



UNIVERSITAT DE
BARCELONA

Facultat de Matemàtiques
i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

Processos estocàstics de
ramificació i aplicacions a
problemes de la COVID 19

Autora: Júlia Loncà Castellnou

Director: Dra. Marta Sanz-Solé

Realitzat a: Departament de Matemàtiques i Informàtica

Barcelona, 22 de gener de 2021

Abstract

Epidemiological models have been used by scientists for decades to calculate the number of people infected with diseases that affect large populations.

The aim of this report is to introduce us to the world of stochastic processes and, in particular, of branching processes.

We will apply the theory seen in this report in a study on the control of Covid-19 spread on university campuses, both in the rooms and in the classrooms based on the article *Controlling the spread of COVID-19 on college campuses*[1].

Resum

Els models epidemiològics han estat utilitzats pels científics durant dècades per calcular el nombre d'infectats en malalties que afecten a poblacions grans.

L'objectiu d'aquesta memòria és el de introduir-nos en el món dels processos estocàstics i, en particular, en els processos de ramificació.

Aplicarem la teoria vista en aquesta memòria en un estudi sobre el control de la propagació de la Covid-19 en els campus universitaris, tant en les habitacions com en les aules basat en l'article *Controlling the spread of COVID-19 on college campuses*[1].

Agraïments

En primer lloc, agrair a la Dra. Marta Sanz-Solé per acceptar dirigir aquest treball, ajudar-me en tot moment i per les hores dedicades a corregir fins a l'últim detall.

Finalment, vull donar les gràcies a la meva família i la meva parella per creure en mi durant tota la carrera i sobretot a Laura, per corregir-me el treball tot i no entendre ni una paraula.

Índex

1	Introducció	1
2	Processos estocàstics	3
2.1	Cadenes de Markov	3
2.1.1	Distribució límit i estacionària	5
2.1.2	Exemples	7
2.2	Procés de Poisson	10
3	Processos de Ramificació	14
3.1	Mitjana i variància de la distribució de la descendència	14
3.2	Probabilitats d'extinció	16
4	Preliminars relatius a l'anàlisi del cas d'estudi	18
4.1	Nombre bàsic de reproducció	18
4.2	Polymerase Chain Reaction	19
4.3	Funció de Lambert	21
5	Cas d'estudi	22
5.1	Introducció	23
5.2	Model de bombolles i aplicació a residències universitàries	24
5.2.1	Model de bombolla	24
5.2.2	Habitacions individuals	25
5.2.3	Habitacions dobles	27
5.3	Propagació de la infecció en classes universitàries	29
5.3.1	Les mides de classe iguals redueixen la propagació	30
5.3.2	Model estocàstic per als escenaris 1 i 2	31
5.3.3	Fer classes online	32
6	Conclusions	34

1 Introducció

El projecte

Durant la pandèmia de la Covid-19 s'ha parlat sobre diversos termes tècnics, com R_0 i propagació del virus, el que no s'explica és que aquests termes formen part dels models epidemiològics, utilitzats per intentar descriure, replicar i predir la conducta dels agents infecciosos causants de la pandèmia.

Hi ha una gran quantitat de models epidemiològics per descriure la pandèmia mundial de la Covid-19, tant de tipus deterministes com estocàstics.

Els models deterministes són utilitzats per tractar malalties que afecten a poblacions grans i infecten als individus de manera aleatòria. Un dels models deterministes més famosos és el model *SIR*, creat l'any 1927 pels científics W. O. Kermack i A. G. McKendrick, el qual relaciona les variacions de les tres poblacions (Susceptible, Infectada i Recuperada) per mitjà de la taxa d'infecció i el període infecció promig. Aquest model *SIR* ha estat ampliat en diverses ocasions per donar lloc a models com *SIRS*, on els individus recobrats poden perdre la immunitat a la malaltia i tornar a formar part del grup de susceptibles i el model *SEIR*, que incorpora una nova població *E*, que són tots els individus de la població portadors de la malaltia però que al trobar-se en període d'incubació no presenten símptomes i poden o no estar en condició d'infectar a altres.

La transmissió de la Covid-19 és un fenomen complex, estocàstic i amb un fort component d'agrupació perquè el contagi és una probabilitat i no un factor determinista. Intentar descriure el seu desenvolupament amb xifres deterministes o fixes seria contraproduent. Els científics han desenvolupat durant dècades models estocàstics per descriure el procés de contagi.

Estructura de la Memòria

L'objectiu principal d'aquesta memòria ha estat l'estudi de l'article *Controlling the spread of COVID-19 on college campuses*², el qual es basa en estudiar estratègies a prendre per tal de minimitzar la difusió de la pandèmia en el model de residència universitària, tant a les residències com a les aules.

Per aquesta raó, la memòria està dividida en dues parts principals. En la primera enunciam i demostrarem la teoria que s'utilitza en l'article en els processos estocàstics i en els processos de ramificació. En la segona part, estudiarem més detingudament l'article esmentat, nomenant en cada pas la teoria aplicada per obtenir el resultat.

La primera part l'hem estructurada en tres seccions. En la primera secció donem els ingredients fonamentals dels processos estocàstics on introduïm els processos de Markov. Seguidament veiem un exemple important de Procés de Markov com és el Procés de Poisson.

En la segona secció introduïm els Processos de Ramificació, on veiem les propietats fonamentals. També definim les probabilitats d'extinció que utilitzarem en l'article i veiem alguns exemples.

En la tercera secció, definim algunes definicions necessàries per fer l'estudi de l'article, com és el nombre bàsic de reproducció, la funció de Lambert i una explicació del procés de *Polymerase Chain Reaction*, més conegut com *PCR*.

La segona part de la memòria es troba en la quarta secció, on es realitza l'estudi de l'article científic. Aquesta secció està dividida en diverses etapes. En la primera etapa es realitza una breu descripció del model matemàtic que s'utilitza, anomenat el model de bombolles. Després s'aplica aquest model a un model amb habitacions simples i habitacions dobles. En la segona etapa s'estudia la propagació de la infecció a les aules.

²ARXIV:2008.07293v1: BOROWIAK, M., NING, F., PEI, J., ZHAO, S., TUNG, H. AND DURRET, R. (2020) Controlling the spread of COVID-19 on college campuses

2 Processos estocàstics

Un procés estocàstic és una família de variables aleatòries $\{X_t, t \in T\}$ definides sobre un espai de probabilitat (Ω, A, P) , on T és un conjunt de índexs que, generalment, representarà el temps.

T pot ser continu si és un interval (el nombre dels seus valors és no numerable) i es denota com $\{X_t : t \geq 0\}$; es diu que el procés és discret si T és numerable (o finit) i generalment es denota com $\{X_t : t = 0, 1, 2, \dots\}$.

En general, les variables aleatòries d'un procés estocàstic no són independents. Aquesta és la situació més freqüent i en la que ens centrarem.

2.1 Cadenes de Markov

Una Cadena de Markov és un tipus especial de procés estocàstic en el qual la probabilitat que esdevingui un event depen únicament del passat immediat. Anem a definir-ho en el cas de temps discret i temps continu.

Definició 2.1. Una Cadena de Markov a temps discret $\{X_t, t \in \mathbb{N}\}$ és un procés estocàstic que compleix:

$$P(X_{n+1} = j | X_0 = i_0, \dots, X_n = i_n) = P(X_{n+1} = j | X_n = i_n). \quad (2.1)$$

Per tots els índexs n i tots els estats $i_0, \dots, i_n, j \in \mathbb{N}$.

Anem a definir ara una Cadena de Markov a temps continu.

Definició 2.2. Una Cadena de Markov a temps continu $\{X_t, t \in \mathbb{R}\}$ és un procés estocàstic tal que per qualsevol $0 \leq s_0 < s_1 < \dots < s_n < s$ es té:

$$P(X_{t+s} = j | X_s = i, X_{s_n} = i_n, \dots, X_{s_0} = i_0) = P(X_t = j | X_0 = i). \quad (2.2)$$

Per tots els possibles estats $i_0, \dots, i_n, i, j \in \mathbb{R}$.

Definició 2.3. S'anomena probabilitat de transició a la probabilitat que X_{n+1} estigui en l'estat j , donat que X_n es troba en l'estat i . És a dir:

$$p_{ij}^{n,n+1} = P(X_{n+1} = j | X_n = i). \quad (2.3)$$

Quan les probabilitats de transició són independents de la variable temps n , la Cadena de Markov ens diu que té probabilitats de transició estacionàries i escriurem $p_{ij}^{n,n+1} = p_{ij}$.

Atès que la gran majoria de cadenes de Markov amb què ens trobarem tenen probabilitats de transició estacionàries, limitem la nostra discussió a aquest cas.

Proposició 2.4. Els valors p_{ij} es poden combinar per formar una matriu de transició T de rang $n \times n$ on

$$T = [p_{ij}] = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{pmatrix} \quad (2.4)$$

La i -ésima fila de T per $i = 0, 1, \dots$ és la distribució condicional de X_{n+1} donat que $X_n = i$. És immediat que

$$\begin{aligned} p_{ij} &= P(X_{n+1} = j | X_n = i) \geq 0, & \text{per } i, j = 0, 1, \dots \\ \sum_{j=0}^n p_{ij} &= \sum_{j=0}^n P(X_{n+1} = j | X_n = i) = 1, & \text{per } i = 0, 1, \dots \end{aligned} \quad (2.5)$$

De manera que cada fila de la matriu representa una distribució de probabilitat.

Definició 2.5. El producte de dues matrius estocàstiques $n \times n$ és també estocàstic.

Demostració. Suposem dues matrius $A = [a_{ij}]$ i $B = [b_{ij}]$ estocàstiques de rang $n \times n$. Per la regla de multiplicació de matrius:

$$\begin{aligned} C &= [c_{ij}] = [a_{ij}] \cdot [b_{ij}] = \left[\sum_{k=1}^n a_{ik} b_{kj} \right] \\ \sum_{j=1}^n c_{ij} &= \sum_{j=1}^n \sum_{k=1}^n a_{ik} b_{kj} = \sum_{k=1}^n \sum_{j=1}^n a_{ik} b_{kj} = \sum_{k=1}^n a_{ik} \sum_{j=1}^n b_{kj} = \sum_{k=1}^n a_{ik} \cdot 1 = 1. \end{aligned}$$

□

Per aprofundir en la definició de cadenes de Markov, anem a veure'n un exemple senzill.

Exemple 2.6. En una certa regió, el temps atmosfèric segueix la següent seqüència: Es denomina un dia *solejat* (S) si hi ha sol més de la meitat del dia, i es denomina *nuvolat* (N) en cas contrari. Sabem que si un dia és nuvolat hi ha la mateixa probabilitat que l'endemà sigui solejat com nuvolat. Si un dia és solejat hi ha una probabilitat de $\frac{2}{3}$ que l'endemà sigui solejat.

Veiem clarament que el temps atmosfèric d'un dia en concret depen únicament del temps del dia anterior, per tant aquest és un exemple de Cadena de Markov.

1. Calculem la matriu de transició M del procés.

Les probabilitats són:

$$\begin{aligned} P[N|N] &= \frac{1}{2} & P[S|N] &= \frac{1}{2} \\ P[N|S] &= \frac{1}{3} & P[S|S] &= \frac{2}{3} \end{aligned}$$

Llavors la matriu de transició és:

	N	S
N	$\frac{1}{2}$	$\frac{1}{2}$
S	$\frac{1}{3}$	$\frac{2}{3}$

És a dir,

$$T = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} \quad (2.6)$$

2. Si avui és un dia nuvolat, quina probabilitat hi ha que en tres dies també sigui nuvolat?

Comencem els passos a partir del dia actual:

$$P^{(0)} = \begin{pmatrix} p_1^{(0)} & p_2^{(0)} \end{pmatrix} = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

D'aquesta manera, si avui està nuvolat en tres dies tindriem:

$$P^{(1)} = P^{(0)}T$$

$$P^{(2)} = P^{(1)}T = P^{(0)}T^2$$

$$\begin{aligned} P^{(3)} &= P^{(2)}T = P^{(0)}T^3 = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}^3 = \\ &= \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{29}{72} & \frac{43}{72} \\ \frac{43}{108} & \frac{65}{108} \end{pmatrix} = \begin{pmatrix} \frac{29}{72} & \frac{43}{72} \end{pmatrix} = (0.403 \quad 0.597) \end{aligned}$$

Així, la probabilitat que hi hagi un dia nuvolat en el tercer dia és $p_1^{(3)} = \frac{29}{72}$.

2.1.1 Distribució límit i estacionària

Partint d'un procés estocàstic, no sempre podrem calcular la probabilitat que es trobi en un estat concret en cada moment, per això buscarem una probabilitat d'equilibri que no depengui del moment inicial. A llarg termini, la distribució dependrà de l'estat on arribi, no del que hem partit. Anomenem a aquesta la distribució límit.

Definició 2.7. Una distribució límit d'una Cadena de Markov en temps discret $\{X_n\}_n$ és una successió $\{\pi_j\}_j$ que compleix:

$$\pi_j > 0, \forall j \geq 0, \quad \sum_{j=0}^{\infty} \pi_j = 1,$$

i

$$\lim_{n \rightarrow \infty} P(X_n = j | X_0 = i) = \pi_j, \forall j \in \mathbb{N} \quad (2.7)$$

Aquesta convergència significa que quan n tendeix a infinit, la probabilitat de trobar la Cadena de Markov en l'estat j és aproximadament π_j , sense importar en quin estat es trobava la cadena en el moment inicial.

Definició 2.8. Es diu que una Cadena de Markov $\{X_n\}_n$ té distribució estacionària si existeix una successió $\{\pi_j\}_j$ que compleix:

$$\pi_j > 0, \forall j \geq 0, \quad \sum_{j=0}^{\infty} \pi_j = 1,$$

i

$$\pi_j = \sum_{k=0}^{\infty} \pi_k P_{kj}, \forall j \in \mathbb{N} \quad (2.8)$$

Definició 2.9. Una matriu de transició és regular si existeix un nombre natural n tal que la potència n -ésima de la seva matriu de transició T té tots els seus elements positius.

Teorema 2.10. *Sigui T una matriu de probabilitat de transició regular en els estats $0, 1, \dots, N$. Llavors la distribució límit $\pi = (\pi_0, \pi_1, \dots, \pi_N)$ és també estacionària, és a dir, és l'única solució no negativa de les equacions*

$$\begin{aligned} \pi_j &= \sum_{k=0}^N \pi_k p_{kj}, & j &= 0, 1, \dots, N \\ \sum_{k=0}^N \pi_k &= 1. \end{aligned} \tag{2.9}$$

Demostració. Sabent que la Cadena de Markov és regular, tenim una distribució límit, $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \pi_j$, per al qual $\sum_{k=0}^N \pi_k = 1$. Escrivim T^n com al producte de la matriu $T^{n-1}T$ en la forma

$$p_{ij}^{(n)} = \sum_{k=0}^N p_{ik}^{(n-1)} p_{kj}, \quad j = 0, 1, \dots, N \tag{2.10}$$

i ara deixem $n \rightarrow \infty$. Aleshores $p_{ij}^{(n)} \rightarrow \pi_j$, mentre que $p_{ik}^{(n-1)} \rightarrow \pi_k$, i (2.10) passa a $\pi_j = \sum_{k=0}^N \pi_k p_{kj}$ tal com volíem demostrar.

Queda per demostrar que la solució és única. Suposem que x_0, x_1, \dots, x_N es resol

$$x_j = \sum_{k=0}^N x_k p_{kj}, \quad j = 0, 1, \dots, N \tag{2.11}$$

i

$$\sum_{k=0}^N x_k = 1. \tag{2.12}$$

Volem mostrar que $x_j = \pi_j$. Comencem multiplicant (2.11) a la dreta per p_{jl} , i a continuació, sumem j per obtenir

$$\sum_{j=0}^N x_j p_{jl} = \sum_{j=0}^N \sum_{k=0}^N x_k p_{kj} p_{jl} = \sum_{k=0}^N x_k p_{kl}^{(2)}, \tag{2.13}$$

Però per (2.11), tenim $x_l = \sum_{j=0}^N x_j p_{jl}$, d'on esdevé (2.13)

$$x_l = \sum_{k=0}^N x_k p_{kl}^{(2)}, \quad l = 0, 1, \dots, N.$$

Repetint aquest argument n vegades deduïm

$$x_l = \sum_{k=0}^N x_k p_{kl}^{(n)}, \quad l = 0, 1, \dots, N.$$

i després passant al límit en n i fent servir $p_{kl}^{(n)} \rightarrow \pi_l$, veiem que

$$x_l = \sum_{k=0}^N x_k \pi_l, \quad l = 0, 1, \dots, N.$$

Però per (2.12) tenim $\sum_k x_k = 1$, d'on $x_l = \pi_l$, tal com volíem demostrar. □

2.1.2 Exemples

S'utilitzen cadenes de Markov en teoria de genètica de poblacions per descriure el canvi de freqüències gèniques en una població petita amb generacions discretes. Anem a veure un exemple.

Exemple 2.11. *Exemple de cadenes de Markov en genètica*³

Anem a investigar la fluctuació de la freqüència genètica sota la influència de la mutació i la selecció.

Comencem descrivint el model haploide simple de reproducció aleatòria, sense tenir en compte les mutacions ni les forces selectives. Suposem que estem davant un cas inicial amb un nombre fix de gens $2N$, compostos per gens de tipus a i tipus A . La composició de la pròxima generació està determinada per una distribució de Bernoulli de la següent manera: si la generació consta de j gens a i $2N - j$ gens A , llavors cada assaig resulta en a o A amb probabilitats

$$p_j = \frac{j}{2N} \quad \text{i} \quad q_j = 1 - \frac{j}{2N},$$

respectivament. Les seleccions repetides es fan amb la substitució. Mitjançant aquest procediment generem una Cadena de Markov $\{X_n\}$ on X_n és el nombre de gens a en la n -èsima generació entre una mida de població constant de $2N$ individus. La matriu de probabilitat de transició es calcula segons la distribució binomial com

$$P(X_{n+1} = k | X_n = j) = p_{jk} = \binom{2N}{k} p_j^k q_j^{2N-k}, \quad \forall j, k = 0, 1, \dots, 2N \quad (2.14)$$

Fixem-nos que, si $X_n = 0$ (o $2N$) llavors $X_{n+k} = 0$ (o $2N$, respectivament) per a tots els $k \geq 0$.

Una de les qüestions d'interès és determinar la probabilitat que, amb la condició $X_0 = i$, la població assoleixi la fixació, és a dir, que es converteixi en una població pura composta únicament per gens a o gens A . També és interessant determinar la velocitat d'aproximació a la fixació. Anem a examinar aquestes qüestions.

Un model més complet té en compte les pressions de mutació. Suposem que abans de la formació de la nova generació cada gen té la possibilitat de mutar, és a dir, de convertir-se en un gen de l'altre tipus. Concretament, suposem que per cada gen es produeix la mutació $a \rightarrow A$ amb probabilitat α , i $A \rightarrow a$ amb probabilitat β . De nou, suposem que la composició de la següent generació està determinada per lleis binomials $2N$. Els valors rellevants de p_j i q_j quan la generació inicial està formada per j gens de tipus a , es considera que les probabilitats són:

$$\begin{aligned} p_j &= \frac{j}{2N}(1 - \alpha) + \left(1 - \frac{j}{2N}\right) \beta, \\ \text{i} \\ q_j &= \frac{j}{2N}\alpha + \left(1 - \frac{j}{2N}\right) (1 - \beta). \end{aligned} \quad (2.15)$$

³PINSKY, M., KARLIN, S., An Introduction to Stochastic Modeling, (3.3.3):90-91, 2011

La justificació és la següent: suposem que les pressions de mutació operen en primer lloc, i després s'escull un nou gen per selecció aleatòria de la població. Ara, la probabilitat de seleccionar un gen a després que les forces de mutació hagin actuat és tan sols $\frac{1}{2N}$ el nombre de gens a presents; per tant, la probabilitat mitjana respecte a les possibles mutacions és simplement $\frac{1}{2N}$ el nombre mitjà de gens a després de la mutació. Aquest nombre mitjà és clarament $j(1 - \alpha) + (2N - j)\beta$, que condueix alhora a (2.15).

Les probabilitats de transició de la Cadena Markov associada es calculen mitjançant (2.14) utilitzant els valors de p_j i q_j , donats a (2.15).

Si $\alpha\beta > 0$, llavors la fixació no es produirà en cap estat. En canvi, com n tendeix a infinit, la funció de distribució d' X_n s'acostarà a una distribució d'estat estacionari d'una variable aleatòria ξ , on $P(\xi = k) = \pi_k$ ($k = 0, 1, 2, \dots, 2N$) ($\sum_{k=0}^n \pi_k = 1, \pi_k > 0$). La funció de distribució ξ s'anomena distribució de freqüència de gens estacionaris.

Tornem al model simple d'aparellament aleatori i anem a veure el concepte d'una força de selecció que opera a favor, per exemple, dels gens a . Suposem que volem imposar un avantatge selectiu per als gens a sobre els A de manera que el nombre relatiu de descendents tingui expectatives proporcionals a $1 + s$ i 1 , respectivament, on s sigui petit i positiu. Substituïm $p_j = \frac{j}{2N}$ i $q_j = 1 - \frac{j}{2N}$ per les següents equacions i construïm la següent generació mitjançant mostreig binomial.

$$p_j = \frac{(1+s)j}{2N+sj}, \quad q_j = 1 - p_j.$$

Si la població principal estava formada per gens a aleshores, a la propera generació, les mides de població esperades dels gens a i dels gens A , respectivament, són

$$p_j = 2N \frac{(1+s)j}{2N+sj}, \quad q_j = 2N \frac{2N-j}{2N+sj}.$$

La proporció de la mida esperada de la població dels gens a a gens A en la $(n + 1)$ -èssima generació és

$$\frac{1+s}{1} \times \frac{j}{2N-j} = \left(\frac{1+s}{1} \right) \left(\frac{\text{nombre de gens } a \text{ en la } n\text{-èssima generació}}{\text{nombre de gens } A \text{ en la } n\text{-èssima generació}} \right).$$

Exemple 2.12. Passejada aleatòria.

Quan parlem de passejades aleatòries, és útil per a la intuïció parlar de l'estat del sistema com la posició d'una 'partícula' en moviment.

$$T = \begin{pmatrix} r_0 & p_0 & 0 & \cdots & 0 & \cdots \\ q_1 & r_1 & p_1 & \cdots & 0 & \cdots \\ 0 & q_2 & r_2 & \cdots & 0 & \cdots \\ & \ddots & & & \ddots & \\ \cdots & 0 & q_i & r_i & p_i & 0 & \cdots \\ & & \ddots & & & \ddots & \end{pmatrix}$$

on $p_i > 0$, $q_i > 0$, $r_i \geq 0$, i $q_i + r_i + p_i = 1, i = 1, 2, \dots (i \geq 1)$, $p_0 \geq 0$, $r_0 \geq 0$, $r_0 + p_0 = 1$. Específicament, si $X_n = i$, llavors per $i \geq 1$,

$$P(X_{n+1} = i + 1 | X_n = i) = p_i,$$

$$P(X_{n+1} = i - 1 | X_n = i) = q_i,$$

$$P(X_{n+1} = i | X_n = i) = r_i,$$

l'estat N , començant per $X_0 = i$. Les probabilitats de ruïna són:

$$u_i = pu_{i+1} + qu_{i-1}, \quad \text{per } i = 1, \dots, N - 1$$

juntament amb les condicions límit evidents:

$$u_0 = 1 \quad \text{i} \quad u_N = 0$$

Les probabilitats de ruïna del jugador corresponents a la matriu de probabilitat de transició (2.17) són

$$u_i = P(X_n \text{ arriba a l'estat } 0 \text{ abans que l'estat } N | X_0 = i) = \begin{cases} \frac{N-i}{N} & \text{quan } p = q = \frac{1}{2}, \\ \frac{(q/p)^i - (q/p)^N}{1 - (q/p)^N} & \text{quan } p \neq q. \end{cases} \quad (2.18)$$

Les probabilitats de ruïna u_i donades per (2.18) tenen la següent interpretació. En un joc on el jugador A comença amb una fortuna inicial de i unitats i el jugador B comença amb $N - i$ unitats, la probabilitat que el jugador A perdi tots els seus diners abans que el jugador B perdi el joc ve donada per u_i , on p és la probabilitat que el jugador A guanyi en un sol concurs. Si el jugador B és infinitament ric ($N \rightarrow \infty$), passa al límit de (2.18) i fa servir $(q/p)^N \rightarrow \infty$ si $p < q$, mentre que $(q/p)^N \rightarrow 0$ si $p > q$, veiem que les probabilitats de ruïna són:

$$u_i = \begin{cases} 1 & \text{si } p \leq q, \\ \left(\frac{q}{p}\right)^i & \text{si } p > q. \end{cases} \quad (2.19)$$

(En passar al límit, el cas $p = q = \frac{1}{2}$ s'ha de tractar per separat.) Veiem que la ruïna és certa ($u_i = 1$) contra un adversari infinitament ric quan el joc és desfavorable ($p < q$), i fins i tot quan el joc és just ($p = q$). En un joc favorable ($p > q$), començant per la fortuna inicial i , es produeix la ruïna del jugador A amb probabilitat $(q/p)^i$. Aquesta probabilitat de ruïna disminueix a mesura que augmenta la fortuna inicial. En un joc favorable contra un oponent infinitament ric, amb probabilitat $1 - (q/p)^i$ la fortuna del jugador A augmenta, a la llarga, sense límit.

2.2 Procés de Poisson

El procés de Poisson és un procés estocàstic en temps continu que consisteix en un recompte d'esdeveniments amb probabilitat baixa, anomenats esdeveniments "rars", que tenen lloc al llarg del temps.

El temps entre cada parell d'esdeveniments consecutius té una distribució exponencial amb paràmetre λ , on cadascun d'aquests temps són independents entre si.

En la teoria de la probabilitat i la estadística, la distribució de Poisson és una distribució de probabilitat discreta. La distribució de Poisson es produeix sovint en fenòmens naturals i s'utilitza en situacions on es desconeix el nombre total de possibles resultats.

Definició 2.13. *Una variable aleatòria X té distribució de Poisson de paràmetre $\lambda > 0$ amb probabilitat donada per:*

$$p_k = P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad \text{per } k = 0, 1, \dots \quad (2.20)$$

Teorema 2.14. *Siguin X i Y variables aleatòries independents amb distribucions de Poisson amb paràmetres λ i ν , respectivament. Llavors, la suma $X + Y$ té distribució de Poisson amb paràmetre $\lambda + \nu$.*

Demostració. Comencem calculant la probabilitat de $X + Y$ segons la llei de les probabilitats totals, i sabent que X i Y són independents,

$$\begin{aligned}
 P[X + Y = n] &= \sum_{k=0}^n P[X = k, Y = n - k] = \\
 &= \sum_{k=0}^n P[X = k]P[Y = n - k] = \\
 &= \sum_{k=0}^n \left(\frac{\lambda^k e^{-\lambda}}{k!} \right) \left(\frac{\nu^{n-k} e^{-\nu}}{(n-k)!} \right) = \\
 &= \frac{e^{-(\lambda+\nu)}}{n!} \sum_{k=0}^n \frac{n!}{k!(n-k)!} \lambda^k \nu^{n-k}.
 \end{aligned} \tag{2.21}$$

La expressió binomial de $(\lambda + \nu)^n$ és:

$$(\lambda + \nu)^n = \sum_{k=0}^n \frac{n!}{k!(n-k)!} \lambda^k \nu^{n-k}.$$

Aplicant aquesta expressió a (2.21):

$$P[X + Y = n] = \frac{e^{-(\lambda+\nu)}(\lambda + \nu)^n}{n!}, \quad n = 0, 1, \dots$$

Obtenint així la distribució de Poisson amb paràmetre $\lambda + \nu$. □

Proposició 2.15. *La distribució binomial s'aproxima a la distribució de Poisson quan n tendeix a infinit i p prou petita, de forma que $\lambda = \lim_{n \rightarrow \infty} np$, amb λ constant. Si això passa, la distribució binomial s'aproxima a una distribució de Poisson de paràmetre $\lambda = \lim_{n \rightarrow \infty} np$.*

Demostració. Suposem que tenim la distribució binomial

$$P(x) = \binom{n}{x} p^x (1-p)^{n-x}, \tag{2.22}$$

Substituïm a la distribució binomial $p = \frac{\lambda}{n}$, tindrem la següent expressió:

$$\begin{aligned}
P(x) &= \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} = \\
&= \frac{n!}{x!(n-x)!} \frac{\lambda^x}{n^x} \left(1 - \frac{\lambda}{n}\right)^{n-x} = \\
&= \frac{1 \cdot 2 \cdot 3 \cdots n}{x!(1 \cdot 2 \cdot 3 \cdots (n-x))} \frac{\lambda^x}{n^x} \left(1 - \frac{\lambda}{n}\right)^{n-x} = \\
&= \frac{n(n-1)(n-2) \cdots (n-x+1)}{x!} \frac{\lambda^x}{n^x} \left(1 - \frac{\lambda}{n}\right)^{n-x} = \\
&= \frac{n(n-1)(n-2) \cdots (n-x+1)}{n^x} \frac{\lambda^x}{x!} \left(1 - \frac{\lambda}{n}\right)^{n-x} = \\
&= \frac{n}{n} \left(\frac{n-1}{n}\right) \cdots \left(\frac{n-x+1}{n}\right) \frac{\lambda^x}{x!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x} = A \cdot G \\
\text{on } A &= \frac{n}{n} \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{x-1}{n}\right) \text{ i } G = \frac{\lambda^x}{x!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x}
\end{aligned}$$

Calculant ara els límits de A i G obtenim:

$$\begin{aligned}
\lim_{n \rightarrow \infty} A &= \lim_{n \rightarrow \infty} \frac{n}{n} \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{x-1}{n}\right) = 1 \cdots 1 = 1, \\
\lim_{n \rightarrow \infty} G &= \lim_{n \rightarrow \infty} \frac{\lambda^x}{x!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x} = \frac{\lambda^x}{x!} \cdot e^{-\lambda} \cdot 1.
\end{aligned}$$

Fent ara $\lim_{n \rightarrow \infty} A \cdot G = \frac{\lambda^x}{x!} e^{-\lambda}$ que és la funció de distribució de Poisson. \square

Definició 2.16. *Sigui H una variable aleatòria amb valors enters no negatius $0, 1, 2, \dots$ definim la funció generatriu de probabilitat de H com*

$$G_H(z) = E[z^H] = \sum_{k=0}^{\infty} z^k p_k.$$

Teorema 2.17. *Siguin X i Y variables aleatòries independents amb funcions generatrius ϕ_X i ϕ_Y , respectivament. Llavors, la suma $X + Y$ té funció generatriu $\phi_X \phi_Y$.*

Demostració. Si X i Y variables aleatòries independents amb funcions generatrius ϕ_X i ϕ_Y , respectivament, llavors la funció generatriu de la suma $X + Y$ ve donada per:

$$\phi_{X+Y} = E[z^{(X+Y)}] = E[z^X z^Y] = E[z^X] E[z^Y] = \phi_X \phi_Y, \quad (2.23)$$

Aquest resultat s'extén a la suma de n variables aleatòries independents: si $S_n = X_1 + \dots + X_n$,

$$\phi_{S_n} = \prod_{i=1}^n E[z^{X_i}] = \prod_{i=1}^n \phi_{X_i},$$

\square

Exemple 2.18. La funció generatriu de la distribució de Poisson amb valor esperat λ d'una variable aleatòria X és:

$$G_\lambda(z) = E[z^X] = \sum_{k=0}^{\infty} z^k p_k = \sum_{k=0}^{\infty} z^k \frac{\lambda^k e^{-\lambda}}{k!} = e^{\lambda(z-1)}. \quad (2.24)$$

El Procés de Poisson s'utilitza bàsicament per modelar els anomenats processos de cues. En ells es poden incloure molts processos: cotxes que arriben al peatge d'una autopista, clients que arriben a un banc, etc.

A continuació, anem a veure la definició del Procés de Poisson.

Definició 2.19. Un Procés de Poisson amb paràmetre $\lambda > 0$, és un procés estocàstic amb valors enters $\{X(t); t \geq 0\}$ pels quals:

1. $X(t)$ té increments independents, és a dir, per qualsevol punt del temps $t_0 = 0 < t_1 < t_2 < \dots < t_n$, les variables aleatòries

$$X(t_1) - X(t_0), X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1}),$$

són independents.

2. Donats $s \geq 0$ i $t > 0$, la variable aleatòria $X(s+t) - X(s)$ té la distribució de Poisson amb paràmetre λt ,

$$P[X(s+t) - X(s) = k] = \frac{(\lambda t)^k e^{-\lambda t}}{k!}.$$

3. $X(0) = 0$.

3 Processos de Ramificació

Els processos de ramificació ofereixen un model matemàtic per a l'estudi de poblacions en els quals els individus viuen, es reproduïen i moren independentment els uns dels altres. Amb aquesta eina es poden respondre diferents preguntes, com ara: quines poblacions s'extingeixen? Com es comporten aquelles que no ho fan?, etc.

El model de Galton-Watson, també anomenat *Procés de Ramificació Simple*, és el model més antic i simple dels processos de ramificació en temps discret. Aquest procés modela l'evolució d'una població en la qual cada individu té descendència de manera independent seguint la mateixa llei de probabilitat. Les poblacions estudiades en el Procés de Galton-Watson no tenen en compte les generacions anteriors, només tenen un tipus d'individu i el temps es mesura de manera discreta.

Definim les variables aleatòries $\{\xi_j^{(n)} : n = 0, 1, \dots; j = 0, 1, \dots\}$ independents i idènticament distribuïdes (v.a.i.i.d.) amb valors enters no negatius i amb distribució $\{p_k\}_{k \geq 0}$. Aquestes variables representen el nombre de descendents que té l'individu j en la generació n :

$$P(\xi_j^{(n)} = k) = p_k, \quad \text{per } k = 0, 1, 2, \dots \quad (3.1)$$

on $p_k \geq 0$ i $\sum_{k=0}^{\infty} p_k = 1$.

Definició 3.1. *El Procés de Galton-Watson és un procés estocàstic $\{X_n\}_{n \leq 0}$ que es defineix de manera recursiva de la manera següent: en la n -èsima generació X_n , els individus donen lloc a un nombre independent de descendents $\xi_1^{(n)}, \xi_2^{(n)}, \dots, \xi_{X_n}^{(n)}$ i, per tant, el nombre acumulat produït per la següent generació $(n+1)$ és*

$$X_{n+1} = \xi_1^{(n)} + \xi_2^{(n)} + \dots + \xi_{X_n}^{(n)}. \quad (3.2)$$

Aquesta equació caracteritza l'evolució del Procés de Ramificació com a sumes aleatòries successives de variables aleatòries.

Definició 3.2. *La funció generatriu d'un Procés de Galton-Watson ve determinada per la funció:*

$$\phi(s) = \sum_{k=0}^{\infty} s^k p_k, \quad |s| \leq 1$$

on s és una variable complexa.

3.1 Mitjana i variància de la distribució de la descendència

Definició 3.3. *Sigui $\mu = E[\xi]$ i $\sigma^2 = \text{Var}[\xi]$ la mitjana i la variància respectivament de la distribució de la descendència (3.1). Sigui $M(n)$ i $V(n)$ la mitjana i la variància de X_n sota la condició inicial $X_0 = 1$. Llavors, la mitjana i la variància respecte la suma aleatòria (3.2) dóna les igualtats següents:*

$$M(n+1) = \mu M(n), \quad (3.3)$$

$$V(n+1) = \sigma^2 M(n) + \mu^2 V(n), \quad (3.4)$$

Podem obtenir aquestes equacions en l'apartat (2.3.2) del llibre d'en Pinsky i en Karlin⁴.

Proposició 3.4. *En general, la mitjana i la variància de la distribució de descendència es poden expressar de la manera següent:*

$$M(n) = \mu^n, \quad \text{per } n = 0, 1, \dots$$

$$V(n) = \sigma^2 \mu^{n-1} \times \begin{cases} n & \text{si } \mu = 1 \\ \frac{1-\mu^n}{1-\mu} & \text{si } \mu \neq 1 \end{cases}, \quad \text{per } n = 0, 1, \dots$$

Demostració. La condició inicial $X_0 = 1$ inicia les equacions (3.3) i (3.4) a $M(0)=1$ i $V(0)=0$. Llavors, a partir de (3.3) obtenim:

$$M(1) = \mu,$$

$$M(2) = \mu M(1) = \mu^2,$$

$$M(3) = \mu M(2) = \mu^2 M(1) = \mu^3,$$

i, en general,

$$M(n) = \mu^n, \quad \text{per } n = 0, 1, \dots$$

Per tant, la mida mitjana de la població:

- augmenta geomètricament quan $\mu > 1$,
- es manté constant quan $\mu = 1$,
- disminueix geomètricament quan $\mu < 1$.

Aquest fet implica que quan $\mu > 1$, hi ha un creixement exponencial en la mitjana de la distribució de descendència. Aquests càlculs són utilitzats, per exemple, en fenòmens com una pandèmia.

A continuació, substituint $M(n) = \mu^n$ en l'equació (3.4) ens queda l'expressió $V(n+1) = \sigma^2 \mu^n + \mu^2 V(n)$, i calculant-ho de forma recursiva amb $V(0) = 0$ produeix

$$V(1) = \sigma^2$$

$$V(2) = \sigma^2 \mu + \mu^2 V(1) = \sigma^2 \mu + \sigma^2 \mu^2,$$

$$V(3) = \sigma^2 \mu^2 + \mu^2 V(2) = \sigma^2 \mu^2 + \sigma^2 \mu^3 + \sigma^2 \mu^4,$$

i, en general,

$$V(n) = \sigma^2 [\mu^{n-1} + \mu^n + \dots + \mu^{2n-2}] =$$

$$= \sigma^2 \mu^{n-1} [1 + \mu + \dots + \mu^{n-1}] =$$

$$= \sigma^2 \mu^{n-1} \times \begin{cases} n & \text{si } \mu = 1 \\ \frac{1-\mu^n}{1-\mu} & \text{si } \mu \neq 1 \end{cases}$$

Per tant, la variància de la mida de la població:

- augmenta geomètricament quan $\mu > 1$,
- es manté constant quan $\mu = 1$,
- disminueix geomètricament quan $\mu < 1$. □

⁴PINSKY, M.A. AND KARLIN, S. (2011) *An Introduction to STOCHASTIC MODELING*, (4th edition), (2.3.2)

3.2 Probabilitats d'extinció

En moltes ocasions i per diferents motius, les espècies d'animals comencen a desaparèixer, i això es coneix com a extinció.

Les causes de l'extinció d'una espècie poden ser molt variades, com la destrucció del seu hàbitat natural, la caça o captura d'animals exòtics, la contaminació, la desforestació o el canvi climàtic, que són conseqüència de la intervenció dels éssers humans. També pot relacionar-se amb diferents fenòmens naturals com inundacions, incendis o l'arribada de malalties.

Si una malaltia en particular s'elimina a tot el món, es considera erradicada. L'única malaltia declarada oficialment erradicada fins el dia d'avui, és la verola. Aquest fet es va aconseguir gràcies a les vacunes i a l'aïllament de les persones infectades i les que hi havien estat en contacte.

Definició 3.5. *L'extinció de la població es produeix en el moment en que la seva mida es redueix a zero, en conseqüència, si aquest moment no arriba mai, l'extinció de la població no es produeix. El temps aleatori d'extinció és el primer cop que X_n assoleix el valor 0 i després, obviament $X_k = 0$, per tots els $k > n$.*

En la terminologia de la cadena de Markov, 0 és un estat absorbent, i podem calcular la probabilitat d'extinció invocant una anàlisi del primer pas. Sigui

$$u_n = P(X_n = 0), \quad (3.5)$$

la probabilitat d'extinció en la n -èsima generació o abans, començant per un únic $X_0 = 1$. Suposem que l'únic representat per $X_0 = 1$ dóna lloc a $\xi_1^{(0)} = k$ descendents. Al seu torn, cadascun d'aquests descendents generarà una població dels seus propis descendents i, si la població original ha d'extingir-se en n generacions, cada una d'aquestes línies de descendència ha de desaparèixer en $n - 1$ generacions. Podem trobar aquest anàlisi en la Figura 1.

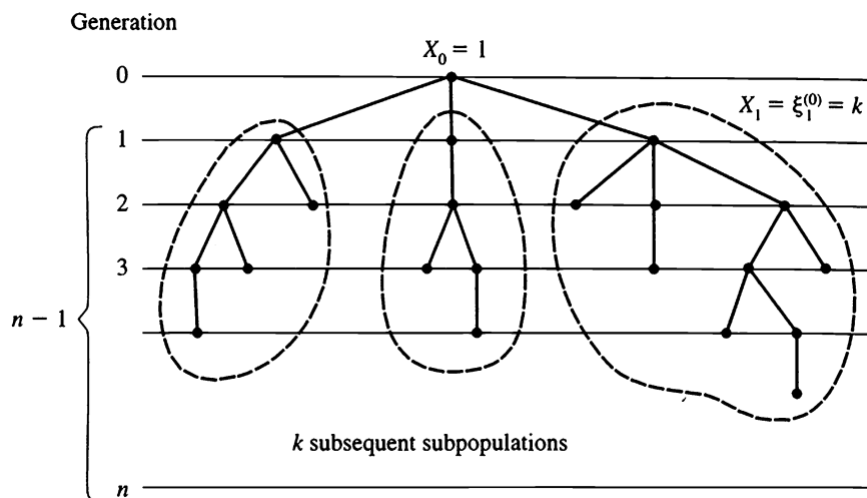


Figura 1: El diagrama il·lustra que si la població original ha de morir en la generació n , les subpoblacions generades per descendents inicials diferents han de desaparèixer en $n - 1$ generacions.

Ara, les k subpoblacions generades per diferents descendents del X_0 original són independents i tenen les mateixes propietats estadístiques que la població original. Per tant, la probabilitat que un d'ells desaparegui en $n - 1$ generacions és u_{n-1} per definició, i la probabilitat que totes les k subpoblacions desapareguin en $n - 1$ generacions és la k -èsima potència $(u_{n-1})^k$ perquè són independents. En ponderar aquest factor per la probabilitat de k descendents i sumar segons la llei de la probabilitat total, obtenim

$$u_n = \sum_{k=0}^{\infty} p_k (u_{n-1})^k. \quad (3.6)$$

Partint de $u_0 = 0$ i $u_1 = p_0$ (la probabilitat que el X_0 original no tingui descendència).

Considerem un Procés de Ramificació, la mida de la població de la qual es denota n en l'etapa X_n .

Proposició 3.6. *Suposem que la distribució de descendència $p_k = P(X_n = k)$ té la funció generatriu $\phi(s) = E[s^{\xi}] = \sum_k s^k p_k$. Si $u_n = P(X_n = 0)$ és la probabilitat d'extinció per l'etapa n , llavors la recursió (3.6) en termes de funcions generatrius és:*

$$u_n = \sum_{k=0}^{\infty} p_k (u_{n-1})^k = \phi(u_{n-1}),$$

és a dir, coneixent la funció generatriu $\phi(s)$, podem calcular successivament les probabilitats d'extinció u_n començant per $u_0 = 0$ i després $u_1 = \phi(u_0)$, $u_2 = \phi(u_1)$, etc.

4 Preliminars relatius a l'anàlisi del cas d'estudi

Abans de començar amb el cas d'estudi, necessitem tenir uns coneixements bàsics sobre com s'identifiquen les persones infectades amb una malaltia. En el cas de la Covid-19, s'utilitza la *Polymerase Chain Reaction*, més coneguda amb les sigles PCR.

Per entendre l'article necessitem definir el nombre bàsic de reproducció (R_0), ja que aquest valor s'utilitza per determinar el procés de l'epidèmia. Veurem que si $R_0 > 1$ l'epidèmia creix, si $R_0 < 1$ l'epidèmia decreix i si $R_0 = 1$ l'epidèmia s'estanca.

També definirem la Funció de Lambert, una funció utilitzada en l'article i necessària per entendre el procediment que es dur a terme.

4.1 Nombre bàsic de reproducció

Definició 4.1. *En epidemiologia, el nombre bàsic de reproducció R_0 d'una infecció és la mitjana de casos esperats generats directament per un cas en una població. La definició suposa que tots els individus són susceptibles a la infecció, és a dir, cap altre individu està infectat, la població no està ni immunitzada ni vacunada.*

El nombre de reproducció R_0 és un dels paràmetres per reflectir la infecció que té una malaltia i determinar quina proporció de la població s'ha d'immunitzar mitjançant la vacunació per erradicar una malaltia. També s'utilitza aquest valor per avaluar intervencions alternatives per controlar un brot.

El nombre reproductiu bàsic es veu afectat per diversos factors:

- La taxa de contactes a la població d'acollida
- La probabilitat de transmissió d'infecció durant el contacte
- La durada de la infecció

El càlcul de R_0 depen del model probabilístic que s'agafi. En els models d'infecció d'ús habitual, quan $R_0 > 1$ la infecció podrà començar a propagar-se en una població i el virus s'extendria exponencialment per la població, però en canvi si $R_0 < 1$, cada persona infectada infectaria menys d'una persona de mitjana, i el virus s'extingiria. En general, com més gran sigui el valor de R_0 , més difícil serà de controlar l'epidèmia. Aquest fet es justifica amb l'estudi del Procés de Galton-Watson vist anteriorment.

Per evitar una propagació sostinguda de la infecció, la proporció de la població que ha de ser immunitzada (per no ser susceptible a la infecció) ha de ser superior a R_0 .

La *immunitat col·lectiva* s'assoleix quan una proporció significativa de la població ha estat vacunada (o és immune per qualsevol altre mecanisme), la qual cosa produeix protecció per a individus susceptibles (per exemple, individus no vacunats). Com més gran sigui el nombre de persones immunes en una població, menor serà la probabilitat que una persona susceptible entri en contacte amb la infecció. És més difícil que les malalties es propaguin entre individus si un gran nombre ja és immune a mesura que es trenca la cadena d'infecció.

El llindar d'immunitat del col·lectiu és la proporció d'una població que necessita ser immune per tal que una malaltia infecciosa es mantingui estable en aquesta comunitat. Si s'arriba a aquest moment, per exemple mitjançant la vacunació, cada cas condueix a un nou infectat ($R_0 = 1$) i la infecció es mantindria estable dins la població.

La secció a continuació considera un exemple important d'un Procés de Ramificació que descriu una de les eines més importants en biologia molecular, la reacció en cadena de la polimerasa (PCR).

4.2 Polymerase Chain Reaction

La PCR, acrònim en anglès de 'Polymerase Chain Reaction', és un sistema experimental per generar grans quantitats de material genètic a partir d'una mostra inicial petita per tal de detectar un fragment del material genètic d'un patogen. El resultat de la PCR és una gran ampliació d'una seqüència d'ADN en concret. En la pandèmia del coronavirus, s'utilitza per determinar si una persona està infectada o no amb coronavirus.

A més de detectar virus en una mostra, la PCR pot estimar la quantitat de virus present o càrrega viral al cos d'un individu. En malalties com l'hepatitis C o les infeccions pel virus de la immunodeficiència humana (VIH), la càrrega viral és una bona indicació de la malaltia que pot tenir, del funcionament del medicament o del tractament d'un individu. Amb aquesta informació els metges poden determinar quan començar el tractament i la seva resposta, creant així un tractament personalitzat per cada individu.

Abans d'iniciar la PCR, s'ha d'aïllar l'ADN d'una mostra. L'extracció d'ADN és un procés de diversos passos que es pot realitzar manualment o amb un instrument que prepara mostres de forma automàtica, sense la intervenció humana.

També es necessita tenir el fragment de l'ADN del patogen, el qual anomenarem *primer* (trossos curts d'ADN sintètic).

Després de la preparació de la mostra, comença el procés de la PCR en tres passos per cicle.

1. Separació de l'ADN diana: desnaturalització

En el primer pas de la PCR anomenat desnaturalització, el tub que conté la mostra d'ADN s'escalfa a més de 90°C (194°F), per separar l'ADN bicatenari en dues cadenes independents. La temperatura elevada trenca els enllaços relativament febles entre els nucleòtids que formen el codi d'ADN.

2. Fixació dels *primers* a la seqüència d'ADN: hibridació

El tub es refreda entre 40°C i 60°C graus centígrads (104°F i 140°F). En aquesta temperatura la PCR utilitza els *primers*, que s'uneixen o s'hibriden a les cadenes monocatenàries d'ADN en la regió que se li ha especificat. En el primer cicle s'utilitzen dos *primers*, un per cada cadena d'ADN individual recentment separada. En aquest moment les dues cadenes estan a punt per copiar-se.

3. Fer una còpia: extensió

En la tercera fase de la reacció, anomenada extensió, la temperatura augmenta fins aproximadament 72°C (161,5°F). Els nucleòtids s'afegeixen a les regions marcades pels *primers* per crear una nova cadena d'ADN complementària a cadascuna de les cadenes individuals. Després de completar l'extensió, s'han creat dues còpies idèntiques de l'ADN original.

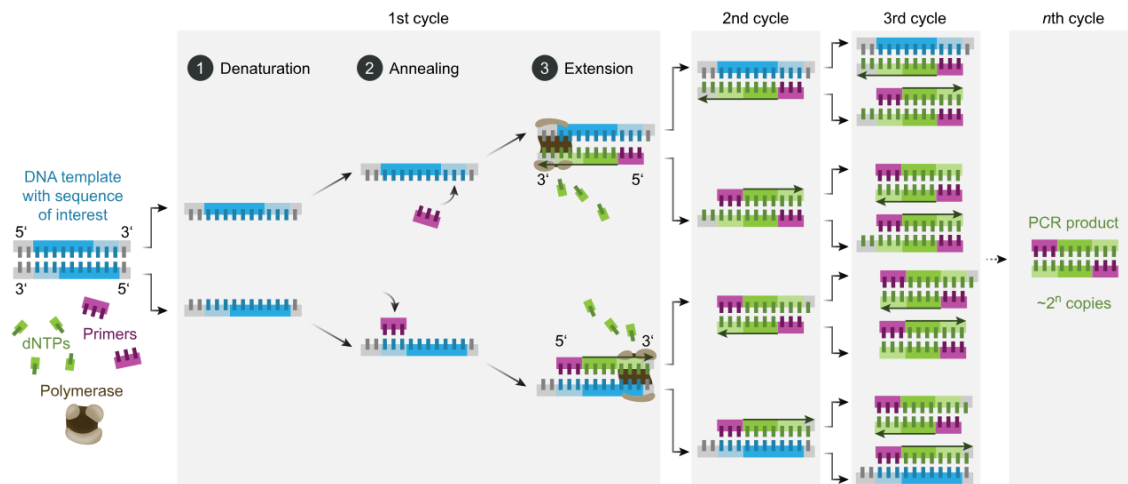


Figura 2: Imatge de n cicles de la PCR.

Desnaturalització: La cadena inicial bicatenària es divideix en dues cadenes monocatenàries (90°C).

Hibridació: els primers s'uneixen a les cadenes monocatenàries ($40^{\circ}\text{C} - 60^{\circ}\text{C}$).

Extensió: els nucleòtids s'uneixen a la cadena i el primer per crear una còpia exacta de l'ADN inicial (72°C).

Després del primer cicle de PCR en el qual s'han generat dues còpies de l'ADN original, el cicle torna a començar. Cada duplicat crea dues còpies noves i, després d'aproximadament 30 cicles de la PCR, s'hauran creat més de 1.000 milions de còpies del segment d'ADN original. Com el procés de la PCR és automàtic, es pot completar en poques hores. Els resultats obtinguts proporcionen informació als clínics i altres proveïdors d'assistència sanitària per orientar el tractament.

A través de la PCR, es localitza i s'amplifica un fragment de material genètic, que en el cas del coronavirus és una molècula d'ARN.

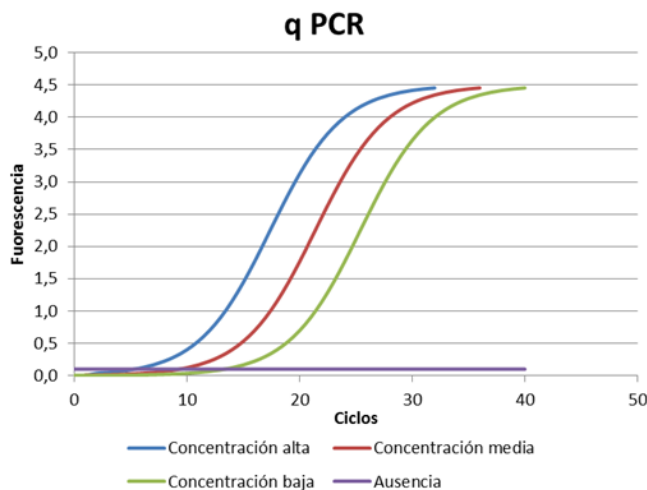


Figura 3: Exemple de resultats de quatre PCR's amb diferent càrrega viral.

Si després d'analitzar-lo en un laboratori la prova detecta l'ARN del virus, el resultat és positiu i es confirma que aquesta persona està infectada amb la malaltia. Que la tècnica de la PCR no detecti el material genètic del virus no implica que la persona no estigui infectada per aquest, vol dir que no hi ha suficient material genètic del virus en la mostra d'ADN per afirmar que la persona estigui infectada. Per aquesta raó quan hi ha una sospita clínica important, es recomana repetir la prova uns dies després per assegurar que el pacient no està infectat pel virus.

A diferència de la PCR, les proves ràpides d'antígens no identifiquen l'ARN del virus, sinó que detecten les proteïnes del virus (antígens) presents en mostres respiratòries d'exsudat nasofaringi.

Una altra forma de detectar si la persona està o ha estat infectada és a través d'una mostra de sang, en la qual es detecta si existeixen anticossos produïts contra el virus en la mostra.

4.3 Funció de Lambert

Al resoldre problemes de física, química o enginyeria, és habitual trobar-se amb equacions transcendents com

$$ye^y = a,$$

on y és la incògnita i a una constant. Aquestes equacions no poden resoldre en termes de funcions elementals com polinomis, exponencials, logaritmes o funcions trigonomètriques. Com es procedeix en aquests casos? Doncs amb la Funció de Lambert.

Definició 4.2. *La funció W de Lambert, també coneguda com funció Omega o log producte, és la inversa de la funció $f(w) = we^w$, on e^x és la funció exponencial natural i w és qualsevol nombre complex. La funció es defineix mitjançant W . Aquestes funcions tenen la propietat següent: si z i w són nombres complexos, llavors:*

$$z = W(z)e^{W(z)}.$$

En l'article que s'estudia en la secció a continuació, apareixen equacions transcendents en les quals es necessita la Funció de Lambert per resoldre-les.

5 Cas d'estudi

La Covid-19 és la malaltia infecciosa causada pel coronavirus que s'ha descobert més recentment. El primer brot d'aquest nou virus va donar-se a Wuhan (Xina) el desembre de 2019. Actualment la Covid-19 és una pandèmia que afecta la majoria de països a tot el món.

Els símptomes més habituals de la Covid-19 són febre, tos seca i cansament. Altres símptomes menys freqüents que afecten alguns pacients són congestió nasal, mal de cap, conjuntivitis, mal de coll, diarrea i pèrdua del gust o l'olfacte. La majoria de pacients solen tenir símptomes lleus i comencen gradualment. Altres pacients són asimptomàtics.

Una persona pot contraure la Covid-19 inhalant les partícules procedents d'una persona infectada pel virus. Les gotes amb el virus també es poden dipositar en diferents superfícies i objectes, que al tocar-los es poden adherir a les mans, el que és potencialment infecciós si la persona després amb aquesta mateixa mà es toca la boca, el nas o els ulls. Per aquesta raó la Universitat de Duke, igual que altres universitats de tot el món a l'inici de la pandèmia, es plantejaven com començar el semestre de tardor del 2020 de manera segura.

Les universitats d'arreu del món van optar per diferents formats de classe el curs 2020/2021. Mentre la Universitat de Harvard va optar per un model de classe 100% online, França i Espanya apostaven per un model mixt. En canvi Itàlia va anunciar que les classes tornarien a ser totalment presencials.

A les universitats amb *colleges* el tema es complicava, ja que els estudiants estan obligats a viure a les residències universitàries a la vegada que assisteixen a cursos al mateix campus. Les universitats amb aquesta obligatorietat es van haver de plantejar un canvi en la seva manera habitual de treballar per tal de reduir l'exposició dels estudiants i dels docents.

L'article que anem a analitzar a continuació, es centra en utilitzar models matemàtics per avaluar estratègies per suprimir la propagació del virus al campus de la Universitat de Duke, especialment a les bombolles i a les aules.

Per a les residències, el següent article confirma amb un estudi quantitatiu la impressió lògica que utilitzar habitacions individuals en lloc de dobles redueix substancialment la propagació del virus.

Per les aules, veurem que l'article dona evidències que, canviar les classes amb una mida superior a algun tipus de llindar de modalitat presencial a modalitat online pot fer disminuir la reproducció bàsica número $R_0 < 1$, ajudant així a l'extinció de l'epidèmia. El llindar dependrà del contagi de la malaltia a les aules.

5.1 Introducció

La pregunta més freqüent entre els estudiants i docents d'universitats de tot el món durant les vacances de primavera de 2020 va ser com les universitats afrontarien el semestre de tardor. Moltes van canviar el format de classes a un model a distància, restringint així el nombre d'estudiants al campus. Altres van modificar els horaris del semestre de tardor per no fer descansos a mig semestre, reduint així el contagi entre els alumnes que viuen als campus i els que marxen a casa durant les vacances.

Per controlar encara més la propagació del virus al campus, van canviar el sistema de menjador per un de tipus *take away* per fomentar el distanciament social, a més d'informes de símptomes diaris, augment de les proves, així com seguiment de contactes d'infectats i quarantena de persones infectades.

L'article que analitzem es divideix en dues seccions importants:

- La **primera secció** es centra en les residències dels alumnes i respon a la qüestió: Quant es reduiria la propagació si totes les habitacions dels estudiants fossin de tipus individual?

Aquesta és una qüestió interessant ja que en la Duke University, és obligatori que els estudiants visquin al campus tres dels seus quatre anys de formació. Això implica que en qualsevol moment, més del 80% dels estudiants viuen en les residències. Mitjançant l'ús d'una generalització del model de bombolla⁵, es veurà que la reducció pot ser substancial.

- En la **segona secció** centrem la nostra atenció a les aules i fem una pregunta en dues parts inspirada en Gressman i Peck⁶, els quals van analitzar un model complex basat en agents estocàstics per determinar l'impacte de tenir classes amb un número elevat d'alumnes només en línia.

- La primera part de la qüestió és: Com afecta la mida de classe en la propagació del virus? Per afrontar aquesta qüestió, es consideraran dos escenaris amb 1000 estudiants on cadascun d'ells assisteix a 3 classes. En el primer escenari, totes les classes tenen 30 alumnes. En el segon escenari, $\frac{1}{4}$ de les classes hi ha 60 estudiants i $\frac{3}{4}$ són de 20. En el segon cas demostrarem que el nombre de reproducció bàsica R_0 és un 48% més gran en l'escenari 2 a causa de l'augment de la propagació en classes més grans.

- Sabent que les classes més grans afavoreixen a la propagació del virus, és normal moure les classes que superin un determinat llindar d'alumnes inscrits, per aquesta raó, la segona part de la qüestió és: Quin hauria de ser el llindar per traslladar classes a no presencials? Per respondre a aquesta qüestió, partim d'un escenari amb 2405 estudiants, on cadascun d'ells assisteix a 3 classes i les classes de diferents mides, des de 10 fins a 120 alumnes inscrits. Aquest fet ens permetrà investigar una àmplia gamma de models. Es veurà que el llindar per moure les classes a model online per reduir el nombre de reproducció bàsica a valors $R_0 < 1$ i, per tant, aturar la propagació de l'epidèmia, depèn de la probabilitat bàsica d'infecció p del model.

⁵BALL, F., MOLLISON, D., AND SCALIA-TOMBA, G. (1997) Epidemics with two levels of mixing.

⁶GRESSMAN, P.T., AND PECK, J.R. (2020) Simulating COVID-19 in a University Environment
ARXIV:2006.03175

5.2 Model de bombolles i aplicació a residències universitàries

En aquesta secció es respon la pregunta: Quant es reduiria la propagació si totes les habitacions dels estudiants fossin de tipus individual?

Per respondre a aquesta qüestió, utilitzarem els resultats desenvolupats pel model de bombolla⁷. El model de bombolla millora lleugerament a un model de mescla homogeni dividint la població en diverses bombolles, on la probabilitat d'infectar un individu sigui més gran pels individus de la mateixa bombolla.

A continuació es veurà la teoria necessària per desenvolupar el model de bombolla, per després aplicar-ho a habitacions individuals i habitacions dobles.

5.2.1 Model de bombolla

Partim d'una població de mida $N = mn$ amb m bombolles de mida n . Per simplicitat, considerarem fix el temps d'infecció.

Utilitzarem dos nomenclatures, H per indicar la bombolla (*home* en anglès) i G per designar global.

Definim p_H com la probabilitat que un individu i infecti un individu j de la mateixa bombolla. Definim també p_G com la probabilitat que un individu i infecti un individu j que no és de la mateixa bombolla.

Pas 1. Començarem connectant cada parell d'individus de la mateixa bombolla amb probabilitat p_H .

π_k indica la probabilitat que un individu de la població pertanyi a un grup de mida k .

$\mu = \sum_k k\pi_k$ indica el valor mig del nombre d'infectats en les bombolles.

$G_H(z) = \sum_{k=1}^{\infty} \pi_k z^k$ és la funció generatriu de la mida de les bombolles.

Pas 2. Fem ara connexions a llarga distància entre individus de diferents bombolles amb probabilitat p_G .

Una persona té probabilitat p_G d'infectar a una altra. Definim $Y_1 = \text{Binomial}(N, p_G)$ com la variable aleatòria que dona el nombre d'infeccions causades per N individus amb probabilitat p_G , on la mida de la població N és gran i el paràmetre p_G s'espera que sigui un valor baix. Per això és d'esperar que Np_G sigui un valor més o menys constant i aplicant la Proposició 2.15, podem saber que el nombre d'infeccions a llarga distància causades per una persona és $Y_1 \approx \text{Poisson}(Np_G)$.

Com s'ha vist en el Exemple 2.18, la funció generatriu de Y_1 amb el paràmetre de Poisson $\lambda = Np_G$ és $G_\lambda(z) = e^{Np_G(z-1)}$.

Amb la informació vista fins al moment, recordant que μ indica el valor mig del nombre d'infectats en les bombolles i aplicant la definició que es dona del R_0 en aquest tipus de models que podem trobar en l'article *Epidemics with two levels of mixing*⁸,

$$R_0 = \mu N p_G. \quad (5.1)$$

⁷BALL, F., MOLLISON, D., AND SCALIA-TOMBA, G. (1997) Epidemics with two levels of mixing.

⁸BALL, F., MOLLISON, D., AND SCALIA-TOMBA, G. (1997) Epidemics with two levels of mixing. *Annals Applied Probability*. 7, p.48

Aplicant el Teorema 2.17 per un grup de mida k amb la mateixa funció generatriu $G_\lambda(z) = e^{Np_G(z-1)}$, la funció generatriu d'infectats resultant és $(e^{Np_G(z-1)})^k$, i podem veure que el nombre d'individus infectats per un individu triat a l'atzar és

$$\sum_{k=0}^{\infty} \pi_k \left(e^{Np_G(z-1)} \right)^k = G_H(G_\lambda(z)).$$

Un resultat per als processos de ramificació que podem trobar en l'article *Epidemics with two levels of mixing*⁹ és que la probabilitat d'una gran epidèmia és la més gran de les solucions de l'equació:

$$1 - \zeta = G_H(G_\lambda(\zeta + 1)). \quad (5.2)$$

Quan apliquem el model de residències als nostres models d'habitacions veient una habitació com una bombolla, G_H variarà entre models, ja que l'elecció de les habitacions individuals o dobles resulta en una distribució local lleugerament alterada dins d'una habitació. D'altra banda, atès que la propagació global de la infecció no es veu afectada per la mida de l'habitació, G_λ serà la mateixa per als dos escenaris d'habitacions.

5.2.2 Habitacions individuals

En adaptar la propagació de la infecció del model de bombolla a un model de residència universitària amb habitacions individuals, s'escriurà el p_H com p_D , ja que és més interessant veure la propagació en una residència (*dorm* en anglès) que en una bombolla. De la mateixa manera, s'escriurà el G_H com G_D^1 , amb el superíndex 1 indicant habitacions individuals.

Primer calculem R_0 .

Considerem una mida de població $N = mn$, amb m residències i n habitacions individuals. Si n és gran, el nombre total d'individus infectats en una residència com a conseqüència d'un infectat és la descendència total d'un procés de ramificació. Com el nombre d'estudiants en una residència és n , el procés de ramificació tindrà la descendència d'una distribució de *Poisson*(λ_1), on $\lambda_1 = np_D$. Suposem $np_D < 1$ de manera que la mitjana de la descendència total, aplicant la Proposició 3.4, és

$$\mu_1 = 1 + \lambda_1 + \lambda_1^2 + \dots = \frac{1}{1 - \lambda_1}. \quad (5.3)$$

Cada individu infectat en la residència infecta una mitjana d'individus Np_G fora de la residència, de manera que el nombre bàsic de reproducció R_0 és

$$R_0 = \mu_1 Np_G = \frac{Np_G}{1 - np_D}, \quad (5.4)$$

com a resultat de multiplicar la mitjana d'individus per el valor mig de la variable aleatòria que dona el nombre d'individus infectats per un individu.

⁹BALL, F., MOLLISON, D., AND SCALIA-TOMBA, G. (1997) Epidemics with two levels of mixing. *Annals Applied Probability*. 7, p. 53

Anem a obtenir la funció generatriu de la mida dels grups G_D^1 . Aplicant l'Exemple 2.18 de la funció generatriu, i sabent que $\lambda_1 = np_D$, podem afirmar que la funció generatriu de la propagació de la infecció en la residència és $G_b^1(z) = e^{np_D(z-1)}$.

Desglossant segons el nombre d'individus infectats directament pel primer individu, la funció generatriu de la mida de l'epidèmia en la residència causada per un infectat és

$$G_D^1(z) = z \sum_{k=0}^{\infty} e^{-\lambda_1} \frac{\lambda_1^k}{k!} G_D^1(z)^k = z e^{\lambda_1(G_D^1(z)-1)} = z G_b^1(G_D^1(z)), \quad (5.5)$$

on z representa la proporció inicial d'alumnes infectats en la residència.

En el cas de les habitacions individuals, es pot trobar una solució explícita per a G_D^1 prenent la derivada de l'equació (5.5) per obtenir l'equació diferencial

$$G_D^{1'}(z) = \frac{G_D^1(z)}{z} + \lambda_1 z G_D^1(z) G_D^{1'}(z),$$

amb la condició límit $G_D^1(1) = 1$. L'equació diferencial $G_D^1(z)$ té solució

$$G_D^1(z) = \frac{-W(zW^{-1}(-\lambda_1))}{\lambda_1}, \quad (5.6)$$

on W és la funció de Lambert W la qual està definida com $W^{-1}(z) = ze^z$.

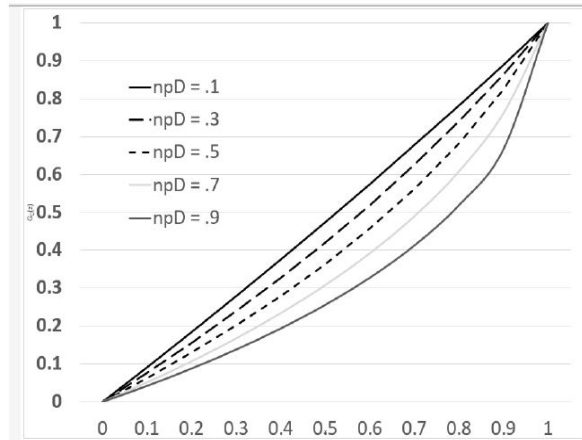


Figura 4: Funció generatriu $G_D^1(z)$ per a $np_D = 0.1, 0.3, 0.5, 0.7, 0.9$. Resolt per iteració amb l'equació (5.6). z representa la proporció inicial d'alumnes infectats al dormitori.

Per (5.2), la probabilitat d'una gran epidèmia és la solució més gran de

$$1 - \zeta = G_D^1(e^{-Np_G\zeta}). \quad (5.7)$$

Aplicant l'equació diferencial G_D^1 en funció de la funció de Lambert W , obtenim les igualtats següents:

$$\begin{aligned}
1 - \zeta &= G_D^1(e^{-Np_G\zeta}) = \\
&= \frac{-W(e^{-Np_G\zeta}W(-\lambda_1))}{\lambda_1} = \frac{-W(e^{-Np_G\zeta}(-\lambda_1 e^{-\lambda_1}))}{\lambda_1} = \\
&= \frac{-W(-\lambda_1 e^{-Np_G\zeta - \lambda_1})}{\lambda_1} = \frac{-W(-\lambda_1 e^{-Np_G - \lambda_1\zeta + \lambda_1\zeta - \lambda_1})}{\lambda_1} = \\
&= \frac{-W(-\lambda_1 e^{-(Np_G + \lambda_1)\zeta} e^{\lambda_1(\zeta - 1)})}{\lambda_1}
\end{aligned}$$

és a dir,

$$1 - \zeta = \frac{-W(-\lambda_1 e^{-(Np_G + \lambda_1)\zeta} e^{\lambda_1(\zeta - 1)})}{\lambda_1}.$$

Multipliquem per λ_1 ,

$$\lambda_1(1 - \zeta) = -W(-\lambda_1 e^{-(Np_G + \lambda_1)\zeta} e^{\lambda_1(\zeta - 1)}),$$

a continuació, apliquem W^{-1} a les dues parts

$$\begin{aligned}
W^{-1}(\lambda_1(1 - \zeta)) &= W^{-1}\left(-W(-\lambda_1 e^{-(Np_G + \lambda_1)\zeta} e^{\lambda_1(\zeta - 1)})\right) \\
\lambda_1(\zeta - 1)e^{\lambda_1(\zeta - 1)} &= -\lambda_1 e^{-(Np_G + \lambda_1)\zeta} e^{\lambda_1(\zeta - 1)},
\end{aligned}$$

Recordant que $\lambda_1 = np_D$ i simplificant, obtenim

$$(1 - \zeta) = e^{(Np_G + np_D)\zeta},$$

que és la probabilitat d'una gran epidèmia d'un procés de ramificació amb una distribució de Poisson ($R_0 = Np_G + np_D$).

5.2.3 Habitacions dobles

Considerem ara una població de mida $N = 2m_1n_1$, amb m_1 residències i n_1 habitacions dobles. Per fer una millor comparació suposarem que, independentment de si les residències tenen únicament habitacions individuals o dobles, el nombre de persones en una residència sigui el mateix, és a dir, $n = 2n_1$.

Definim p_L com la probabilitat d'infectar al company d'habitació, i p_D la probabilitat d'infectar un individu dins de la mateixa residència.

Per trobar la funció generatriu de la propagació de la infecció en la residència (l'anàleg de G_b^1 en l'escenari anterior) hem de pensar que el progenitor es divideix en dues parts amb probabilitat p_L , una la reproducció dins l'habitació amb probabilitat $1 - p_L$ i l'altra dins la residència en la qual es formen connexions de llarga distància. La distribució de la descendència és ara una barreja de dos processos de Poisson:

$$(1 - p_L)Poisson(2n_1p_D) + p_LPoisson(2 \cdot 2n_1p_D).$$

La primera serà una Poisson amb paràmetre $2n_1p_D$, el nombre d'infeccions a llarga distància causades per una persona en una habitació.

La segona Poisson amb paràmetre $2 \cdot 2n_1p_D$ correspon al nombre d'infeccions a llarga distància causades per una persona en una residència.

El nombre mitjà d'infeccions en una generació, λ_2 , serà la suma de les mitjanes de les dues distribucions de Poisson de manera que $\lambda_2 = (1 + p_L)2n_1p_D$.

Imitant càlculs emprats en les habitacions individuals, si $\lambda_2 < 1$, la mitjana de la descendència total del procés de ramificació és

$$\mu_2 = 1 + \lambda_2 + \lambda_2^2 + \dots = \frac{1}{1 - \lambda_2}, \quad (5.8)$$

i el nombre bàsic de reproducció R_0 és

$$R_0 = \mu_2 N p_G = \frac{N p_G}{1 - (1 + p_L)2n_1p_D}. \quad (5.9)$$

Per calcular G_D^2 (anàleg de G_D^1 en l'escenari d'habitació individual) observem que la funció generatriu del nombre d'infeccions en una generació és

$$G_b^2(z) = (1 - p_L)e^{2n_1p_D(z-1)} + p_L \left(e^{2n_1p_D(z-1)} \right)^2, \quad (5.10)$$

així, G_D^2 satisfà:

$$G_D^2(z) = z G_b^2(G_D^2(z)), \quad (5.11)$$

i G_D^2 es pot resoldre amb el mètode d'iteració descrit en l'apartat d'habitacions individuals.

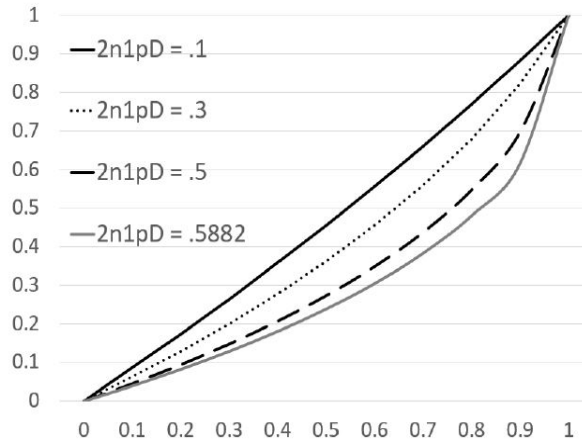


Figura 5: Funció generatriu $G_D^2(z)$ per a $2n_1p_D = 0.1, 0.3, 0.5, 0.5882$ i $p_L = 0.7$. Resolt per iteració amb l'equació (5.11). Quan $p_L = 0.7$ i $2n_1p_D > 0.5882$, hi ha una gran epidèmia en la residència. El gràfic de $G_D^2(z)$ augmenta independentment del valor de $2n_1p_D$.

Un cop calculat el $G_D^2(z)$, podem calcular la probabilitat d'una epidèmia gran com en el cas anterior resolent numèricament

$$1 - \zeta = G_D^2(e^{-Np_G\zeta}). \quad (5.12)$$

En la Figura 6, es pot veure una comparació de les probabilitats d'epidèmia ζ per residències amb habitacions dobles i individuals. S'ha de tenir en compte que en el cas d'habitació doble, quan el valor de $2n_1p_D$ és 0.7 o 0.9, hi ha una probabilitat positiva d'una epidèmia gran dins un dormitori individual amb $np_D=2n_1p_D$, per la qual les corbes són positives quan $Np_G = 0$. Quan el valor de $2n_1p_D$ és 0.3 o 0.5, p_G ha de superar el llindar positiu perquè hi hagi una epidèmia en una universitat amb dormitoris dobles, però aquests llindars són més elevats quan només hi ha dormitoris individuals.

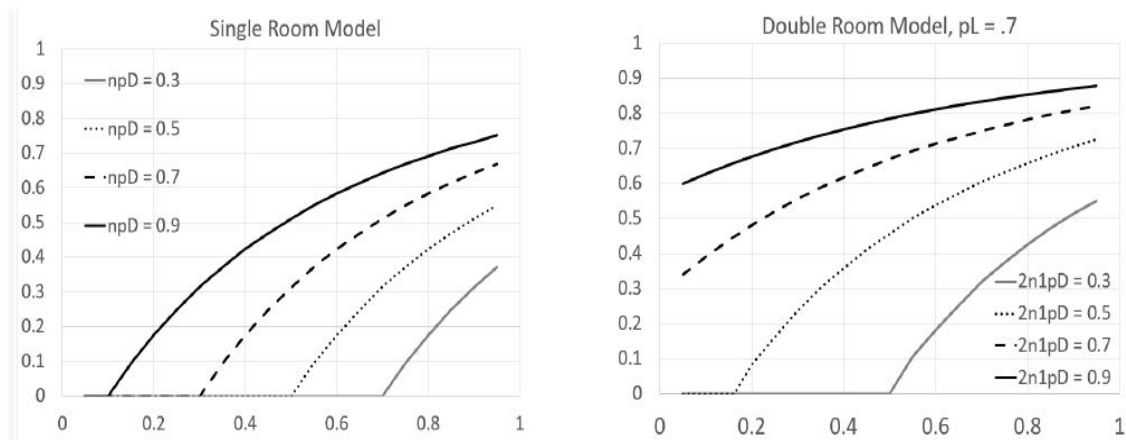


Figura 6: Mida de les epidèmies en escenaris d'habitacions individuals i dobles del model de residència quan $np_D = 2n_1p_D = 0.3, 0.5, 0.7, 0.9$.

5.3 Propagació de la infecció en classes universitàries

En aquesta secció abordarem dues qüestions. La primera qüestió és: Com afecta la distribució de la mida de les classes en la propagació del virus?, i la segona és: Quin hauria de ser el llindar per traslladar classes a no presencials?

Per abordar aquest problema, suposem que cada alumne assisteix a tres classes, i tenim en total n estudiants i m classes, cadascuna de mida c_i , $1 \leq i \leq m$. Es suposa que els estudiants es barregen de forma homogènia a les aules.

Els tres membres de les classes que li corresponen a un mateix alumne s'anomenen els seus *alter egos*.

Un individu infectat en la classe i en un moment t , alhora afegeix els seus dos *alter egos* a la població infectada en el mateix moment. Després, es produeixen infeccions en el moment $t + 1$ en les tres classes de l'individu.

5.3.1 Les mides de classe iguals redueixen la propagació

Per demostrar que la mida de les classes implica propagació, compararem dos escenaris per una univèrsitat amb 1000 alumnes, on cadascun d'ells assisteix a tres classes.

Escenari 1 Partim de 100 classes de mida 30. Un individu infectat a la classe i en el moment 0 produirà un nombre de Poisson($29p$) de nous individus infectats en la classe i en el moment 1. Els seus dos alter egos són igualment susceptibles de ser de les classes 2-100, i cada alter ego produirà infectats de Poisson($29p$) en qualsevol classe en què es trobi. Per tant, si m_{ij} és el nombre mitjà d'infectats de la classe j al moment 1 d'una sola persona infectada de la classe i , llavors

$$m_{i,i} = 29p \quad m_{i,j} = \frac{2}{99} \cdot 29p \quad \text{quan } j \neq i,$$

A partir d'això veiem que un infectat al moment 0 produeix una mitjana de $3 \cdot 29p$ infectats al moment 1, ja que cada individu infectat assisteix i infecta tres classes amb $29p$ de mitjana en cadascuna. Llavors,

$$R_0 = 3 \cdot 29p = 87p. \quad (5.13)$$

Escenari 2. Partim de 25 classes de mida 60 i 75 classes de mida 20. Definim $m_{i,j}$ com el nombre esperat d'infeccions a la classe j causats per un estudiant infectat en la classe i . Definim c_i i c_j , que representen les mides de la classe i i j respectivament, i n_i i n_j el nombre de classes de mida i i j respectivament. En aquest escenari, la probabilitat que un individu de la classe i tingui un alter ego en una classe j diferent és $\frac{2c_j}{(3000-c_i)}$, on c_i és la mida de la classe i . El 2 prové dels dos alter ego de l'estudiant i el 3000 prové del nombre total de classes per el nombre d'estudiants en cadascuna. Així, tenim

$$c_i = \begin{cases} 60 & \text{si } 1 \leq i \leq 25 \\ 20 & \text{si } 26 \leq i \leq 100 \end{cases}, \quad n_i = \begin{cases} 25 & \text{si } 1 \leq i \leq 25 \\ 75 & \text{si } 26 \leq i \leq 100 \end{cases},$$

$$m_{i,j} = \begin{cases} (c_i - 1)p & \text{si } i = j \\ 2 \cdot \frac{c_j}{3000-c_i} \cdot (c_j - 1)p & \text{si } i \neq j \end{cases},$$

Com que el vector propi x associat al valor propi màxim tindrà $x_i = a$ per $1 \leq i \leq 25$ i $x_i = b$ per $26 \leq i \leq 100$, el problema del valor propi per a la matriu 100×100 es pot reduir al problema del valor propi per a la matriu 2×2 amb entrades

$$\begin{bmatrix} (c_i - 1) + (n_i - 1) \cdot 2 \cdot \frac{c_i}{3000-c_i} \cdot (c_i - 1) & n_j \cdot 2 \cdot \frac{c_j}{3000-c_j} \cdot (c_j - 1) \\ n_i \cdot 2 \cdot \frac{c_i}{3000-c_i} \cdot (c_i - 1) & (c_j - 1) + (n_j - 1) \cdot 2 \cdot \frac{c_j}{3000-c_j} \cdot (c_j - 1) \end{bmatrix} p =$$

$$\begin{bmatrix} 59 + 24 \cdot 2 \cdot \frac{60}{2940} \cdot 59 & 75 \cdot 2 \cdot \frac{20}{2940} \cdot 19 \\ 25 \cdot 2 \cdot \frac{60}{2980} \cdot 59 & 19 + 74 \cdot 2 \cdot \frac{20}{2980} \cdot 19 \end{bmatrix} p =$$

$$\begin{bmatrix} 116.795 & 19.388 \\ 59.396 & 37.783 \end{bmatrix} p,$$

Resolent el problema dels valors propis de la matriu obtenim l'equació

$$(\lambda^2 - 154.579\lambda + 3261.33362)p = 0,$$

De la qual obtenim que la matriu té valors propis $129.380p$ i $25.288p$. Així,

$$R_0 = 129.38p. \quad (5.14)$$

Comparant aquest resultat amb la R_0 obtinguda en l'escenari 1 ($R_0 = 87p$), veiem que la R_0 del segon escenari és un 48% més gran que per l'escenari 1. Aquest fet suggereix que les classes més grans augmenten la transmissió del virus.

5.3.2 Model estocàstic per als escenaris 1 i 2

Per obtenir una visió més detallada dels dos escenaris, passem a un model estocàstic. Com en la secció anterior, el model té lloc en temps discret. Cada pas de temps és aproximadament una reunió de classe, de manera que dos passos de temps són aproximadament equivalents a una setmana en temps real. Una persona infectada d'una classe infecta als altres amb probabilitat $p = 0.01$ i té una probabilitat d'entrar en quarantena amb 0.5.

Utilitzant (5.13), veiem que en l'escenari 1, $R_0 = 2(0.87) = 1.74$.

Aplicant (5.14), veiem que en l'escenari 2, $R_0 = 2(1.2938) = 2.5876$.

Veiem clarament que la R_0 de l'escenari 1 és un 32.76% més baixa que en l'escenari 2.

Les següents gràfiques d'infectats per temps confirmen que l'escenari 2 és pitjor que l'escenari 1.

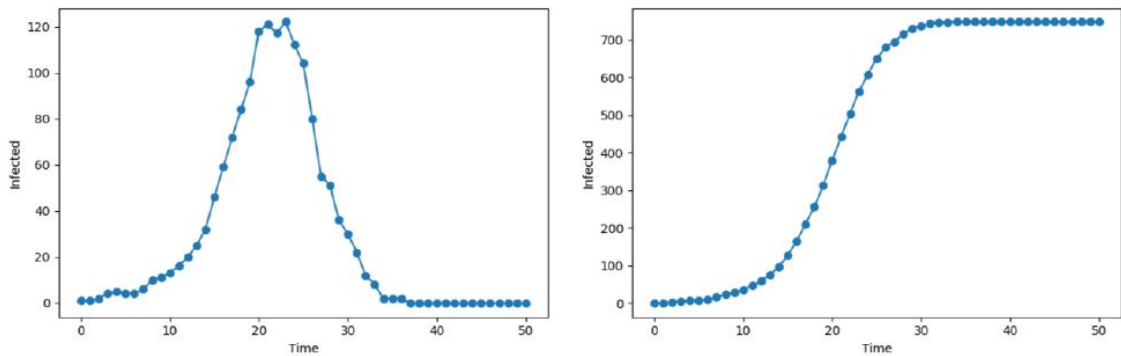


Figura 7: Una realització típica de l'epidèmia quan totes les classes tenen la mida 30. El panell esquerra mostra el nombre actual d'infeccions. A la dreta el nombre acumulatiu. En aquesta situació, l'epidèmia triga aproximadament 11 setmanes a assolir el seu punt màxim. A la setmana 15, gairebé totes les persones del campus han resultat infectades.

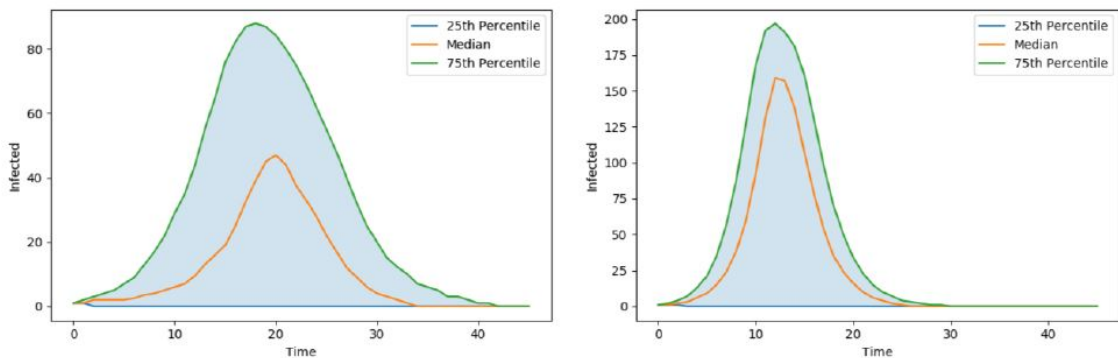


Figura 8: Mostra d'epidèmies generades per Python. El model va ser executat 1000 cops per cada escenari. El tauler esquerra dona el resultat de l'escenari 1, el de la dreta el de l'escenari 2. La línia superior (verda) és el percentil 75, la següent línia (taronja) és la mediana, la paral·lela a l'eix (blava) és el percentil 25.

En ambdós escenaris, podem veure en la Figura 9 que el brot generalment es descontrola en poques setmanes i infecta la majoria dels estudiants. A l'escenari 2, els brots tendeixen a desenvolupar-se més ràpidament, tenen un pic més alt i varien menys en mida i dinàmica. L'escenari 2 també produeix una major probabilitat que es produeixi un brot important d'un individu i mides epidèmiques més grans. Per tant, és evident que la presència de classes més grans, fins i tot si la mida mitjana de les classes continua sent la mateixa, augmenta el risc per als estudiants del campus.

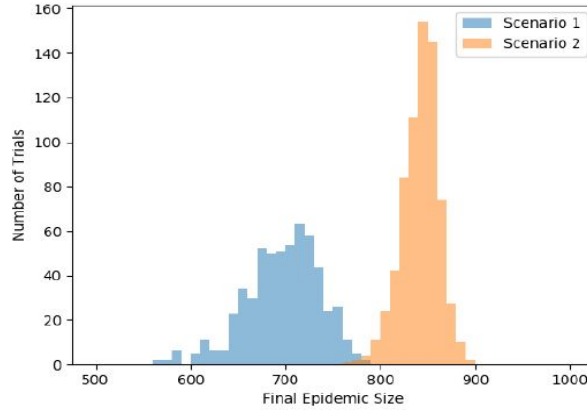


Figura 9: Distribucions de la mida de l'epidèmia final atès que es produeix un brot important. Les dades són les mateixes que es mostren a la Figura 8.

5.3.3 Fer classes online

Anem a considerar els efectes de moure les classes per sobre d'una mida determinada a tipus online. Per fer-ho, considerem una universitat amb 111 classes, cadascuna de mides que oscil·len entre 10 i 120 alumnes per classe. La mida total de les classes és de $10 + 11 + \dots + 120 = 7215$ i hi ha $\frac{7215}{3} = 2405$ estudiants. Per començar, primer ens plantegem quan totes les classes són presencials.

Després dels càlculs de l'escenari 2 de la secció 5.3.1, veiem que el nombre mitjà de persones infectades a l'aula j per un individu infectat a l'aula i és

$$m_{i,j} = \begin{cases} (c_i - 1)p & \text{si } i = j \\ 2 \cdot \frac{c_j}{7215 - c_i} \cdot (c_j - 1)p & \text{si } i \neq j, \end{cases}$$

De manera similar als escenaris anteriors d'aquesta secció, hem de trobar el valor propi més gran de la matriu anterior per determinar R_0 . En l'escenari 2 hem pogut calcular-ho manualment perquè només hi havia dos mides de classes, en canvi en aquest escenari la matriu és complicada calcular-la manualment per tenir totes les classes de diferents mides. Per aquesta raó s'ha utilitzat Python per trobar el valor propi analíticament.

$$R_0 = 251.5p.$$

Quan $p = 0.01$, $R_0 = 2.515$

Observem ara la situació en què la universitat mou totes les classes de mida $> k$ a tipus virtual. Els virus no es pot propagar a través de classes en línia:

$$m_{i,j} = 0, \text{ quan } c_j > k,$$

Els altres valors de $m_{i,j}$ es mantenen sense canvis. En canvi en R_0 depenent de k i de diverses p es pot veure en la Figura 10, que deixa clar que establir un límit pot reduir dràsticament la propagació del virus a través de les classes. Tot i que Gressman i Peck van suggerir un límit de 30 basat en les seves simulacions estocàstiques per a la difusió a les universitats, els nostres resultats en la Figura 10 suggereixen que el límit òptim depèn en gran mesura de p . Per un p suficientment petit, inclús aules de fins a 70 alumnes poden ser efectives.

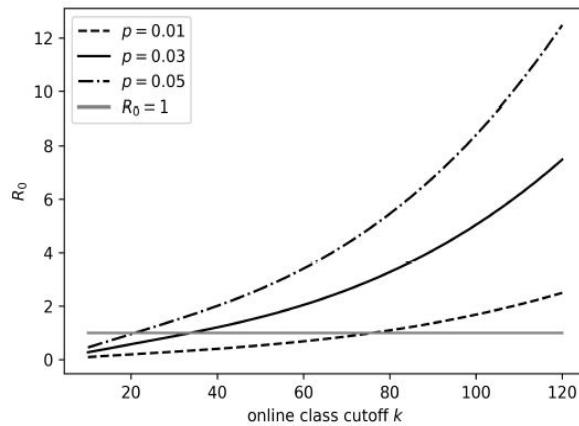


Figura 10: Número bàsic de reproducció de Covid-19 després de moure classes de mida superior a k en línia per a valors variables de p . Establir un límit suficientment petit per moure classes en línia pot fer que $R_0 < 1$ eviti un brot en un ampli rang de p . El que es considera suficientment petit depèn molt de la p .

6 Conclusions

Aquesta memòria m'ha permès introduir-me en els processos estocàstics, una branca de les matemàtiques desconeguda fins al moment per mi. La predisposició de la Dra. Marta Sanz-Solé a ajudar-me ha resultat clau a l'hora de realitzar la memòria, tant en aconseguir-me l'article estudiat en aquesta memòria com en la teoria necessària per entendre'l.

En l'article ens hem plantejat diverses preguntes. En la primera secció ens preguntem: Quant es reduiria la propagació si totes les habitacions dels estudiants fossin de tipus individual?

Hem arribat a la conclusió que, després d'un petit estudi, les habitacions individuals redueixen la propagació del virus en comparació amb les dobles. Podem veure un resultat gràfic d'aquest en la Figura 6.

La primera pregunta de la segona secció és: Com afecta la mida de classe en la propagació del virus?

Hem vist que en l'escenari on tenim 100 classes de mida 30 i cada alumne assisteix a tres classes, la $R_0 = 87p$.

En l'escenari on partim de 25 classes de mida 60 i 75 classes de mida 20 i cada alumne assisteix també a tres classes, la $R_0 = 129.38p$.

Comparant les R_0 dels dos escenaris, veiem que la R_0 del segon escenari és un 48% més gran que en l'escenari 1. Aquest fet suggereix que les classes més grans augmenten la transmissió del virus.

La segona pregunta de la segona secció és: Quin hauria de ser el llindar per traslladar les classes a no presencials?

La Figura 10 deixa clar que establir un límit pot reduir dràsticament la propagació del virus a través de les aules. El límit òptim depèn en gran mesura de la probabilitat de transmissió del moment. Per un probabilitat de transmissió suficientment petita, inclús aules de fins a 70 alumnes poden ser efectives.

Després de realitzar aquesta memòria i veure els resultats obtinguts en el cas d'estudi, podem afirmar que, per reduir la propagació del virus a les residències, les universitats haurien d'optar per habitacions individuals en lloc d'habitacions dobles.

Les universitats també haurien de reduir els nombre d'alumnes a les aules per contenir la propagació del virus. Una altra solució seria canviar a tipus online les classes més grans i totes les classes quan la $R_0 > 1$.

Tot i les precaucions que puguin prendre les universitats, la millor forma de reduir els contagis, i per tant la propagació del virus, és que tota la població segueixi les instruccions establertes: mascareta, distància de seguretat, reduir el contacte únicament amb la bombolla familiar i, sobretot, aïllament de les persones que hagin estat en contacte amb un infectat.

Referències

- [1] BOROWIAK, M., NING, F., PEI, J., ZHAO, S., TUNG, H. AND DURRET, R. (2020) Controlling the spread of COVID-19 on college campuses ARXIV:2008.07293v1
- [2] PINSKY, M.A. AND KARLIN, S. (2011) *An Introduction to STOCHASTIC MODELING*, ELSEVIER (4th edition)
- [3] KIMMEL, M., AND AXELROD, D.E. (2011) *Branching Processes in Biology*, Springer (Volume 19)
- [4] BALL, F., MOLLISON, D., AND SCALIA-TOMBA, G. (1997) Epidemics with two levels of mixing. *Annals Applied Probability*. 7, 46-89
- [5] GRESSMAN, P.T., AND PECK, J.R. (2020) Simulating COVID-19 in a University Environment ARXIV:2006.03175
- [6] MARÍN, J.M. (2014) <http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/PEst/tema4pe.pdf>
- [7] ORTEGA, J.
<https://www.cimat.mx/~jortega/MaterialDidactico/probabilidad17/Tema4.pdf>
<https://www.cimat.mx/~jortega/MaterialDidactico/probabilidad17/Tema5.pdf>
<https://www.cimat.mx/~jortega/MaterialDidactico/modestoI10/Poisson2.pdf>
- [8] *The Reproduction Number* <https://www1.health.gov.au/internet/publications/publishing.nsf/Content/mathematical-models~mathematical-models-models.htm~mathematical-models-2.2.htm>
- [9] MARTÍNEZ, A. (2014) *Procesos de ramificación*, Universidad de Zaragoza
- [10] BLANCO, L. (Nueva Serie, Volumen III (1996), pp.43-50) *Que es un proceso de ramificación?* <https://core.ac.uk/download/pdf/289980232.pdf>
- [11] FILELLA, I. (2016) *Processos de Ramificació*
<http://diposit.ub.edu/dspace/bitstream/2445/97880/2/memoria.pdf>