



UNIVERSITAT DE  
BARCELONA

Facultat de Matemàtiques  
i Informàtica

## GRAU DE MATEMÀTIQUES

Treball final de grau

---

# Epidemiologia estocàstica

---

Autor: Isaac Martínez i Escribano

Director: Dr. Carles Rovira Escofet

Realitzat a: Departament de Matemàtiques i Informàtica

Barcelona, 15 de gener de 2025

## Abstract

This work examines discrete compartmental models used in stochastic epidemiological studies. The central focus of the study is the SIR model and the subsequent extensions derived from it in a markovian chain. In this context, fundamental concepts of the theory of stochastic processes and the main associated analytical tools are first introduced. Concretely, Itô's integral and stochastic differential equations. Next, the study explores the nondeterministic theory of disease propagation. Diffusion approximations that describe some of the SIR-associated models are derived. The  $R_0$  value is also computed. Finally, a Covid-19 outbreak in a nursing home is modeled using the SEIHR model, which constitutes an extension of the original SIR model.

## Resum

En aquest treball s'estudien els models de compartimentació discrets emprats en l'estudi estocàstic epidemiològic. L'eix vertebrador de l'estudi és el model SIR i les extensions posteriors derivades d'aquest seguint una cadena de Màrkov. Per aquest motiu, primer es presenten conceptes fonamentals de la teoria de processos estocàstics i les principals eines d'anàlisi associades. Concretament, la integral d'Itô i les equacions diferencials estocàstiques. Segonament, s'aprofundeix en la teoria indeterminista de la propagació de malalties. S'obtenen aproximacions de difusió que descriuen alguns dels models associats al SIR. També es calcula el valor de  $R_0$ . Finalment, es modela un brot de Covid-19 en una residència utilitzant el model SEIHR que constitueix una ampliació del model SIR inicial.

## Agraïments

M'agradaria donar les gràcies al meu tutor, el Dr. Carles Rovira Escofet, per la seva predisposició per abordar un treball sobre aquesta temàtica i per la seva dedicació al llarg del desenvolupament d'aquest treball. La seva orientació i atenció als dubtes i detalls que han anat sorgint durant l'elaboració de la memòria han estat essencials. El TFG ha estat una experiència molt enriquidora en la que he gaudit i après molt.

També voldria estendre l'agraïment a la meua família i als meus amics no només pel seu recolzament durant l'elaboració d'aquest treball sinó pel seu suport incondicional al llarg de tot el grau.

A tots vosaltres, moltes gràcies.

# Índex

|                                                                     |           |
|---------------------------------------------------------------------|-----------|
| <b>Introducció</b>                                                  | <b>1</b>  |
| <b>1 Processos Estocàstics</b>                                      | <b>3</b>  |
| 1.1 Processos Estocàstics discrets i a temps continu . . . . .      | 3         |
| 1.1.1 Cas particular: El moviment brownià . . . . .                 | 3         |
| 1.2 Procés de Màrkov . . . . .                                      | 5         |
| <b>2 Anàlisi estocàstica</b>                                        | <b>8</b>  |
| 2.1 La integral estocàstica d'Itô . . . . .                         | 8         |
| 2.2 Càlcul numèric d'equacions diferencials estocàstiques . . . . . | 14        |
| <b>3 Models epidemiològics estocàstics</b>                          | <b>19</b> |
| 3.1 Primers desenvolupaments. Model SIR . . . . .                   | 19        |
| 3.1.1 SIR amb $R=\infty$ . . . . .                                  | 22        |
| 3.1.2 SIR amb $R < \infty$ . Model general. . . . .                 | 25        |
| 3.1.3 SIR amb nombre de contactes diaris variable . . . . .         | 27        |
| 3.2 El model SEIHR . . . . .                                        | 32        |
| 3.2.1 El model SEIHRS . . . . .                                     | 36        |
| 3.2.2 El model SEIAHCRD . . . . .                                   | 37        |
| 3.3 Aplicació del model SEIHR . . . . .                             | 41        |
| <b>4 Conclusions</b>                                                | <b>44</b> |
| <b>Bibliografia</b>                                                 | <b>45</b> |
| <b>Annex A: Codi cas pràctic</b>                                    | <b>47</b> |

# Introducció

## El projecte

Segons [1], l'**epidemiologia** és “l'estudi de l'ocurrència i distribució d'esdeveniments, estats i processos relacionats amb la salut en poblacions especificades, incloent l'estudi dels determinants que influeixen en aquests processos i l'aplicació d'aquest coneixement per controlar problemes de salut rellevants.” És clar que, en aquest sentit, el modelatge matemàtic de l'evolució d'una malaltia infecciosa en funció de diferents paràmetres és un element cabdal, atès que aquest estudi sobre una població determinada permet, juntament amb coneixements sanitaris, anticipar els efectes de la malaltia numèricament i els impactes dels agents mitigans o excitatoris sobre aquesta. Són aquests coneixements els que permeten decidir les accions a prendre per tal de mitigar o extingir una epidèmia.

Per tal de tractar aquest problema, un estudi important va ser el publicat per Kermack i McKendrick[2] el 1927, on dividien la població tancada estudiada en diferents subgrups: aquelles persones que no havien estat contagiades, les malaltes i aquelles que havien superat la malaltia (i, per tant, no eren susceptibles de contagiar-se de nou) o havien mort. Aquestes aproximacions s'anomenen models compartimentals i permeten, en major o menor mesura, introduir simplificacions que facilitin l'estudi de poblacions grans. Una observació adequada al respecte seria, per exemple, considerar epidèmies en què les persones curades puguin tornar a patir la malaltia. En aquest cas, el model compartimental es reduiria a dos subgrups si la malaltia no fos mortal. Clarament, un coneixement precís permet modificar els subgrups per descriure amb més detall el procés en qüestió.

Tornant al primer cas, aquest model és abreujat com a SIR ja que es refereix als subgrups “susceptibles”, “infectats” i “recuperats”. És un model senzill i àmpliament estudiat, i amb el qual els autors van aconseguir, mitjançant equacions diferencials acoblades, relacionar els valors mínims de densitat de població perquè una infecció es propagui per a determinades taxes d'infecció, recuperació i mort.

Finalment, l'últim pas d'aquesta introducció es troba en articles com el de Kendall[3], on es discuteix l'apropament determinista dels primers estudis del model i es proposa l'estocàstica, en un procés markovià discret amb un nombre finit d'estats, per estudiar en detall el model. En concret, aquest nou enfocament permet tractar de forma més precisa poblacions petites on el nombre d'individus no assoleix el límit termodinàmic on l'aproximació determinista és vàlida i, alhora, proporciona resultats físicament plausibles en el nombre d'individus en cada subgrup, doncs les transicions del procés de Màrkov són només entre estats (nombre d'individus en un cert subgrup) que representen nombres enters.

Aquest breu resum històric pretén explicar i motivar el lector sobre la teoria explicada en aquest treball. L'estocàstica és, doncs, una construcció matemàtica que sorgeix en

fenòmens naturals. L'estudi d'aquesta facilita examinar conceptes profunds del comportament indeterminista que s'observa en processos físics i caldrà analitzar-la per descobrir els seus fonaments en l'aplicació a models epidemiològics. Particularment útil serà el càlcul estocàstic que definirà la construcció d'equacions diferencials dins d'aquesta teoria, conceptes clau per les dependències entre quantitat i creixement en dinàmiques poblacionals.

Es pretén comprendre en detall els avenços del model emprant, com a guia, l'estudi dut a terme per Tuckwell i Williams[4] en el qual, el 2007, van realitzar un estudi detallat dels efectes de la variació de les diferents variables del model SIR. Finalment, també es detallarà el model SEIHR proposat per Ferrante, Ferraris i Rovira[5] el 2016 (Susceptibles, malalts latents, malalts infecciosos, malalts no infecciosos i recuperats) que sofisticava l'anterior model adequant-se millor a certs casos epidèmics. En tots dos casos, utilitzarem altres articles per ampliar conceptes de cadascun dels estudis i obtenir una visió més àmplia dels avenços teòrics en els models. Finalment, tractarem de compendiar els resultats específics dels models epidemiològics per aplicar-los a un cas d'estudi que satisfaci les condicions exigides pels primers.

## Estructura de la Memòria

L'objectiu d'aquest treball és comprendre els models actuals d'epidemiologia estocàstica. A tal efecte, es farà una introducció teòrica a aquells conceptes i resultats rellevants de la teoria de processos estocàstics que sigui emprada pels models o que aportï context útil a l'estudi. En particular, es descriuran els processos estocàstics markovians i les seves propietats, es caracteritzaran els processos de difusió i s'aprofundirà en els aspectes teòrics més rellevants de l'integral estocàstica d'Itô i les equacions diferencials estocàstiques.

La segona part del treball consistirà a desenvolupar els models epidemiològics SIR i SEIHR partint de les bases teòriques anteriors seguint els estudis comentats en l'introducció juntament amb altres publicacions que les complementin. També es modelarà un brot de Covid-19 seguint els conceptes teòrics del model SEIHR, en concret l'aproximació de difusió del procés, per tal de visualitzar les aplicacions pràctiques dels models estudiats fent simulacions numèriques.

# Capítol 1

## Processos Estocàstics

### 1.1 Processos Estocàstics discrets i a temps continu

Començarem per introduir els processos estocàstics, on farem una revisió dels diferents tipus, per tal de dotar de context la decisió d'abordar el problema del treball amb unes o altres eines matemàtiques. Trobem una definició intuïtiva d'aquests a [1] segons la qual representen una "seqüència temporal amb elements aleatoris", és a dir, una seqüència de variables aleatòries que trobaran una aplicació directa en la modelització de fenòmens amb una dependència parcial o total amb l'atzar. En són exemples clàssics el moviment d'una partícula de pols en un fluid o les dinàmiques poblacionals. Tal i com podem trobar a diferents textos com [6] i [7] entre d'altres:

**Definició 1.1.1.** *Un **procés estocàstic** amb un espai d'estats  $S$  és una família  $\{X_t, t \in T\}$  de variables aleatòries  $X_t : \Omega \rightarrow S$  en un espai comú de probabilitats  $(\Omega, A, P)$ .*

Si anomenem  $T$  l'espai de paràmetres, distingirem entre *processos discrets* quan  $T$  sigui numerable (per exemple  $T = \mathbb{N}$ ) i *processos continus* quan  $T$  sigui no numerable (per exemple  $T = \mathbb{R}_+$ ). En el nostre cas particular, i sovint en general, associarem el paràmetre  $t$  al temps transcorregut.

De la mateixa manera, anomenarem  $E$  l'espai d'estats (els diferents valors que pot prendre la variable aleatòria  $X_t$ ) i distingirem entre espais d'estats discrets o numerables i continus (no numerables).

**Definició 1.1.2.** *Fixat  $\omega \in \Omega$ , anomenem la funció:*

$$t \rightarrow X_t(\omega)$$

*definida a l'espai de paràmetres  $T$  com una **realització** o **trajectòria** del procés.*

#### 1.1.1 Cas particular: El moviment brownià

Per tal d'exemplificar un procés estocàstic a temps continu introduïrem el procés paradigmàtic les propietats del qual seràn útils per als propers capítols. De fet, és el procés associat al moviment erràtic de les partícules suspeses en un fluid. El moviment brownià modelitza el moviment aleatori continu.

**Definició 1.1.3.** *El **procés de Wiener** (o moviment brownià) és un procés estocàstic  $W_t$  amb  $t \in [0, \infty)$  tal que:*

1.  $W_0 = 0$  quasi segurament.
2. Per qualsevol  $s_1 \leq t_1 \leq s_2 \leq t_2 \leq \dots \leq s_n \leq t_n$  les variables aleatòries  $W_{t_1} - W_{s_1}, \dots, W_{t_n} - W_{s_n}$  són independents.
3. Per qualsevol  $s \leq t$ , la variable aleatòria  $W_t - W_s$  té una distribució normal amb mitjana 0 i variància  $(t - s)\sigma^2$ , és a dir,  $W_t - W_s \sim N(0, (t - s)\sigma^2)$ .
4. Les trajectòries són contínues, és a dir, la funció  $t \rightarrow W_t$  és una funció contínua de  $t$ .

**Observació 1.1.4.** Direm que un procés satisfà la condició (2) si té **increments independents**.

**Observació 1.1.5.** Anomenarem **moviment brownià estàndard** a aquell procés de Wiener amb  $\sigma^2 = 1$ .

**Proposició 1.1.6.** Per qualsevol  $s, t \geq 0$ ,  $E(W_s W_t) = \min(s, t)$ .

*Prova:* Considerem  $s < t$ :

$$E(W_s W_t) = E(W_s(W_t - W_s)) + E(W_s^2) = E(W_s)E(W_t - W_s) + s = s = \min(s, t),$$

emprant les propietats d'increments independents i els resultats de variància i esperança dels processos.

**Proposició 1.1.7.** Les trajectòries del moviment brownià estàndard són quasi segurament contínues però no diferenciables per  $t > 0$ .

Aquesta propietat descriptiva del procés serà útil per comprendre l'abast de les aproximacions numèriques d'aquest.

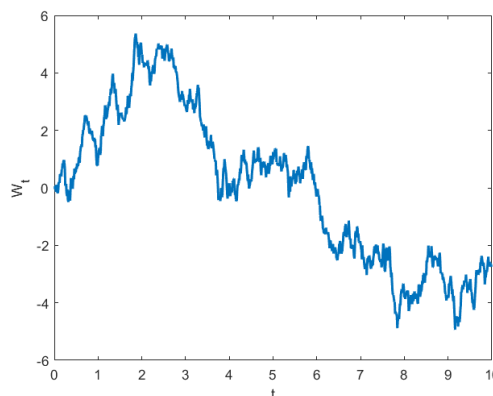


Figura 1.1: Simulació amb el programari *MATLAB* d'una trajectòria del moviment brownià estàndard. Les unitats d'espai ( $W_t$  associat a una coordenada espacial) i temps són arbitràries.

## 1.2 Procés de Màrkov

També resultarà adient l'estudi dels processos de Màrkov ja que podrem entendre fàcilment la seva relació amb l'objectiu d'aquest estudi. En trets generals, aquests processos tenen la propietat de "no tenir memòria", és a dir, que per realitzar prediccions sobre estats futurs, només és necessari conèixer l'estat present. Afegirem conceptes del capítol 1 i 6 de [8] a més dels continguts de les fonts principals anteriors.

**Definició 1.2.1.** Un **procés de Màrkov** és un procés estocàstic  $X_t$ ,  $t \in [0, T]$ , tal que per a tot  $s, t \geq 0$  i  $j \in S$ :

$$P(X_{t+s} = j | X_u, u \leq t) = P(X_{t+s} = j | X_t)$$

i anomenem **Cadena de Màrkov** al procés markovià que pren valors en un espai d'estats numerable  $S$ . Distingirem, doncs, entre temps discrets i continus.

**Observació 1.2.2.** Anomenem **propietat de Màrkov** la propietat d'un procés segons la qual l'estat futur del qual només depèn del seu estat present, això és, en un procés discret:

$$P(X_{n+1} = i_{n+1} | X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = P(X_{n+1} = i_{n+1} | X_n = i_n).$$

**Definició 1.2.3.** Un procés **homogeni en el temps** és aquell en que, fixat  $t \in T$ ,  $P(X_{s+t} = j | X_s = i) = P(X_{u+t} = j | X_u = i) \forall i, j \in E$  i  $\forall t, s, u \in T$ .

**Definició 1.2.4.** Anomenem **Probabilitat de transició** d'un procés homogeni en el temps a la funció  $p_t : t \rightarrow p_t(i, j)$  on  $i, j \in E$  i  $p_t(i, j) = P(X_{s+t} = j | X_s = i)$  i **funció de transició** a la matriu  $P_t = (p_t(i, j))_{(i, j) \in E \times E}$ .

**Lema 1.2.5.** En una cadena de Màrkov a temps discret homogènia  $X_n$ , podem obtenir la matriu de transicions  $n$ -èsima,  $P_n$ , de coeficients  $p_n(i, j) = P(X_n = j | X_0 = i)$ , si coneixem  $P$  de coeficients  $p(i, j) = P(X_1 = j | X_0 = i)$ :

$$P_n = P^n.$$

D'aquests desenvolupaments es desprenen equacions relacionades amb la propietat de Màrkov i que permeten descriure el procés futur i passat.

**Proposició 1.2.6. Equació de Chapman-Kolmogòrov.** Sigui  $X_n$  una cadena de Màrkov a temps discret homogènia, en un espai d'estats  $S$  i siguin  $n, m \in \mathbb{N}$ , aleshores se satisfà:

$$p_{n+m}(i, j) = \sum_{k \in S} p_n(i, k) p_m(k, j)$$

És fàcil comprovar aquesta propietat seguint l'enfocament matricial del Lema 1.2.5:

$$P_{n+m} = P^{n+m} = P^n P^m = P_n P_m.$$

**Definició 1.2.7.** Un **procés de difusió** és un procés de Màrkov continu, amb matriu de transició  $p(s, x; t, y)$  (equivalent continu:  $P(X(t) = y | X(s) = x)$ ) que satisfà l'existència dels següents límits per a tot  $\epsilon > 0, s \geq 0$ :

$$\lim_{t \rightarrow s} \frac{1}{t - s} \int_{|y-x| > \epsilon} p(s, x; t, y) dy = 0,$$

$$\lim_{t \rightarrow s} \frac{1}{t-s} \int_{|y-x| < \epsilon} (y-x)p(s, x; t, y) dy = a(s, x),$$

$$\lim_{t \rightarrow s} \frac{1}{t-s} \int_{|y-x| < \epsilon} (y-x)^2 p(s, x; t, y) dy = b^2(s, x).$$

Anomenem  $a(s, x)$  el terme de desplaçament o "drift" i  $b(s, x)$  el coeficient de difusió a temps  $s$  i "posició"  $x$ .

**Observació 1.2.8.** Clarament, els processos de difusió són aquells que segueixen una trajectòria marcada per un terme determinista, el de desplaçament, i un d'aleatori, el difusiu. En són exemples el **procés de Wiener estàndard**, presentat a l'Observació 1.1.5, amb coeficients  $a(s, x) = 0$ ,  $b(s, x) = 1$  i el d'**Ornstein-Uhlenbeck** amb  $a(s, x) = -x$ ,  $b(s, x) = \sqrt{2}$ .

**Lema 1.2.9.** Del resultat anterior es desprèn una expressió més intuïtiva:

$$a(s, x) = \lim_{t \rightarrow s} \frac{1}{t-s} E(X(t) - X(s) | X(s) = x),$$

$$b^2(s, x) = \lim_{t \rightarrow s} \frac{1}{t-s} E((X(t) - X(s))^2 | X(s) = x).$$

A més, el primer límit evita que el procés tingui salts instantanis.

**Lema 1.2.10. Equació de Chapman-Kolmogòrov contínua.** Sigui  $X(t)$  una cadena de Màrkov a temps continu homogènia, en un espai d'estats  $S$  i siguin  $0 \leq s \leq \tau \leq t$ , aleshores se satisfà:

$$p(s, x; t, y) = \int_{-\infty}^{\infty} p(s, x; \tau, z) p(\tau, z; t, y) dz.$$

**Proposició 1.2.11. Kolmogorov Forward and Backward equations.** Els processos de difusió satisfan:

$$\frac{\partial p}{\partial t} + \frac{\partial}{\partial y} (a(t, y)p) - \frac{1}{2} \frac{\partial^2}{\partial y^2} (b^2(t, y)p) = 0,$$

Kolmogorov forward equation (KFE).

$$\frac{\partial p}{\partial t} + a(s, x) \frac{\partial p}{\partial x} + \frac{1}{2} b^2(s, x) \frac{\partial^2 p}{\partial x^2} = 0,$$

Kolmogorov backward equation (KBE).

Existeix una formulació general de les equacions de Kolmogòrov per processos més generals, els processos d'Itô, però escau comprendre amb profunditat la diferenciabilitat estocàstica.

**Observació 1.2.12.** Els processos de difusió són una eina útil en les aplicacions de la teoria d'equacions diferencials estocàstiques, doncs permeten aproximar processos markovians a temps discret a processos de difusió, és a dir, continus. Aquest tractament s'anomena **Aproximació de difusió** i és d'utilitat quan es volen fer servir resultats d'equacions diferencials estocàstiques en Cadenes de Màrkov.

**Exemple 1.2.13.** Una manera simple d'aproximar una seqüència de Cadenes de Màrkov mitjançant un procés de difusió és particionar els passos de temps en  $n$  instants:  $t_i = i/n$  per  $i = 0, 1, \dots, n$ , on prenem la unitat com a pas de temps. Així, el procés  $X(t)$  esdevé  $X_i^n$  en el qual es genera una passejada aleatòria entre instants de temps on cada pas és de mida  $b(X_i^n)n^{-1/2}$  i s'afegeix un terme fix  $a(X_i^n)n^{-1}$ . Anomenem  $\chi_i$  a la variable aleatòria que pren els valors  $-1$  i  $1$  equiprobablement a cada instant  $i$  (i són independents i idènticament distribuïdes. Així:

$$X_{i+1}^n = X_i^n + a(X_i^n)\frac{1}{n} + b(X_i^n)\frac{1}{\sqrt{n}}\chi_i.$$

Aquest exemple convergirà en distribució a un procés de difusió si les funcions  $a(x), b(x)$  satisfan les propietats de les funcions homònimes del procés de difusió o si convergeixen a aquestes últimes en  $L^1$ . D'aquesta manera, podem formar l'equació diferencial consistent:

$$dX_t = a(X_t)dt + b(X_t)dW_t$$

de la qual en parlarem en detall, properament.

## Capítol 2

# Anàlisi estocàstica

La motivació d'aquest capítol és comprendre com es construeixen les equacions diferencials estocàstiques, unes construccions que sorgeixen naturalment en la recerca de processos que satisfan certes relacions entre el creixement o decreixement i el propi procés. Per tal de comprendre les complexitats de l'adequació de les equacions diferencials a l'estocàstica, estudiarem els treballs del Dr. Kiyosi Itô en aquest àmbit. Els següents resultats són de fàcil abast en diferents textos sobre l'integral d'Itô entre els que es troben [6] i [8].

### 2.1 La integral estocàstica d'Itô

Similarment a l'integral de Riemann, la construcció de l'integral es durà a terme mitjançant l'aproximació per sumes, és a dir, siguin  $0 = t_0 < t_1 < t_2 < \dots < t_{n-1} < t_n = T$  i  $s_j \in [t_j, t_{j+1}]$  :

$$\sum_{j=0}^{n-1} f(s_j) (W(t_{j+1}) - W(t_j)).$$

Ara bé, contràriament a l'integral de Riemann, el resultat d'aquesta suma dependrà de la tria que es faci del punt  $s_j$ . La integral d'Itô soluciona aquesta divergència de resultats demanant que  $f(s_j)$  estigui adaptat a la filtració  $\mathcal{F}_t$ , per tant, prenent  $s_j = t_j$ .

Recordem que l'integral d'Itô s'aproxima per seqüències de sumes de variables aleatòries convergents a  $L^2$  i no a  $\mathbb{R}$ .

**Observació 2.1.1.** Una successió de variables aleatòries a valors reals  $X_0, X_1, \dots$  convergeix a  $X$  en  $L^2$  si  $E(|X_n|^2)$  i  $E(|X|^2)$  existeixen i

$$\lim_{n \rightarrow \infty} E(|X_n - X|^2) = 0.$$

**Definició 2.1.2.** Anomenarem  $f(t)$  amb  $t \geq 0$ , un **procés escalonat aleatori** si existeix una seqüència finita de nombres  $0 = t_0 < t_1 < \dots < t_n$  i variables aleatòries de quadrat integrable  $\eta_0, \eta_1, \dots, \eta_{n-1}$  on  $\eta_j$  és  $\mathcal{F}_{t_j}$ -mesurable per tot  $j = 0, 1, \dots, n-1$ :

$$f(t) = \sum_{j=0}^{n-1} \eta_j \mathbf{1}_{[t_j, t_{j+1})}(t).$$

Denotarem per  $M_{step}^2$  al conjunt de processos aleatoris escalonats.

**Definició 2.1.3.** *L'integral estocàstica d'un procés aleatori escalonat  $f \in M_{step}^2$  està definida per:*

$$I(f) = \sum_{j=0}^{n-1} \eta_j (W(t_{j+1}) - W(t_j)).$$

Una propietat que es desprèn fàcilment de la definició és la següent:

**Proposició 2.1.4.** *Per a qualsevol procés aleatori escalonat  $f \in M_{step}^2$ , la integral estocàstica  $I(f)$  és una variable aleatòria de quadrat integrable, és a dir,  $I(f) \in L^2$ , tal que*

$$E(|I(f)|^2) = E\left(\int_0^\infty |f(t)|^2 dt\right).$$

*Prova:*

Sigui  $\Delta_j W = W(t_{j+1}) - W(t_j)$  i  $\Delta_j t = t_{j+1} - t_j$ . Aleshores:

$$E(\Delta_j W) = 0 \quad \text{i} \quad E((\Delta_j W)^2) = \Delta_j t.$$

Primer, calculem l'esperança de

$$I(f)^2 = \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} \eta_j \eta_k \Delta_j W \Delta_k W = \sum_{j=0}^{n-1} \eta_j^2 (\Delta_j W)^2 + 2 \sum_{j < k} \eta_j \eta_k \Delta_j W \Delta_k W.$$

Donat que  $\eta_j$  i  $\Delta_j W$  són independents:

$$E(\eta_j^2 \Delta_j^2 W) = E(\eta_j^2) E(\Delta_j^2 W) = E(\eta_j^2) \Delta_j t.$$

Si  $k < j$ , aleshores  $\eta_j \eta_k \Delta_k W$  i  $\Delta_j W$  són independents:

$$E(\eta_j \eta_k \Delta_j W \Delta_k W) = E(\eta_j \eta_k \Delta_k W) E(\Delta_j W) = 0.$$

Per tant:

$$E(I(f)^2) = \sum_{j=0}^{n-1} E(\eta_j^2) \Delta_j t.$$

D'aquesta manera, es dedueix que  $I(f) \in L^2$ , ja que  $\eta_0, \eta_1, \dots, \eta_{n-1} \in L^2$ . D'altra banda:

$$I(f)^2 = \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} \eta_j \eta_k 1_{[t_j, t_{j+1}]}(t) 1_{[t_k, t_{k+1}]}(t) = \sum_{j=0}^{n-1} \eta_j^2 1_{[t_j, t_{j+1}]}(t).$$

Cosa que implica que:

$$E\left(\int_0^\infty |f(t)|^2 dt\right) = \sum_{j=0}^{n-1} E(\eta_j^2) \Delta_j t.$$

Finalment:

$$E(|I(f)|^2) = E\left(\int_0^\infty |f(t)|^2 dt\right).$$

**Definició 2.1.5.** *Denotem  $M^2$  la classe de processos estocàstics  $f(t)$ ,  $t \geq 0$  tals que:*

$$E\left(\int_0^\infty |f(t)|^2 dt\right) < \infty$$

i existeix una seqüència  $f_1, f_2, \dots \in M_{step}^2$  de processos escalonats aleatoris tals que:

$$\lim_{n \rightarrow \infty} E \left( \int_0^\infty |f(t) - f_n(t)|^2 dt \right) = 0.$$

Aleshores diem que la seqüència de processos escalonats aleatoris  $f_1, f_2, \dots$  aproximen  $f$  a  $M^2$ .

**Definició 2.1.6.** Anomenem  $I(f) \in L^2$  la **integral estocàstica d'Itô** (de 0 a  $\infty$ ) de  $f \in M^2$  si

$$\lim_{n \rightarrow \infty} E (|I(f) - I(f_n)|^2) = 0$$

per a qualsevol seqüència  $f_1, f_2, \dots \in M_{step}$  de processos escalonats aleatoris que aproximen  $f$  en  $M^2$ . També escriurem

$$\int_0^\infty f(t) dW(t)$$

en lloc d' $I(f)$ .

**Proposició 2.1.7.** L'integral estocàstica de tota funció  $f \in M^2$ ,  $I(f) \in L^2$  existeix i és única (quasi segurament, en tant que element d' $L^2$ ) i satisfà:

$$E (|I(f)|^2) = E \left( \int_0^\infty |f(t)|^2 dt \right).$$

*Prova:*

Sigui  $\|f\|_{M^2} = \sqrt{E(\int_0^\infty |f(t)|^2 dt)}$  i  $\|\eta\|_{L^2} = \sqrt{E(\eta^2)}$  les normes de  $f$  i  $\eta$  en  $M^2$  i  $L^2$  respectivament, per qualsevol  $f \in M^2$  i  $\eta \in L^2$ . Sigui  $f_1, f_2, \dots \in M_{step}^2$  una seqüència de processos escalonats aleatoris aproximant  $f \in M^2$ , aleshores se satisfà:

$$\lim_{n \rightarrow \infty} E \left( \int_0^\infty |f(t) - f_n(t)|^2 dt \right) = \lim_{n \rightarrow \infty} \|f(t) - f_n(t)\|_{M^2}^2 = 0.$$

Per tant:

$$\lim_{n \rightarrow \infty} \|f(t) - f_n(t)\|_{M^2} = 0.$$

Aleshores, donat que  $\forall \epsilon > 0$ ,  $\exists N$  tal que  $\|f - f_n\|_{M^2} < \frac{\epsilon}{2} \forall n > N$ , podem veure que  $I(f_1), I(f_2), \dots$  és una successió de Cauchy en  $L^2$ :

$$\|I(f_m) - I(f_n)\|_{L^2} = \|I(f_m - f_n)\|_{L^2} = \|f_m - f_n\|_{M^2} \leq \|f - f_m\|_{M^2} + \|f - f_n\|_{M^2} < \frac{\epsilon}{2} + \frac{\epsilon}{2} < \epsilon$$

per qualsevol  $m, n > N$ .

Donat que  $L^2$  amb la norma és un espai complet, tota successió de Cauchy té límit. Per tant,  $I(f_1), I(f_2), \dots$  té límit a  $L^2$  per qualsevol seqüència de processos escalonats aleatoris  $f_1, f_2, \dots \in M_{step}^2$  aproximant  $f$ . Ara bé, suposem que dues seqüències de processos escalonats aleatoris  $f_1, f_2, \dots$  i  $g_1, g_2, \dots$  aproximen  $f$  i tenen diferent límit. Aleshores, la seqüència conjunta  $f_1, g_1, f_2, g_2, \dots$  aproxima  $f$  i té límit. Les subseqüències tenen el mateix límit, per tant, aquest és igual per  $f_1, f_2, \dots$  i  $g_1, g_2, \dots$ . Per tant:

$$\lim_{n \rightarrow \infty} \|I(f) - I(f_n)\|_{L^2} = 0.$$

Però sabem que:

$$\|I(f_n)\|_{L^2} = \|I(f_n)\|_{M^2}.$$

Prenent límits a banda i banda:

$$\|I(f)\|_{L^2} = \|I(f)\|_{M^2}.$$

**Observació 2.1.8.** Podem albirar que els increments (diferencials) dels processos de Wiener estan relacionats amb els increments (diferencials) de temps per potències. Entendrem millor aquest concepte en el desenvolupament de la Fòrmula d'Itô.

**Definició 2.1.9.**  $\forall T > 0$  anomenem  $M_T^2$  l'espai de tots els processos estocàstics  $f(t)$ ,  $t \geq 0$  tals que  $1_{[0,T)}f \in M^2$ . L'integral estocàstica (de 0 a  $T$ ) de  $f \in M_T^2$  es defineix com  $I_T(f) = I(1_{[0,T)}f)$  o, alternativament:

$$\int_0^T f(t)dW(t)$$

**Exemple 2.1.10.** Calcularem la integral estocàstica d'Itô per la funció  $f(t) = W_t$ :

$$\int_0^T W_t dW_t$$

Considerem el procés escalonat aleatori:

$$f_n(t) = \sum_{i=0}^{n-1} W_{t_j^n} 1_{[t_j^n, t_{j+1}^n)}$$

on  $0 = t_0 < t_1^n < \dots < t_{n-1}^n < T = t_n$ ,  $t_j^n = \frac{jT}{n}$ .

Aquest procés aproxima  $f(t) = W_t$  en  $M^2$ :

$$\begin{aligned} \lim_{n \rightarrow \infty} E \left( \int_0^T |f(t) - f_n(t)|^2 dt \right) &= \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} \int_{t_j^n}^{t_{j+1}^n} E(|W_t - W_{t_j^n}|^2) dt \\ &= \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} \int_{t_j^n}^{t_{j+1}^n} (t - t_j^n) dt = \lim_{n \rightarrow \infty} \frac{1}{2} \sum_{i=0}^{n-1} (t_{j+1}^n - t_j^n)^2 = \lim_{n \rightarrow \infty} \frac{1}{2} \frac{T}{n} = 0. \end{aligned}$$

Ens cal, doncs, calcular el límit de la integral del procés escalonat en  $L^2$ :

$$\begin{aligned} \int_0^T W_t dW_t &= \lim_{n \rightarrow \infty} I(f_n) = \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} W_{t_j^n} (W_{t_{j+1}^n} - W_{t_j^n}) \\ &= \lim_{n \rightarrow \infty} \left( \frac{1}{2} \sum_{i=0}^{n-1} (W_{t_{j+1}^n}^2 - W_{t_j^n}^2) - \frac{1}{2} \sum_{i=0}^{n-1} (W_{t_{j+1}^n} - W_{t_j^n})^2 \right) = \frac{1}{2} W_t^2 - \frac{1}{2} T. \end{aligned}$$

**Observació 2.1.11.** Comprovem que el resultat intuïtiu riemannianà  $\int_0^t W_s ds = \frac{W_t^2}{2} - \frac{W_0^2}{2}$  no és correcte, tot i que un estudi de les esperances d'ambdós costats ja constatava aquest fet. El terme esquerre té esperança 0 i el terme dret  $\frac{t^2}{2}$ .

**Definició 2.1.12.** Anomenem **Equació Diferencial Estocàstica** a una equació diferencial que descriu un procés influenciat per una força de caràcter indeterminista que pot o no tenir coeficients aleatoris. Són equacions diferencials en forma però representen equacions integrals formalment. Siguin  $a(t, X_t)$ ,  $b(t, X_t)$  qualssevol, l'equació diferencial estocàstica d'un procés  $X_t$  és:

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t,$$

que representa simbòlicament l'equació integral:

$$X_t = X_{t_0} - \int_{t_0}^t a(s, X_s)ds + \int_{t_0}^t b(s, X_s)dW_s,$$

on l'última integral és estocàstica, és a dir, una integral d'Itô.

**Lema 2.1.13. Fòrmula d'Itô.** Sigui  $f$  una funció amb dues derivades contínues i  $W_t$  és un moviment brownià estàndard, aleshores:

$$f(W_t) - f(W_0) = \int_0^t f'(W_s)dW_s + \frac{1}{2} \int_0^t f''(W_s)ds.$$

Equivalentment, en forma diferencial:

$$df(W_t) = f'(W_t)dW_t + \frac{1}{2}f''(W_t)dt.$$

L'intuïció rere aquest resultat es troba estretament relacionada amb la fòrmula de Taylor. Un càlcul senzill anterior permet deduir:

$$E(dW_t)^2 = E((W_{t+dt} - W_t)^2) = dt.$$

Emprant la fòrmula de Taylor sabem que:

$$f(W_t) = f(W_{t_0}) + f'(W_{t_0})(W_t - W_{t_0}) + \frac{1}{2}f''(W_{t_0})(W_t - W_{t_0})^2 + \dots$$

que expressada en forma diferencial resulta:

$$df(W_t) = f'(W_t)dW_t + \frac{1}{2}f''(W_t)(dW_t)^2 + \dots$$

i relacionant-la amb el resultat anterior, utilitzant la fòrmula de Taylor a primer ordre, ens adonem que hem d'incloure la segona derivada (amb la relació estreta entre diferencials) per incloure la dependència en dues variables que té el terme  $f(W_t)$ :

$$df(W_t) = f'(W_t)dW_t + \frac{1}{2}f''(W_t)dt + \dots$$

D'aquesta manera, veiem el terme afegit depenent de  $t$  que sorgeix del càlcul estocàstic en considerar les dues variables de les que depèn una funció  $f(W_t)$ , la part determinista depenent de  $t$  i la part estocàstica (de soroll) depenent de  $W_t$ .

És interessant veure que la resta de termes són d'ordre superior:  $dt dW_t = O(dt^{\frac{3}{2}})$  i que;

$$E(dt dW_t) = dt E(dW_t) = 0.$$

**Definició 2.1.14.** Anomenem **Procés d'Itô** a un procés estocàstic  $\xi_t$  que satisfà l'equació diferencial:

$$d\xi_t = a(t)dt + b(t)dW_t$$

on  $a(t)$  és  $L_t^1$  i  $b(t)$  és  $M_t^2$  per tot  $t \geq 0$

**Lema 2.1.15. Fòrmula d'Itô, cas general.** Sigui  $f$  una funció amb dues derivades contínues i  $\xi_t$  és un procés d'Itô aleshores:

$$f(t, \xi_t) - f(t_0, \xi_0) = \int_0^t (f'_t(t, \xi_t) + f'_x(t, \xi_t)a(t) + \frac{1}{2}f''_{xx}(t, \xi_t)b(t)^2)dt + \int_0^t f'_x(t, \xi_t)a(t)dW_t.$$

Equivalentment, en forma diferencial:

$$df(t, \xi_t) = (f'_t(t, \xi_t) + f'_x(t, \xi_t)a(t) + \frac{1}{2}f''_{xx}(t, \xi_t)b(t)^2)dt + f'_x(t, \xi_t)a(t)dW_t.$$

**Observació 2.1.16.** D'aquests resultats es desprèn que podem escriure un procés de difusió  $X(t)$  amb  $a(s, x)$ ,  $b^2(s, x)$  com aquell procés que satisfà l'equació diferencial:

$$dX(t) = a(x, t)dt + b(x, t)dW_t.$$

## 2.2 Càlcul numèric d'equacions diferencials estocàstiques

Sabem que les equacions diferencials estocàstiques representen una extensió de les deterministes i que aquestes últimes, sovint, no tenen una solució analítica. És aleshores quan ens preguntem si l'equació té solució i emprem els mètodes numèrics per aproximar aquestes solucions si existeixen.

De la mateixa manera que hem parlat dels processos de difusió i les aproximacions de difusió com una eina útil per aproximar processos discrets com a processos continus, ara veurem com la seva contrapart, les aproximacions discretes, o discretitzacions de processos continus, permeten solucionar equacions diferencials estocàstiques numèricament.

Per això, primer, caldrà trobar quines equacions tenen solució, per bé que no sigui analítica. En aquest sentit, ens emmirallarem en els teoremes d'existència i unicitat per equacions amb coeficient Lipschitz que trobem a [9] i continuarem amb els mètodes numèrics de [8]:

**Definició 2.2.1.** *Sigui  $f : X \rightarrow Y$  una funció contínua i  $X, Y$  dos espais normats. Es diu que  $f$  és **Lipschitz contínua** si  $\forall x, y \in X \exists K > 0$  tal que:*

$$|f(x) - f(y)| \leq K|x - y|.$$

**Definició 2.2.2.** *Sigui  $f : X \rightarrow Y$  una funció contínua i  $X, Y$  dos espais normats. Es diu que  $f$  és **localment Lipschitz contínua** si  $\forall |x|, |y| < N \exists K_N > 0$  tal que:*

$$|f(x) - f(y)| \leq K_N|x - y|.$$

La segona propietat Lipschitz, menys forta, ens permet obtenir solucions per a casos concrets, com en les equacions diferencials ordinàries. En el cas estocàstic, també haurérem de restringir els coeficients per trobar solucions generals. Distingirem dos tipus de solucions:

**Definició 2.2.3.** *Direm que una solució de l'equació diferencial és **forta** si té solució per qualsevol procés de Wiener,  $W_t$ , triat i **feble** quan aquesta tingui solució per un parell  $X_t, W_t$  concret.*

**Definició 2.2.4.** *Definirem la **unicitat** d'una solució de l'equació diferencial estocàstica si, donats dos processos estocàstics solució  $X, \hat{X} \in L^2$ , aleshores:*

$$P(X(t) = \hat{X}(t), \forall t, 0 \leq t \leq T) = 1.$$

**Teorema 2.2.5. Existència i unicitat.** *Direm que el problema del valor inicial:*

$$\begin{cases} dX(t) = a(t, X_t)dt + b(t, X_t)dW_t, & 0 \leq t \leq T \\ X(0) = X_0 \end{cases}$$

*associat a una equació diferencial estocàstica té solució forta i és única si es donen simultàniament:*

- a)  $a(t, X_t), b(t, X_t) \in M^2$  són funcions Lipschitz contínues
- b)  $\begin{cases} |a(t, X_t)| \leq L(1 + |X_t|) \\ |b(t, X_t)| \leq L(1 + |X_t|) \end{cases} \quad 0 \leq t \leq T, \text{ per una certa constant } L > 0.$
- c)  $E(X_0) < \infty$  i  $X_0$  és una variable aleatòria independent de  $W_t$ .

*Prova:*

Primer, demostrarem la unicitat donades dues solucions com les descrites i després demostrarem que existeixen aquestes solucions.

*Unicitat:*

Suposem que existeixen dues solucions  $X, \hat{X} \in L^2$  del problema del valor inicial. Aleshores:

$$X_t - \hat{X}_t = \int_0^t a(s, X_s) - a(s, \hat{X}_s) ds + \int_0^t b(s, X_s) - b(s, \hat{X}_s) dW_s$$

Com  $(a + b)^2 \leq 2a^2 + 2b^2$ :

$$E(|X_t - \hat{X}_t|^2) = 2E(|\int_0^t a(s, X_s) - a(s, \hat{X}_s) ds|^2) + 2E(|\int_0^t b(s, X_s) - b(s, \hat{X}_s) dW_s|^2)$$

Emprant la desigualtat de Cauchy-Schwarz ( $|\int_0^t f ds|^2 \leq t \int_0^t |f|^2 ds$  amb  $t > 0$ ) i la propietat Lipschitz:

$$E(|\int_0^t a(s, X_s) - a(s, \hat{X}_s) ds|^2) \leq TE(\int_0^t |a(s, X_s) - a(s, \hat{X}_s)|^2 ds) \leq L^2 T \int_0^t E(|X_s - \hat{X}_s|^2) ds.$$

Per la integral d'Itô, tenim la proposició 2.1.7 i la propietat Lipschitz:

$$E(|\int_0^t b(s, X_s) - b(s, \hat{X}_s) dW_s|^2) = E(\int_0^t |b(s, X_s) - b(s, \hat{X}_s)|^2 ds) \leq L^2 \int_0^t E(|X_s - \hat{X}_s|^2) ds.$$

Per tant, trobem que existeix una constant  $C > 0$  tal que:

$$E(|X_t - \hat{X}_t|^2) \leq C \int_0^t E(|X_s - \hat{X}_s|^2) ds \text{ per a tot } 0 \leq t \leq T.$$

Reanomenant  $\phi(t) = (|X_t - \hat{X}_t|^2)$ , podem veure que  $\phi(t)$  satisfà les premisses del Lema de Gronwall:

$$0 \leq \phi(t) \leq C_0 + C \int_0^t \phi(s) ds \text{ per a tot } 0 \leq t \leq T.$$

Amb el que obtenim que  $\phi(t) \leq 0$  ja que  $C_0 = 0$  i, finalment:

$$P(X(t) = \hat{X}(t), \forall t, 0 \leq t \leq T) = 1.$$

*Existència:*

Emprarem l'esquema següent: 
$$\begin{cases} X^0(t) = X_0 \\ X^{n+1}(t) = X_0 + \int_0^t a(s, X^n(s)) ds + \int_0^t b(s, X^n(s)) dW_s \end{cases}$$
 per  $n = 0, 1, \dots$  i  $0 \leq t \leq T$ .

Definim també  $d^n(t) = E(|X^{n+1}(t) - X^n(t)|^2)$ .

Primer, volem veure que:

$$d^n(t) \leq \frac{(Mt)^{n+1}}{(n+1)!}$$

se satisfà per tot  $n = 0, 1, \dots$  i  $0 \leq t \leq T$ , per alguna constant  $M$  funció de  $L, T$  i  $X_0$ .

Per  $n=0$ :

$$\begin{aligned} d^0(t) &= E(|X^1(t) - X^0(t)|^2) = E(|\int_0^t a(s, X_0)ds + \int_0^t b(s, X_0)dW_s|^2) \\ &\leq 2E(|\int_0^t L(1 + |X_0|)ds|^2) + 2E(|\int_0^t L^2(1 + |X_0|)ds|^2) \leq tM \end{aligned}$$

per una  $M$  prou gran.

Farem el pas inductiu suposant que l'afirmació és certa per  $n - 1$ . Aleshores:

$$\begin{aligned} d^n(t) &= E(|X^{n+1}(t) - X^n(t)|^2) \\ &= E(|\int_0^t a(s, X^n(s)) - a(s, X^{n-1}(s))ds + \int_0^t b(s, X^n(s)) - b(s, X^{n-1}(s))dW_s|^2) \\ &\leq 2TL^2E(\int_0^t |X^n(s) - X^{n-1}(s)|^2ds) + 2L^2E(\int_0^t |X^n(s) - X^{n-1}(s)|^2ds) \\ &\leq 2L^2(1 + T) \int_0^t \frac{M^n s^n}{n!} ds \leq \frac{M^{n+1}t^{n+1}}{(n+1)!} \end{aligned}$$

triant  $M \geq 2L^2(1 + T)$ .

Clarament, hem emprat  $(a + b)^2 \leq 2a^2 + 2b^2$ , la desigualtat de Cauchy-Schwarz, la propietat Lipschitz i la propietat més forta del teorema i finalment hem introduït la hipòtesi d'inducció.

Ara, un cop demostrada la desigualtat, podem veure que:

$$\begin{aligned} &E\left(\max_{0 \leq t \leq T} |X^{n+1}(t) - X^n(t)|^2\right) \\ &\leq 2TL^2 \int_0^T E(|X^n(s) - X^{n-1}(s)|^2)ds + 8L^2 \int_0^T E(|X^n(s) - X^{n-1}(s)|^2)ds \\ &\leq C \frac{(MT)^n}{n!}. \end{aligned}$$

Podem, doncs, utilitzar el Lemma de Borel Cantelli doncs:

$$\begin{aligned} &P\left(\max_{0 \leq t \leq T} |X^{n+1}(t) - X^n(t)| > \frac{1}{2^n}\right) \\ &\leq 2^{2n} E\left(\max_{0 \leq t \leq T} |X^{n+1}(t) - X^n(t)|^2\right) \leq 2^{2n} \frac{C(MT)^n}{n!}, \end{aligned}$$

i

$$\sum_{n=1}^{\infty} 2^{2n} \frac{(MT)^n}{n!} < \infty.$$

Per tant:

$$P\left(\max_{0 \leq t \leq T} |X^{n+1}(t) - X^n(t)| > \frac{1}{2^n} i.o.\right) = 0$$

on *i.o.* indica que succeeix "infinitament sovint".

Finalment, per quasi tots els camins  $\omega$ :

$$X^n(t) = X^0(t) + \sum_{j=0}^{n-1} (X^{j+1}(t) - X^j(t))$$

convergeix uniformement en  $[0, T]$  a un procés  $X(t) = X_t \in L^2$ . Prenent límits en la definició d' $X^{n+1}(t)$  per veure que:

$$X(t) = X_0 + \int_0^t a(s, X_s)ds + \int_0^t b(s, X_s)dW_s$$

per  $0 \leq t \leq T$ . Equivalentment, en forma diferencial:

$$\begin{cases} dX(t) = a(t, X_t)dt + b(t, X_t)dW_t, & 0 \leq t \leq T \\ X(0) = X_0 \end{cases} \quad \text{per } 0 \leq t \leq T.$$

Resta per demostrar  $X_t \in L^2$ :

$$\begin{aligned} & E(|X^{n+1}(t)|^2) \\ & \leq CE(|X_0|^2) + CE(|\int_0^t a(s, X^n(s))ds|^2) + CE(|\int_0^t b(s, X^n(s))dW_s|^2) \\ & \leq C(1 + E(|X_0|^2)) + C \int_0^t E(|X_0|^2)ds \end{aligned}$$

on  $C$  és una constant adequada.

Per inducció:

$$E(|X^{n+1}(t)|^2) \leq \left[ C + C^2 + \dots + C^{n+2} \frac{t^{n+1}}{(n+1)!} \right] (1 + E(|X_0|^2)).$$

amb el que:

$$E(|X^{n+1}(t)|^2) \leq C(1 + E(|X_0|^2))e^{Ct}.$$

Per  $n \rightarrow \infty$ :

$$E(|X(t)|^2) \leq C(1 + E(|X_0|^2))e^{Ct}$$

per tot  $0 \leq t \leq T$ . Veiem que  $X_t \in L^2$ .

**Observació 2.2.6.** Existeixen altres resultats per solucions febles que consideren propietats menys restrictives i que seràn emprats per assegurar l'existència de solucions dels models epidemiològics. En particular, es pot trobar un resultat per funcions localment Lipschitz contínues a [8], Capítol 4.

**Definició 2.2.7.** El *mètode d'Euler-Maruyama* per discretitzar un procés d'Itô  $\xi_t$ ,  $t \in [0, T]$  que satisfà l'equació diferencial:

$$d\xi_t = a(t, \xi_t)dt + b(t, \xi_t)dW_t$$

amb valor inicial  $\xi_0$ , consisteix a prendre una discretització de l'interval de temps considerat:  $0 = \tau_0 < \tau_1 < \dots < \tau_n = T$ , on  $\tau_i = i\frac{T}{n}$ ,  $i = 0, 1, \dots, n$  tal que obtenim un procés discret  $X_i$  que satisfà l'esquema iteratiu:

$$X_{i+1} = X_i + a(\tau_i, X_i)(\tau_{i+1} - \tau_i) + b(\tau_i, X_i)(W_{\tau_{i+1}} - W_{\tau_i})$$

per  $i = 0, 1, \dots, n$  amb valor inicial  $X_0 = \xi_0$  i  $X_i = X(\tau_i)$  per al valor de l'aproximació en el temps discret  $\tau_i$ .

Si reanomenem convenientment  $\Delta t = \frac{T}{n}$ ,  $\Delta W_i = W_{\tau_{i+1}} - W_{\tau_i}$  (aquest últim terme segueix una distribució normal  $N(0, \Delta t)$ ) obtenim l'expressió més comuna:

$$X_{i+1} = X_i + a(\tau_i, X_i)\Delta t + b(\tau_i, X_i)\Delta W_i.$$

**Observació 2.2.8.** Existeixen mètodes més precisos que consideren termes d'ordre superior en  $\Delta t$  i que poden ser més adequats per termes estocàstics grans o altament no lineals com Milstein o Runge-Kutta estocàstic però el mètode d'Euler-Maruyama és senzill d'implementar i intuïtiu.

**Definició 2.2.9.** Una *interpolació constant per trams* permet calcular el valor aproximat del procés original en un temps no contingut per la discretització.

Sigui  $X(t)$  el valor desitjat, considerem  $i_t = \max\{i = 0, 1, \dots, n : \tau_i \leq t\}$ . Aleshores:

$$X(t) = X_{i_t} + \frac{t - \tau_{i_t}}{\tau_{i_t+1} - \tau_{i_t}}(X_{i_t+1} - X_{i_t}).$$

**Observació 2.2.10.** Clarament, les interpolacions no preserven la no diferenciabilitat dels processos d'Itô, del qual en forma part el moviment brownià, per a tot  $t$ . És per això que el focus de l'estudi ha de residir en les aproximacions en els valors discretitzats.

Una aplicació directa de les interpolacions és el càlcul numèric d'equacions diferencials estocàstiques amb endarreriment.

**Definició 2.2.11.** Una *Equació diferencial estocàstica amb endarreriment* és una equació diferencial estocàstica de la forma:

$$d\xi_t = a(t, \xi_t, \xi_{t-i_1}, \xi_{t-i_2}, \dots, \xi_{t-i_m})dt + a(t, \xi_t, \xi_{t-i_1}, \xi_{t-i_2}, \dots, \xi_{t-i_m})dW_t$$

amb  $i_1, i_2, \dots, i_m \in [0, T]$ .

**Observació 2.2.12.** Les especificitats de l'estocàstica amb endarreriment no seràn tractades en aquest treball per motius d'extensió però es poden comprovar les propietats que ens permetran emprarles en el càlcul numèric dels models epidemiològics a [10].

**Definició 2.2.13.** El mètode d'Euler-Maruyama per equacions amb endarreriment, en les mateixes condicions que anteriorment, s'expressa com:

$$X_{i+1} = X_i + a(\tau_i, X_i, X_{i-j_1}, X_{i-j_2}, \dots, X_{i-j_m})\Delta t + a(\tau_i, X_i, X_{i-j_1}, X_{i-j_2}, \dots, X_{i-j_m})\Delta W_i$$

amb la diferència que s'han de considerar els valors de la funció en l'interval  $[-l_{max}, 0]$  on  $l_{max} = \max\{j_l : l = 1, \dots, m\}$  i, si el temps  $t - i_l$  no coincideix amb un múltiple de  $\Delta t$  caldrà obtenir  $X_{i-j_l}$ , les aproximacions dels valors endarrerits, per interpolació.

**Observació 2.2.14.** L'ús d'aquest mètode per aquestes equacions introdueix errors més importants numèricament per les interpolacions i l'emmagatzematge de més dades suposa un major temps de computació.

## Capítol 3

# Models epidemiològics estocàstics

### 3.1 Primers desenvolupaments. Model SIR

Tal i com s'ha esmentat a l'introducció, l'objectiu d'aquest treball és analitzar models actuals d'epidemiologia estocàstica i, a tal efecte, començarem pel model compartimental més estès: l'SIR.

**Definició 3.1.1.** *El model **SIR** (Susceptibles, Malalts i Recuperats) a estudiar [4] es basa en les següents consideracions:*

1. *Donat que les dades epidemiològiques s'actualitzen (com a molt) cada dia, el model estocàstic descriu un procés markovià a temps discret amb un dia com a pas de temps.*
2. *La població és tancada i en ella el nombre d'individus a cada un dels tres subgrups és mutable a cada interval de temps.*
3. *Cada individu interacciona amb un nombre fix sumat a un altre d'aleatori d'altres individus a cada pas  $i$ .*
4. *El contagi es dona en les interaccions entre individus malalts i individus susceptibles i la probabilitat de contagi és la mateixa en tots els contactes.*
5. *La probabilitat de que un individu malalt contagiï un individu susceptible és " $p$ " fixe i la probabilitat de contagiar-se en contacte amb " $y$ " malalts és de  $1 - (1 - p)^y$ .*
6. *Un individu contagiad durant el dia " $n$ " romandrà susceptible fins el dia " $n + 1$ " on esdevindrà malalt, per tant, contagiós.*
7. *Un individu malalt romandrà en aquest estat durant " $R$ " dies (fixos), després adoptarà l'estat "Recuperat" indefinidament.*

**Observació 3.1.2.** El model probabilístic d'interacció entre individus és coneguda com una **Cadena Binomial**. És un model dinàmic que assumeix que les interaccions es donen en unitats discretes de temps, produint una cadena d'infeccions seguint la distribució probabilística binomial:  $f(k, n, p) = \binom{n}{k} p^k (1 - p)^{n-k}$  a cada pas de temps canviant, però, les variables en funció de l'evolució del nombre d'infectats.

**Observació 3.1.3.** Es poden fer diversos tractaments respecte a la capacitat de contagi d'un individu a cada pas de temps però els més extesos en són dos: el de Greenwood i el de Reed-Frost. El de **Greenwood** que assumeix que la probabilitat de contagi és constant si es té contacte amb un malalt o amb més i el de **Reed-Frost** que suposa que la probabilitat de contagi augmenta en funció de la quantitat d'individus infectats amb el que un susceptible té contacte.

Totes dues aproximacions responen a raonaments mèdics: la primera, segueix la intuïció de que en una sala, l'ambient s'omple de partícules infeccioses tant si hi ha un individu malalt com si n'hi ha més (l'exposició a un individu malalt és igual que l'exposició a més individus) i la segona, assumeix que les interaccions amb malalts són independents (per tant a més interaccions amb malalts, més probabilitat de contagi). En aquest estudi se seguirà el segon apropament.

Trobar la probabilitat de contagi d'un individu és clau ja que, en el context d'una cadena de Màrkov a temps discret, dona resposta a les probabilitats de transició entre estats d'un individu  $i$ , en conseqüència, de tota la població, com veurem en els propers resultats.

**Proposició 3.1.4.** *Segui  $y$  el nombre de malalts,  $N_i(t)$  el nombre d'interaccions de l'individu  $i$ -èsim i  $n$  el nombre d'individus, la probabilitat de que l'individu es trobi amb  $j$  malalts és, considerant  $n$  prou gran:*

$$P_i^j(y, N_i(t); n) \simeq \binom{N_i(t)}{j} \left( \frac{y}{n-1} \right)^j \left( 1 - \frac{y}{n-1} \right)^{N_i(t)-j}.$$

Considerant que la probabilitat de contagiar-se donades  $j$  interaccions amb individus malalts és de  $p_j = 1 - (1-p)^j$ , podem veure que la probabilitat de contagi és:

$$P_{\text{contagi}} = \sum_{j=1}^{N_i(t)} p_j P_i^j(y, N_i(t); n) \simeq 1 - \left( 1 - \frac{py}{n-1} \right)^{N_i(t)}.$$

*Prova:*

Podem observar que:

$$P_{\text{contagi}} = \sum_{j=1}^{N_i(t)} p_j P_i^j(y, N_i(t); n) = \sum_{j=1}^{N_i(t)} (1-(1-p)^j) \binom{N_i(t)}{j} \left( \frac{y}{n-1} \right)^j \left( 1 - \frac{y}{n-1} \right)^{N_i(t)-j}.$$

D'on podem distingir dos termes: el sumatori de  $P_i^j$  sobre tots els possibles  $j$  excepte  $j = 0$ , és a dir,  $1 - (1 - \frac{y}{n-1})^{N_i(t)}$ , i el sumatori de  $(1-p)^j P_i^j$ :

$$\begin{aligned} & \sum_{j=1}^{N_i(t)} (1-p)^j \binom{N_i(t)}{j} \left( \frac{y}{n-1} \right)^j \left( 1 - \frac{y}{n-1} \right)^{N_i(t)-j} \\ &= \sum_{j=1}^{N_i(t)} \binom{N_i(t)}{j} \left( (1-p) \frac{y}{n-1} \right)^j \left( 1 - \frac{y}{n-1} \right)^{N_i(t)-j} \\ &= \sum_{j=1}^{N_i(t)} \binom{N_i(t)}{j} (a)^j (b)^{N_i(t)-j} \end{aligned}$$

on  $a = (1-p)\frac{y}{n-1}$ ,  $b = 1 - \frac{y}{n-1}$ , amb el que trobem una clara relació amb el binomi de Newton:  $(a+b)^n = \sum_{j=0}^n \binom{n}{j} a^j b^{n-j}$ . Caldrà considerar el terme  $j=0$  per assolir el càlcul final:

$$\sum_{j=1}^{N_i(t)} \binom{N_i(t)}{j} (a)^j (b)^{N_i(t)-j} = (a+b)^{N_i(t)} - b^{N_i(t)} = \left(1 - p\frac{y}{n-1}\right)^{N_i(t)} - \left(1 - \frac{y}{n-1}\right)^{N_i(t)}.$$

Per tant, prenent l'aproximació de la probabilitat de trobada d'un individu  $i$  amb  $j$  individus malalts considerada anteriorment:

$$P_{contagi} \simeq 1 - \left(1 - \frac{py}{n-1}\right)^{N_i(t)}.$$

**Observació 3.1.5.** Clarament estem simplificant el model assumint que no hi ha interaccions més favorables que d'altres i, a més, es pot donar el cas paradoxal que un individu interactui amb un altre i el darrer no interactui amb el primer. Per  $n$  petit, les interaccions entre individus no poden simplificar-se com independents idènticament distribuïdes, doncs les interaccions d'un determinen les de la resta.

**Proposició 3.1.6.** *Sigui  $X(t)$ ,  $Y(t)$  i  $Z(t)$  el nombre d'individus susceptibles, malalts i recuperats a temps  $t$  respectivament, considerem els processos markovians  $X^i(t)$ ,  $Y_j^i(t)$ ,  $\forall j \in \{0, \dots, R-1\}$  i  $Z^i(t)$  (les  $R$  variables "Y", representen els  $R$  dies que un individu roman malalt) que només prenen dos valors possibles: 0 si l'individu  $i$ -èsim no es troba en l'estat associat amb aquest procés a temps  $t$  (susceptible per  $X^i(t)$ , primer dia infectat per  $Y_0^i(t)$ , ...) i 1 si sí s'hi troba, és a dir, l'individu sempre pren el valor 1 en algun estat i 0 a la resta. Aleshores, aquests estats conformen una cadena de Màrkov a temps discret:*

$$\begin{aligned} Y_j^i(t) = 1 &\rightarrow Y_{j+1}^i(t+1) = 1, & j \leq R-2, \\ Y_{R-1}^i(t) = 1 &\rightarrow Z^i(t+1) = 1, \\ Z^i(t) = 1 &\rightarrow Z^i(t+1) = 1, \end{aligned}$$

de manera unívoca, amb l'única transició no trivial (considerant l'anterior proposició), la de susceptible a susceptible i de susceptible a malalt per primer dia:

$$\begin{aligned} P(X^i(t+1) = 1 | X^i(t) = 1) &= \left(1 - \frac{py}{n-1}\right)^{N_i(t)}, \\ P(Y_0^i(t+1) = 1 | X^i(t) = 1) &= 1 - \left(1 - \frac{py}{n-1}\right)^{N_i(t)}. \end{aligned}$$

**Lema 3.1.7.** *De l'anterior proposició, podem definir un nou procés markovià que denoti si l'individu  $i$ -èsim es troba malalt o no (independentment dels dies que porti en aquest estat) en un determinat temps  $t$ :  $Y^i(t)$ . Així, deduïm:*

$$Y^i(t) = \sum_{j=0}^{R-1} Y_j^i(t), \quad X(t) = \sum_{i=0}^n X^i(t), \quad Y(t) = \sum_{i=0}^n Y^i(t), \quad Z(t) = \sum_{i=0}^n Z^i(t).$$

**Lema 3.1.8.** *Podem calcular la probabilitat que l'infecció no es transmeti a nous individus "Q<sub>0</sub>" fàcilment si considerem el nombre d'interaccions constant  $n_i$  i un nombre inicial de contagiats  $y_0$ :*

$$Q_0 = \left[ \prod_{i=1}^{n-y_0} \left(1 - \frac{py}{n-1}\right)^{n_i} \right]^R.$$

Perquè no hi hagi contagi, cadascun dels individus susceptibles a temps  $t$ , amb el seu identificador "i" respectivament, han de trobar-se en el cas  $X^i(t+1) = 1$  de la Proposició 3.1.6. Per tant, la probabilitat de que aquests fets succeeixin simultàniament, éssent independents, és la multiplicació de probabilitats.

### 3.1.1 SIR amb $R=\infty$

Un cas particular de l'SIR és aquell on els malalts no es recuperen mai de l'afecció, això és,  $R = \infty$ . Aquest cas, útil per comprendre extrems del model, té exemples coneguts com el VIH, herpes i certes hepatitis en les quals el malalt només disposa de pal·liatius i l'infecció es pot seguir produint.

Tal i com veurem, mitjançant la recurrent aproximació segons la qual els individus interactuen amb un nombre fixat d'altres individus, podrem estudiar l'esperança i variància de variables relacionades amb el nombre de malalts fet que resultava massa complex dins del marc d'estudi per al model més general. De retruc, aquests resultats ens permetran endinsar-nos en la teoria de l'estocàstica diferencial amb la qual emprarem una aproximació de difusió.

**Proposició 3.1.9.** *Segui  $N$  el nombre d'interaccions de cada individu fixe i comú, aleshores:*

$$E(Y(t+1) - Y(t) | Y(t) = y) = (n - y) \left[ 1 - \left( 1 - \frac{py}{n-1} \right)^N \right],$$

$$Var(Y(t+1) - Y(t) | Y(t) = y) = (n - y) \left[ 1 - \left( 1 - \frac{py}{n-1} \right)^N \right] \left( 1 - \frac{py}{n-1} \right)^N.$$

*Prova:*

Per l'homogeneïtat dels contactes, la probabilitat de que un individu es contagiï en un cert temps "t" donats  $y$  individus malalts és:

$$\hat{p} = 1 - \left( 1 - \frac{py}{n-1} \right)^N.$$

Per tant:

$$P(Y(t+1) = y + \hat{y} | Y(t) = y) = \binom{n-y}{\hat{y}} \left( 1 - \left( 1 - \frac{py}{n-1} \right)^N \right)^{\hat{y}} \left( 1 - \frac{py}{n-1} \right)^{N(n-y-\hat{y})}.$$

Llavors, donat que ens trobem davant d'una distribució binomial, el tractament és estàndard:

$$\begin{aligned} E(Y(t+1) - Y(t) | Y(t) = y) &= \sum_{\hat{y}=0}^{n-y} \hat{y} \binom{n-y}{\hat{y}} \left( 1 - \left( 1 - \frac{py}{n-1} \right)^N \right)^{\hat{y}} \left( 1 - \frac{py}{n-1} \right)^{N(n-y-\hat{y})} \\ &= (n-y) \left[ 1 - \left( 1 - \frac{py}{n-1} \right)^N \right]. \end{aligned}$$

Igualment per  $Var(Y(t+1) - Y(t) | Y(t) = y)$ .

Els resultats de l'anterior proposició poden fer pensar que per una població suficientment gran (per tant un alt nombre d'estats) amb una probabilitat de transmissió moderada tal que  $nNp$  sigui de mida moderada, podrem aproximar el procés per un de difusió i, per tant, continu. Vegem-ho:

**Definició 3.1.10.** *Segui  $\hat{Y}^n(t) = \frac{Y([nt])}{n}$  per tot  $t \geq 0$ , on  $[\cdot]$  és la part sencera del nombre, aleshores el procés  $\hat{Y}^n(t)$  proporciona a fracció de la població infectada en un temps  $[nt]$ .*

**Proposició 3.1.11.** *Segui  $\theta = nNp$  de mida moderada, aleshores el procés  $\hat{Y}^n(t) = \frac{Y([nt])}{n}$  es pot aproximar per un procés de difusió  $\hat{Y}(t)$ .*

*Prova:*

Podem emprar l'aproximació  $1 - (1 - x)^N \simeq Nx$  per  $x$  properes a 0 i, sabent que  $\Delta t = \frac{1}{n}$  i  $t = 0, \frac{1}{n}, \frac{2}{n}; \dots$ :

$$E(\hat{Y}^n(t + \Delta t) - \hat{Y}^n(t) | \hat{Y}^n(t) = \hat{y}) \simeq (1 - \hat{y})Np\hat{y} = \theta\hat{y}(1 - \hat{y})\Delta t,$$

$$Var(\hat{Y}^n(t + \Delta t) - \hat{Y}^n(t) | \hat{Y}^n(t) = \hat{y}) \simeq (1 - \hat{y})Np\hat{y}(1 - Np\hat{y}) \simeq \frac{\theta}{n}\hat{y}(1 - \hat{y})\Delta t.$$

Aleshores és clar que els termes del procés de difusió existeixen i són:  $a(t, \hat{y}) = \theta\hat{y}(1 - \hat{y})$  i  $b^2(t, \hat{y}) = \frac{\theta}{n}\hat{y}(1 - \hat{y})$ .

**Lema 3.1.12.** *Per ser un procés de difusió amb terme de desplaçament  $a(t, \hat{y}) = \theta\hat{y}(1 - \hat{y})$  i coeficient de difusió  $b^2(t, \hat{y}) = \frac{\theta}{n}\hat{y}(1 - \hat{y})$  satisfà l'equació diferencial:*

$$d\hat{Y}(t) = \theta\hat{Y}(t)(1 - \hat{Y}(t))dt + \sqrt{\frac{\theta}{n}\hat{Y}(t)(1 - \hat{Y}(t))}dW(t)$$

i la KBE:

$$\frac{\partial p}{\partial t} = -\theta \frac{\partial}{\partial \hat{y}}(\hat{y}(1 - \hat{y})p) + \frac{\theta}{2n} \frac{\partial^2}{\partial \hat{y}^2}(\hat{y}(1 - \hat{y})p).$$

**Observació 3.1.13.** Podem comprovar l'existència de solucions observant que són funcions localment Lipschitz contínues.

Per tal de simplificar l'equació diferencial anterior, podem considerar que l'efecte del soroll no és rellevant i, per tant, reduir-nos a una equació diferencial clàssica.

**Proposició 3.1.14.** *La mitjana de  $\hat{Y}^n(t)$  es pot aproximar, en un apropament determinista, menyspreant el terme de soroll (factor  $dW_t$ ), per:*

$$m(t) = \frac{1}{1 + \frac{1-m_0}{m_0} \exp(-\theta t)} \quad \forall t \geq 0$$

on  $m_0 = E(\hat{Y}^n(0)) = \frac{1}{n}E(Y(0))$ .

Per tant:

$$E(Y(t)) = nE\left(\hat{Y}^n\left(\frac{t}{n}\right)\right) \simeq nm\left(\frac{t}{n}\right) = \frac{n}{1 + \frac{n-y_0}{y_0} \exp(-pNt)}$$

on  $y_0 = E(Y(0))$  Prova:

Eliminant el terme de soroll (factor associat a  $dW_t$ ) i reanomenant  $\hat{Y}(t) = m(t)$  a l'equació del lema 3.1.12, ens queda:

$$dm(t) = \theta m(t)(1 - m(t))dt$$

La solució de la qual, sabent que  $m(0) = m_0$  és directa.

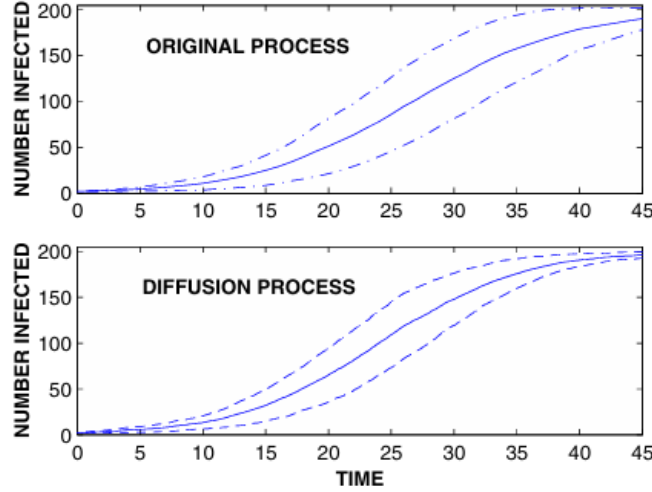


Figura 3.1: Gràfics de l'evolució del nombre d'infectats respecte el temps per una població de  $n = 200$  individus, una probabilitat de contagi de  $p = 0.05$  per contacte,  $N = 4$  contactes diaris i  $Y(0) = 2$  individus infectats inicialment. En línia contínua es representa la mitjana de les simulacions i en línia discontinua la desviació típica. A dalt, el gràfic de la mitjana de 50 simulacions del procés markovià discret original. A baix, el gràfic de la mitjana de 50 simulacions de l'aproximació de difusió del model. *Imatges extretes de [4], Fig. 4.*

Se'n desprèn un resultat interessant aïllant el temps de l'equació primera de la Proposició 3.1.14:

**Lema 3.1.15.** *El temps  $t_\alpha$  que es necessita per assolir un cert percentatge (fracció de la població) de malalts  $\alpha$  dins de la població total sota aquesta aproximació:*

$$t_\alpha = \frac{1}{pN} \ln \left( \frac{\frac{n}{y_0} - 1}{\frac{1}{\alpha} - 1} \right)$$

aïllant  $t_\alpha$  de l'equació  $m\left(\frac{t_\alpha}{n}\right) = \alpha$ .

**Observació 3.1.16.** D'aquesta aproximació logarítmica es desprenen resultats intuïtius sobre l'evolució de l'epidèmia estudiada:

1. Per una mida de població, un nombre de contactes diaris i un nombre de malalts inicials fixats, el temps que triga un cert percentatge de la població a infectar-se és inversament proporcional a la probabilitat d'infecció per contacte malalt susceptible.

2. Així mateix, fixant la probabilitat d'infecció per contacte malalt-susceptible juntament amb la mida de la població i el nombre de malalts inicial, veiem que el temps a assolir un cert percentatge de malalts també és inversament proporcional al nombre de contactes diaris.
3. Suposant que la mida de la població és molt més gran que el nombre d'infectats inicial, podem aproximar:

$$t_\alpha \simeq \frac{1}{pN} \left( \ln \left( \frac{n}{y_0} \right) - \ln \left( \frac{1}{\alpha} - 1 \right) \right)$$

i així, fixant el nombre de contactes, la probabilitat de contagi per contacte malalt-susceptible i la mida de la població, el temps que triga la meitat de la població a contagiar-se és proporcional al logaritme de l'invers del nombre inicial de malalts.

### 3.1.2 SIR amb $R < \infty$ . Model general.

Un afegit incorporat a [5] és el càlcul d'un valor rellevant descriptiu de l'evolució de l'epidèmia:

**Definició 3.1.17.**  $R_0$  és el valor que denota quants individus susceptibles contagia de mitjana un malalt durant la fase infecciosa de la malaltia. Es diu que una malaltia és persistent si  $R_0 > 1$ .

En aquest treball, centrarem l'estudi al valor d'aquest coeficient a l'inici de la malaltia, és a dir, veurem quants individus contagia, de mitjana, l'individu infeccios inicial dins una població, amb l'excepció d'aquest, plenament susceptible.

**Proposició 3.1.18.** Sigui  $R_0^R$  el coeficient  $R_0$  per un període de latència de  $R$  dies i considerem  $n - 1$  individus susceptibles i 1 individu infeccios a  $t = 0$ , aleshores:

$$\begin{aligned} R_0^1 &= (n - 1) \left( 1 - \left( 1 - \frac{py}{n - 1} \right)^N \right) = pN + O \left( \frac{1}{n - 1} \right), \\ R_0^2 &= (n - 1) \left( 1 - \left( 1 - \frac{py}{n - 1} \right)^{2N} \right) = 2pN + O \left( \frac{1}{n - 1} \right), \end{aligned}$$

i per  $R > 2$  amb  $n$  prou gran:

$$R_0^R = RpN + O \left( \frac{1}{n - 1} \right).$$

*Prova:*

Per  $R = 1$ , només cal comptar els contagis de l'únic individu infeccios, en una única iteració. Recordem que el nombre de contagiats per tots els individus infecciosos a cada pas de temps és  $Y_0(t)$ . Donat que coneixem l'esperança del nombre de contagis a  $t + 1$  donats  $y$  contagis a  $t$ , anomenem  $C(1)$  el nombre de contagiats per l'individu en el primer pas de temps:

$$R_0^1 = E(C(1)) = (n - 1)p(1) = (n - 1) \left( 1 - \left( 1 - \frac{p}{n - 1} \right)^N \right) = pN + O \left( \frac{1}{n - 1} \right).$$

Per  $R=2$ , l'individu contagiarà en dos temps i donarà lloc a  $C(1)$  i  $C(2)$  contagis.

$$E(C(2)) = E(E(C(2)|Y_0(1))) = E((n-1-Y_0(1))p(1)) = (n-1)(p(1)-p^2(1)),$$

$$\begin{aligned} R_0^2 &= E(C) = E(C(1)) + E(C(2)) = (n-1)(2p(1)-p^2(1)) \\ &= (n-1)\left(1 - \left(1 - \frac{p}{n-1}\right)^{2N}\right) = 2pN + O\left(\frac{1}{n-1}\right). \end{aligned}$$

Clarament fem  $p(1)$  a  $E(C(2))$  per no considerar els contagiats pels infectats en el primer pas de temps.

Finalment, per  $R=3$ :

$$\begin{aligned} E(C(3)) &= E(E(C(3)|Y_0(2), Y_0(1))) = E((n-1-Y_0(1)-Y_0(2))p(1)) = \\ &= (n-1)p(1) - p(1)E(C(1)) - p(1)E(C(2)) = (n-1)p(1) + O\left(\frac{1}{n-1}\right) \end{aligned}$$

on el nombre d'infectats a la segona iteració és el total (no només el del primer individu contagiós). Per tant s'ha de considerar  $p(Y_0(1)+1)$  per tal de calcular  $E(Y_0(2))$ :

$$E(Y_0(2)) = E(E(Y_0(2)|Y_0(1))) = E((n-1-Y_0(1))p(Y_0(1)+1)).$$

Ara bé, com  $p(y) = 1 - (1 - \frac{py}{n-1})^N \leq \frac{Npy}{n-1}$  considerant el primer terme de l'expansió de Taylor  $(1-x)^n$  per  $x$  prou petit:

$$\begin{aligned} p(1)E(Y_0(1)) &= (n-1)p(1)^2 \leq \frac{p^2N^2}{n-1}, \\ p(1)E(Y_0(2)) &\leq p(1)(n-1)E(p(Y_0(1)+1)) \leq p(1)(n-1)\frac{pNE(Y_0(1)+1)}{n-1} \\ &= pN(p(1)E(Y_0(1)) + p(1)) \leq \frac{pN}{n-1}(p^2N^2 + pN). \end{aligned}$$

Amb el que arribem a:

$$R_0^3 = 3(n-1)p(1) + O\left(\frac{1}{n-1}\right) = 3pN + O\left(\frac{1}{n-1}\right).$$

La generalització es troba calculant  $E(C(k))$  per  $1 \leq k \leq r$  on  $r > 3$ :

$$\begin{aligned} E(C(k)) &= E(E(C(k)|Y_0(k-1), \dots, Y_0(1))) = E((n-1 - \sum_{l=1}^{k-1} Y_0(l))p(1)) \\ &= (n-1)p(1) - p(1) \sum_{l=1}^{k-1} E(Y_0(l)) = (n-1)p(1) + O\left(\frac{1}{n-1}\right) \end{aligned}$$

ja que els termes  $p(1)E(Y_0(l))$  per  $l \geq 3$  donen resultats similars als d' $l=1$  i  $l=2$  amb una dependència amb  $n-1$ .

Per tant:

$$R_0^r = rpN + O\left(\frac{1}{n-1}\right).$$

### 3.1.3 SIR amb nombre de contactes diaris variable

En aquest apartat, veurem els resultats teòrics que sorgeixen en no fixar el nombre de contactes diaris i fer-lo aleatori i dependent del temps. La sofisticació d'aquesta característica present a [4], és descrita en profunditat a [11] i descriu comportaments socials observables com el canvi d'hàbits entre estacions (més activitat interpersonal a l'estiu i major reclusió durant l'hivern) o la conscienciació sobre el risc de contagi per contactes que es duu a terme, progressivament, durant l'evolució d'una malaltia.

**Definició 3.1.19.** *El model SIR amb nombre de contactes variable pren la notació i les consideracions de la Definició 3.1.1 però afegeix:*

1. *Les variables  $N^i(t)$ , que representen el nombre d'individus aleatoris que interactuen amb l'individu  $i$ -èsim en el període de temps  $(t, t + 1]$ , són variables independents entre elles i independents de l'estat de la població.*
2. *Es considera el mateix nombre de trobades per tots els individus, això és:  $N^i(t) = N(t) \forall i$ , donat que no existeixen distincions entre individus tret de la pertinença a les diferents classes del model.*
3.  *$N(t)$  pot prendre només un conjunt finit de valors amb probabilitat canviant amb el temps, això és:*

$$P(N^i(t) = n_k) = p_k(t) \quad \forall i, t$$

$$\text{on } k = 1, \dots, m \text{ i } \sum_{k=1}^m p_k(t) = 1 \quad \forall t.$$

El següent pas donada la característica cadena de Màrkov a temps discret és estudiar la transició no trivial  $P(Y_0^i(t+1) = 1 | X^i(t) = 1)$ , és a dir, quina és la probabilitat d'infecció d'un individu susceptible a temps  $t$  donats  $Y(t) = y$  infectats, una probabilitat de transmissió per contacte  $p$ , un paràmetre d'aïllament, o quarantena,  $\lambda \in [0, 1]$  dels infectats i el nombre de contactes  $N(t)$ .

**Proposició 3.1.20.** *La probabilitat d'infecció d'un individu "i" susceptible a temps  $t$  per  $Y(t) = y$ ,  $p$  i  $\lambda$  és:*

$$p(t, y) = 1 - \sum_{k=1}^m \left(1 - \frac{p\lambda y}{n-1}\right)^{n_k} p_k(t)$$

$$\text{per } y < n-1 \text{ i } p(t, n-1) = 1 - \sum_{k=1}^m (1-p)^{n_k} p_k(t).$$

*Prova:*

Primer calcularem la probabilitat de trobar  $j$  individus infectats fixat un nombre de contactes diaris  $N(t) = n_k$ . Assumint  $n > N(t)$ , prendrem l'aproximació binomial, les conseqüències de la qual ja hem tractat a 3.1.5:

$$p_j^i(y, n_k; n) \simeq \binom{n_k}{j} \left(\frac{\lambda y}{n-1}\right)^j \left(1 - \frac{\lambda y}{n-1}\right)^{n_k-j}$$

quan  $y < n-1$  i  $p_{n_k}^i(y, n_k; n) = 1$ ,  $p_j^i(y, n_k; n) = 0$  per  $j = 0, \dots, n_k-1$  quan  $y = n-1$ .

Així, donat que la probabilitat de contagi si s'ha estat en contacte amb  $j$  individus infectats és  $p_j = 1 - (1-p)^j$ , la probabilitat de contagi serà la suma de les probabilitats

de contagi per cada nombre de contactes diaris possible ponderats. Vegem-ho:

$$\begin{aligned}
p(t, y) &= P(Y_0^i(t+1) = 1 \mid X^i(t) = 1, Y(t) = y) \\
&= \sum_{k=1}^m \sum_{j=0}^{n_k} p_j^i(y, n_k; n) p_j p_k(t) \\
&= \sum_{k=1}^m \sum_{j=0}^{n_k} \binom{n_k}{j} \left( \frac{\lambda y}{n-1} \right)^j \left( 1 - \frac{\lambda y}{n-1} \right)^{n_k-j} [1 - (1-p)^j] p_k(t) \\
&= \sum_{k=1}^m p_k(t) \sum_{j=0}^{n_k} \binom{n_k}{j} \left( \frac{\lambda y}{n-1} \right)^j \left( 1 - \frac{\lambda y}{n-1} \right)^{n_k-j} \\
&\quad + \sum_{k=1}^m p_k(t) \sum_{j=0}^{n_k} \binom{n_k}{j} \left( \frac{(1-p)\lambda y}{n-1} \right)^j \left( 1 - \frac{\lambda y}{n-1} \right)^{n_k-j} \\
&= 1 - \sum_{k=1}^m \left( 1 - \frac{p\lambda y}{n-1} \right)^{n_k} p_k(t).
\end{aligned}$$

per  $y < n-1$  i, similarment,  $p(t, n-1) = 1 - \sum_{k=1}^m (1-p)^{n_k} p_k(t)$ .

En un apropament semblant al realitzat a l'SIR amb  $R = \infty$ , volem trobar una variable que ens permeti descriure el comportament global de la malaltia i estudiar-ne propietats. Calcularem l'evolució conjunta del nombre d'individus infectats  $Y(t)$  i recuperats  $Z(t)$ , és a dir, el nombre d'individus no susceptibles a cada instant de temps  $t$ ,  $V(t) = Y(t) + Z(t)$ . Clarament, els individus que hagin estat no susceptibles durant, com a mínim,  $R$  dies, seràn recuperats i la resta infecciosos. Així:

$$Y(t) = V(t) - V(t-R), \quad Z(t) = V(t-R)$$

i el nombre de nous infectats a temps  $t+1$  és  $V(t+1) - V(t)$ .

**Proposició 3.1.21.** *L'esperança i la variància de  $V(t+1) - V(t)$  són respectivament:*

$$\begin{aligned}
E[V(t+1) - V(t) \mid Y(t) = y, Z(t) = z] &= (n - y - z) \left[ 1 - \sum_{k=1}^m \left( 1 - \frac{p\lambda y}{n-1} \right)^{n_k} p_k(t) \right] \\
\text{Var}[V(t+1) - V(t) \mid Y(t) = y, Z(t) = z] \\
&= (n - y - z) \left[ 1 - \sum_{k=1}^m \left( 1 - \frac{p\lambda y}{n-1} \right)^{n_k} p_k(t) \right] \left[ \sum_{k=1}^m \left( 1 - \frac{p\lambda y}{n-1} \right)^{n_k} p_k(t) \right]
\end{aligned}$$

quan  $y < n-1$ ,

$$\begin{aligned}
E[V(t+1) - V(t) \mid Y(t) = n-1, Z(t) = 1] &= 0 \\
\text{Var}[V(t+1) - V(t) \mid Y(t) = n-1, Z(t) = 1] &= 0
\end{aligned}$$

per  $y = n-1$ ,  $z = 1$  i

$$\begin{aligned}
E[V(t+1) - V(t) \mid Y(t) = n-1, Z(t) = 0] &= p(t, n-1) \\
\text{Var}[V(t+1) - V(t) \mid Y(t) = n-1, Z(t) = 0] &= (1 - p(t, n-1))p(t, n-1).
\end{aligned}$$

per  $y = n-1$ ,  $z = 0$ , éssent aquests dos últims els casos extrems.

*Prova:*

Donat que:

$$\begin{aligned} P(V(t+1) = w + y + z | Y(t) = y, Z(t) = z) \\ = P(V(t+1) = w + y + z | V(t) - V(t-R) = y, V(t-R) = z) \\ = \binom{n-y-z}{w} p(t, y)^w (1-p(t, y))^{n-y-z-w}, \end{aligned}$$

podem veure que el nombre de nous infectats a cada pas de temps segueix una llei binomial:

$$V(t+1) - V(t) | Y(t) = y, Z(t) = z \sim \text{Binom}(n - y - z, p(t, y)).$$

Els resultats d'esperances i variàncies se'n deriven naturalment.

Tal i com en els altres apartats, el càlcul de l'esperança i variància del model, indiquen la possibilitat de fer una aproximació de difusió però, en aquest cas, amb la variable del total de no-susceptibles  $V(t)$  donat que el període de permanença a la classe  $I$  és finit.

**Definició 3.1.22.** *Sigui  $\hat{V}^n(t) = \frac{V([nt])}{n}$  per tot  $t \geq 0$ , on  $[\cdot]$  és la part sencera del nombre, aleshores el procés  $\hat{V}^n(t)$  proporciona a fracció de la població no-susceptible en un temps  $[nt]$ .*

**Proposició 3.1.23.** *Sigui  $\theta(t) = npE[N([nt])]$  de mida moderada, aleshores el procés  $\hat{V}^n(t) = \frac{V([nt])}{n}$  es pot aproximar per un procés de difusió.*

*Prova:*

Podem emprar de nou l'aproximació  $1 - (1-x)^N \simeq Nx$  i que  $\Delta t = \frac{1}{n}, t = 0, \frac{1}{n}, \frac{2}{n}, \dots$ :

$$\begin{aligned} E[\hat{V}^n(t + \Delta t) - \hat{V}^n(t) | \hat{V}^n(t) - \hat{V}^n(t - \frac{R}{n}) = \hat{y}, \hat{V}^n(t - \frac{R}{n}) = \hat{z}] \\ = \frac{1}{n} E[V([nt] + 1) - V([nt]) | V([nt]) - V([nt] - r) = n\hat{y}, V([nt] - r) = n\hat{z}] \\ = \frac{1}{n} (n - n\hat{y} - n\hat{z}) \left[ 1 - \sum_{k=1}^m \left( 1 - \frac{\lambda np \hat{y}}{n-1} \right)^{n_k} p_k([nt]) \right] \\ \simeq (1 - \hat{y} - \hat{z}) \frac{\lambda np \hat{y}}{n-1} E[N([nt])] \simeq \lambda \theta(t) \hat{y} (1 - \hat{y} - \hat{z}) \Delta t \end{aligned}$$

i de la mateixa manera:

$$\begin{aligned}
& Var[\hat{V}^n(t + \Delta t) - \hat{V}^n(t)] | \hat{V}^n(t) - \hat{V}^n(t - \frac{R}{n}) = \hat{y}, \hat{V}^n(t - \frac{R}{n}) = \hat{z}] \\
&= \frac{1}{n^2} Var[V([nt] + 1) - V([nt]) | V([nt]) - V([nt] - r) = n\hat{y}, V([nt] - r) = n\hat{z}] \\
&= \frac{1}{n^2} (n - n\hat{y} - n\hat{z}) \left[ 1 - \sum_{k=1}^m \left( 1 - \frac{\lambda np \hat{y}}{n-1} \right)^{n_k} p_k([nt]) \right] \left[ \sum_{k=1}^m \left( 1 - \frac{\lambda np \hat{y}}{n-1} \right)^{n_k} p_k([nt]) \right] \\
&\simeq \frac{1}{n} (1 - \hat{y} - \hat{z}) \frac{\lambda np \hat{y}}{n-1} E[N([nt])] \left( 1 - \frac{\lambda np \hat{y}}{n-1} E[N([nt])] \right) \\
&\simeq \frac{1}{n} (1 - \hat{y} - \hat{z}) \frac{\lambda np \hat{y}}{n-1} E[N([nt])] \\
&\simeq \frac{1}{n} \lambda \theta(t) \hat{y} (1 - \hat{y} - \hat{z}) \Delta t.
\end{aligned}$$

**Lema 3.1.24.** *L'aproximació de difusió  $\hat{V}(t)$  del procés  $\hat{V}^n(t)$  satisfà l'equació diferencial estocàstica amb endarreriment:*

$$d\hat{V}(t) = \lambda \theta(t) \left( \hat{V}(t) - \hat{V}(t - \tau) \right) \left( 1 - \hat{V}(t) \right) dt + \left[ \frac{\lambda \theta(t)}{n} (\hat{V}(t) - \hat{V}(t - \tau)) (1 - \hat{V}(t)) \right]^{\frac{1}{2}} dW(t).$$

**Observació 3.1.25.** De nou, afirmem l'existència de solucions en aquest cas pels resultats de [10] esmentat a 2.2.12.

Seguint els resultats de [11], ens centrarem al cas paradigmàtic de la variació anual de contactes comentat a l'inici de l'apartat. Per això, considerarem que el nombre de contactes diaris,  $N(t)$ , segueix una distribució binomial  $N(t) \sim Binom(N, p(t))$  amb un paràmetre  $p(t)$  l'evolució del qual és sinusoidal:

$$\begin{aligned}
\text{Primavera: } p(t) &= 0.3 \sin\left(\frac{2\pi t}{365}\right) + 0.5. \\
\text{Estiu: } p(t) &= 0.3 \cos\left(\frac{2\pi t}{365}\right) + 0.5. \\
\text{Tardor: } p(t) &= -0.3 \sin\left(\frac{2\pi t}{365}\right) + 0.5. \\
\text{Hivern: } p(t) &= -0.3 \cos\left(\frac{2\pi t}{365}\right) + 0.5.
\end{aligned}$$

Així, prenent una població de  $n = 10000$  individus i un únic individu infectat inicialment,  $Y(0) = 1$ , una probabilitat de transmissió per contacte de  $p = 0.1$  i una duració de la malaltia de  $R = 5$  dies.

|                         | Primavera     | Estiu        | Tardor        | Hivern        |
|-------------------------|---------------|--------------|---------------|---------------|
| Duració malaltia (dies) | 43.54 (12.77) | 35.33 (5.05) | 30.83 (21.92) | 29.50 (42.08) |
| Població infectada (%)  | 86% (28%)     | 96% (14%)    | 61% (23%)     | 14% (31%)     |

Taula 3.1: Taula dels resultats mitjans de 10000 simulacions de 365 dies de la simulació numèrica del model per al nombre de contactes propi de cada estació de l'any i la desviació típica respectiva en parèntesi. *Dades extretes de [11], Taula 1.*

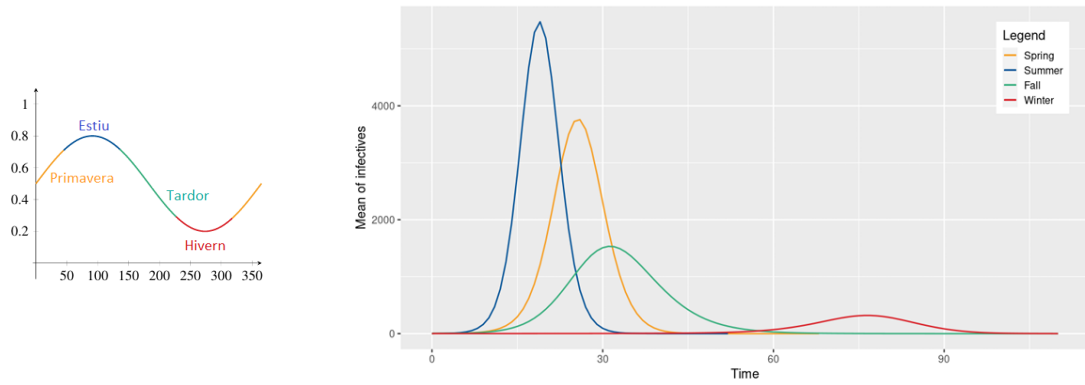


Figura 3.2: Gràfics característics del comportament del model per cada estació de l'any. A l'esquerra, gràfic de l'evolució del paràmetre  $p(t)$  respecte el temps (dies) al llarg de l'any. A la dreta, gràfic del nombre d'individus infectius (classe I) respecte el temps (en dies) de les mitjanes de les simulacions considerades a la Taula 3.1. En groc, primavera; blau, estiu; verd, tardor i vermell, hivern. *Imatges extretes de [11], imatge sense identificador i Figura 1.*

Podem observar com afecta l'augment en el nombre de contactes diari característic de l'estiu i èpoques de l'any d'augment de la calor i la disminució de contactes que comporta el fred i l'hivern. També podem notar el nombre simultani d'individus infectius en èpoques de calor. Un augment en el nombre de contactes diari no només augmenta el nombre total de població infectada un cop acabat el brot sinó que augmenta el nombre d'individus infectius simultanis. Donat que la fracció d'individus malalts greus no depèn del nombre de contactes diaris, observarem una tensió hospitalària (nombre de pacients atesos alhora pels serveis mèdics) major en el cas de l'estiu per atendre tots els pacients alhora.

Amb tot, una característica significativa de la duració de la malaltia és la desviació típica de les dades de l'hivern donat que es produeixen casos extrems de llargues duracions. També és rellevant notar la diferència entre el nombre total d'infectats a la tardor i a la primavera, éssent el primer considerablement menor per una evolució del paràmetre  $p(t)$  simètrica.

## 3.2 El model SEIHR

Després de l'estudi de models simplificats com han estat l'SIR i l'SI (amb  $R = \infty$ ), hem pogut comprendre heurísticament com esperem que els diferents paràmetres principals que descriuen l'epidèmia afectin el nombre de casos i el temps que es triga a contagiar-los.

Ara bé, podem fer un pas més en aquesta descripció i afegir conceptes estesos de les malalties actuals al model. Són ben conegudes per la població general, les fases de latència (on l'individu ha estat infectat però la malaltia no s'ha desenvolupat a l'organisme i no és infecciosa) i asintomatologia (on la persona és infecciosa però no mostra símptomes que permetin identificar-la com a tal). Si bé s'ha pogut comprovar que existeixen certs virus asintomàtics per certa part de la població i amb una simptomatologia clara per d'altres, en aquest estudi ens restringirem al cas on la malaltia és asintomàtica en els primers estadis i simptomàtica posteriorment.

**Definició 3.2.1.** *El model **SEIHR** a estudiar[5] conté les següents característiques:*

1. *El model parteix de les premisses de l'SIR però subdivideix la classe "I" de malalts en dues classes: malalts infecciosos asintomàtics "I" i malalts hospitalitzats "H".*
2. *La probabilitat de trobar individus de la classe "I" i de la classe "H" poden ser diferents, doncs introduïrem un paràmetre  $\lambda \in [0, 1]$  de confinament a aquesta última, amb  $\lambda = 0$  per un confinament total dels malalts amb símptomes i  $\lambda \simeq 1$  per malalties lleus que no requereixen aïllament dels malalts.*
3. *El model també afegeix la classe "E" d'individus infectats que encara no encomanen la malaltia doncs la tenen latent. Situem aquesta classe entre la "S" i la "I".*
4. *El temps d'estada a cadascuna de les fases de la malaltia serà fixe:  $r_E$  per "E",  $r_I$  per "I" i  $r_H$  per "H".*
5. *Ajustarem les probabilitats de trobada al paràmetre  $\lambda$  segons el nombre d'individus total "n", infecciosos classe "H", " $y_H$ ", i classe "I", " $y_I$ ":*

$$p_H = \lambda \frac{y_H}{n-1}, \quad p_I = \frac{n-1-\lambda y_H}{n-1-y_H} \cdot \frac{y_I}{n-1},$$

$$p_N = (n - y_H - y_I) = \frac{n-1-\lambda y_H}{n-1-y_H} \cdot \frac{n-1-y_H-y_I}{n-1},$$

on  $p_N$  és la probabilitat de trobar un individu de qualsevol de la resta de classes. Aquests resultats són vàlids per  $y_H < n-1$ , si  $y_H = n-1$  aleshores  $p_H = 1$ ,  $p_I = 0$ ,  $p_N = 0$ .

**Proposició 3.2.2.** *Segui  $j_{iI}$ ,  $j_{iH}$  les trobades de l'individu  $i$ -èsim en un cert temps  $t$ , amb individus de les classes "I" i "H" respectivament, la probabilitat de trobar aquesta proporció d'individus és:*

$$P_{j_{iI}, j_{iH}}^i(y_I, y_H, N_i(t); n) = \frac{N_i(t)!}{j_{iI}! j_{iH}! j_{iN}!} p_I^{j_{iI}} p_H^{j_{iH}} p_N^{j_{iN}}$$

On  $N_i(t)$  és, de nou, el nombre de contactes de l'individu  $i$   $j_{iN} = N_i(t) - j_{iI} - j_{iH}$ .

**Proposició 3.2.3.** *Siguin  $q_H, q_I$  la probabilitat de transmissió de la malaltia per contacte entre individus susceptibles i individus del grup infecciosos en qüestió, aleshores la probabilitat de contagi d'un individu "i" susceptible que hagi estat en contacte amb  $j_{iH}$  individus del grup "H" i  $j_{iI}$  de l' "I" és:*

$$p_{j_{iH}, j_{iI}} = 1 - ((1 - q_H)^{j_{iH}} (1 - q_I)^{j_{iI}}).$$

**Proposició 3.2.4.** *Prenent els dos resultats anteriors podem calcular la probabilitat de contagi d'un individu a temps  $t + 1$ :*

$$\begin{aligned} \beta_i &= P(Y_0^i(t+1) = 1 | N_i(t), X^i(t) = 1, Y_I(t) = y_I, Y_H(t) = y_H) \\ &= \sum_{j_{iI}, j_{iH}=1}^{N_i(t)} p_{j_{iH}, j_{iI}} \frac{N_i(t)!}{j_{iI}! j_{iH}! j_{iN}!} p_I^{j_{iI}} p_H^{j_{iH}} p_N^{j_{iN}} \\ &= 1 - \sum_{j_{iI}, j_{iH}=1}^{N_i(t)} \frac{N_i(t)!}{j_{iI}! j_{iH}! j_{iN}!} (p_I(1 - q_I))^{j_{iI}} (p_H(1 - q_H))^{j_{iH}} p_N^{j_{iN}}. \end{aligned}$$

*I substituïnt els valors de  $p_H, p_I, p_N$ :*

$$\beta_i = 1 - \left( 1 - \frac{1}{n-1} \left( \frac{y_I(n-1-\lambda y_H)}{n-1-y_H} q_I + \lambda y_H q_H \right) \right)^{N_i(t)}$$

*quan  $y_H < n-1$  i per  $y_H = n-1$ ,  $\beta_i = 1 - (1 - q_H)^{N_i(t)}$ .*

De nou, posem el focus en el càlcul de la probabilitat de contagi d'un individu. En aquest cas, per anar més enllà en el tractament teòric, el resultat portarà la definició d'una nova classe genèrica (com la classe  $I$  per a l'SIR amb  $R = \infty$  i la classe  $\hat{V}$  del model SIR amb contactes diaris variable) que estableixi una dicotomia entre susceptibilitat i no-susceptibilitat.

**Proposició 3.2.5.** *Sigui el nombre de contactes de cada individu constant  $N_i(t) = N$ , per tant  $\beta_i = \beta$  i considerem  $V(t)$  el nombre d'individus no susceptibles a cada instant  $t$ , aleshores:*

$$V_{t, y_E, y_I, y_H, y_R} = \{V(t) - V(t - r_E) = y_E, V(t - r_E) - V(t - r_E - r_I) = y_I,$$

$$V(t - r_E - r_I) - V(t - r_E - r_I - r_H) = y_H, V(t - r_E - r_I - r_H) = y_R\},$$

*descriu el nombre d'individus de cada classe en un determinat instant de temps. On  $V(t) - V(t - r_E)$  dóna el nombre d'individus a la classe  $E$  a temps  $t$ ,  $V(t - r_E) - V(t - r_E - r_I)$  els de la classe  $I$ ,  $V(t - r_E - r_I) - V(t - r_E - r_I - r_H)$  els de la classe  $H$  i  $V(t - r_E - r_I - r_H)$  els de la classe  $R$ .*

**Observació 3.2.6.** Clarament, els individus que s'afegeixen a la classe de no susceptibles primer entren com a latents i després transicionen a infecciosos, hospitalitzats i, finalment, recuperats. Els individus afegits en els últims  $r_E$  passos de temps (podem assumir dies per simplicitat), que primerament són latents, encara no han superat aquest estadi (no han passat  $r_E$  dies). Els individus afegits entre  $r_E$  i  $r_I$  passos de temps abans de l'actual, ja han transicionat a malalts infecciosos però encara no han transicionat a hospitalitzats. (La resta de classes són anàlogues.)

**Proposició 3.2.7.** *Amb l'anterior resultat, podem obtenir la probabilitat d'augmentar el nombre de no susceptibles un cert valor  $y$  en un pas de temps:*

$$\begin{aligned} P(V(t+1) = y_E + y_I + y_H + y_R + y | V_{t,y_E,y_I,y_H,y_R}) \\ = \binom{n - y_E - y_I - y_H - y_R}{y} \beta^y (1 - \beta)^{n - y_E - y_I - y_H - y_R - y} \end{aligned}$$

per  $y = 0, 1, 2, \dots$

**Proposició 3.2.8.** *Similarment al cas SIR amb  $R = \infty$ , si considerem la classe de no susceptibles  $V(t)$  al model SEIHR obtenim:*

$$\begin{aligned} E(V(t+1) - V(t) | V_{t,y_E,y_I,y_H,y_R}) &= (n - y_E - y_I - y_H - y_R)\beta, \\ \text{Var}(V(t+1) - V(t) | V_{t,y_E,y_I,y_H,y_R}) &= (n - y_E - y_I - y_H - y_R)\beta(1 - \beta). \end{aligned}$$

Tornem a situar-nos en un entorn propici per parlar de processos de difusió, després del proper resultat, veurem en quines condicions podem establir un marc teòric similar al de l'equació diferencial estocàstica de l'SIR amb  $R = \infty$ .

Seguint els desenvolupaments portats a terme pel model SIR, calcularem el coeficient  $R_0$  per al model SEIHR. Per això, sabem que el nombre de dies de latència, infecció i hospitalització han de ser fixos per fer un estudi detallat de les transicions d'estats.

**Lema 3.2.9.** *Si considerem  $r_I = r_H = 1$ , aleshores el període infecció d'un individu és  $r_I + r_H = 2$ . Considerant  $r_E \geq r_I + r_H$  i que el procés comença amb  $X(0) = n - 1$ ,  $E(0) = 0$ ,  $Y_I(0) = 1$ ,  $Y_H(0) = 0$  i  $R(0) = 0$ , aleshores el nombre d'individus latents al següent pas de temps segueix una binomial de paràmetres  $X(t)$  i  $p(Y_I(t), Y_H(t), \lambda)$  amb:*

$$p(y_I(t), y_H(t), \lambda) = 1 - \left( 1 - \frac{1}{n-1} \left( \frac{y_I(n-1-\lambda y_H)}{n-1-y_H} q_I + \lambda y_H q_H \right) \right)^N$$

per  $y_H < n - 1$ ,

$$p(0, n-1, \lambda) = 1 - (1 - q_H)^N$$

per  $y_H = n - 1$ .

**Proposició 3.2.10.** *El coeficient  $R_0$  del model SEIHR sota les premisses de l'anterior lema és:*

$$R_0 = (n-1) [p(1, 0, \lambda) + p(0, 1, \lambda) - p(1, 0, \lambda)p(0, 1, \lambda)].$$

*Prova:*

De la mateixa manera que per al model SIR, calcularem  $E(C(1))$  i  $E(C(2))$  donat que el malalt només contagia durant les fases infecciosa i d'hospitalització (aquesta última en funció del paràmetre  $\lambda$ ), d'1 dia de duració ambdues. Així:

$$E(C(1)) = E(E(1) | Y_I(0) = 1) = (n-1)p(1, 0, \lambda).$$

On cal notar la diferència entre el nombre d'individus latents  $E(t)$  i l'esperança d'una variable aleatòria  $E(X)$ . Per a  $t = 1$ , s'hauràn contagiat  $e_1$  individus i el recompte varia:  $X(1) = n - 1 - e_1$ ,  $E(1) = e_1$ ,  $Y_I(1) = 0$ ,  $Y_H(1) = 1$ ,  $R(1) = 0$  i  $C(2) | (E(1) = e_1, Y_I(1) = 0, Y_H(1) = 1)$  és una binomial de paràmetres  $X(1)$ ,  $p(0, 1, \lambda)$ , per tant:

$$E(C(2)) = E(E(C(2) | E(1) = e_1, Y_I(1) = 0, Y_H(1) = 1))$$

$$= E((n-1-E(1))p(0,1,\lambda)) = (n-1)(p(0,1,\lambda) - p(1,0,\lambda)p(0,1,\lambda)).$$

De nou, observem que, per aquest model, també podem fer una aproximació de difusió per tal de poder emprar les propietats dels processos markovians a temps continu. En aquest cas, la població també haurà de ser gran i, les probabilitats  $q_I$  i  $q_H$  tals que el producte  $nN(q_I + q_H)$  sigui de mida moderada.

Així, obtenim la versió contínua del nombre d'individus no susceptibles com  $\hat{V}^n(t) = \frac{V(\lfloor nt \rfloor)}{n}$  per tot  $t > 0$ , emprant les mateixes aproximacions que en el cas anterior.

**Proposició 3.2.11.** *Segui  $\theta_I = nNq_I$ ,  $\theta_H = nNq_H$  de mida moderada, aleshores el procés  $\hat{V}^n(t) = \frac{V(\lfloor nt \rfloor)}{n}$  s'aproxima a un procés de difusió.*

*Prova:*

$$\begin{aligned} E(\hat{V}^n(t + \Delta t) - \hat{V}^n(t) | \hat{V}_{t, \hat{y}_E, \hat{y}_I, \hat{y}_H, \hat{y}_R}) \\ \simeq (1 - \hat{y}_E - \hat{y}_I - \hat{y}_H - \hat{y}_R) N \left( \frac{1 - \lambda \hat{y}_H}{1 - \hat{y}_H} q_I \hat{y}_I + \lambda q_H \hat{y}_H \right) \\ = (1 - \hat{y}_E - \hat{y}_I - \hat{y}_H - \hat{y}_R) \left( \frac{1 - \lambda \hat{y}_H}{1 - \hat{y}_H} \theta_I \hat{y}_I + \lambda \theta_H \hat{y}_H \right) \Delta t, \\ Var(\hat{V}^n(t + \Delta t) - \hat{V}^n(t) | \hat{V}_{t, \hat{y}_E, \hat{y}_I, \hat{y}_H, \hat{y}_R}) \simeq (1 - \hat{y}_E - \hat{y}_I - \hat{y}_H - \hat{y}_R) \frac{1}{n} \left( \frac{1 - \lambda \hat{y}_H}{1 - \hat{y}_H} \theta_I \hat{y}_I + \lambda \theta_H \hat{y}_H \right) \Delta t \end{aligned}$$

amb  $\Delta t = \frac{1}{n}$  i  $t \in \{0, \frac{1}{n}, \frac{2}{n}, \dots\}$  i, clarament,

$$\begin{aligned} \hat{V}_{t, \hat{y}_E, \hat{y}_I, \hat{y}_H, \hat{y}_R} = \{ \hat{V}^n(t) - \hat{V}^n(t - \frac{r_E}{n}) = \hat{y}_E, \hat{V}^n(t - \frac{r_E}{n}) - \hat{V}^n(t - \frac{r_E}{n} - \frac{r_I}{n}) = \hat{y}_I, \\ \hat{V}^n(t - \frac{r_E}{n} - \frac{r_I}{n}) - \hat{V}^n(t - \frac{r_E}{n} - \frac{r_I}{n} - \frac{r_H}{n}) = \hat{y}_H, \hat{V}^n(t - \frac{r_E}{n} - \frac{r_I}{n} - \frac{r_H}{n}) = \hat{y}_R \}. \end{aligned}$$

Aquesta última descripció presenta una dificultat afegida: el càlcul de l'esperança i de la variància prenen valors de temps anteriors del procés  $\hat{V}^n$ .

Mentre que al càlcul pel model SIR,  $\hat{Y}^n$  es prenien sempre en un mateix instant de temps, ara estem considerant  $t, t - \frac{r_E}{n}, t - \frac{r_E}{n} - \frac{r_I}{n}, t - \frac{r_E}{n} - \frac{r_I}{n} - \frac{r_H}{n}$ , això impedeix, en primera instància, un tractament simple amb equacions diferencials del procés. Primerament, reanomenarem per simplicitat:  $\tau_E = \frac{r_E}{n}$ ,  $\tau_{EI} = \tau_E + \frac{r_I}{n}$  i  $\tau_{EIH} = \tau_{EI} + \frac{r_H}{n}$ .

**Lema 3.2.12.** *El procés de difusió  $\hat{V}^n(t)$  satisfà l'equació diferencial:*

$$\begin{aligned} d\hat{V}_t = (1 - \hat{V}_t) \left( \theta_I \frac{1 - \lambda(\hat{V}_{t-\tau_{EI}} - \hat{V}_{t-\tau_{EIH}})}{1 - (\hat{V}_{t-\tau_{EI}} - \hat{V}_{t-\tau_{EIH}})} (\hat{V}_{t-\tau_E} - \hat{V}_{t-\tau_{EI}}) + \theta_H \lambda (\hat{V}_{t-\tau_{EI}} - \hat{V}_{t-\tau_{EIH}}) \right) dt \\ + \left[ (1 - \hat{V}_t) \left( \theta_I \frac{1 - \lambda(\hat{V}_{t-\tau_{EI}} - \hat{V}_{t-\tau_{EIH}})}{1 - (\hat{V}_{t-\tau_{EI}} - \hat{V}_{t-\tau_{EIH}})} (\hat{V}_{t-\tau_E} - \hat{V}_{t-\tau_{EI}}) + \theta_H \lambda (\hat{V}_{t-\tau_{EI}} - \hat{V}_{t-\tau_{EIH}}) \right) \right]^{\frac{1}{2}} dW_t. \end{aligned}$$

Trobem una equació diferencial estocàstica amb endarreriment i sabem de l'existència de solucions pels resultats de [10].

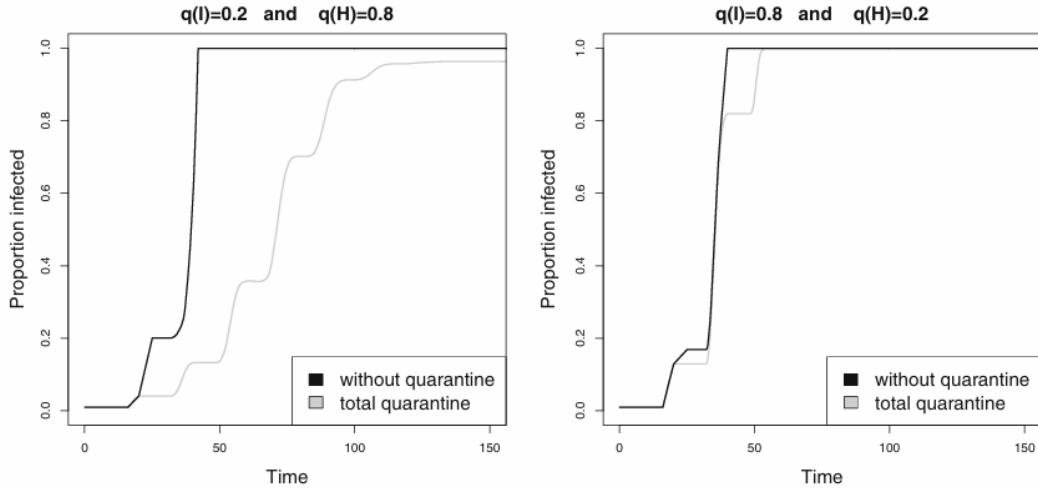


Figura 3.3: Gràfics de la mitjana de 100 simulacions de l'aproximació de difusió del model SEIHR sense quarantena ( $\lambda = 1$ ) en negre i amb quarantena total ( $\lambda = 0$ ) en gris pels paràmetres  $r_E = 16$ ,  $r_I = 4$ ,  $r_H = 5$ ,  $N = 5$ ,  $n = 100$ . S'observa un alentiment en l'assoliment del total d'infectats per quarantenes estrictes, concepte clau per alleugerir la tensió hospitalària. A l'esquerra:  $q_I = 0.2$  i  $q_H = 0.8$ . A la dreta,  $q_I = 0.8$  i  $q_H = 0.2$ . *Imatges extrems de [5], Figura 5.*

### 3.2.1 El model SEIHRS

Finalment, l'última consideració teòrica en el desenvolupament del model SEIHR és considerar com aquest varia quan la malaltia té un caràcter cíclic, és a dir, quan els individus no assolixen un estat invariant en el temps. És el cas de malalties tals com el refredat comú per a les quals els anticossos desenvolupats tant per vacunació com per infecció perden efectivitat degut a mutacions del virus que el provoca que, a nivell teòric, consisteix en un desplaçament dels individus de la classe o estat "Recuperats" a "Susceptibles" (si no considerem mortalitat en el procés) en l'instant de temps en què es produeix la mutació.

Seguint amb el tractament anterior, considerarem un temps fixe  $r_R$  en el que els individus estàn a la classe  $R$  fins tornar a la classe  $S$ . Clarament, la probabilitat de contagi a temps  $t + 1$  d'un individu susceptible a temps  $t$  és la mateixa que abans.

Un canvi notable és l'evolució en el nombre de no susceptibles. Mentre que el procés  $\hat{V}^n$  acumula el nombre de no susceptibles en l'interval de temps  $[0, t]$ , caldrà restar a aquest valor el nombre d'individus que han tornat a ser susceptibles, és a dir, que han transicionat passats els períodes fixats de temps entre classes, per tant, aquells que havien a temps anteriors a  $r = r_E + r_I + r_H + r_r$ . Així, l'única variació de l'equació diferencial anterior serà en el nombre d'individus no susceptibles a temps  $t$ .

**Lema 3.2.13.** *El procés de difusió  $\hat{V}^n(t)$  satisfà l'equació diferencial:*

$$\begin{aligned} d\hat{V}_t = & (1 - \hat{V}_t - \hat{V}_{t-\tau_{EIHR}})(\theta_I \frac{1 - \lambda(\hat{V}_{t-\tau_{EI}} - \hat{V}_{t-\tau_{EIH}})}{1 - (\hat{V}_{t-\tau_{EI}} - \hat{V}_{t-\tau_{EIH}})}(\hat{V}_{t-\tau_E} - \hat{V}_{t-\tau_{EI}}) \\ & + \theta_H \lambda(\hat{V}_{t-\tau_{EI}} - \hat{V}_{t-\tau_{EIH}}))dt \\ & + [(1 - \hat{V}_t - \hat{V}_{t-\tau_{EIHR}})(\theta_I \frac{1 - \lambda(\hat{V}_{t-\tau_{EI}} - \hat{V}_{t-\tau_{EIH}})}{1 - (\hat{V}_{t-\tau_{EI}} - \hat{V}_{t-\tau_{EIH}})}(\hat{V}_{t-\tau_E} - \hat{V}_{t-\tau_{EI}}) \\ & + \theta_H \lambda(\hat{V}_{t-\tau_{EI}} - \hat{V}_{t-\tau_{EIH}}))]^{\frac{1}{2}} dW_t \end{aligned}$$

on  $\tau_{EIHR} = \tau_{EIH} + \frac{\tau_R}{n}$  i  $1 - \hat{y}_E - \hat{y}_I - \hat{y}_H - \hat{y}_R = 1 - \hat{V}_t - \hat{V}_{t-\tau_{EIHR}}$ .

De nou, ens adrecem a [10] per detalls més avançats de les propietats d'aquestes equacions amb endarreriment i l'existència de solucions.

### 3.2.2 El model SEIAHCRD

Com a extensió del model SEIHR (el qual ja suposa una sofisticació del model clàssic SIR) podem considerar l'existència d'un nombre major de classes que descriguin millor les etapes de la malaltia en l'individu. En aquesta secció, es pretén reflexar el ventall d'opcions teòriques que suposa el model discret SIR i com es poden afegir classes i característiques a un model a mesura que es coneixen en profunditat les particularitats biològiques de les malalties.

El cas que ens ocupa és el de la Covid-19 en un context general. Més endavant farem una descripció detallada de les seves característiques, però s'ha observat l'existència de diversos tipus de contagiats que recollirem en el model SEIAHCR proposat per [12]:

**Definició 3.2.14.** *El model **SEIAHCR** conté les següents classes:*

1. "S" referent als individus susceptibles.
2. "E" referent als individus contagiats però latents, per tant ni malalts ni infecciosos.
3. "I" referent als individus infecciosos però no malalts que sí desenvoluparan la malaltia.
4. "A" referent als individus asimptomàtics, és a dir, infecciosos però no malalts que no desenvoluparan la malaltia.
5. "H" referent als individus malalts amb símptomes lleus.
6. "C" referent als individus malalts amb símptomes greus
7. "D" referent als individus difunts.
8. "R" referent als individus recuperats, que no tornen a ser susceptibles per la immunitat adquirida.

Com en el model anterior, tots els individus partiran de la classe "S" a excepció d'un, que partirà de la classe "I" o "A".

**Definició 3.2.15.** *Les transicions entre estats seràn, majoritàriament, indeterministes amb probabilitats fixades i només es donaran les següents:*

1. *Els individus de la classe "S" mantindrà un nombre de trobades  $N_i(t)$  determinista i, el resultat de trobar-se amb un individu de la classe "I" o "A", representarà la transmissió de la malaltia (i per tant, la transició a la classe "E", amb probabilitat  $q_I$  i  $q_A$ , respectivament. Els individus contagiosos de la resta de classes, estaran confinats i no participaran dels contactes diaris.*
2. *Els individus de les classes "E", "I" i "A" romandran en aquestes classes durant  $r_E$ ,  $r_I$  i  $r_A$  dies, respectivament. Els individus de la classe "E" transicionaran a la classe "I" amb probabilitat  $\mu$  i a l'"A" amb  $1 - \mu$ .*
3. *Els individus de la classe "I" transicionaran a la classe "H" i els d'"A" a "R" unívocament.*
4. *Els individus que entrin a la classe "H" hi romandran  $r_{HC}$  dies, i per tant transicionaran a la classe "C" en finalitzar aquest període, amb probabilitat  $\alpha$  i  $r_{HR}$ , transicionant consegüentment, amb probabilitat  $1 - \alpha$ .*
5. *Els individus de la classe "C" hi romandran durant  $r_{CD}$  amb probabilitat  $\lambda$  i  $r_{CR}$  amb probabilitat  $1 - \lambda$ , anàlogament.*
6. *Els individus de les classes "R" i "C" es mantenen en les respectives classes indefinidament.*

**Observació 3.2.16.** Podem identificar les probabilitats esmentades:  $\lambda$  representa la mortalitat de la malaltia per casos severs,  $\alpha$  la probabilitat de que un cas amb simptomatologia lleu esdevingui greu i  $\mu$  la probabilitat d'emmalaltir un cop contagiats. És interessant notar que les transicions són biològicament suaus doncs es passa de no presentar símptomes ("I") a presentar-ne de lleus ("H") a, en alguns casos, presentar-ne de greus ("C"). De la mateixa manera, es distingeix entre un cas asimptomàtic durant tot el procés ("A") a un asimptomàtic temporal ("I").

Seguidament, estudiarem els paràmetres que hem anat seguint en aquest desenvolupament teòric: la probabilitat de contagi d'un individu susceptible en un temps  $t + 1$  donat i el coeficient  $R_0$  del model. Per aquest propòsit, considerarem el nombre d'infectats de cada classe com  $Y_X(t)$  on  $X = I, A, H, C, D$  és la classe concreta a estudiar. Igualment, considerarem la probabilitat de contagi per contacte  $q_I$  si el contacte es dona amb un individu classe "I" i  $q_A$  si el contacte es dona amb un de la classe "A".

**Proposició 3.2.17.** *La probabilitat de contagi d'un individu susceptible a temps  $t + 1$  és:*

$$\beta_i = 1 - \left( 1 - \frac{q_i y_i - q_A y_A}{n - y_H - y_C - y_D - 1} \right)^{N_i(t)}$$

*per  $y_H + y_C + y_D < n - 1$  i  $\beta_i = 0$  altrament.*

*Prova:*

Podem veure que existeixen quatre camins a seguir per un individu infectat:

1. Segueix la transició d'estats  $E, I, H, C, D$  amb temps d'estada  $r_E, r_I, r_{HC}$  i  $r_{CD}$  dies, mort en la transició a la classe  $D$ .
2. Segueix la transició d'estats  $E, I, H, C, R$  amb temps d'estada  $r_E, r_I, r_{HC}$  i  $r_{CR}$  dies, recuperant-se (i al seu torn esdevenint immune) en la transició a la classe  $R$ .
3. Segueix la transició d'estats  $E, I, H, R$  amb temps d'estada  $r_E, r_I$  i  $r_{HR}$  dies, recuperant-se en la transició a la classe  $R$ .
4. Segueix la transició d'estats  $E, A, R$  amb temps d'estada  $r_E$  i  $r_A$  dies, recuperant-se en la transició a la classe  $R$ .

Cadascun d'ells amb una probabilitat de  $\mu\alpha\lambda$ ,  $\mu\alpha(1-\lambda)$ ,  $\mu(1-\alpha)$  i  $1-\mu$ , respectivament.

Per les consideracions anteriors, els individus de les classes  $H, C$  (pel confinament),  $D$  (pel decés dels mateixos) no participen en els contactes diaris. De fet, els individus de la classe  $D$  són exclosos del sistema). La resta de la població interacciona d'igual forma independentment de la seva classe.

Així, les probabilitats de trobar un individu de les classes  $I, A$  i no contagiosos,  $V$ , són, respectivament:

$$\begin{aligned} p_I &= \frac{y_I}{n - y_H - y_C - y_D - 1}, \\ p_A &= \frac{y_A}{n - y_H - y_C - y_D - 1}, \\ p_V &= 1 - p_I - p_A, \end{aligned}$$

quan  $y_H + y_C + y_D < n - 1$ , mentre que  $p_I = p_A = 0$  altrament.

De nou, la probabilitat que un individu susceptible  $i$  trobi  $j_I$  individus de classe  $I$ ,  $j_{iI}$ , i  $j_A$  individus de classe  $A$  a temps  $t$  donat un nombre  $N_i(t)$  de trobades diàries és:

$$\frac{N_i(t)!}{j_{iI}!j_{iA}!(N_i(t) - j_{iI} - j_{iA})!} p_I^{j_{iI}} p_A^{j_{iA}} p_V^{N_i(t) - j_{iI} - j_{iA}},$$

i la probabilitat de contagi de l'individu  $i$  donats aquest nombre de contactes de cada classe és:

$$p_{j_{iI}, j_{iA}} = 1 - (1 - q_I)^{j_{iI}} (1 - q_A)^{j_{iA}}.$$

Amb el que arribem al càlcul de la probabilitat de contagi de l'individu  $i$  realitzant un sumatori de probabilitats condicionals de contagi per a un cert nombre de trobades amb cada classe, el sumatori recorre la suma dels valors  $j_{iI} + j_{iA}$  d'1 a  $N_i(t)$  de totes les formes possibles:

$$\begin{aligned} \beta_i &= \sum_{j_{iI}, j_{iA}}^{N_i(t)} p_{j_{iI}, j_{iA}} \frac{N_i(t)!}{j_{iI}!j_{iA}!(N_i(t) - j_{iI} - j_{iA})!} p_I^{j_{iI}} p_A^{j_{iA}} p_V^{N_i(t) - j_{iI} - j_{iA}} \\ &= 1 - \sum_{j_{iI}, j_{iA}}^{N_i(t)} \frac{N_i(t)!}{j_{iI}!j_{iA}!(N_i(t) - j_{iI} - j_{iA})!} ((1 - q_I)p_I)^{j_{iI}} ((1 - q_A)p_A)^{j_{iA}} p_V^{N_i(t) - j_{iI} - j_{iA}} \end{aligned}$$

Similarment al binomi de Newton, el teorema multinomial permet simplificar aquesta expressió:

$$\begin{aligned}\beta_i &= 1 - ((1 - q_I)p_I + (1 - q_A)p_A + p_V)^{N_i(t)} \\ &= 1 - (1 - q_I p_I - q_A p_A)^{N_i(t)} \\ &= 1 - \left(1 - \frac{q_I y_I - q_A y_A}{n - y_H - y_C - y_D - 1}\right)^{N_i(t)}\end{aligned}$$

per  $y_H + y_C + y_D < n - 1$  i  $\beta_i = 0$  altrament.

**Proposició 3.2.18.** *El coeficient  $R_0$  per una població amb  $n - 1$  individus susceptibles i 1 individu infeccios (classe I o A) amb a un nombre fixat de contactes diaris independent de l'individu i el temps és:*

$$R_0 = \mu r_I q_I N + (1 - \mu) r_A q_A N + O\left(\frac{1}{N}\right)$$

L'individu es troba amb  $N$  individus (primerament susceptibles i després susceptibles i latents) i els contagia amb probabilitat  $q_x$ , i això durant  $r_x$  dies, en funció del tipus d'infectat ( $A$  o  $I$ ) amb probabilitat  $1 - \mu$  d'ésser del primer tipus i  $\mu$  del segon. Els termes superiors reflexen la variació del nombre disponible de susceptibles amb el pas dels dies i l'augment de contagiats, però donat que  $n \gg N$ , menysprearem els termes, calculables considerant l'esperança de contagiats a cada iteració.

### 3.3 Aplicació del model SEIHR

En aquesta secció, volem comparar el model SEIHR amb dades experimentals de la propagació d'una malaltia de característiques similars a les tractades pel model. A tal efecte, considerarem l'estudi [13], que estudia la propagació del virus SARS-CoV-2 que va tenir lloc durant el febrer i el març de 2020 a un mòdul d'una residència de gent gran al comtat de King County, a l'estat de Washington (Estats Units d'Amèrica). Després de l'identificació el 28 de febrer d'un cas positiu de Covid-19 dins del mòdul referit anteriorment, diverses institucions mèdiques van procedir a fer un estudi del cas en el qual es van seguir els contactes, es va posar en quarantena les persones exposades i aïllar els casos confirmats o sospitosos.

La Covid-19, és una malaltia provocada pel coronavirus SARS-CoV-2 que va ser considerada pandèmia l'11 de març del 2020 per l'OMS (Organització Mundial de la Salut) després de l'aparició del primer cas el 31 de desembre de 2019 a la Xina. El primer cas diagnosticat de Covid-19 als Estats Units d'Amèrica va ser el 20 de gener de 2020 a l'estat de Washington, d'on s'han obtingut les dades anteriors. Tal i com recull [14], la malaltia té una fase de latència d'aproximadament 5-6 dies i un període infecció que pot començar entre 2-3 dies abans de l'aparició de simptomatologia. Les dades de [13] indiquen que els malalts eren confinats un cop detectada la infecció (i, per tant, existeix un període posterior a la fase infecciosa i anterior a la recuperació), el que constituiria un període d'hospitalització. D'aquesta manera, podem observar l'existència de 5 fases distingibles dins de la malaltia que encaixen amb les descripcions de les classes del model SEIHR. Per les característiques de l'estudi, tant en el tractament indistingible dels contagiats com en la subdivisió binària d'infectats i susceptibles en els resultats, prenem el model SEIHR, del qual coneixem d'afegit el model de difusió que l'aproxima, com a més proper a la malaltia que models a priori més detallats com [12] però que descriuen la malaltia en un context més general.

Referent a l'estudi que ens ocupa, fixarem les variables del model seguint l'anterior discussió i afegirem  $r_I = 2$ ,  $r_H = 8$  barrejant l'aïllament descrit al centre i la prevalència del virus descrita a [14], les variables  $q_I = 0.06$  i  $q_H = 0.06$  dins dels marges dels resultats obtinguts per [15] i fixant  $\lambda = 0$  donat l'aïllament estricte i  $n = 200$  considerant que l'article inicial parla del personal del centre com un ens mòbil en el qual es distribuïa el personal total entre les sales i  $N = 1$  perquè es buscava limitar els contactes a la residència. D'aquesta manera, obtenim:

| Variable  | Valor |
|-----------|-------|
| $n$       | 300   |
| $N$       | 1     |
| $r_E$     | 6     |
| $r_I$     | 2     |
| $r_H$     | 8     |
| $q_I$     | 0.06  |
| $q_H$     | 0.06  |
| $\lambda$ | 0     |

Taula 3.2: Valors de les variables del model.

Ajustarem el dia inicial de les dades experimentals com el dia 5 de la simulació per tal de reflectir el període en el qual l'individu inicial va ser infecció però no detectat. Podem

considerar-lo un cas atíptic en el nombre de dies de contagi per ser l'individu inicial.

Procedint de manera similar a [5], portarem a terme simulacions numèriques de l'aproximació de difusió del model per observar el comportament del model junt a les dades experimentals. Farem ús del programari *MATLAB* a tal efecte. El codi emprat per realitzar aquestes simulacions, utilitzant el mètode d'Euler-Maruyama descrit a la Definició 2.2.13, es pot trobar a l'Annex A d'aquest treball.

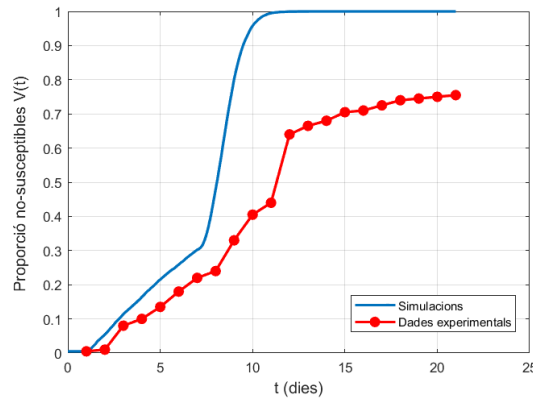


Figura 3.4: Gràfic de la proporció de no-susceptibles respecte el temps (dies) de la mitjana de 100 simulacions numèriques de l'aproximació de difusió del model SEIHR amb les dades de la Taula 3.3 i les dades experimentals de no-susceptibles de l'estudi.

Els comportaments similars de l'evolució del nombre de no-susceptibles tant en les dades experimentals com en les simulacions donen evidència de com el model descriu l'evolució de virus que encaixen amb les hipòtesis d'aquest marcant un període de contagi moderat, un amb una alta taxa de contagi i un tercer d'estancament. Malgrat tot, l'alta capacitat de contagi teòrica del SARS-CoV-2 extreta dels paràmetres dels estudis emprats en aquesta anàlisi resulten en l'assoliment eventual del contagi de tota la població, fet que no succeeix a les dades experimentals.

La discussió, doncs, podria centrar-se en les dades i no en el model. Les generalitats associades a les dades dels articles utilitzades en la fixació dels paràmetres, no contemplen les casuístiques concretes del cas de la residència i així, no recullen la disminució significativa en el nombre de contagis que potencialment pot succeir en un entorn petit i controlat o la mitigació de la capacitat de contagi del virus en la presència d'un ambient estèril o especialment preparat per evitar-ne la propagació. Aquestes especulacions, però, es poden veure reflexades en la Figura 3.5 i argumenten les conjectures presentades en aquest paràgraf. D'igual manera, es poden esmentar d'altres aspectes inferits en la presa de paràmetres com un nombre menor de treballadors al mòdul, que constituiria el valor efectiu d'aquest, o una millor segmentació del procés del virus en l'individu.

Podem observar com les variacions dels paràmetres a la Figura 3.5 no modifiquen el comportament general: existeixen tres etapes diferenciades dins de la corba d'individus no-susceptibles i els paràmetres regulen la pendent i longitud d'aquests trams. Així mateix, es pot observar un comportament similar a les Figures 3.4 i 3.5 respecte a les dades experimentals.

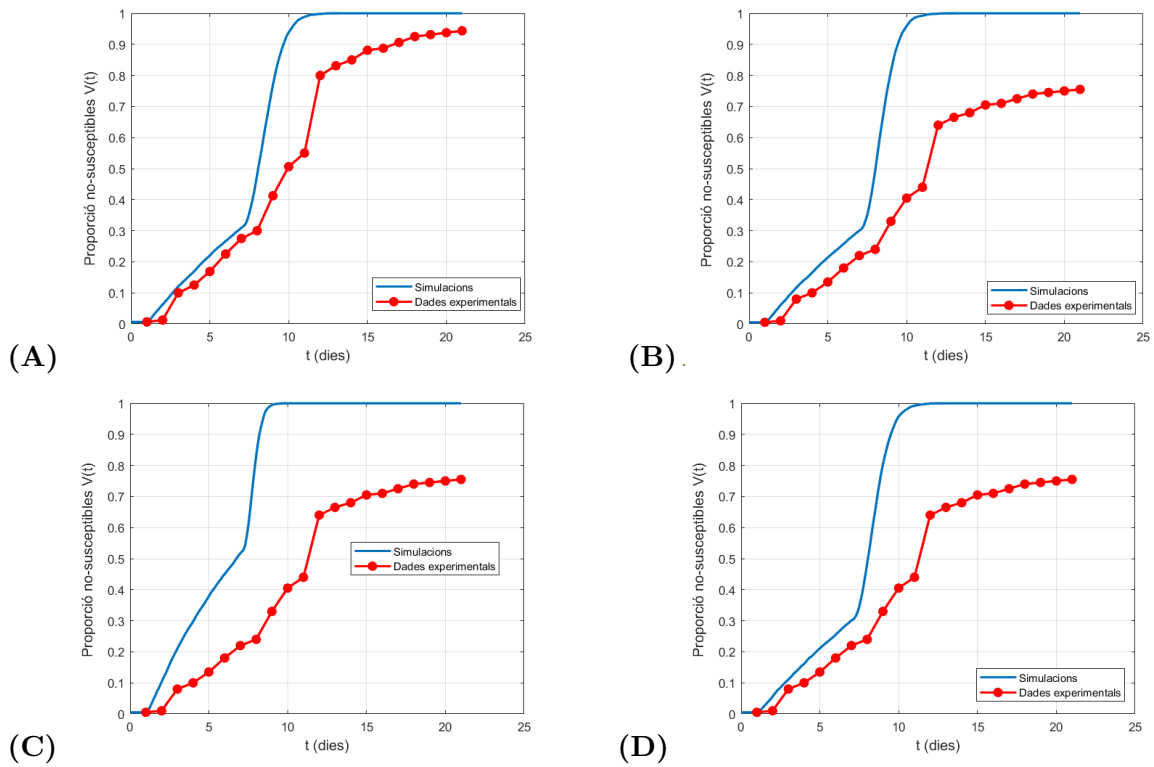


Figura 3.5: Gràfics de la proporció d'individus no susceptibles respecte el temps (dies) de la mitjana de 100 simulacions numèriques de l'aproximació de difusió del model SEIHR amb variacions concretes de les dades de la Taula 3.3 i les dades experimentals de no-susceptibles de l'estudi. (A) Disminució del nombre total d'individus a 160, fent consideracions extremes sobre el personal involucrat al mòdul. (B) Permetre contagi dels aïllats;  $\lambda = 0.2$ . (C) Augmentar el nombre de contactes diaris  $N = 2$ . (D) Augmentar el període infecció a  $r_I = 3$  dies.

## Capítol 4

# Conclusions

El treball realitzat ha acomplert els objectius establerts a l'introducció d'aquesta memòria. La revisió teòrica realitzada en aquest treball ens ha permès comprendre els processos markovians i en quines condicions es poden aproximar cadenes de Màrkov com a processos de difusió, establint una relació entre els processos a temps discret i els de temps continu. També, hem establert una relació inversa quan hem estudiat el mètode d'Euler-Maruyama que estableix una relació entre un procés a temps continu i una aproximació discreta. Per arribar a aquest últim punt, hem hagut d'estudiar les propietats de l'integral d'Itô que és el resultat d'integrar una funció respecte un procés de Wiener. Amb aquesta eina, hem presentat les equacions diferencials estocàstiques i vist sota quines condicions tenen solució.

Les equacions diferencials estocàstiques, resultat final de l'introducció teòrica, han permès donar una expressió analítica a les aproximacions, mitjançant processos de difusió, de l'evolució del nombre de contagis o individus no-susceptibles. D'aquesta manera, i amb el càlcul numèric exposat, hem pogut observar les singularitats de cada model i veure l'impacte de les variables. També hem pogut utilitzar un dels models teòrics considerats per descriure un brot de Covid-19 a una residència.

Respecte al desenvolupament dels models epidemiològics, la compartimentació inherent en el model SIR ha estat ampliada per incloure conceptes com asintomatologia, aïllament, fases lleus i greus de la malaltia i, inclús, l'aparició de variants dels patògens que situen els individus recuperats, passat un temps, com a susceptibles novament. L'evolució indeterminista del nombre de contactes diaris amb el temps ha donat resposta als canvis de comportament social com els observats entre estacions o per una major conscienciació sobre la malaltia.

# Bibliografia

- [1] Porta, M. et al.: *A Dictionary of Epidemiology*, 6a edició, Oxford University Press, New York, 2014.
- [2] Kermack, W. O.; McKendrick, A. G.: A contribution to the mathematical theory of epidemics, *Proceedings of the Royal Society of London. Series A*, 115:700-721, 1927.
- [3] Kendall, D. G.: Deterministic and Stochastic Epidemics in Closed Populations, *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, 4:149-166, 1956.
- [4] Tuckwell, H. C.; Williams, R. J.: Some properties of a simple stochastic epidemic model of SIR type, *Mathematical Biosciences*, 208:76-97, 2007.
- [5] Ferrante, M; Ferraris, E; Rovira, C: On a stochastic epidemic SEIHR model and its diffusion approximation, *TEST*, 25:482-502, 2016.
- [6] Brzeźniak, Z.; Zastawniak, T.: *Basic stochastic processes : a course through exercises*, 7a edició, Springer undergraduate mathematics series, 2005.
- [7] Lawler, G. F.: *Introduction to Stochastic Processes*, 2a edició, Chapman & Hall/CRC, Boca Raton, 2006.
- [8] Kloeden, P. E.; Platen, E.: *Numerical Solution of Stochastic Differential Equations*, 3a edició, Springer, 1999.
- [9] Evans, L. C.: *An Introduction to Stochastic Differential Equations*, 1a edició, AMS, Berkeley, 2013.
- [10] Salah-Eldin, M. A.: Stochastic differential systems with memory: Theory, examples and applications, *Stochastic Analysis and related topics VI*, Springer Science, 1996.
- [11] Besalu, M.; Binotto, G.: Time-dependent non-homogeneous stochastic epidemic model of SIR type, *AIMS Mathematics*, 8(10):23218-23246, 2023.
- [12] Bardina, X; Ferrante, M; Rovira, C.: A stochastic epidemic model of COVID-19 disease, *AIMS Mathematics* 5(6):7661-7677, 2020.
- [13] McMichael, M. T. et al.: Epidemiology of Covid-19 in a Long-Term Care Facility in King County, Washington, *The New England Journal of Medicine*, 382(21):2005-2011, 2020.
- [14] He, X. et al.: Temporal dynamics in viral shedding and transmissibility of COVID-19, *Nature Medicine*, 26:672-675, 2020.

- [15] van Boven, M. et al.: Estimation of introduction and transmission rates of SARS-CoV-2 in a prospective household study, *PLoS Computational Biology* 20(1):e1011832, 2024.

# Annex A: Codi cas pràctic

Codi de la simulació numèrica del model SEIHR aplicat al cas de Covid-19 a la residència estudiada a [13]. Els paràmetres segueixen els valors discutits a la Taula 3.3. S'observa una iteració general sobre el nombre de simulacions i una iteració niuda sobre els passos de temps de cada simulació. Es segueix el mètode d'Euler-Maruyama.

```
1 % Paràmetres del model
2 n = 130 + 70; % Població total
3 N = 1; % Trobades diàries
4 r_E = 6; % Temps d'exposició
5 r_I = 2; % Temps d'infecció
6 r_H = 8; % Temps d'hospitalització
7 q_I = 0.06; % Probabilitat de transmissió infectat
8 q_H = 0.06; % Probabilitat de transmissió hospitalitzat
9 t_max = r_E + r_I + r_H + 26; % Duració en dies
10 lambda = 0; % Paràmetre de restricció
11 numsim = 100; % Número de simulacions
12 pas_t = 10; % Resolució temporal
13 dt = 1 / pas_t; % Pas temporal
14
15 % Temps en passos discrets
16 t_E = round(r_E * pas_t);
17 t_EI = t_E + round(r_I * pas_t);
18 t_EIH = t_EI + round(r_H * pas_t);
19
20 % Paràmetres de transmissió
21 theta_I = n * N * q_I;
22 theta_H = n * N * q_H;
23
24 % Matriu de resultats
25 V_all = zeros(numsim, t_max * pas_t);
26
27 % Proporció inicial de no-susceptibles
28 V0 = 1 / n;
29
30 % Dades experimentals
31 contagis_dades = [1, 2, 16, 20, 27, 36, 44, 48, 66, 81, 88, 128,
133, 136, 141, 142, 145, 148, 149, 150, 151];
32
33 % Proporció de no-susceptibles experimental
34 susceptible_data = contagis_dades / n;
35
36
```

```

37 % Iteració sobre el nombre de simulacions
38 for sim = 1:numsim
39     % Inicialitzar el vector de no-susceptibles
40     V = zeros(1, t_max * pas_t + 1); % Vector inicial amb zeros
41     V(t_EIH + 1) = V0; % Un únic infectat en t = t_EIH + 1
42
43     % Simulació
44     for t = t_EIH + 1 : t_max * pas_t
45         susceptible_factor = (1 - V(t)) * dt;
46         term_I = theta_I * (1 - lambda * (V(t - t_EI) - V(t -
            t_EIH))) / (1 - (V(t - t_EI) - V(t - t_EIH))) * (V(t -
            t_E) - V(t - t_EI));
47         term_H = theta_H * (V(t - t_EI) - V(t - t_EIH));
48
49         random_value = 2 * (rand < 0.5) - 1;
50         noise = sqrt(max(0, susceptible_factor * (term_I +
            term_H))) * random_value;
51
52         % Evolució de V
53         V(t + 1) = V(t) + susceptible_factor * (term_I + term_H) +
            noise / pas_t;
54         V(t + 1) = min(max(V(t + 1), 0), 1); % Limitar V a [0, 1]
55     end
56
57     % Guardar resultats
58     V_all(sim, :) = V(1:t_max * pas_t);
59 end
60
61 % Calcular mitjana de totes les simulacions
62 V_mean = mean(V_all, 1);
63
64 % Vector de temps ajustat
65 adjusted_time = (1:t_max * pas_t) * dt - (t_EIH + 5 * pas_t) * dt;
66
67 % Filtrar valors amb temps no negatius
68 non_negative_indices = adjusted_time >= 0;
69 adjusted_time_non_negative = adjusted_time(non_negative_indices);
70 V_mean_non_negative = V_mean(non_negative_indices);
71
72 % Graficar resultats sense temps negatius
73 figure;
74 plot(adjusted_time_non_negative, V_mean_non_negative, 'LineWidth',
    2); % Simulacions
75 hold on;
76
77 % Ajustar temps experimentals a 1 dia de distància
78 adjusted_time_data = (1:length(contagis_dades));
79 plot(adjusted_time_data, susceptible_data, 'ro-',
    'MarkerFaceColor', 'r', 'LineWidth', 2); % Dades experimentals
80
81 xlabel('t (dies)', 'FontSize', 12);
82 ylabel('Proporció no-susceptibles V(t)', 'FontSize', 12);
83 grid on;
84 legend('Simulacions', 'Dades experimentals', 'Location', 'Best');

```