



UNIVERSITÀ DEGLI STUDI DI PADOVA
Dipartimento di Fisica e Astronomia “Galileo Galilei”
Corso di Laurea in Fisica

Tesi di Laurea

Studio Stocastico di Modelli Epidemici

Relatore

Prof./Dr. Samir Simon Suweis

Correlatore

Prof./Dr. Enzo Orlandini

Laureando

Andrea Zanola

Anno Accademico 2019/2020

*"Chi tace e chi piega la testa muore ogni volta che lo fa,
chi parla e chi cammina a testa alta muore una volta sola."
Giovanni Falcone*

Indice

1	Motivazioni	3
2	Introduzione	3
3	Modelli Epidemiologici via EDO	4
3.1	Introduzione	4
3.1.1	Classi Costituenti	4
3.1.2	Demografia	5
3.1.3	Immunità e Mortalità	5
3.2	Descrizione Matematica	6
3.3	Simulazione via EDO	7
4	Modelli Epidemiologici Stocastici	10
4.1	Introduzione	10
4.2	Rivisitazione Probabilistica	10
4.3	Algoritmo di Gillespie	13
4.4	Simulazioni con SSA	15
5	Modello con Immunità e Mortalità Indotta	21
5.1	Descrizione Matematica	21
5.1.1	Immunità	22
5.1.2	Mortalità	22
5.2	Simulazione via SSA	23
6	Conclusioni	25
7	Sviluppi Futuri	25
8	Bibliografia	25

1 Motivazioni

Per mesi le vite e le libertà di tutti sono state appese ad una curva, curva di cui si ha iniziato ad apprezzarne la natura, la forma, una curva che ha fatto molto parlare di sè.

Per settimane ci si continuò a chiedere alla fine come sarebbe andata a finire, si guardò con angoscia quella curva, di cui poco la popolazione ancora sapeva e che per molti giorni non fece altro che salire.

Ebbene, è giunto ora il momento di studiarla a dovere.

2 Introduzione

L'intento dello studio dei modelli epidemiologici è quello di, sulla base di considerazioni statistiche e fenomenologiche, riuscire a descrivere quella che è l'evoluzione di un patogeno attraverso una società di possibili *suscettibili*, sia riferendosi al numero di persone *infette*, sia al numero di persone *guarite* o *decedute*; questi numeri daranno dunque una qualche informazione su quello che è e che sarà l'impatto sulla società, potendo dunque discutere eventuali metodi per contrastare la diffusione.

Conoscere l'evoluzione temporale di una epidemia, e quindi la possibile estensione e durata del periodo di esposizione della società al virus, risulta infatti molto importante, a fini preventivi, per predire e ridurre il più possibile il picco dei contagi.

Si evidenzia fin da subito che lo studio dei modelli epidemiologici non ha quindi la presunzione di descrivere microscopicamente la struttura del patogeno, nè dare una spiegazione della aggressività di questo, nè tanto meno di dare la soluzione medica definitiva.

Una qualsiasi formulazione teorica di questi modelli avrà come fine quindi, sulla base di pochi parametri caratterizzanti, quello di descrivere l'epidemiologia di un patogeno che poi, con i dati medici si andrà a validare o a falsificare; se quindi si fosse in presenza di un'infezione, la cui evoluzione sia ben descrivibile con il modello utilizzato (ovverosia i dati ospedalieri confermano le simulazioni numeriche di questo), allora si potrebbe inferire su di esso attribuendo valori a grandezze di particolare interesse che verranno descritte in seguito.

In particolare in questo lavoro ci si concentrerà sulla simulazione di modelli epidemiologici, inizialmente di tipo deterministico lavorando nel continuo (cioè considerando concentrazioni e non un numero discreto di individui), per poi sviluppare un modello stocastico probabilistico discreto, andando a studiare le eventuali differenze o somiglianze fra i due approcci.

3 Modelli Epidemiologici via EDO

3.1 Introduzione

3.1.1 Classi Costituenti

Il modello epidemiologico di base da cui si può partire per svilupparne poi sempre di più accurati e precisi, è il modello SIR.

SIR sta per *Susceptible - Infected - Recovered* tripartendo gli individui di una società quindi in suscettibili, infetti e guariti.

I suscettibili sono tutti coloro che sono potenzialmente sensibili al patogeno ossia che, se ne venissero a contatto, potrebbero contrarre la malattia e in tal caso un suscettibile diventerebbe un infetto; gli infetti sono quindi tutti coloro che sono stati a contatto con il patogeno e che al momento di un ideale conteggio risultano positivi cioè ospitano l'infezione. Se essi presentano sintomi allora si diranno *sintomatici*, altrimenti si diranno *asintomatici*, ma non vi è quindi qui alcuna distinzione; infine i guariti sono tutti coloro che hanno con successo superato l'infezione e che non presentano più la malattia, non potendo più essere contagiosi.

Il vero punto chiave nella modellizzazione dell'epidemia sarà quello di comprendere come un individuo si sposti da una classe all'altra e come più individui di classi diverse possano interagire fra loro, dunque comprendere la risposta di tutta la popolazione.

La transizione da S a I, dunque l'atto di trasmissione, è governata da tre fattori: la quantità di infetti, la probabilità di trasmissione a seguito di un contatto e la probabilità di interazione fra S e I, cioè la struttura di contatti soggiacente per una data società. Con struttura di contatti si intende quanto un insieme di individui interagisca fra loro, questo dovuto sia alla rete di relazioni e contatti potenziali, ma anche da quanto l'ambiente in cui essi vivono glielo permettono. Il parametro che descrive questa transizione è β il quale risulta essere il prodotto di frequenza di contatto e probabilità di trasmissione (Es: $\beta = 1.5/\text{giorno}$).

Un altro importante parametro che si definisce è la *forza di infezione* definita come:

$$\lambda = \beta Y/N \quad (1)$$

$$\lambda = \beta Y \quad (2)$$

dove Y è il numero di infetti, mentre N è la popolazione in esame.

La prima formulazione viene detta *frequenza dipendente*, la seconda *densità dipendente*.

La frequenza dipendente riflette quella situazione in cui il numero di contatti che un individuo può avere è indipendente dal numero totale di individui nella popolazione.

La densità indipendente assume invece che, aumentando la popolazione, aumenti il rate di contatto (generalmente per i patogeni si predilige la frequenza dipendente). Questo rispecchia il fatto che, benché una città possa essere più popolosa di un'altra, i contatti sociali che una persona ha risultino mediamente gli stessi. In parole semplici si può pensare che, se si abitasse a Londra o a New York la cerchia di persone con la quale si interagisce giornalmente in modo consistente sia sempre la stessa.

La transizione da I a R invece, da un punto di vista modellistico dipenderà da quanto tempo questo è stato nella classe I: se un soggetto è stato a lungo nella classe I, allora ha avuto

una piccola probabilità di spostarsi alla classe R. Questo parametro, che viene detto *tasso di recupero* γ , è interpretabile come l'inverso del tempo necessario di guarigione (Es: $1/\gamma = 3$ giorni).

3.1.2 Demografia

Il modo più semplice di costruire il modello è quello di assumere una popolazione chiusa, ossia senza morti, nascite o migrazioni. Questa approssimazione si configura bene con un' infezione abbastanza veloce da rendere insensibili le variazioni di nascite e decessi.

Le morti, sostanzialmente, sottraggono individui da tutte le classi con un rate di decessi *naturali* definito dal parametro μ e, proprio perché naturali, μ non ha nulla a che vedere con l'aggressività del virus e rappresenta l'inverso di una aspettativa di vita (Es: $\frac{1}{\mu} = 80$ anni).

Questo è favorevole, per quanto spiacevole, all'arresto della pandemia, visto che da una parte la morte di un infetto arresta una catena di trasmissione e dall'altra, la morte di un suscettibile toglie potenziali futuri infetti rallentando la diffusione del patogeno.

D'altra parte si può introdurre un rate di nascite ν (Es: $\nu = 50/\text{mese}$) che andrà a fornire nuovi soggetti per la classe dei suscettibili. In genere una possibile immunità ereditata dalla madre viene persa entro i 6 mesi, che è un periodo breve rispetto alla tipica età in cui un bambino può essere esposto a contrarre un'infezione che è 4-5 anni per i paesi sviluppati (per approfondimenti vedi [AndMay]); questo fa intuire perché sia ragionevole assumere ogni nuovo nato come suscettibile.

È molto importante sottolineare che ν è sostanzialmente diverso da μ , poiché mentre il primo è un rate assoluto di nuovi nati associato al numero di individui di una società (N), μ è un rate (inteso come inverso di un tempo) che va moltiplicato per l'attuale numero di componenti di una data classe.

Dunque, il vincolo per avere una popolazione chiusa deve essere riscritto come:

$$S + I + R = N \rightarrow N(t) \quad (3)$$

Questo è sempre applicabile, perché un individuo deve obbligatoriamente appartenere ad una e una sola classe e la somma delle classi corrisponde all'intera popolazione N che qui si considera costante. In altri termini, si ipotizza che il numero di morti sia mediamente pari a quello dei nuovi nati e in tal caso si parlerà di popolazione chiusa, i.e.,

$$\frac{dN}{dt} = \nu - \mu N = 0 \rightarrow N = \frac{\nu}{\mu} \quad (4)$$

3.1.3 Immunità e Mortalità

Nel modello SIR attualmente descritto non è presente alcun termine di immunità o di mortalità, ossia gli infetti non muoiono a causa della malattia e una volta che sono guariti preservano una immunità per un tempo indefinito.

Analogamente a quanto fatto precedentemente per considerare la mortalità indotta, occorre introdurre un nuovo parametro ρ che rappresenti la probabilità che un individuo muoia, prima che guarisca o prima che muoia per morte naturale (Es: $\rho = 0.2$). Se $\rho \rightarrow 1$ allora la morte del paziente diventa via via più certa e questo comporta conseguentemente un calo di infetti come verrà mostrato in seguito (Es: Ebola presenta una mortalità che può arrivare al 90%).

Per quanto riguarda l'immunità si possono studiare due diversi scenari limiti: che l'infezione

immunizzi a vita o che l'immunità non si presenti mai; dunque si può introdurre un parametro w che rappresenti il rate con la quale un individuo perde l'immunità e dalla classe R dei guariti si spostati alla classe S dei suscettibili e quindi si presenti una sorta di retroazione $S \rightarrow I \rightarrow R \rightarrow S$. Se la perdita di immunità avviene immediatamente dopo la guarigione allora occorrerà utilizzare un modello tipo SIS e si avrà $w \rightarrow \infty$; viceversa si parlerà di modello SIR e si avrà $w \rightarrow 0$. In effetti dal momento che w è un rate, se $w \rightarrow 0$ allora il tempo affinché l'individuo perda l'immunità è infinitamente grande, ossia sostanzialmente non la perde mai.

In verità non si può definire una immunizzazione a vita dal momento che gli anticorpi sono specifici ad una forma esatta di patogeno, il cui genoma negli anni generalmente può evolvere e quindi infettare nuovamente l'individuo non preparato alla nuova forma mutata; d'altra parte non è semplice definire cosa voglia dire perdere l'immunità dal momento che in genere quello che si verifica è una diminuzione della concentrazione di anticorpi ma mai una loro totale assenza.

3.2 Descrizione Matematica

Definite le classi costitutive del modello e definiti i parametri fondamentali, si può descrivere il modello epidemiologico SIR intuitivamente come diagramma:

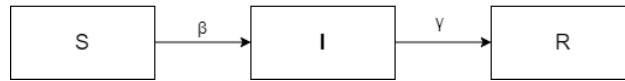


Figura 1: *Diagramma a Blocchi SIR base*

oppure più formalmente via equazioni:

$$\frac{dS}{dt} = -\beta SI \quad (5)$$

$$\frac{dI}{dt} = \beta SI - \gamma I \quad (6)$$

$$\frac{dR}{dt} = \gamma I \quad (7)$$

con condizioni iniziali:

$$S(0) > 0, I(0) > 0, R(0) = 0 \quad (8)$$

con la notazione:

$$S = \frac{X}{N}, I = \frac{Y}{N}, R = \frac{Z}{N}. \quad (9)$$

dove (X, Y, Z) sono il numero assoluti di individui invece (S, I, R) sono le proporzioni (o concentrazioni) di individui in ogni classe.

Si osserva dunque che la struttura del modello è un sistema di equazioni differenziali accoppiate autonome, in cui dunque il tempo non gioca un ruolo esplicito. Nonostante l'apparente semplicità, il sistema descritto non è risolubile in maniera esatta ed ecco perché il ruolo delle simulazioni diventa fondamentale.

Da questa formulazione si può notare che nelle equazioni (5)-(6) non appare un termine in R : questo è compatibile con il fatto che i guariti non hanno un ruolo attivo nella dinamica endemica dal momento che i guariti non possono transitare più in nessun'altra classe.

Si precisa fin da subito che l'indicazione $S \rightarrow I$ in Figura (1), in realtà è imprecisa poiché

come si vede dalle equazioni (5)-(6), l'evoluzione da S a I è descritta dal termine βSI , perché ovviamente ci deve essere una mutua interazione fra gli individui delle due classi.

Nonostante non si possa risolvere in generale e analiticamente in forma chiusa la soluzione per la dinamica di $I(t)$ e $S(t)$, per avere una comprensione del problema si può applicare lo studio dei sistemi dinamici per trovare gli equilibri del sistema, i.e:

$$\frac{dS}{dt} = 0, \frac{dI}{dt} = 0, \frac{dR}{dt} = 0 \quad (10)$$

In questo caso vi è un unico equilibrio banale che risulta essere $(S^*, I^*, R^*) = (1, 0, 0)$, ovvero in assenza di infetti, dunque uno stato in cui l'epidemia persiste indefinitamente non esiste.

La derivazione delle equazioni (5)-(6)-(7) a partire dalle transizioni microscopiche non verrà presentato qui, ma si può derivare esplicitamente da principi primi (vedi [KeeRoh]). Questa descrizione tuttavia giustifica, o almeno sottolinea, l'utilità di aver definito i parametri β, γ, w come rate, che ben si configura con la struttura di EDO.

Con la definizione di β e γ , che sono i due parametri fondamentali per la descrizione del modello SIR base, si può introdurre un nuovo parametro (ormai famoso) definito come:

$$R_0 = \frac{\beta}{\gamma} \quad (11)$$

universalmente così rappresentato, che viene detto *rapporto di riproduzione di base* definito come (vedi [Allen]):

"il numero medio di casi secondari emergenti da un caso primario in una popolazione di interi suscettibili."

Esso sostanzialmente rappresenta il massimo potenziale di riproduzione per una malattia infettiva.

Nella equazione (11) si noti come β , che rappresenta il rate di trasmissione cioè quanto facile è contrarre l'infezione, viene diviso per γ che è l'inverso del tempo medio di appartenenza di un infetto alla classe degli infetti; la ragione è che se γ è grande (dunque c'è un elevato rate di guarigioni), allora si stanno togliendo velocemente membri dalla classe degli infetti. In questo caso, dunque, R_0 è piccolo e quindi l'epidemia non esplode anzi, è ben noto in letteratura il così detto *Threshold Phenomenon*, che accade quando $R_0 > 1$: affinché l'epidemia avvenga, il tasso di riproduzione deve essere maggiore di uno, ossia un infetto deve almeno mediamente infettare un altro individuo perché l'infezione avanzi.

3.3 Simulazione via EDO

Dal momento che l'intento è simulare le curve epidemiche, cioè risolvere il problema differenziale associato, serve una strategia numerica per affrontare le EDO; in questo lavoro si è usato il metodo di Eulero esplicito con incremento temporale $dt = 1h$.

Dalla simulazione numerica ($N = 1000, R_0 = 10, N_{infetti} = 1$) si ottiene:

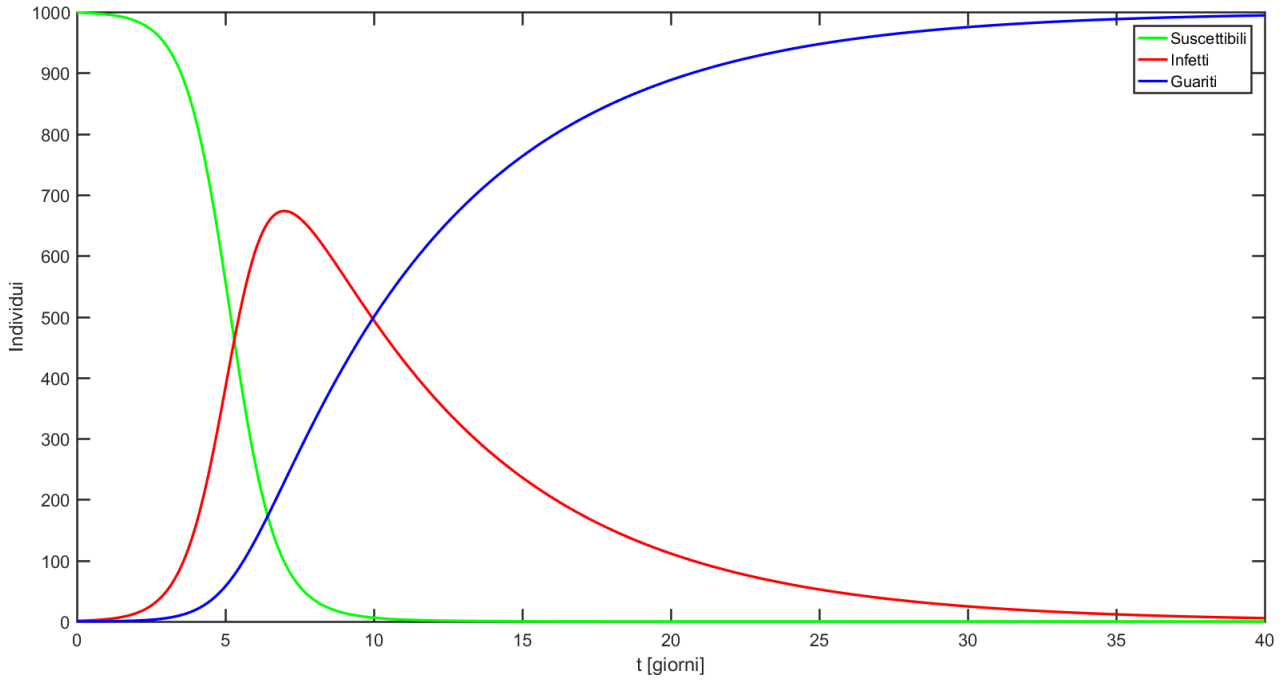


Figura 2: Simulazione numerica del modello SIR base descritto via EDO.

in cui in particolare si vede che le curve epidemiche $S(t)$ e $R(t)$ sono strettamente monotone e che dunque non vi può essere un equilibrio endemico, come trovato precedentemente.

In secondo luogo si potrebbe cercare di visualizzare il threshold phenomenon; in effetti la Figura (3) permette sia di verificare che l'epidemia si avvia quando $R_0 > 1$ (iperbole nera), ma sia anche che il numero massimo di infetti è invariante per coppie (β, γ) e dunque sarà descritto da curve di livello anch'esse iperboliche.

Per una derivazione formale di questo, basta osservare che:

$$\frac{dI}{dt} = 0 \rightarrow S^* = \frac{\gamma}{\beta} = \frac{1}{R_0} \quad (12)$$

$$\frac{dS}{dt} \frac{dt}{dR} = -\frac{\beta S}{\gamma} \rightarrow S(t) = S(0)e^{-R(t)R_0} \rightarrow R^* = \frac{1}{R_0} \ln(R_0 S(0)) \quad (13)$$

$$I^* = 1 - \frac{1}{R_0} \left[1 + \ln(R_0 S(0)) \right]. \quad (14)$$

Dunque noto R_0 e le condizioni iniziali (vedi (8)), il numero massimo di infetti è predeterminato. Chiaramente da (12) si comprende perché $R_0 > 1$ abbia senso come vincolo matematico affinché l'epidemia avvenga, dal momento che il contrario porterebbe al risultato assurdo $S^* > 1$, che è chiaramente non accettabile.

Questo però, implica solo che il numero massimo di infetti è invariante per scelta di (β, γ) , ma non implica che le curve epidemiche siano le stesse. In effetti si può ricavare una forma alternativa di (5)-(6)-(7) come:

$$\frac{dS}{\gamma dt} = -R_0 SI \quad \frac{dI}{\gamma dt} = R_0 SI - I \quad \frac{dR}{\gamma dt} = I \quad (15)$$

e definendo:

$$dt' = \gamma dt \quad (16)$$

permettere di concludere che la variazione di γ agisce sulle curve $S(t), I(t), R(t)$ come una dilatazione o una contrazione temporale.

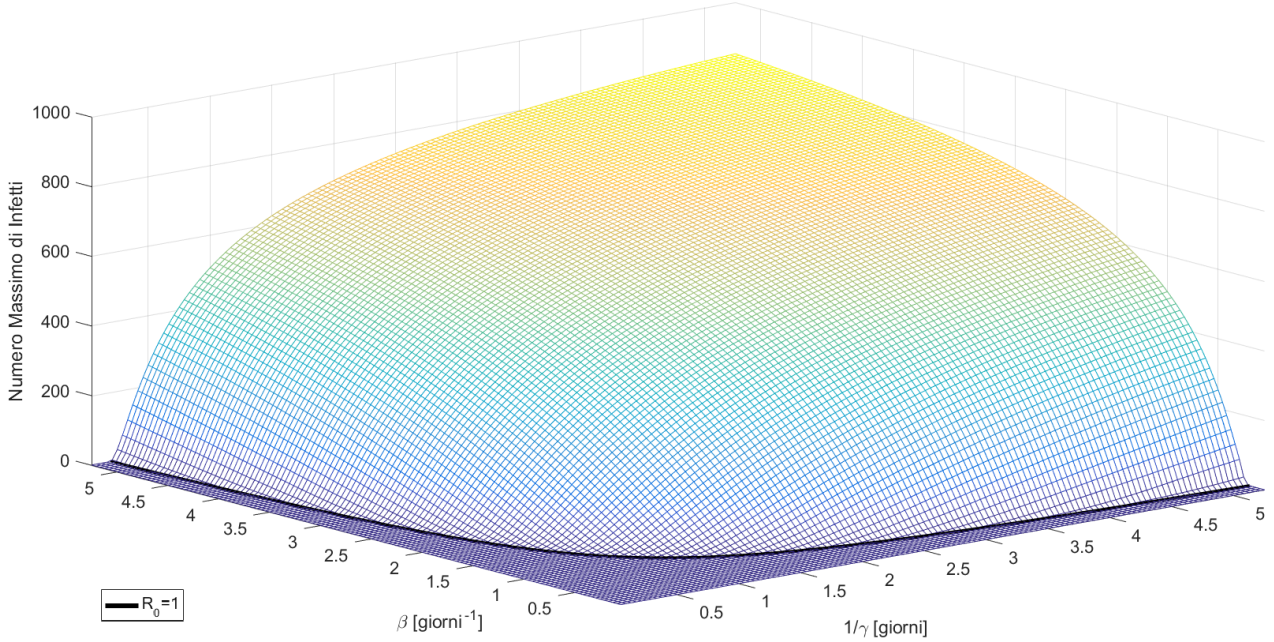


Figura 3: *Studio della variazione del massimo numero di infetti variando β e γ , $N=1000$. In linea nera il valore di soglia $R_0=1$.*

Si sottolinea infine che la catena di trasmissione può essere fermata non per una mancanza totale di suscettibili, cosa sperabilmente evitabile, ma da una possibile diminuzione degli infetti, che si riflette nella richiesta che $R_0 < 1$. Questo è ottenibile sia se molti infetti guariscono (si aumenta γ) e si sviluppa quella che viene detta immunità di gregge, sia se molti infetti vengono isolati (si abbassa β) imponendo ad esempio la quarantena e l'isolamento.

4 Modelli Epidemiologici Stocastici

4.1 Introduzione

Le simulazioni viste nel paragrafo precedente sono state ottenute risolvendo numericamente il sistema di equazioni differenziali accoppiate che definisce il modello epidemiologico. Queste equazioni sono però deterministiche e non considerano gli effetti che intrinseche fluttuazioni statistiche, dovute a fattori ambientali o demografici, possono avere sull'evoluzione dell'epidemia. Per una descrizione più accurata della situazione reale, soprattutto nel caso di onset dell'epidemia dove solo pochi individui sono coinvolti, è pertanto opportuno considerare anche una versione stocastica dei modelli introdotti in precedenza.

Questo è naturale se si riflette sul significato di tutta le serie di parametri introdotti nella sezione precedente: per esempio per affermare che vi sia un rate μ di morti vuol dire che *mediamente* per cause non legate alla malattia una persona muore dopo $\frac{1}{\mu}$ anni; questo è un comportamento medio, statistico, ma non vuol dire naturalmente che tutti gli individui muoiano dopo esattamente $\frac{1}{\mu}$ anni.

L'altro aspetto importante a cui non si vuole rinunciare è la discretizzazione del numero dei soggetti, ovvero si vuole tenere conto del fatto che la morte di un individuo è un evento discreto: o avviene o non avviene. Allo stesso modo l'atto di trasmissione dell'infezione o avviene con successo oppure fallisce, i.e. è un evento discreto.

4.2 Rivisitazione Probabilistica

Si supponga di avere un sistema con N specie ($X_1 \dots X_N$), le quali possono interagire in M modi o reazioni ($R_1 \dots R_M$) e si assuma che le N specie siano ben mescolate nello spazio Ω .

Supposto questo, lo scopo sarà descrivere la traiettoria delle concentrazioni $\bar{X}(t)$ date le opportune condizioni iniziali $\bar{X}(0)$, dove con traiettoria si intende una possibile evoluzione che l'epidemia potrebbe sviluppare.

Sostanzialmente nell'atto di trasmissione può accadere che un infetto interagisca con un suscettibile ma che non sempre riesca ad infettarlo, dunque da un punto di vista computazionale c'è stato un contatto ma non vi è stata alcuna reazione, dunque il vettore degli stati $\bar{X}(t)$ risulta invariato. Questo è computazionalmente costoso visto che per la maggior parte delle volte l'interazione fra due soggetti non dà alcuna variazione, basti pensare che le interazioni I-I, S-S, R-R, S-R, I-R, sono tutte $\bar{X}(t)$ invarianti, dunque si vorrebbe trovare un metodo per ovviare a ciò (stiamo assumendo per definizione che $w = 0$).

Una reazione o *canale di reazione* R_j è caratterizzato matematicamente da due quantità: il vettore di cambiamento di stato o *state change vector* e la funzione di propensione o *propensity function*.

Lo state change vector di una reazione $\bar{v}_j = (v_{1j} \dots v_{Nj})$, è composto da termini v_{ij} che rappresentano la variazione della classe X_i nella reazione R_j .

Per comprendere cosa rappresenti $\bar{v}_j$, lo si applica direttamente al modello SIS il quale è caratterizzato da due reazioni:



La prima reazione descrive l'infezione di un suscettibile a seguito del contatto con un infetto, dunque il numero di infetti aumenta (+1) e quello dei suscettibili diminuisce (-1); la seconda invece descrive la guarigione di un infetto e dunque gli infetti perdono un individuo (-1), mentre i suscettibili ne guadagnano uno (+1), quindi i corrispettivi state change vector sono:

$$\bar{v}_1 = (-1, +1) \quad (19)$$

$$\bar{v}_2 = (+1, -1) \quad (20)$$

Si può notare che le due reazioni sono simmetriche e questo non deve sorprendere, come d'altronde non deve sorprendere la simmetria fra le equazioni differenziali del modello SIS:

$$\frac{dS}{dt} = -\beta SI + \gamma I \quad (21)$$

$$\frac{dI}{dt} = +\beta SI - \gamma I. \quad (22)$$

Entrambe le rappresentazioni devono infatti soggiacere al vincolo $S + I = N$ (in caso di popolazione chiusa come si sta ora assumendo), dunque la somma delle variazioni di tutte le classi deve essere zero, visto che esplicitamente:

$$N = S + I = \text{const} \rightarrow \frac{dN}{dt} = \frac{dS}{dt} + \frac{dI}{dt} = 0. \quad (23)$$

La propensity function a_j invece è definita come:

$$a_j dt = \text{"la probabilità che, dato } \bar{X}(t), \text{ una reazione } R_j \text{ accadrà nell'intervallo } [t, t + dt)\text{"}$$

Dunque data questa definizione, il modello stocastico dovrà seguire inevitabilmente le leggi di conservazione della probabilità.

Se si analizza ora una reazione monospecie, la probabilità che la classe X effettui la reazione R_j , sarà legata alla popolosità di quella classe, al tempo a disposizione che si dà ai soggetti per interagire e ad una eventuale costante c_j e infatti:

$$\Delta t a_j = c_j X \Delta t. \quad (24)$$

Se invece si ha una reazione bispecie si avrà che:

$$\Delta t a_j \propto c_j X_1 X_2 \Delta t. \quad (25)$$

e nella generalizzazione a N specie quello che si trova è che occorre normalizzare rispetto al volume Ω , potendo scrivere:

$$\Delta t a_j = c_j \frac{X_1 X_2 \dots}{\Omega^{\# \text{specie} - 1}} \Delta t \quad (26)$$

Se si osserva la struttura di quest'ultima formula si vede che è identica ai singoli termini che entrano nei modelli via equazioni differenziali.

Una rappresentazione intuitiva di quanto detto è la *rappresentazione a flussi*, ossia si può immaginare lo stato attuale di una classe $X_j(t)$ come il bilancio di un flusso uscente Φ_{out} , che

rappresenterà tutte le possibili reazioni che possono avvenire (si ricordi che la popolazione è well-mixed dunque la probabilità che una reazione avvenga è solamente descritta dalla probabilità intrinseca della reazione) e che porterà ad una riduzione dei soggetti della classe e di un flusso entrante Φ_{in} , che rappresenterà tutte le possibili reazioni che porteranno ad un incremento dei soggetti della classe in esame, tutto questo a istante fissato.

In particolare i due flussi saranno descritti come:

$$\Phi_{in} = \sum_{j=1}^M a_j(\bar{X}(t) - \bar{v}_j) P(\bar{X} - \bar{v}_j, t | \bar{X}(t_0), t_0) \quad (27)$$

$$\Phi_{out} = \sum_{j=1}^M a_j(\bar{X}(t)) P(\bar{X}, t | \bar{X}(t_0), t_0) \quad (28)$$

dove

- $P(\bar{X} - \bar{v}_j, t | \bar{X}(t_0), t_0)$ rappresenta la probabilità condizionata che un punto precedente $\bar{X}(t) - \bar{v}_j$, date condizioni iniziali $\bar{X}(t_0)$ si sia mosso e, tramite una reazione R_j , che avrà uno state change vector $+\bar{v}_j$, sia giunto in $\bar{X}(t)$.

- $P(\bar{X}, t | \bar{X}(t_0), t_0)$ invece, rappresenta la probabilità condizionata che l'attuale stato $\bar{X}(t)$, date condizioni iniziali $\bar{X}(t_0)$, si muova in un futuro stato $\bar{X}(t) + \bar{v}_j$ a seguito di una reazione R_j con state change vector $+\bar{v}_j$.

La variazione dunque dei due flussi, entrante e uscente, dovrà dare la variazione di probabilità di stare nello *stato attuale* $\bar{X}(t)$. Notando che per convenzione il flusso entrante è positivo, quanto detto porta alla definizione della *Master Equation (ME)*:

$$\frac{\partial P(\bar{X}, t | \bar{X}(t_0), t_0)}{\partial t} = \sum_{j=1}^M a_j(\bar{X} - \bar{v}_j, t) P(\bar{X} - \bar{v}_j, t | \bar{X}(t_0), t_0) - a_j(\bar{X}) P(\bar{X}, t | \bar{X}(t_0), t_0), \quad (29)$$

il cui lo stato stazionario è naturalmente dato da:

$$\frac{\partial P(\bar{X}, t | \bar{X}(t_0), t_0)}{\partial t} = 0. \quad (30)$$

Benché questo approccio sia intrinsecamente diverso da quello con le equazioni differenziali, non si deve pensare che il risultato poi sia completamente diverso da quello che si aveva già precedentemente trovato.

Si supponga ad esempio (si veda [Gill]) di avere una reazione $R : X \rightarrow \emptyset$, ossia i membri della classe X si estinguono e che si possa quindi scrivere $a(X) = cX$, con state change vector $v = -1$.

Da un punto di vista di EDO, questo è un processo simile al decadimento:

$$X(t) = X(0)e^{-ct} \quad (31)$$

Dal punto di vista della ME, il processo si scrive come:

$$\frac{\partial P(X, t | X_0, 0)}{\partial t} = -a(X)P(X, t | X_0, 0) \quad (32)$$

dal momento che non vi è flusso entrante previsto per X .

Così posta allora, si può dimostrare, che la ME è risolubile esattamente e porta ad una PDF binomiale:

$$P(X, t|X_0, 0) = \frac{X_0!}{X!(X_0 - X)!} e^{-cXt} (1 - e^{-ct})^{X_0 - X} \quad (33)$$

Se ora si calcola il valore di aspettazione di questa PDF (che sarà la traiettoria mediamente più probabile), si trova:

$$\langle X(t) \rangle = X(0)e^{-ct} \quad (34)$$

Questo esempio, benché sia un caso specifico particolarmente semplice, fornisce l'idea che sta dietro alla ME, ossia descrivere probabilisticamente le traiettorie del nostro sistema le quali, mediamente, vengono descritte dalle traiettorie trovate con le EDO.

Da ciò non si deve concludere che la ME sia un'inutile complicazione delle EDO, poiché $P(X, t|X_0, 0)$ descrive la probabilità che una qualsiasi traiettoria, fra quelle possibili, si manifesti: è dunque un livello decisamente superiore di informazione.

Dal momento che in generale la ME risulta di difficile risoluzione sia da un punto di vista analitico, che da un punto di vista numerico, occorre sviluppare un algoritmo che possa almeno simularla. La strategia dunque non è risolvere la ME, oltretutto ricavare $P(\bar{X}, t|\bar{X}(t_0), t_0)$, ma simulare tramite Metodi Monte Carlo (MCM) possibili traiettorie stocastiche $\bar{X}(t)$ le cui proprietà statistiche (ad esempio valore di aspettazione, varianza) obbediscano alla PDF che risolve la ME.

4.3 Algoritmo di Gillespie

Per poter sviluppare una simulazione stocastica occorre prima formulare un algoritmo (detto *Stochastic Simulation Algorithm* SSA) che permetta di fare quanto desiderato.

Il punto chiave degli SSA non è simulare una $\bar{X}(t)$ partendo dalla $P(\bar{X}, t|\bar{X}(t_0), t_0)$, ma è sfruttare una nuova probabilità definita come (vedi [Gill]):

$$P(\tau, j|\bar{X}(t), t)d\tau = \text{"probabilità che, dato } \bar{X}(t), \text{ la prossima reazione che accadrà nell'intervallo di tempo } [t + \tau, t + \tau + d\tau) \text{ sarà } R_j\text{."}$$

Formalmente la nuova scrittura rappresenta la probabilità congiunta di avere determinati valori delle due variabili casuali (τ, j) , dato $\bar{X}(t)$.

Il grande vantaggio di questa definizione sta nel pensare al tempo sezionato in intervalli di lunghezza variabile τ dopo le quali, in un piccolo intervallo $d\tau$, accadrà *certamente* la reazione R_j ; questo infatti risolve proprio il problema che era stato esposto nel paragrafo precedente relativo all'efficienza della simulazione.

τ e j sono variabili casuali con PDF associate derivabili dalla definizione di $P(\tau, j|\bar{X}(t), t)d\tau$; per esplicitare quest'ultima si pensi all'intervallo τ sezionato in k intervalli infinitesimi $d\tau$ nei quali non avviene alcuna reazione, allora la probabilità che *nessuna* reazione avvenga in questo intervallo τ sarà dunque

$$P(\tau|\bar{X}(t), t) = \left[1 - d\tau \sum_{j=1}^M a_j(\bar{X}(t)) \right]^k \quad (35)$$

dove $k = \frac{\tau}{d\tau}$.

Se si definisce per semplicità di notazione:

$$a_0(\bar{X}(t)) = \sum_{j=1}^M a_j(\bar{X}(t)), \quad (36)$$

allora la probabilità che una qualsiasi reazione fra le M disponibili non avvenga sarà la produttoria di non avvenuta reazione, tante volte quante il numero di intervalli k . Prendendo il limite di grandi k si ottiene quindi

$$P(\tau|\bar{X}(t), t) = \lim_{k \rightarrow \infty} \left[1 - a_0(\bar{X}(t))d\tau \right]^k = e^{-a_0(\bar{X}(t))\tau}. \quad (37)$$

Poiché dopo τ una reazione j dovrà avvenire sicuramente, si avrà:

$$P(\tau, j|\bar{X}(t), t) = a_j(\bar{X}(t))e^{-a_0(\bar{X}(t))\tau}, \quad (38)$$

che è il prodotto della probabilità che non accada alcuna reazione in τ e la propensione che nel $d\tau$ successivo accada R_j .

Le equazioni sopra scritte danno dunque un'idea di come ricavare la PDF di (τ, j) .

Per τ , invertendo la (37), si ottiene:

$$\tau = \frac{1}{a_0(\bar{X}(t))} \ln \left(\frac{1}{P(\tau|\bar{X}(t), t)} \right). \quad (39)$$

Per j invece, la probabilità che accada la reazione j -esima sarà data da

$$P(j, \bar{X}(t)) = \frac{a_j(\bar{X}(t))}{a_0(\bar{X}(t))}. \quad (40)$$

Queste equazioni sono alla base della formulazione dell'algoritmo di Gillespie, che si costruisce per passi come segue:

1. Si inizializzi il sistema, con opportune condizioni iniziali $\bar{X}(0)$.
2. Si definiscano le reazioni del sistema tramite gli appositi state change vector $\bar{v}_j$.
3. Si valuti le singole a_j per ottenere a_0 usando (36).
4. Si calcoli τ generando un numero casuale r_1 uniformemente distribuito in $[0,1]$ utilizzando (39) ponendo:

$$\tau = \frac{1}{a_0(\bar{X}(t))} \ln \left(\frac{1}{r_1} \right) \quad (41)$$

5. Si calcoli j generando un numero casuale r_2 uniformemente distribuito in $[0,1]$ moltiplicandolo per a_0 e, utilizzando (40), definire j come l'intero più piccolo che renda vera:

$$\sum_{l=1}^j a_l(\bar{X}(t)) > r_2 a_0 \quad (42)$$

6. Aggiornare la traiettoria con i parametri trovati come:

$$\bar{X}(t) \rightarrow \bar{X}(t + \tau) = \bar{X}(t) + \bar{v}_j \quad (43)$$

7. Ripetere l'algoritmo dal punto 4.

Questo algoritmo sfrutta la generazione casuale di r_1 e r_2 al fine di costruire una variabile che sia distribuita con PDF richiesta tramite il metodo della funzione cumulante.

4.4 Simulazioni con SSA

Si passa ora ad applicare in pratica l'algoritmo di Gillespie per simulare delle traiettorie per il modello SIR base stocastico.

La prima cosa che si manifesta (vedi Figura (4)), è come la numerosità del campione influenzi la regolarità delle traiettorie. Queste difatti risultano sempre più lisce e regolari man mano che si aumenta la popolazione N , dal momento che l'incertezza statistica diminuisce seguendo $\frac{1}{\sqrt{N}}$ in accordo con il teorema del limite centrale.

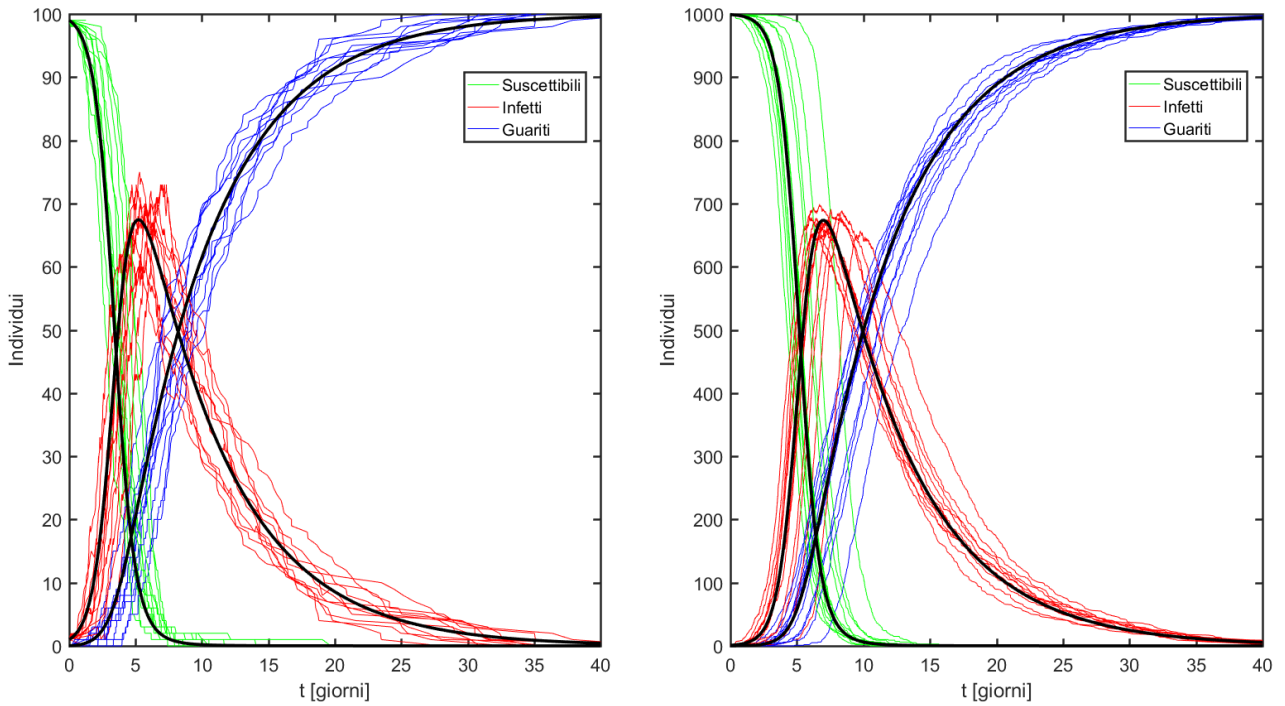


Figura 4: Confronto fra simulazioni stocastiche (a colori) e deterministiche (in nero) del modello SIR con $R_0 = 10$, $N=100$ (sinistra) e $N=1000$ (destra).

Questo dimostra intuitivamente quanto detto nella trattazione teorica della ME.

Tutto lo studio che si farà in seguito è mirato a studiare dal punto di vista stocastico alcune importanti caratteristiche epidemiologiche come ad esempio il tempo di estinzione, definito come l'intervallo di tempo necessario per cui il numero di infetti sia zero mentre, da un punto di vista deterministico come:

$$I(t_{\text{estinzione}}) = \frac{1}{N} \quad (44)$$

Innanzitutto si simula (3 run) diverse curve epidemiche variando il numero di infetti iniziali per R_0 fissato (vedi Figura (5)), in cui si vede dapprima che per 1 infetto iniziale, su 3 run solo 2 riescono a dare un'infezione consistente (cioè l'altra traiettoria simulata viene subito soppressa) e poi che in genere le curve stocastiche non seguono una chiara struttura di picco, a volte per niente, non essendo in linea dunque, né da un punto di vista temporale né da uno infettivo, con le predizioni deterministiche.

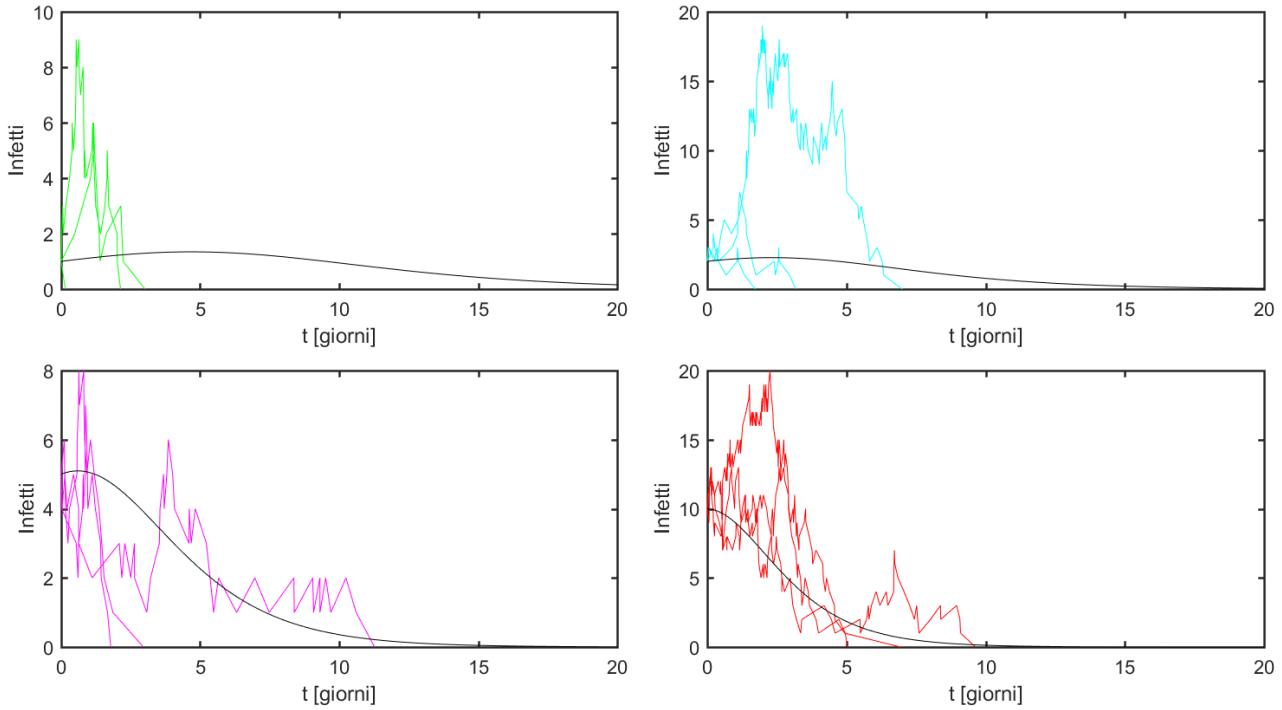


Figura 5: *Studio della dinamica stocastica variando il numero di infetti iniziali (1, 2, 5, 10) con $R_0 = 1.1$, $N=100$. La linea nera rappresenta la corrispondente simulazione deterministica.*

Dal momento che alcune traiettorie terminano più velocemente rispetto il modello EDO, si è pertanto interessati a costruire la PDF del tempo di estinzione dell'epidemia (vedi Figura (6)) ottenuta tramite MMC, svolgendo 10^5 run e visualizzate con bin width di 1 giorno, (questi parametri resteranno fissi per ogni altra simulazione MMC).

Quello che si ottiene è che per un numero basso di infetti iniziali (1-2) la distribuzione è piccata sulle prime 12/24 ore dell'epidemia, dunque la probabilità che questa si propaghi temporalmente è molto bassa ad esempio, per $N_{infetti} = 1$, la probabilità che l'epidemia superi le 48 ore è $\sim 30\%$.

Aumentando il numero di infetti, il tempo di estinzione più probabile si sposta dallo zero fino a formare una chiara struttura di picco, passando per un caso particolare che si presenta con $N_{infetti} = 5$, in cui vi è una specie di plateau in cui ogni giorno (dentro il plateau) risulta equiprobabile.

La Figura (6) mostra quindi in conclusione che considerando gli effetti stocastici, le cose vanno tipicamente meglio di quanto descritto dal modello via EDO, ossia stocasticamente l'epidemia termina più velocemente. Allo stesso tempo si può notare come in alcuni (rari) casi le fluttuazioni stocastiche portino ad allungare notevolmente il tempo di estinzione dell'epidemia.

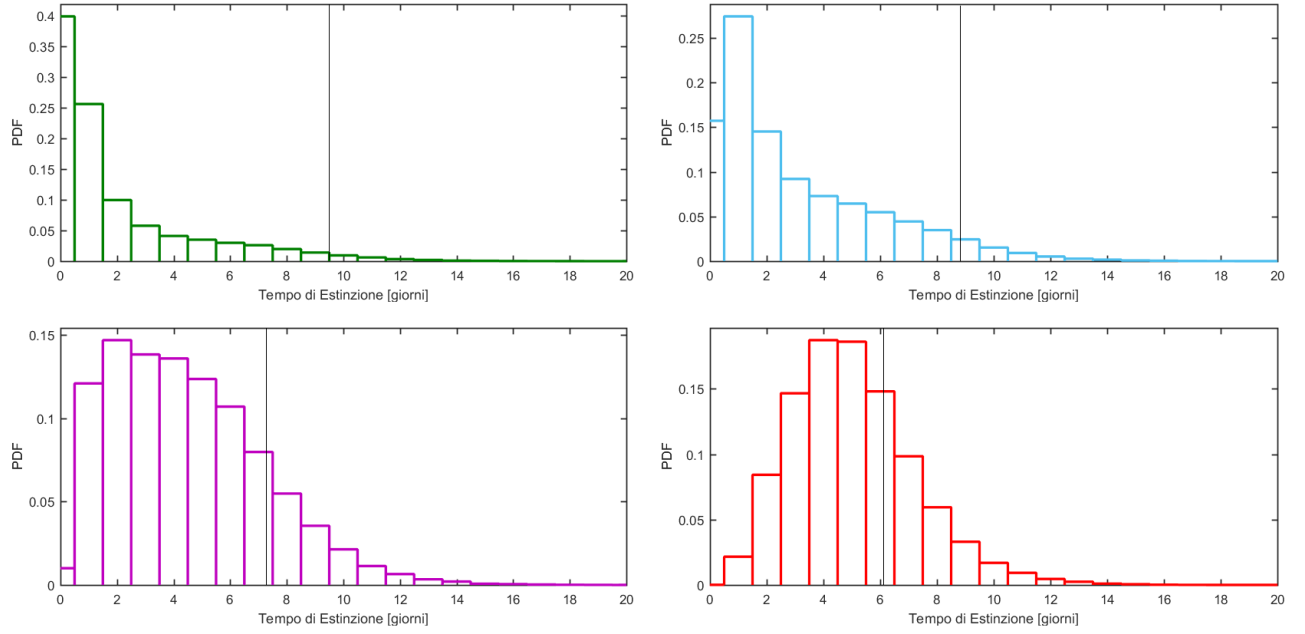


Figura 6: *Distribuzioni di probabilità del tempo di estinzione variando il numero di infetti iniziali (1, 2, 5, 10) con $R_0 = 1.1$, $N=100$. La linea nera rappresenta il corrispettivo tempo di estinzione deterministico.*

D'altra parte si potrebbe studiare l'effetto stocastico della variazione di R_0 con $N_{infetti} = 1$ (6 run), ottenendo:

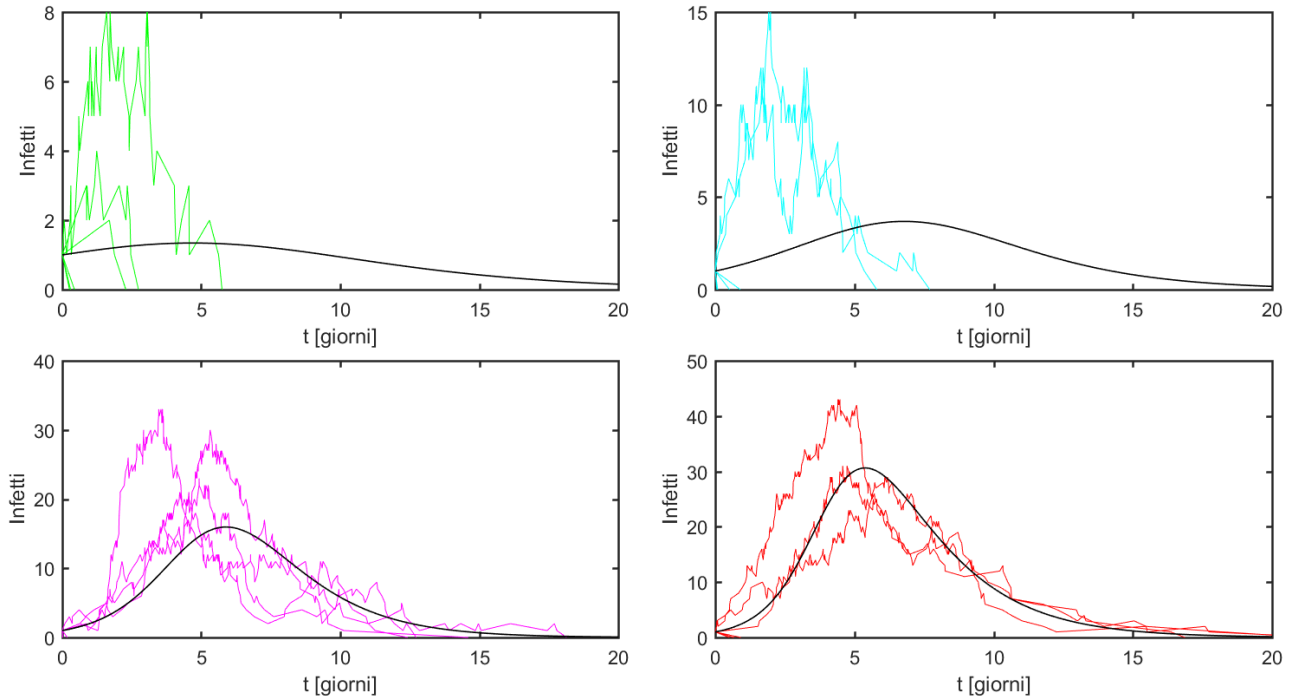


Figura 7: *Studio della dinamica stocastica variando R_0 (1.1, 1.3, 2, 3) con $N_{infetti} = 1$, $N=100$. La linea nera rappresenta la corrispettiva simulazione deterministica.*

La prima cosa evidente che si vede è che per R_0 piccoli (1.1-1.3) il picco degli infetti può essere decisamente più elevato a livello di contagiosità, circa 4-3 volte tanto rispettivamente, ma che allo stesso tempo risulta decisamente più stretto temporalmente; man mano che R_0 aumenta invece, le simulazioni stocastiche si addensano sempre meglio attorno alla curva deterministica. Analogamente a quanto fatto precedentemente, si può costruire la PDF del tempo di estinzione ottenendo:

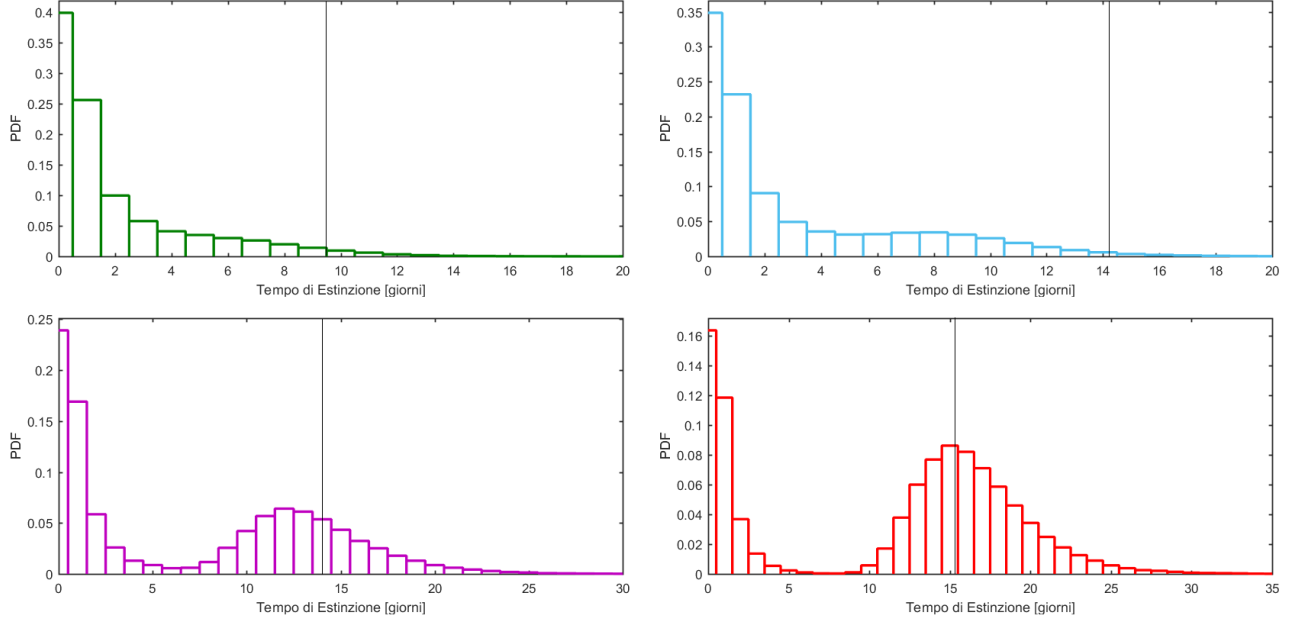


Figura 8: Distribuzioni di probabilità del tempo di estinzione variando R_0 (1.1, 1.3, 2, 3) con $N_{infetti} = 1$, $N=100$. La linea nera rappresenta il corrispettivo tempo di estinzione deterministico.

in cui si vede che facendo crescere R_0 , si forma un picco sempre più pronunciato che sottrae probabilità al bin iniziale corrispondente a $t_{estinzione} < 12h$, rendendo quindi sempre meno probabile il non sviluppo dell'epidemia.

Inoltre si può confrontare quello che è il tempo di estinzione per i modelli stocastici, con quello che è il tempo di estinzione deterministico. Si riportano in Tabella (1) i tempi di estinzione deterministici e il numero massimo di infetti (che sarà utile successivamente) è:

$N_{inizio,infetti}$	R_0	$T_{estinzione,EDO}$	$N_{max,infetti,EDO}$
1	1.1	9d 12h	1.3
2	1.1	8d 21h	2.7
5	1.1	7d 08h	5.1
10	1.1	6d 04h	10.0
1	1.3	14d 06h	3.7
1	2	14d 03h	15.8
1	3	15d 10h	30.4

Tabella 1: Tempi di estinzione e massimo numero di infetti deterministici (d: giorni - h: ore).

Per stimare quindi se il tempo di estinzione via EDO sia compatibile con quello via SSA, serve specificare un approccio per estrarre tale informazione dalle PDF ricavate. Ricordando che le

curve e quindi il tempo di estinzione dipendono esplicitamente dalla scelta fatta di (β, γ) che, in tutte le simulazioni fin'ora eseguite, sono stati scelti di valore $\beta = 1.5$, $\gamma = \beta/R_0$.

Questo a livello medico equivale simulare quali sono le possibili dinamiche endemiche variando il tempo di recupero di $1/\gamma$ giorni; viceversa si potevano studiare le dinamiche endemiche variando β che può essere visto come la facilità con la quale il virus riesce a propagarsi, simulando quindi quali sono i possibili effetti dell'utilizzo di metodi di prevenzione o contrasto (ad esempio DPI e/o quarantena). Entrambe le trattazioni sono utili ma in questo lavoro, per brevità, verrà trattato solamente il primo approccio.

Per R_0 (1.1-1.3), una scelta possibile è prendere la moda dal momento che queste PDF, a differenza di quelle per R_0 (2-3), sono mono-modali; questo comporta che, come visualizzabile in Figura (8), i due tempi stimati da SSA e da EDO non sono compatibili, ottenendo un tempo di estinzione probabile molto minore di quello deterministico. D'altra parte per R_0 (2-3), poiché le PDF sono bimodali, si può concludere che se si considera esclusivamente il picco emergente allora i due tempi stimati sono compatibili e lo sono sempre di più via via che R_0 aumenta.

In conclusione si studia il picco di contagi, cioè il massimo numero di infetti che si può presentare variando R_0 con $N_{infetti} = 1$.

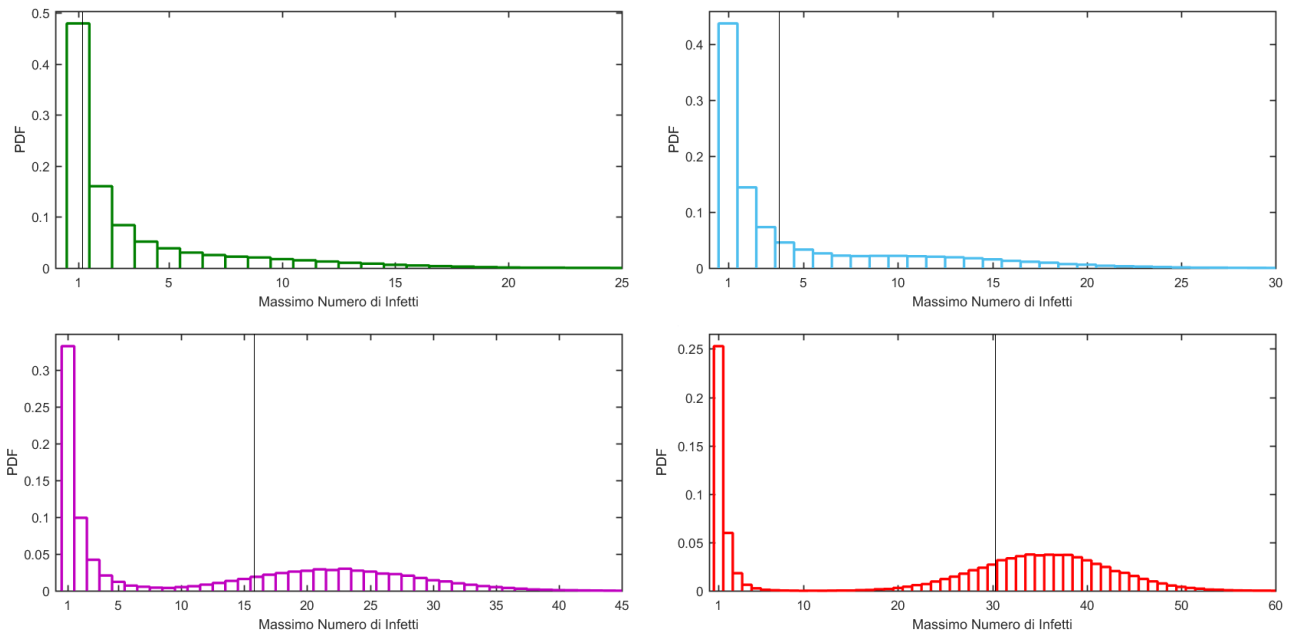


Figura 9: Distribuzioni di probabilità del massimo numero di infetti R_0 (1.1, 1.3, 2, 3) con $N_{infetti} = 1$, $N=100$. La linea nera rappresenta il massimo numero di infetti del modello deterministico calcolato con (14).

Sorprendentemente la PDF risultante ottenuta ha una struttura simile a quella precedentemente trovata per il tempo di estinzione: è dunque naturale domandarsi se vi sia una struttura soggiacente che lega il tempo di estinzione al massimo numero di infetti.

Da un punto di vista deterministico le EDO portano ad una relazione non banale (vedi Figura (10)), in cui con stupore si vede che il tempo di estinzione non è direttamente legato al fattore R_0 da una funzione strettamente crescente, dal momento che si presenta un minimo locale per $R_0 \sim 1.85$.

Infine, da ulteriori simulazioni (qui omesse ma ugualmente eseguite) si vede che anche la PDF

del massimo numero di infetti è invariante per coppie (β, γ) come per il modello via EDO.

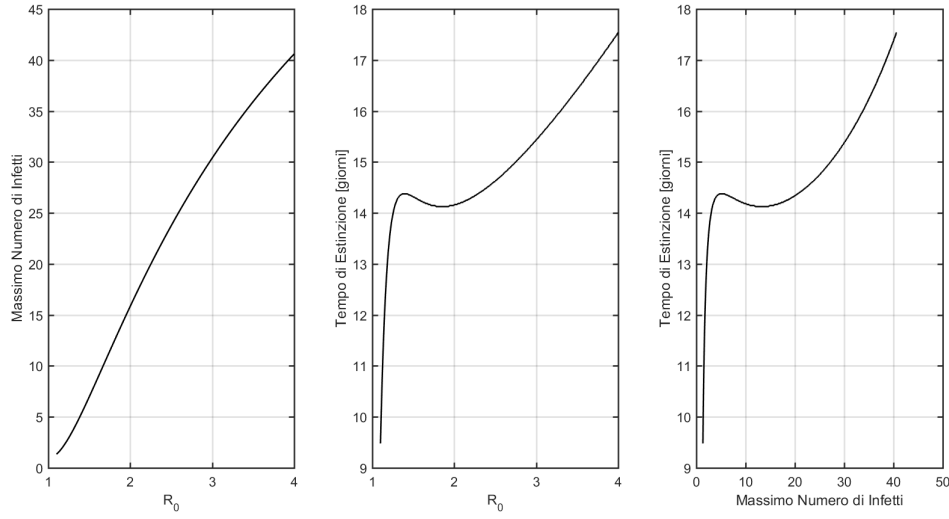


Figura 10: *Dipendenza funzionale tra il massimo numero di infetti, R_0 e il tempo di estinzione, per $N=100$ e $N_{infetti} = 1$, via EDO.*

Analogamente si potrebbe simulare la PDF del massimo numero di infetti rispetto gli infetti iniziali tenendo $R_0 = 1.1$, ottenendo:

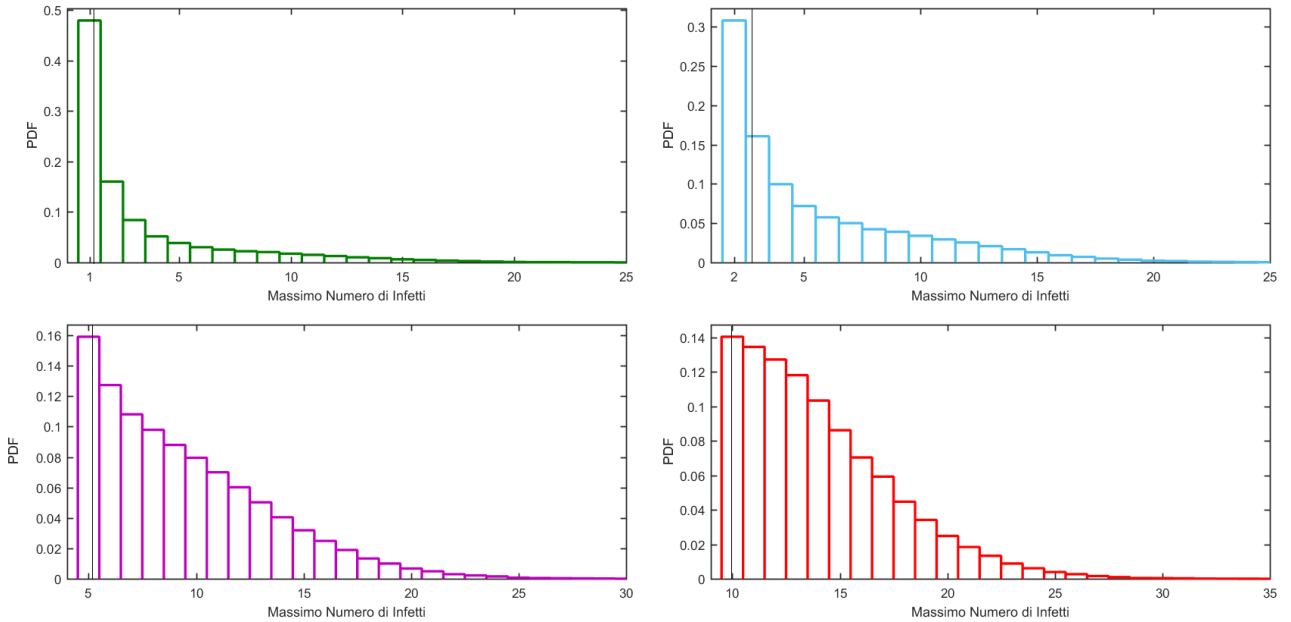


Figura 11: *Distribuzioni di probabilità del massimo numero di infetti (1, 2, 5, 10) con $R_0 = 1.1$, $N=100$. La linea nera rappresenta il massimo numero di infetti ottenuti via EDO usando (14).*

in cui vi è un evidente accordo, tra SSA e EDO. Inoltre si osserva che le PDF ottenute sono one-tail e mono-modali in cui la moda è il minimo valore ottenibile, portando ad una media sicuramente maggiore. Dunque in questo caso, il modello via EDO predice quello che è il minimo valore di infetti ottenibili nella realtà che poi, effettivamente, o sono un numero uguale o un numero più grande. Si verifica che aumentando R_0 la PDF del massimo numero di infetti tende ad una normale con media corrispondente a quella data nel caso deterministico.

5 Modello con Immunità e Mortalità Indotta

5.1 Descrizione Matematica

Se al modello precedente si introducono i parametri legati alla mortalità indotta dall'infezione e alla presenza di immunità, allora si può riformularlo come:

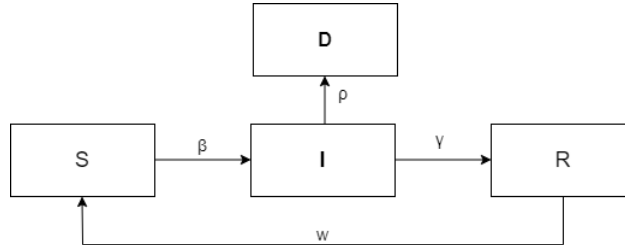


Figura 12: *Diagramma a Blocchi SIR con mortalità e immunità*

$$\frac{dS}{dt} = -\beta SI + wR \quad (45)$$

$$\frac{dI}{dt} = \beta SI - \gamma I - \frac{\rho}{1-\rho} \gamma I \quad (46)$$

$$\frac{dR}{dt} = \gamma I - wR \quad (47)$$

Importante è osservare che il parametro ρ , per come è stato definito nel paragrafo 3.1.3, può solo entrare nella descrizione della classe degli infetti; inoltre, dal momento che si è introdotta la possibilità che l'infetto muoia, per continuare a rispettare il vincolo di N costante, occorre introdurre una nuova classe dei *Deceduti* (D), affinché

$$N = S + I + R + D. \quad (48)$$

Per descrivere la classe dei deceduti via EDO, basti pensare che se si vuole tenere conto di N costante, la descrizione per D sarà:

$$\frac{dD}{dt} = +\frac{\rho}{1-\rho} \gamma I \quad (49)$$

Da un punto di vista di SSA, l'introduzione di una nuova classe e il vincolo di tenere N costante, fanno sì che lo state change vector sia ora 4-dim e che quindi occorra definire una nuova reazione associata all'equazione (49).

Molto importante potrebbe essere anche studiare il numero massimo (o limite) di morti, che può essere ricavato sia analiticamente (sostanzialmente si procede come fatto in (13)), ma sia anche osservando che ρ rappresenta la probabilità che mediamente, dato un numero N di soggetti, muoiano ρN individui. Dunque, dal momento che l'EDO rappresenta il comportamento medio dell'epidemia, si avrà che

$$D_{lim} = \rho N. \quad (50)$$

5.1.1 Immunità

Per studiare l'immunità, conoscendo già i due casi limiti SIS e SIR, si vuole innanzitutto verificare che i due si raccordino con continuità. Svolgendo la simulazione in effetti (vedi Figura (13)) si può apprezzare la transizione dal modello SIS al modello SIR.

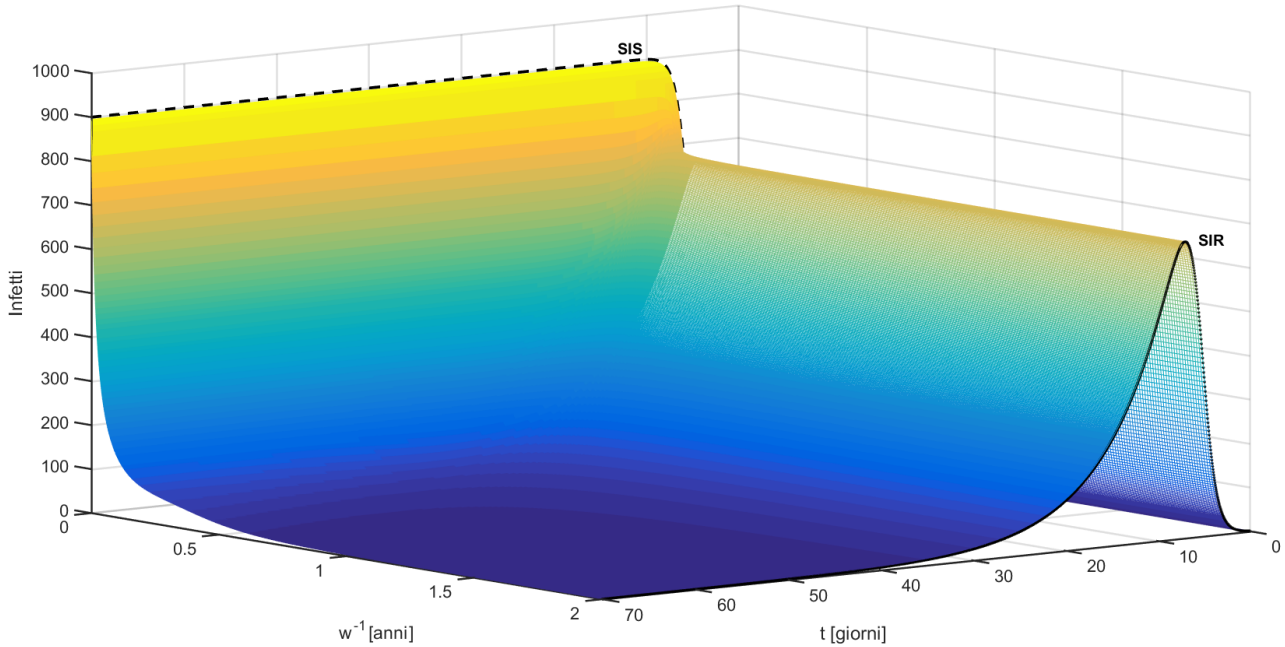


Figura 13: Numero di infetti in funzione del parametro w^{-1} con $R_0 = 10$, $N=1000$. La linea nera tratteggiata rappresenta il modello SIS, mentre la linea nera continua rappresenta il modello SIR.

Tuttavia, a meno che il tempo in cui si perde l'immunità sia paragonabile al tempo di estinzione dell'epidemia, l'immunità non gioca un ruolo apprezzabile nello sviluppo di questa.

Come si vede in Figura (13) in effetti, una modifica sensibile della curva SIR inizia ad aversi attorno 3-6 mesi ma allo stesso tempo, per il modello SIR in esame, il numero di infetti inizia a essere sensibilmente basso attorno a 1 mese, dunque se l'epidemia termina molto prima di una possibile perdita di immunità (solitamente anni), questa non riesce a farla ripartire e sostanzialmente si potrebbe tornare ad avere una società di tutti nuovamente suscettibili, ma con più nessun infetto.

Se ne conclude che, come si ha verificato, è possibile trascurare sia la demografia (trattazione omessa) che l'immunità.

5.1.2 Mortalità

Infine, per comprendere meglio l'impatto della mortalità sull'epidemia, si studia attraverso le simulazioni la variazione di $I(t)$ in funzione di ρ (vedi Figura (14)).

Si nota in particolare che per $\rho \rightarrow 0$ (linea tratteggiata), si ottiene il modello SIR base come predetto e che inoltre, all'aumentare di ρ , il picco degli infetti risulta sempre più basso e sempre più traslato temporalmente: questa è una conseguenza diretta della alta mortalità del patogeno, che fa sempre più fatica a trovare suscettibili per propagarsi.

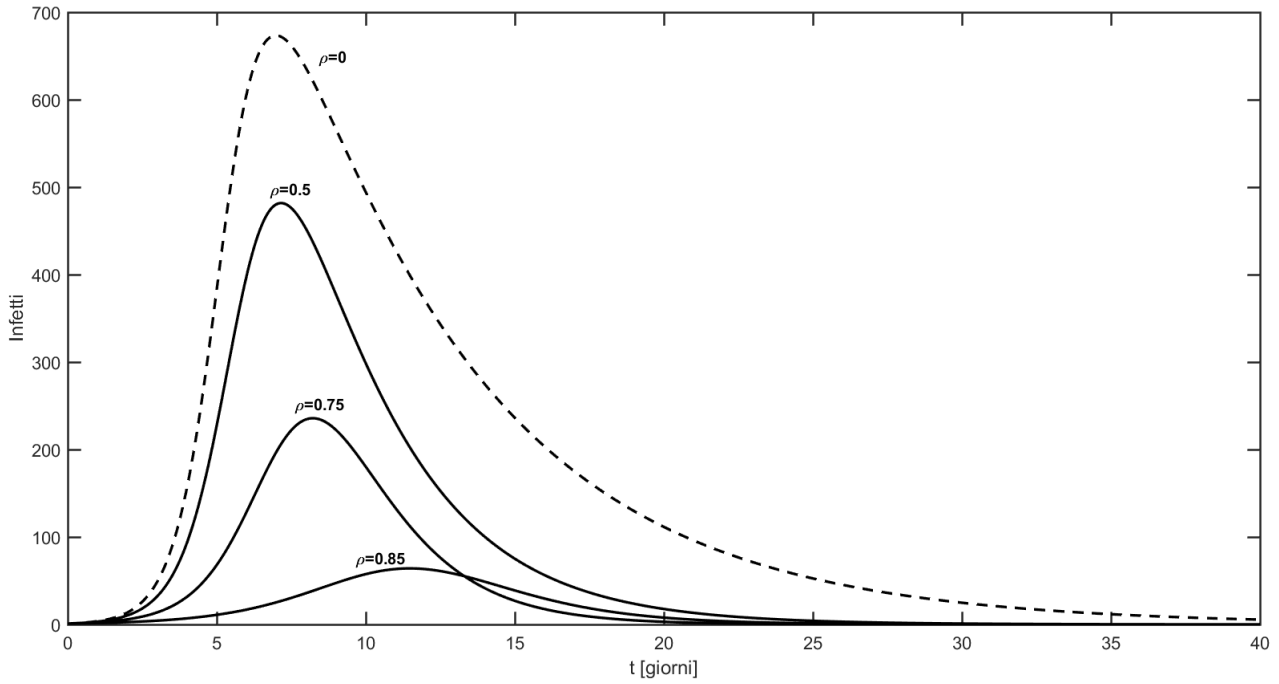


Figura 14: Studio degli infetti in funzione di ρ con modello SIR e mortalità. $R_0 = 10$, $N=1000$. La linea tratteggiata rappresenta il modello SIR base.

5.2 Simulazione via SSA

Come fatto in Figura (4), si simula il modello SIRD confrontando i risultati tra SSA e EDO.

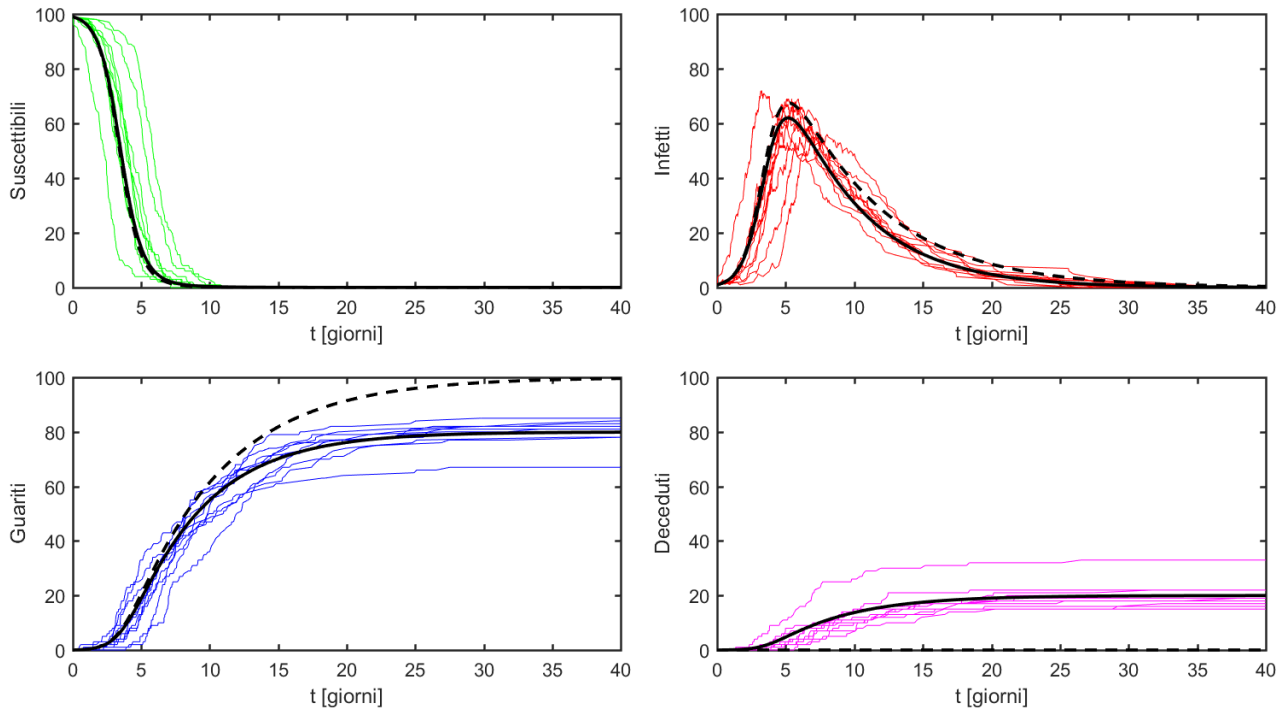


Figura 15: Simulazione modello SIRD con $R_0=10$, $N=100$, $N_{infetti}=1$ e $\rho=0.2$. In nero continuo il modello SIRD, in nero tratteggiato il modello SIR.

Si nota immediatamente che per quanto riguarda infetti e suscettibili, le traiettorie stocastiche si addensano bene sulle corrispettive traiettorie classiche, cosa che invece non accade per niente per ricoverati e deceduti in cui soprattutto al termine dell'epidemia, vi è un intero fascio di traiettorie possibili. È naturale quindi domandarsi se il limite stocastico del numero di infetti e morti sia compatibile con il limite classico del modello via EDO.

Eseguendo una simulazione MC si può costruire la PDF del numero di deceduti, variando ad esempio R_0 :

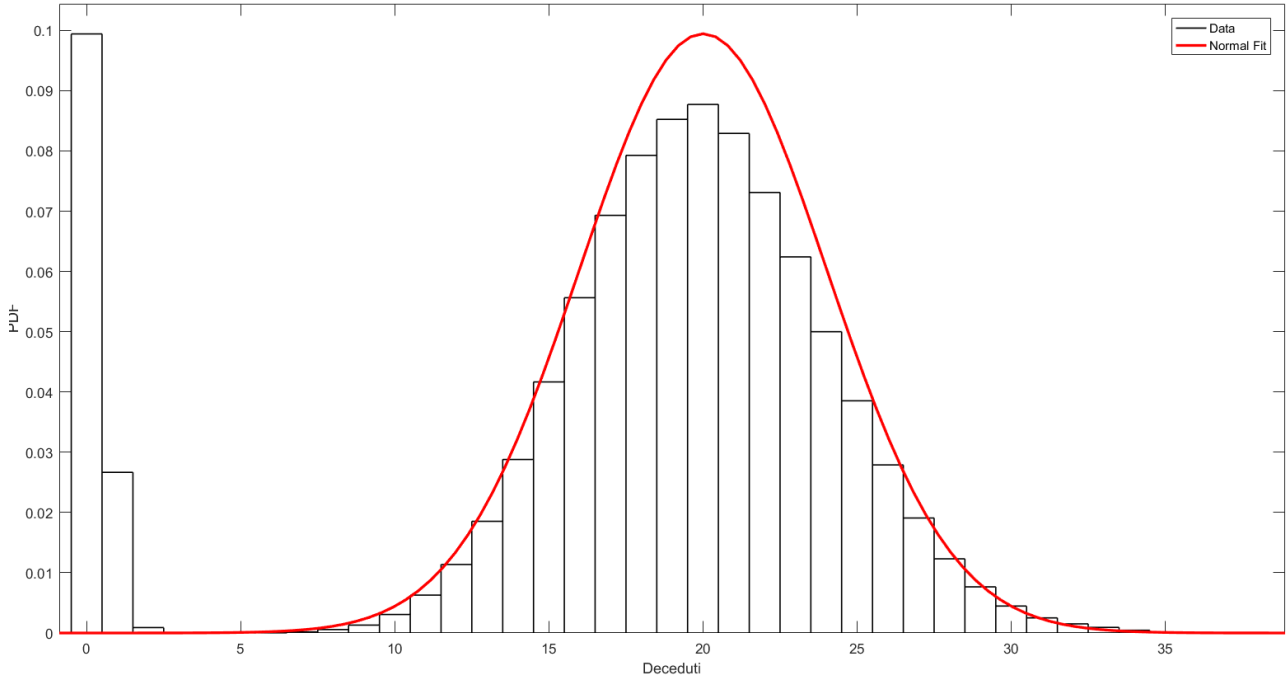


Figura 16: *PDF del numero totale di deceduti con $R_0 = 10$, $N_{infetti} = 1$ e $N=100$. In rosso applicato un fit normale.*

potendo osservare che il valore limite stocastico di defunti si distribuisce secondo una gaussiana attorno al valore $\mu = 19.98 \pm 0.01$ concorde a quanto trovato via EDO di $D_{lim} = 20$.

6 Conclusioni

Si possono quindi riassumere i risultati ottenuti come segue:

- Una simulazione via SSA contiene più informazioni della corrispettiva EDO, anche se con l'aumentare della popolazione N e di R_0 tenderà a quest'ultima.
- Aumentando il numero di infetti iniziali è sempre meno probabile che l'epidemia termini velocemente e/o non si presenti.
- Il modello via EDO in generale sovrastima il tempo di estinzione tendendo a coincidere, per R_0 elevati, alla moda del picco centrale. Considerando però le fluttuazioni stocastiche, in alcuni rari casi il tempo di estinzione può crescere notevolmente rispetto a quello predetto via EDO.
- Il tempo di estinzione non è proporzionalmente collegato a R_0 , ma vi è una dipendenza funzionale non banale, che lega R_0 , tempo di estinzione e massimo numero di infetti.
- La PDF del massimo numero di infetti (e quindi del tempo di estinzione) ha una natura bimodale, con limite alla mono-modalità per $R_0 \sim 1$ e $R_0 \gg 1$.
- Il massimo numero di infetti per $R_0 = 1.1$ ottenuto via EDO, corrisponde alla moda della PDF associata, ma non il suo valore di aspettazione che risulta essere maggiore.
- Il modello con immunità può essere ben approssimato da un modello SIR base, se la perdita di immunità avviene in un tempo molto più lungo (anni) rispetto il tempo di estinzione dell'epidemia (mesi).
- Nel modello con mortalità, limitatamente al caso in esame ($R_0 = 10$), il numero limite di morti si distribuisce normalmente attorno al valore limite calcolato via EDO.

7 Sviluppi Futuri

Vi sono innumerevoli possibili sviluppi futuri di questa tesi. In particolare, in questo lavoro non abbiamo mai analizzato la probabilità degli infetti nella versione multiparametrica, ovvero in funzione congiunta dei parametri R_0 e numero di infetti iniziali. Inoltre, attraverso la ME si potrebbe studiare direttamente la PDF dei tempi di estinzione, oltre ad approfondire il rapporto tra quest'ultima e il massimo numero di infetti, per esempio calcolando la probabilità congiunta $P(t_{estinzione}, N_{maxinfetti})$; Infine sarebbe interessante investigare come la PDF del numero di infetti evolva se consideriamo alcuni parametri quali R_0 come dipendenti dal tempo. Infatti, come vediamo drammaticamente in questo periodo, R_0 non è un parametro fisso, ma la prevenzione e la responsabilità individuale possono modificarlo.

8 Bibliografia

- [Gill] D.T Gillespie: *Stochastic Simulation of Chemical Kinetics*.
 [AndMay] Anderson & May: *Population Biology of Infectious Diseases*.
 [Allen] L.J.S Allen: *An Introduction to Stochastic Processes with Applications to Biology*.
 [KeeRoh] M.J Keeling and P. Rohani: *Introduction to Simple Epidemic Models*.