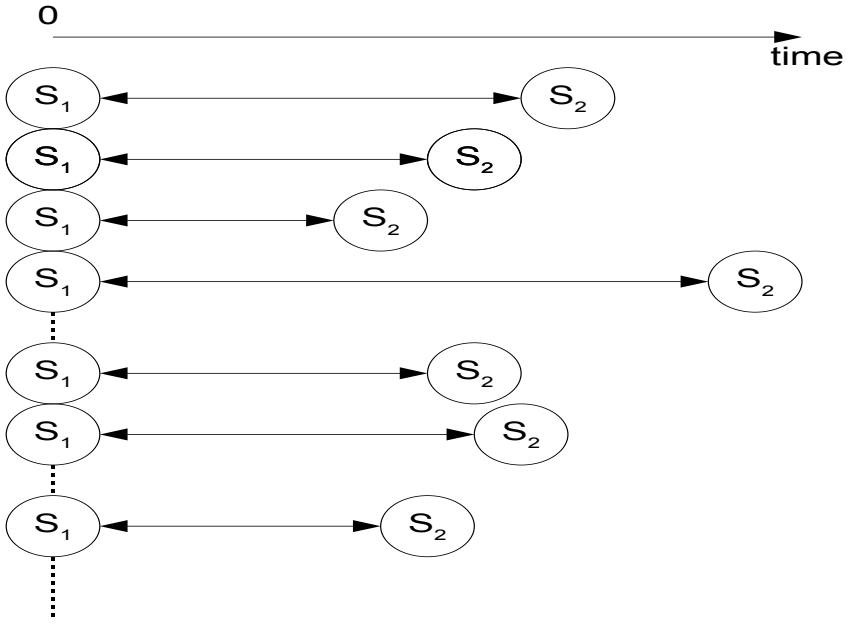
- S_1: working, S_2: broken;
- S_1: infected, S_2: recovered;
- S_1: alive, S_2: dead;
- ...

We take a collection of objects or individuals that are in state S_1 and want some law for the **distribution** of the times spent in S_1, i.e., a law for T.

For example, we make light bulbs and would like to tell our customers that on average, our light bulbs last 200 years..

For this, we conduct an **infinite** number of experiments, and observe the time that it takes, in every experiment, to switch between S_1 and S_2.

From this, we deduce a model, which in this context is called a **probability distribution**.


0
time
S1
S2
S1
S2
S1
S2
S1
S2
S1
S2
S1
S2
S1
S2

Discrete versus continuous random variables

We assume that T is a **continuous** random variable, that is, T takes continuous values. Examples of continuous r.v.:

- height or age of a person (if measured very precisely)
- distance
- time

Another type of random variables are **discrete** random variables, which take values in a denumerable set. Examples of discrete r.v.:

- heads or tails on a coin toss
- the number rolled on a dice
- height of a person, if expressed rounded without subunits, age of a person in years (without subunits)

Probability

A **probability** is a function $\mathcal{P}$, from a probability space to $[0, 1]$.

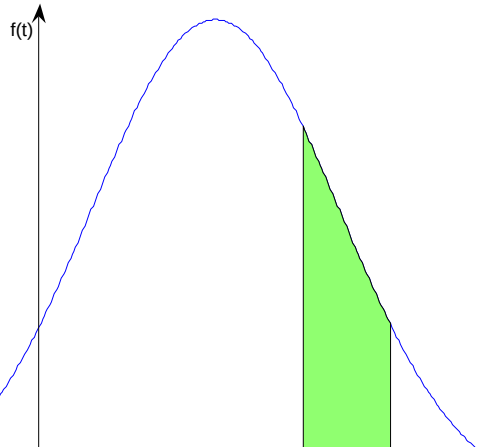
Formally: $(\Omega, \mathcal{F}, \mathcal{P})$ is a probability space, with Ω the **sample** space, $\mathcal{F}$ a σ-algebra of subsets of Ω whose elements are the **events**, and $\mathcal{P}$ a **measure** from $\mathcal{F}$ to $[0, 1]$ such that $\mathcal{P}(E) \geq 0$, $\forall E \subset \Omega$, $\mathcal{P}(\Omega) = 1$ and $\mathcal{P}(E_1 \cup E_2 \cup \cdots) = \sum_i \mathcal{P}(E_i)$.

Gives the likelihood of an event occurring, among all the events that are possible, in that particular setting. For example, $\mathbb{P}\text{getting heads when tossing a coin} = 1/2$ and $\mathbb{P}\text{getting tails when tossing a coin} = 1/2$.

Probability density function

Since T is continuous, it has a continuous **probability density function**, f.

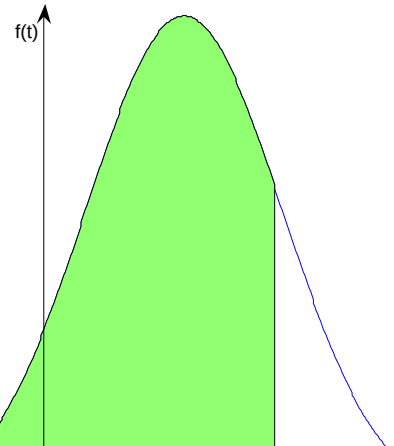
- $f \geq 0$,
- $\int_{-\infty}^{+\infty} f(s)ds = 1$.
- $\mathbb{P}a \leq T \leq b = \int_a^b f(t)dt$.



Cumulative distribution function

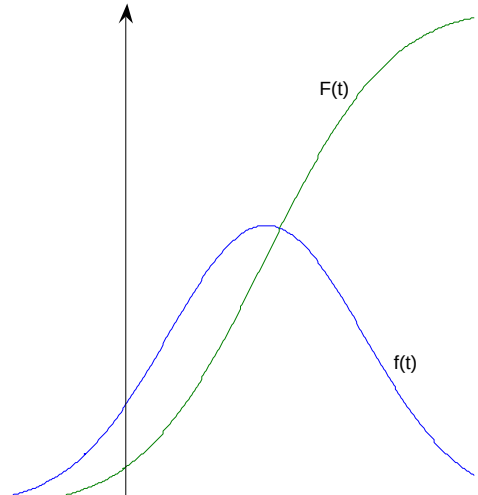
The cumulative distribution function (c.d.f.) is a function $F(t)$ that characterizes the distribution of T, and defined by

$$F(s) = \mathbb{P}T \leq s = \int_{-\infty}^{s} f(x)dx.$$



Properties of the c.d.f.

- Since f is a nonnegative function, F is nondecreasing.
- Since f is a probability density function, $\int_{-\infty}^{+\infty} f(s)ds = 1$, and thus $\lim_{t\to\infty} F(t) = 1$.



Mean value

For a continuous random variable T with probability density function f, the **mean** value of T, denoted $\bar{T}$ or $E(T)$, is given by

$$E(T) = \int_{-\infty}^{+\infty} tf(t)dt.$$

Survival function

Another characterization of the distribution of the random variable T is through the **survival** (or **sojourn**) function.

The survival function of state S_1 is given by

$$\mathcal{S}(t) = 1 - F(t) = \mathbb{P}T > t \tag{1}$$

This gives a description of the **sojourn time** of a system in a particular state (the time spent in the state).

$\mathcal{S}$ is a nonincreasing function (since $\mathcal{S} = 1 - F$ with F a c.d.f.), and $\mathcal{S}(0) = 1$ (since T is a positive random variable).

The **average sojourn time** τ in state S_1 is given by

$$\tau = E(T) = \int_0^\infty t f(t) dt$$

Assuming that $\lim_{t \to \infty} t\mathcal{S}(t) = 0$ (which is verified for most probability distributions),

$$\tau = \int_0^\infty \mathcal{S}(t) dt$$

Time spent in a state

Exponential distribution

A cohort model

Sojourn times in an SIS disease transmission model

The exponential distribution

The random variable T has an **exponential** distribution if its probability density function takes the form

$$f(t) = \begin{cases} 0 & \text{if } t < 0, \\ \theta e^{-\theta t} & \text{if } t \geq 0, \end{cases} \tag{2}$$

with $\theta > 0$. Then the survival function for state S_1 is of the form $\mathcal{S}(t) = e^{-\theta t}$, for $t \geq 0$, and the average sojourn time in state S_1 is

$$\tau = \int_0^\infty e^{-\theta t} dt = \frac{1}{\theta}$$

If on the other hand, for some constant $\omega > 0$,

$$\mathcal{S}(t) = \begin{cases} 1, & 0 \leq t \leq \omega \\ 0, & \omega < t \end{cases}$$

which means that T has a Dirac delta distribution $\delta_\omega(t)$, then the average sojourn time is a constant, namely

$$\tau = \int_0^\omega dt = \omega$$

These two distributions can be regarded as extremes.

Time spent in a state

Exponential distribution

A cohort model

Sojourn times in an SIS disease transmission model

A model for a cohort with one cause of death

We consider a population consisting of individuals born at the same time (a **cohort**), for example, the same year.

We suppose

- At time $t = 0$, there are initially $N_0 > 0$ individuals.
- All causes of death are compounded together.
- The time until death, for a given individual, is a random variable T, with continuous probability density distribution $f(t)$ and survival function $P(t)$.

The model

Denote $N(t)$ the population at time $t \geq 0$. Then

$$N(t) = N_0 P(t). \tag{3}$$

- $N_0 P(t)$ gives the proportion of N_0, the initial population, that is still alive at time t.

Case where T is exponentially distributed

Suppose that T has an exponential distribution with mean $1/d$ (or parameter d), $f(t) = de^{-dt}$. Then the survival function is $P(t) = e^{-dt}$, and (3) takes the form

$$N(t) = N_0 e^{-dt}. \tag{4}$$

Now note that

$$\begin{aligned} \frac{d}{dt}N(t) &= -dN_0 e^{-dt} \\ &= -dN(t), \end{aligned}$$

with $N(0) = N_0$.

$\Rightarrow$ The ODE $N' = -dN$ makes the assumption that the life expectancy at birth is exponentially distributed.

Case where T has a Dirac delta distribution

Suppose that T has a Dirac delta distribution at $t = \omega$, giving the survival function

$$P(t) = \begin{cases} 1, & 0 \leq t \leq \omega, \\ 0, & t > \omega. \end{cases}$$

Then (3) takes the form

$$N(t) = \begin{cases} N_0, & 0 \leq t \leq \omega, \\ 0, & t > \omega. \end{cases} \tag{5}$$

All individuals survive until time ω, then they all die at time ω.

Here, we have $N' = 0$ everywhere except at $t = \omega$, where it is undefined.

Time spent in a state

Exponential distribution

A cohort model

Sojourn times in an SIS disease transmission model

Models of diseases

Consider

- a disease,
- a population of individuals who can be infected by this disease.

Both can be anything:

- a human population subject to influenza,
- an animal population subject to foot and mouth disease,
- a rumor spreading in a human population,
- inovation spreading through businesses,
- a computer virus spreading on the internet,
- ...

Status of individuals

Suppose that individuals can be identified with respect to their epidemiological status:

- susceptible to the disease,
- infected by the disease,
- recovered from the disease,
- ...

These states are clearly of the type we were discussing before.

An SIS model

Consider a disease that confers no immunity. In this case, individuals are either

- **susceptible** to the disease, with the number of such individuals at time t denoted by $S(t)$,
- or **infected** by the disease (and are also **infective** in the sense that they propagate the disease), with the number of such individuals at time t denoted by $I(t)$.

We want to model the evolution with time of S and I (t is omitted unless necessary).

f*Extremelyimportant* : *Stateallyourhypotheses.*

Hypotheses

- Individuals typically recover from the disease.
- The disease does not confer immunity.
- There is no birth or death.
- Infection is of **standard incidence** type

Once your hypotheses are stated, detail them if need be.

Recovery and No immunity

Individuals recover from the disease: the infection is not permanent.

Upon recovery from the disease, an individual becomes susceptible again immediately.

Good description for diseases that confer no immunity, e.g.,

- the cold,
- gonorrhea,
- ...

No birth or death

Suppose that

- the time period of interest is short,
- the population is large enough,

then it is reasonable to assume that the total population is constant, in the absence of disease.

For mild diseases (cold, etc.), there are very little risks of dying from the disease. We assume no disease-induced death.

Hence $N \equiv N(t) = S(t) + I(t)$ is the (constant) total population.

Standard incidence

New infectives result from random contacts between susceptible and infective individuals, described using standard incidence:

$$\beta \frac{SI}{N},$$

- $\beta SI/N$ is a rate (per unit time),
- β is the **transmission coefficient**, giving probability of transmission of the disease in case of a contact, times the number of such contacts made by an infective per unit time.

Recovery

We have not yet stated our hypotheses on the recovery process..

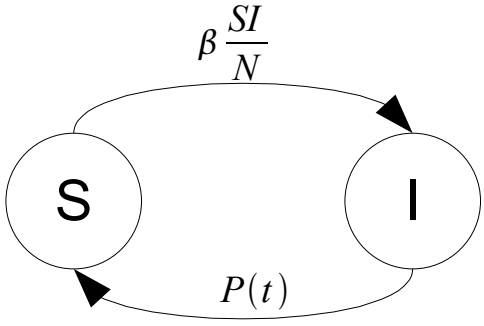
Traditional epidemiological models assume recovery from disease with a rate constant γ.

Here, assume that, of the individuals who have become infective at time t_0, a fraction $P(t - t_0)$ remain infective at time $t \geq t_0$.

Thus, considered for $t \geq 0$, the function $P(t)$ is a survival function.

A flow diagram for the model

This is the **flow diagram** of our model:



It details the flows of individuals between the compartments in the system.

It is extremely useful to rapidly understand what processes are modelled.

Reducing the dimension of the problem

To formulate our model, we would in principle require an equation for S and an equation for I.

But we have

$$S(t) + I(t) = N, \text{ or equivalently, } S(t) = N - I(t).$$

N is constant (equal total population at time $t = 0$), so we can deduce the value of $S(t)$, once we know $I(t)$, from the equation $S(t) = N - I(t)$.

We only need to consider 1 equation. **Do this when possible!** (nonlinear systems are hard, one less equation can make a lot of difference)

Model for infectious individuals

Integral equation for the number of infective individuals:

$$I(t) = I_0(t) + \int_0^t \beta \frac{(N - I(u))I(u)}{N} P(t-u)du \tag{6}$$

- $I_0(t)$ number of individuals who were infective at time $t = 0$ and still are at time t.
  - $I_0(t)$ is nonnegative, nonincreasing, and such that $\lim_{t\to\infty} I_0(t) = 0$.
- $P(t-u)$ proportion of individuals who became infective at time u and who still are at time t.
- $\beta(N - I(u))S(u)/N$ is $\beta S(u)I(u)/N$ with $S(u) = N - I(u)$, from the reduction of dimension.

Expression under the integral

Integral equation for the number of infective individuals:

$$I(t) = I_0(t) + \int_0^t \beta \frac{(N - I(u))I(u)}{N} P(t-u)du \tag{6}$$

The term

$$\beta \frac{(N - I(u))I(u)}{N} P(t-u)$$

- $\beta(N - I(u))I(u)/N$ is the rate at which new infectives are created, at time u,
- multiplying by $P(t-u)$ gives the proportion of those who became infectives at time u and who still are at time t.

Summing over $[0, t]$ gives the number of infective individuals at time t.

Case of an exponentially distributed time to recovery

Suppose that $P(t)$ is such that the sojourn time in the infective state has an exponential distribution with mean $1/\gamma$, *i.e.*, $P(t) = e^{-\gamma t}$.

Then the initial condition function $I_0(t)$ takes the form

$$I_0(t) = I_0(0)e^{-\gamma t},$$

with $I_0(0)$ the number of infective individuals at time $t = 0$. This is obtained by considering the cohort of initially infectious individuals, giving a model such as (3).

Equation (6) becomes

$$I(t) = I_0(0)e^{-\gamma t} + \int_0^t \beta \frac{(N - I(u))I(u)}{N} e^{-\gamma(t-u)} du. \tag{7}$$

Taking the time derivative of (7) yields

$$\begin{aligned} I'(t) &= -\gamma I_0(0) e^{-\gamma t} - \gamma \int_0^t \beta \frac{(N - I(u))I(u)}{N} e^{-\gamma(t-u)} du \\ &\quad + \beta \frac{(N - I(t))I(t)}{N} \\ &= -\gamma \left(I_0(0) e^{-\gamma t} + \int_0^t \beta \frac{(N - I(u))I(u)}{N} e^{-\gamma(t-u)} du \right) \\ &\quad + \beta \frac{(N - I(t))I(t)}{N} \\ &= \beta \frac{(N - I(t))I(t)}{N} - \gamma I(t), \end{aligned}$$

which is the classical logistic type ordinary differential equation (ODE) for I in an SIS model without vital dynamics (no birth or death).

Case of a step function survival function

Consider case where the time spent infected has survival function

$$P(t) = \begin{cases} 1, & 0 \leq t \leq \omega, \\ 0, & t > \omega. \end{cases}$$

i.e., the sojourn time in the infective state is a constant $\omega > 0$.
In this case (6) becomes

$$I(t) = I_0(t) + \int_{t-\omega}^{t} \beta \frac{(N - I(u))I(u)}{N} du. \tag{8}$$

Here, it is more difficult to obtain an expression for $I_0(t)$. It is however assumed that $I_0(t)$ vanishes for $t > \omega$.

When differentiated, (8) gives, for $t \geq \omega$,

$$I'(t) = I_0'(t) + \beta \frac{(N - I(t))I(t)}{N} - \beta \frac{(N - I(t-\omega))\, I(t-\omega)}{N}.$$

Since $I_0(t)$ vanishes for $t > \omega$, this gives the delay differential equation (DDE)

$$I'(t) = \beta \frac{(N - I(t))I(t)}{N} - \beta \frac{(N - I(t-\omega))I(t-\omega)}{N}.$$

Conclusion

- The time of sojourn in classes (compartments) plays an important role in determining the type of model that we deal with.
- All ODE models, when they use terms of the form κX, make the assumption that the time of sojourn in compartments is exponentially distributed.
- At the other end of the spectrum, delay delay differential with discrete delay make the assumption of a constant sojourn time, equal for all individuals.

- Both can be true sometimes.. but reality is often somewhere in between.

Survival function, $\mathcal{S}(t) = \mathbb{P}T > t$, for an exponential distribution with mean 80 years.