

Università degli studi di Padova

CORSO DI LAUREA MAGISTRALE IN SCIENZE STATISTICHE

ANNO ACCADEMICO 2012-2013



Modelli Stocastici in epidemiologia:  
SIR e sue generalizzazioni

*Relatore*

Prof. Marco Ferrante

*Laureanda*

Elisabetta Ferraris



# Indice

<b>1</b>	<b>Modelli epidemiologici SIR</b>	<b>9</b>
1.1	Approccio deterministico . . . . .	10
1.2	Approccio probabilistico: Il modello Reed-Frost . . . . .	14
1.2.1	Modello SIR di Tuckwell e Williams . . . . .	16
<b>2</b>	<b>Estensioni del modello SIR</b>	<b>23</b>
2.1	Modello con R aleatorio . . . . .	24
2.2	Modello SEIMR . . . . .	27
2.3	Modello SEIMRS . . . . .	36
2.4	Modello SEIMS . . . . .	40
<b>3</b>	<b>Il modello SIR a popolazione aggregata</b>	<b>43</b>
3.1	Introduzione . . . . .	43
3.2	Struttura del modello e calcolo della durata media . . . . .	44
<b>4</b>	<b>Applicazioni</b>	<b>53</b>
4.1	Varicella . . . . .	53
4.2	Scarlattina . . . . .	58
<b>A</b>	<b>Esempio di codice R utilizzato</b>	<b>63</b>
	<b>Bibliografia</b>	<b>71</b>

Gente non  
riesco  
a crederci ma con  
zelo e  
iniziativa è arrivata  
e passata  
la laurea..la dedica  
e la citazione sono strane ma  
oh che ci volete fare? Io sono così

# Introduzione

In epidemiologia sono numerosi i modelli utilizzati per descrivere l'andamento delle epidemie. Tra i più diffusi si hanno sicuramente i SIR. L'acronimo SIR indica gli stati in cui è suddivisa la popolazione: *suscettibili, infetti, rimossi*. In questo elaborato, partendo da un approccio stocastico ai modelli SIR proposto da Tuckwell e Williams in [1], sono state realizzate e implementate alcune variazioni. Il modello di riferimento è dato dalla catena di Markov:

$$\mathbf{W}(t) = [X^i(t), Y_0^i(t), Y_1^i(t), \dots, Y_{R-1}^i(t), \quad i = 1, 2, \dots, n]$$

a  $\{0, 1\}^{(R+1)n}$  stati, dove  $R$  è fisso e rappresenta la durata in giorni dell'infezione,  $n$  è la popolazione totale supposta costante,  $X^i(t) \in \{0, 1\}$  indica se l'individuo  $i$ -esimo è suscettibile al tempo  $t$ ,  $Y_l^i(t) \in \{0, 1\}$  indica se l'individuo  $i$ -esimo in  $t$  è infetto da  $l$  unità di tempo. Le assunzioni alla base del modello sono l'indipendenza nel comportamento degli individui (sia nell'ammalarsi sia negli incontri), la possibilità che incontrino, per ogni unità tempo,  $N_i(t)$  persone, l'assenza del tempo di latenza e il fatto che un individuo infettato una volta trascorsi  $R$  giorni diventi rimosso, cioè immune da un ulteriore contagio.

In ogni tempo quindi viene valutata la condizione di ogni singolo individuo e unendo le informazioni sulle singole persone si traggono informazioni sulla popolazione in generale. Sulla base di questo modello è stata inizialmente realizzata una variante in cui la durata dell'infezione è stata considerata aleatoria, le epidemie descritte da questo modello risultano essere leggermente più lunghe delle precedenti ma con lo stesso numero di coinvolti.

È stato poi realizzato un modello SEIMR in cui è gestita la presenza di un tempo di latenza (E) e il periodo di contagio è diviso in due fasi: la prima in cui i soggetti sono infettati e contagiosi ma non manifestano alcun sintomo (I) e una seconda, in cui i soggetti sono contagiosi e manifestano i sintomi (M); una volta terminata questa fase gli individui diventano immuni.

Ulteriori novità introdotte nel modello sono:

- la possibilità di considerare gli incontri non equiprobabili, in quanto è ragionevole ipotizzare che chi manifesta i sintomi della malattia si assenti dalle proprie attività e sia meno facilmente incontrabile;
- la possibilità di utilizzare diverse probabilità di contagio al variare del gruppo di appartenenza degli individui incontrati (se in I o M).

È stata valutata come varia la durata e l'espansione dell'epidemia al variare della probabilità di contagio, della durata del periodo di infezione e del numero di incontri giornalieri.

Sono stati poi valutati altri due modelli varianti del SEIMR: il SEIMRS e il SEIMS.

Nel primo è stata considerata la possibilità che una malattia possa solo in alcuni casi conferire l'immunità e che negli altri i coinvolti tornino suscettibili. La struttura del modello rimane la stessa ma, alla fine del periodo di contagiosità sintomatica, ogni individuo può diventare immune con probabilità  $1 - \gamma$  o tornare suscettibile con probabilità  $\gamma$ . Ovviamente, potendo un individuo riprendere la malattia più volte, il numero di casi è maggiore rispetto al modello precedente e dipende anch'esso da  $\gamma$ . Verranno fatte analisi esplorative per vedere, anche in questo caso, quali sono i fattori che maggiormente influenzano il numero medio di casi e la durata dell'epidemia. Il modello SEIMS invece è adatto a modellare malattie che non danno alcun tipo di immunità e nelle quali, dopo l'infezione, i soggetti tornano ad essere suscettibili.

L'assenza dell'immunità comporta un incremento considerevole del numero di casi che a sua volta dipenderà da altri fattori come la probabilità di contagio nella fase sintomatica (elemento che condiziona anche la durata dell'epidemia). Il modello sembra essere adatto a descrivere patologie con grande diffusione, che hanno periodo di infezione relativamente breve e diverso tipo di contagiosità.

Il modello SEIMR è stato poi applicato a due patologie ben note: la varicella e la scarlattina. Soprattutto nel primo caso i risultati ottenuti sono abbastanza simili all'andamento reale della varicella, infatti è descritto il classico andamento con doppio picco di contagi che rende l'epidemia lunga e molto diffusa. Il SEIMR sembra quindi molto adatto a descrivere questa patologia proprio grazie alla struttura della malattia, che rispetta esattamente tutti gli stati del modello. Con un modello SIR, invece non si riesce a descrivere altrettanto bene l'andamento di questa patologia, in quanto non si può cogliere la differenza tra la fase sintomatica e quella non, che in questo caso hanno durata e probabilità di contagio degli individui molto diverse.

Infine viene proposto un approccio alternativo basato su un modello a popolazione aggregata in cui, per studiare una malattia, in ogni tempo, si valuta il numero totale di persone presenti in ogni stato (suscettibili, malati da un giorno, ecc...). Il modello risulta più difficile da implementare e da gestire numericamente, e per questo motivo sono stati considerati solo due casi relativamente semplici. Nel primo è stata modellata una malattia della durata di un giorno dopo il quale si diventa immuni. Sono stati calcolati i tempi medi di assorbimento, che non sono altro che la durata dell'epidemia partendo da diversi stati iniziali, ed è stata valutata la durata massima raggiunta al variare della probabilità di contagio e della numerosità della popolazione. Nel secondo caso il modello è stato allargato anche a malattie lunghe due giorni e sono state fatte le stesse analisi esplorative.

Per entrambi i modelli ci sono delle pesanti restrizioni circa la numerosità della popolazione  $n$ , infatti per problemi computazionali, non è stato possibile considerare una popolazione molto elevata. Nel primo caso il massimo che è stato possibile valutare è  $n = 100$  mentre nel secondo non è stato possibile utilizzare una popolazione maggiore di 50 unità.





# Capitolo 1

## Modelli epidemiologici SIR

Si definisce epidemia la rapida espansione, o un'improvvisa comparsa, di una malattia, di solito infettiva, in una determinata popolazione o area geografica con tendenza alla diffusione ad un ampio numero di individui. Nei modelli epidemiologici di solito si considerano come variabili non il numero di agenti patogeni (batteri, virus, ecc.) presenti nei contagiati, ma il numero di individui che sono stati infettati. Infatti il numero di virus o batteri presenti all'interno di un individuo infetto cresce molto rapidamente per raggiungere numeri altissimi. Si ritiene in genere che sia più importante, per descrivere l'andamento dell'epidemia, conoscere il numero di individui infetti piuttosto che il numero esatto di agenti patogeni presenti in ognuno di essi. In letteratura sono numerose le tipologie di modelli utilizzati per studiare la diffusione e il decorso delle epidemie, quelli che però risultano essere alla base della maggioranza dei modelli epidemiologici per le malattie infettive sono i modelli SIR.

Il filo conduttore di tutti i modelli basati sui SIR è la suddivisione della popolazione in diverse categorie e lo studio di come avvengono i passaggi di stato. Le categorie principali presenti nella maggioranza dei modelli sono: *suscettibili*, *infettivi* e *rimossi*. In base alle caratteristiche dell'epidemia (conferire o meno l'immunità, tempo di incubazione della malattia ecc..) si ottengono diversi modelli. Ad esempio con un numero maggiore di stati in cui è suddivisa la popolazione, nel caso in cui si consideri il tempo di latenza, o un diverso passaggio di stato, nel caso di assenza di immunità. Il contagio avviene sempre per contatto con una persona contagiosa mentre i passaggi di stato successivi all'infezione avvengono indipendentemente dalle interazioni tra individui. La popolazione considerata è chiusa e quindi la si suppone costante.

I modelli SIR e quelli ad essi legati sono alla base non solo dei modelli deterministici, ma anche di molti stocastico-probabilistici, in cui si inseriscono

delle distribuzioni di probabilità per l'analisi delle fluttuazioni istantanee.

## 1.1 Approccio deterministico

Le prime notizie relative a questi modelli si riscontrano in Kermack e McKendrick [5] dove viene definito per la prima volta un modello SIR, basato su equazioni differenziali, e volto a descrivere il decorso di malattie infettive che conferiscono l'immunità. È stata considerata una popolazione costante  $n$  in ogni tempo  $t$  e suddivisa in tre categorie:

- *Suscettibili*,  $S(t)$
- *Infettivi*,  $I(t)$
- *Rimossi*,  $R(t)$

dove i suscettibili sono i soggetti a rischio, ossia gli individui che non hanno ancora contratto la malattia, gli infettivi o contagiosi sono invece l'insieme degli individui che hanno già contratto la malattia e pertanto in grado di contagiare altre persone, infine la categoria dei rimossi comprende tutti coloro i quali hanno già avuto la malattia e non risultano più contagiosi (immuni o deceduti). Essendo la popolazione costante si ha inoltre:

$$n = S(t) + R(t) + I(t)$$

Il modello si basa su alcune assunzioni:

- l'assenza del tempo di incubazione;
- il contagio avviene per contatto diretto;
- la probabilità di incontro di due qualsiasi individui è uguale.

Considerando quindi la popolazione descritta in precedenza e le assunzioni appena citate, Kermack e McKendrick derivarono le seguenti equazioni:

$$\begin{cases} \frac{dS}{dt} = -\alpha SI \\ \frac{dI}{dt} = \alpha SI - \beta I \\ \frac{dR}{dt} = \beta I \end{cases}$$

dove  $\alpha$  e  $\beta$  rappresentano, rispettivamente, il *tasso d'infezione della malattia* e il *tasso medio di guarigione* (reciproco del tempo medio di guarigione).

Dalla prima equazione si nota come la variazione di suscettibili dipenda da quanto velocemente l'infezione si propaghi, dal numero di infetti e dal numero

di individui ancora infettabili. Essa risulta essere negativa, in quanto casi di nuovi infetti sono presenti fino alla fine dell'epidemia il che comporta la riduzione dei suscettibili.

Per valutare invece la variazione di contagiosi e di rimossi è stato considerato anche il tempo medio di guarigione che condiziona la durata del periodo che un individuo trascorre nella classe degli I.

Dalle equazioni precedenti si può ricavare I in funzione di S, infatti si ha:

$$\begin{aligned}\frac{dI}{dS} &= \frac{\alpha SI - \beta I}{-\alpha SI} \\ &= -1 + \frac{\beta}{\alpha S}\end{aligned}$$

da cui

$$dI = \left[-1 + \frac{\beta}{\alpha S}\right]dS$$

integrando si ottiene

$$I(S) = c_0 - S + \frac{\beta}{\alpha} \log(S) \quad (1.1)$$

da cui si ricava  $c_0$

$$c_0 = I_0 + S_0 - \frac{\beta}{\alpha} \log(S_0) \quad (1.2)$$

Al variare di  $\alpha$  e  $\beta$  e dei valori iniziali  $R_0 S_0$  e  $I_0$  varia il decorso dell'epidemia. Si può notare però che se  $S_0 < \beta/\alpha$  l'infezione non si propaga, ossia il numero di infettati non aumenta, il che non comporta l'assenza di nuovi casi di pazienti infettati, ma il numero di questi sarà inferiore rispetto ai rimossi. Viceversa se  $S_0 > \beta/\alpha$  il numero di infettivi crescerà fino al tempo  $t$  in cui  $S(t) < \beta/\alpha$ .

Nel caso in cui si verificasse un'epidemia, può essere interessante stimare in anticipo la percentuale totale di popolazione colpita e la percentuale massima di persone colpite in un determinato istante,  $I_*/N$ , valori utili al fine dell'organizzazione di misure sanitarie. Per quanto riguarda il primo punto, la percentuale totale di popolazione colpita è la soluzione dell'equazione seguente, in cui si sfrutta la forma della costante di  $c_0$  con  $t=0$  e  $t \rightarrow \infty$  e il fatto che in quest'ultimo caso I sia pari a zero (ossia si è al termine della epidemia)

$$-\left(\frac{\beta}{\alpha}\right)\log(S_\infty) + S_\infty = I_0 + S_0 - \frac{\beta}{\alpha}\log(S_0)$$

che assumendo  $I_0 \ll S_0$  può essere riscritta come

$$\begin{aligned} S_\infty - \left(\frac{\beta}{\alpha}\right)\log(S_\infty) &= S_0 - \frac{\beta}{\alpha}\log(S_0) \\ \log[(S_0/S_\infty)^{(\beta/\alpha)}] &= S_0 - S_\infty \end{aligned}$$

la quantità  $S_\infty/S_0$  rappresenta la frazione di popolazione che non viene coinvolta nell'epidemia. L'equazione che lega  $S_\infty$  e  $S_0$  è però trascendente e va risolta numericamente.

Di grande importanza è, inoltre, la stima dell'ampiezza dell'epidemia nel suo momento di picco,  $I_*$ , valore ottenibile tramite l'equazione (1.2) e sfruttando il fatto che il massimo di  $I$  si raggiunga con  $S = \frac{\beta}{\alpha}$ :

$$I_* = I\left(\frac{\beta}{\alpha}\right) = c_0 - \frac{\beta}{\alpha}\log\left(\frac{\beta}{\alpha}\right) \quad (1.3)$$

che data (1.3) e assumendo  $I_0 \ll S_0$  per cui  $S_0 \simeq I_0 + S_0$  diventa

$$\begin{aligned} I_* &= I\left(\frac{\beta}{\alpha}\right) = S_0 - \frac{\beta}{\alpha}\log(S_0) - \frac{\beta}{\alpha}\log\left(\frac{\beta}{\alpha}\right) \\ &= \left(S_0 - \frac{\beta}{\alpha}\right) - \frac{\beta}{\alpha}\log(S_0/(\beta/\alpha)). \end{aligned} \quad (1.4)$$

Conoscendo, quindi, i valori iniziali,  $\alpha$  e  $\beta$  si può studiare l'andamento dell'epidemia.

Nell'approccio deterministico  $\alpha$  e  $\beta$ , che sono quindi i parametri che regolano la diffusione e la durata dell'epidemia, sono valori costanti, mentre nell'approccio stocastico verranno modellati in modo più complesso, aspetto che verrà approfondito in seguito.

Sempre nell'ambito dell'approccio deterministico, e non solo, esistono molte varianti dei modelli SIR che dipendono dalla numerosità e dalla tipologia di classi in cui viene divisa la popolazione. In alcuni casi, ad esempio, si ipotizza che i soggetti infettati non siano immediatamente contagiosi ma che lo diventino solo dopo un periodo di latenza. Risulta quindi necessario considerare una classe in più in cui suddividere la popolazione: in questo caso si parla di modelli SEIR (dall'inglese *expecting*).

In altri casi si fanno delle assunzioni riguardo l'immunità che, a volte, è considerata temporanea per cui dalla classe dei rimossi si ritorna in quella dei suscettibili (modelli SIRS), altre volte invece è considerata assente, per cui non viene contemplata la possibilità di rimanere immuni alla malattia neanche temporaneamente, come avviene ad esempio per il raffreddore, in questo caso si parla di modelli SI.

Considerando ad esempio quest'ultimo modello, le equazioni che lo descrivono sono leggermente differenti rispetto a quelle dei modelli SIR e sono

$$\begin{cases} \frac{dS}{dt} = -\alpha SI + \beta I \\ \frac{dI}{dt} = \alpha SI - \beta I \end{cases}$$

Essendo la popolazione costante, ci si può ricondurre ad una sola equazione, esprimendo I in funzione di N e S:

$$\begin{aligned} \frac{dS}{dt} &= -\alpha S(N - S) + \beta(N - S) \\ &= -\alpha SN - \alpha S^2 + \beta N - \beta S \\ &= \beta N - (\alpha N + \beta)S - \alpha S^2. \end{aligned} \tag{1.5}$$

la cui soluzione esatta risulta:

$$\frac{\beta e^{\alpha N t} + c_0 N e^{\beta t}}{\alpha e^{\alpha N t} + c_0 e^{\beta t}} \tag{1.6}$$

in cui  $c_0$  ha la forma

$$c_0 = \alpha \left( \frac{S_0 - \beta/\alpha}{N - S_0} \right)$$

Come nei modelli SIR se  $S_0 < \beta/\alpha$  l'epidemia non si diffonde, viceversa se  $S_0 > \beta/\alpha$  l'epidemia si diffonde e il numero di infetti cresce fino al tempo  $t$  in cui si ha  $S(t) < \beta/\alpha$ .

## 1.2 Approccio probabilistico: Il modello Reed-Frost

In passato l'approccio deterministico è stato molto usato, soprattutto per la sua semplicità di calcolo. Un limite però di questo tipo di modelli consiste nel fatto che sono mantenuti inalterati i parametri  $\alpha$ ,  $\beta$  e i valori iniziali, per cui si ottengono sempre gli stessi risultati. Non è possibile quindi cogliere la variabilità intrinseca ad ogni epidemia. Inoltre risulta più intuitivo parlare di probabilità di contagio e di numero atteso di casi, suscettibili e rimossi, piuttosto che di valori certi. Per far questo è stata introdotta nei modelli SIR una componente di variabilità, ossia i tassi sono stati modellati con delle distribuzioni di probabilità, ottenendo così i modelli *probabilistici*. Grande diffusione ha avuto inoltre l'approccio stocastico in cui l'andamento dell'epidemia è modellato con delle catene di Markov. In questi modelli si ipotizza che la suddivisione della popolazione, al tempo  $t$ , dipenda solo dalla suddivisione in  $t - 1$  e dai contagi avvenuti in  $t$ . Questo richiederà un modello strutturato in un particolare modo, che verrà descritto in seguito. I modelli stocastico-probabilistici permettono di considerare quindi le fluttuazioni casuali di popolazione e di tradurli in probabilità. Molti sono i modelli probabilistici esistenti, fra questi uno dei primi che ormai assume un'importanza storica ma che è alla base di un'ampia famiglia di modelli è il Reed-Frost. Reed e Frost proposero il loro modello nel 1928 anche se venne pubblicato solo nel 1952 da Abbay<sup>1</sup>.

La popolazione è supposta costante e pari ad  $n$ , il tempo è discreto e il periodo di incubazione è assente. Le caratteristiche fondamentali del modello sono:

- *durata dell'infezione*, la malattia dura un'unità di tempo nella quale l'individuo è contagioso;
- *probabilità di contagio*, l'infezione si propaga seguendo una distribuzione binomiale (per questo appartiene alla classe dei modelli a catena binomiale).

Quindi il numero di infettati può essere visto come il risultato dell'applicazione a catena di un modello binomiale. Formalizzando queste assunzioni:

$X(t)$  = numero di suscettibili al tempo  $t$

$Y(t)$  = numero di infettivi in  $t$

---

<sup>1</sup>H. Abbay. An examination of the Reed-Frost theory of epidemics, Hum. Biol. 24 (1952).

Si ha quindi <sup>2</sup>:

$$X(t) + \sum_{i=1}^t Y(i) = n.$$

Il numero di infettivi-infetti al tempo  $t$  dipende solo dal numero di suscettibili e di contagiosi al tempo precedente, per cui

$$W(t) = \{X(t-1), Y(t-1)\} \quad (1.7)$$

è una catena di Markov. Inoltre essendo un modello binomiale si ha che la probabilità di contagio diventa:

$$P(Y(t+1) = k | X(t) = x, Y(t) = y) = \binom{x}{k} p(y)^k (1 - p(y))^{x-k} \quad (1.8)$$

dove  $p(y)$  è la probabilità di essere contagiati essendoci  $y$  infettivi al tempo  $t$  ed è data da:

$$\begin{aligned} 1 - p(y) &= (1 - p)^y \\ p(y) &= 1 - (1 - p)^y \end{aligned}$$

dove  $p$  è la probabilità di contagio quando c'è una persona malata; rappresenta quindi la contagiosità della malattia.

Non viene considerato il numero di incontri che ogni individuo fa ogni giorno, ma solo il numero di individui contagiosi presente.

La malattia termina al tempo  $T$  in cui si raggiunge una delle seguenti condizioni:

- $X(T-1) = 0$ , tutte le persone sono già state malate e quindi nessuno al tempo  $T$  può infettarsi.
- $Y(T-1) = 0$ , non c'è nessun infettivo al tempo  $T-1$ , per cui nessuno può essere contagiato e quindi essere malato in  $T$

Sulla base di questo modello, ne sono stati sviluppati molti altri come ad esempio è il modello *Greenwood*, che risulta essere molto simile al Reed-Frost; l'unica differenza è data dalla probabilità di contagio, che non dipende dal numero di infettivi in un determinato tempo ma è costante pari a  $p$ .

---

<sup>2</sup>non vengono menzionati i rimossi che però, durando la malattia solo un'unità di tempo, in  $t$  sono  $\sum_{i=1}^{t-1} Y(i)$

### 1.2.1 Modello SIR di Tuckwell e Williams

Tuckwell e Williams in [1] hanno proposto un modello stocastico che sarà alla base di quelli considerati nel seguito dell'elaborato.

L'intento degli autori dell'articolo è fornire un modello basato sul Reed-Frost ma più realistico e duttile, basato su una struttura markoviana. Valgono le assunzioni di base del modello di riferimento, in più sono state inserite le seguenti novità:

- *definizione di malato*<sup>3</sup>, per ogni individuo  $i$ , con  $i = 1, \dots, n$ , è definito il processo casuale  $Y^i = \{Y^i(t), t = 0, 1, 2, \dots\}$  tale che  $Y^i(t) = 1$  se l'individuo è infetto in  $t$ , altrimenti  $Y^i(t) = 0$ . Il numero totale di infetti al tempo  $t$  sarà:

$$Y(t) = \sum_{i=1}^n Y^i(t);$$

- *numero di incontri giornalieri*, ogni individuo  $i$ , in ogni tempo  $t$ , incontra un numero di persone  $N_i(t) = N + M_i(t)$  dove  $N$  è la componente fissa (numero minimo di incontri), mentre  $M_i(t)$  è la componente casuale che può seguire diverse distribuzioni. La distribuzione di  $M_i(t)$  è la stessa per ogni  $t$ ;
- *durata dell'infezione*, ogni individuo contagiato rimane tale per  $R$  unità di tempo (nei modelli considerati sono giorni), compreso il tempo di infezione.  $R$  è quindi il periodo di malattia: assume un valore intero positivo costante e rappresenta la quantità di tempo in cui un individuo rimane nello stato I.  
Per cui se un individuo viene infettato al tempo  $t$ , rimarrà tale e sarà contagioso nei tempi  $[t, t+1, \dots, t+R-1]$ , dal tempo  $t+R$  l'individuo non sarà più infetto e passerà dallo stato di infettivo a quello di rimosso.

- dato  $Y(t)$ , la probabilità che un individuo, scelto a caso, sia malato al tempo  $t$  è data da:

$$f(t) = \frac{Y(t)}{n}.$$

e ipotizzando che tutte le variabili  $N_i(t)$  siano indipendenti e identicamente distribuiti, allora il modello può essere visto come una catena di Markov a  $(R+1)$  dimensioni (in quanto servono  $R + 1$  valori per descrivere un singolo stato). Infatti siano

- $Y_l(t)$ , con  $l=0,1,2,\dots,R-1$ , il numero di individui che al tempo  $t$  sono infetti esattamente da  $l$  tempi;

---

<sup>3</sup>una definizione simile verrà fatta anche per i suscettibili e i rimossi



- $X(t)$  il numero di individui suscettibili in  $t$ ;
- $Z(t)$  gli individui rimossi prima di  $t$ .

e si assuma che, chi è infetto in  $t = 0$  lo sia appena diventato, per cui

$$Y(0) = Y_0(0), \quad Z(0) = 0 \quad Y_l(0) = 0, \quad l = 1, 2, \dots, R - 1$$

per cui la popolazione allo stato iniziale è definita come

$$X(0) + Y(0) = n \tag{1.9}$$

Allora il modello

$$V(t) = (X(t), Y_0(t), Y_1(t), Y_2(t), \dots, Y_{R-1}(t)), \quad t = 0, 1, 2, \dots \tag{1.10}$$

è una catena di Markov a  $\binom{n+R}{n}$  stati.

Considerando  $(X(t), Y(t), Z(t))$ , si ottiene la descrizione del modello come SIR<sup>4</sup>.

In alternativa si può operare considerando la popolazione descritta individuo per individuo, la via che verrà qui implementata, utilizzando tre processi

$X^i(t), Y^i(t)$  e  $Z^i(t)$  con  $i = 1, 2, \dots, n$

tali che in ogni tempo si possa sapere in quale stato si trova ogni individuo. I tre processi hanno lo stesso comportamento; preso ad esempio quello che modella la condizione di suscettibile, si ha che  $X^i(t) = 1$  se l'individuo  $i$ -esimo è suscettibile al tempo  $t$  e vale 0 altrimenti.

Per ogni tempo, per ogni individuo, si ha:

$$X^i(t) + Y^i(t) + Z^i(t) = 1$$

quindi solo uno dei tre processi allo stesso tempo può valere uno. Si avrà quindi:

$$X(t) = \sum_{i=0}^n X^i(t) \quad Y(t) = \sum_{i=0}^n Y^i(t) \quad Z(t) = \sum_{i=0}^n Z^i(t)$$

Inoltre, anche se le variabili  $N_i(t)$  non fossero tutte identicamente distribuite, si possono considerare comunque i processi  $Y_0^i, Y_1^i, \dots, Y_{R-1}^i$ , con  $i = 1, 2, \dots, n$  in cui  $Y_k^i(t) = 1$  se l'individuo  $i$ -esimo in  $t$  è malato da  $k$  unità di tempo (è stato infettato  $t - k$  tempi prima), vale zero altrimenti. Da cui

$$Y^i(t) = \sum_{k=0}^{R-1} Y_k^i(t).$$

---

<sup>4</sup>Per ogni tempo  $t$ , conoscendo  $V(t)$ ,  $Z(t)$  è noto.

Il nuovo modello è quindi

$$\mathbf{W}(t) = [X^i(t), Y_0^i(t), Y_1^i(t), \dots, Y_{R-1}^i(t), \quad i = 1, 2, \dots, n] \quad (1.11)$$

in cui lo spazio degli stati è  $\{0, 1\}^{(R+1)n}$ . Questo tipo di modello risulta essere più facilmente implementabile e per questo è utilizzato anche quando le variabili  $N_i(t)$  sono identicamente distribuite.

Fissato un individuo  $i$  è interessante studiare

$$W_i(t) = [X^i(t), Y_0^i(t), Y_1^i(t), \dots, Y_{R-1}^i(t)]$$

che è un processo stocastico discreto a  $(R+2)$  stati. Ogni individuo ha una probabilità di contagiarsi  $\beta$ , avvenuto il contagio il decorso della malattia è deterministico, con probabilità pari a 1 si passa allo stato successivo fino a quando, trascorsi  $R$  giorni, l'individuo entra nello stato assorbente dei rimossi; le uniche probabilità di interesse sono, quindi, quella di ammalarsi,  $\beta$  e il suo complementare  $\alpha$ .

Per ogni individuo  $i$  si ha che la matrice che ne descrive l'andamento è

$$P_i = \begin{bmatrix} \alpha & \beta & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & 1 \end{bmatrix}$$

Questa non è una matrice di transizione, inoltre è stato inserito anche lo stato dei rimossi che, nella formulazione della catena di Markov non è presente, ma è necessario per capire l'andamento del singolo individuo. Inoltre sia per il calcolo di  $\alpha$  sia per quello di  $\beta$  servono informazioni riguardanti tutto il resto della popolazione, per poter saper quanti sono gli individui malati, e quindi contagiosi, e quanti sono i sani.

La probabilità di contagio, al tempo  $t$ , dipende da  $y$ , numero di infettati al tempo  $t - 1$ ,  $N_i(t)$  e dalla contagiosità della malattia  $p$ . La probabilità che siano  $j$  le persone infette tra le  $N_i(t)$  incontrate è

$$P_j^i(y, N_i(t); n) = \binom{N_i(t)}{j} \left(\frac{y}{n-1}\right)^j \left(1 - \frac{y}{n-1}\right)^{N_i(t)-j} \quad (1.12)$$

e data  $p$  si ricava  $p_j$ , (probabilità di contagiarsi incontrando  $j$  persone malate) come

$$p_j = 1 - (1 - p)^j \quad (1.13)$$

per cui considerando qualsiasi valore di  $j$ , la probabilità di contagiarsi è

$$\begin{aligned}
 P(Y_0^i(t+1) = 1 | N_i(t), X^i(t) = 1, Y(t) = y) = \alpha &= \sum_{j=1}^{N_i(t)} p_j P_j^i(y, N_i(t); n) \\
 &= 1 - \left(1 - \frac{py}{n-1}\right)^{N_i(t)}
 \end{aligned}
 \tag{1.14}$$

Infatti la realizzazione di  $N_i(t)$ , poichè gli incontri sono indipendenti, può essere vista come un insieme di Bernulli di parametro  $\frac{py}{n-1}$ .

La probabilità che l'individuo  $i$  non si infetti è data quindi dalla probabilità che rimanga sano dopo l'incontro con ogni persona ed è quindi  $(1 - \frac{py}{n-1})^{N_i(t)}$ .

## Osservazioni

Nell'articolo sono stati valutati l'andamento dell'estensione dell'epidemia e della durata, al variare di  $N$  numero di incontri giornaliero,  $p$ , probabilità di contagio, ed  $R$ , durata del periodo di infezione.

Dai risultati ottenuti è stato dedotto che, all'aumentare del numero di incontri giornaliero, cresce il numero di infetti, si raggiunge un picco di malati più alto e in un tempo inferiore, come si può osservare dalla Figura1.1. In particolare le differenze più rilevanti sono osservate per un numero di incontri compreso tra 3 e 7.

La Figura1.2 mostra come anche la durata dell'epidemia, quando  $R > 1$ , aumenti all'aumentare di  $N$ , ma solo fino a una soglia che dipende dalla lunghezza del periodo di infezione. Questo tipo di considerazioni può essere utile per valutare il tipo di intervento farmacologico volto a ridurre  $R$ .

Inoltre non sono emerse differenze considerevoli fra campioni grandi e piccoli, per quanto riguarda la frazione di popolazione che contrae la malattia (in media l'80%), cambia invece la durata che aumenta per campioni più grandi ma comunque in modo proporzionale.

Dalla Figura1.3 si nota inoltre che, all'aumentare di  $R$  (valori maggiori di 2) basta una probabilità dello 0.05 affinché si ammali tutta la popolazione. Quindi, secondo gli autori, per ridurre la diffusione della malattia è necessario che i farmaci riducano  $R$ , che risulta essere uno dei fattori di maggiore influenza sull'espandersi della malattia e sul quale è possibile intervenire.

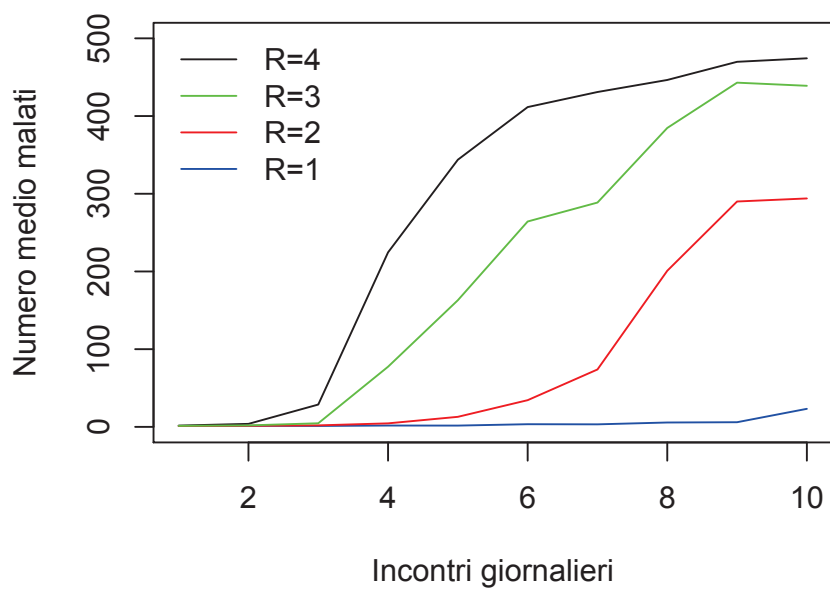


Figura 1.1: Andamento del numero medio di malati per epidemia al variare di N e R

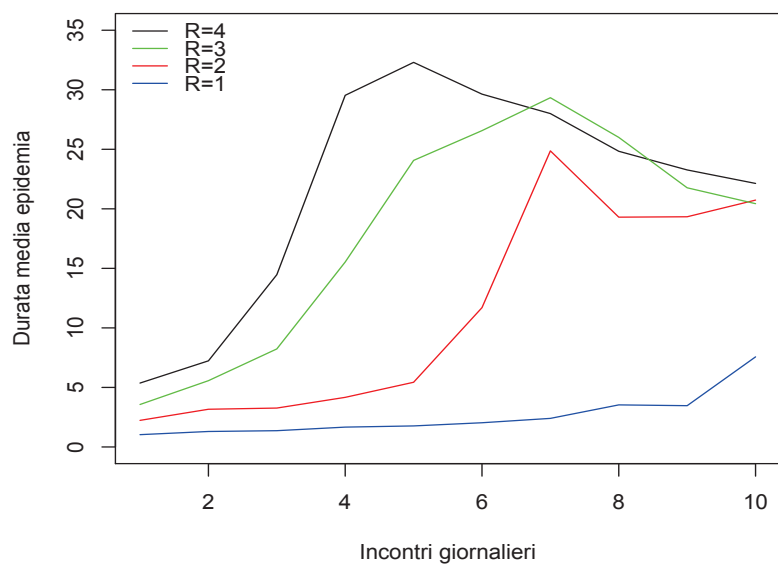


Figura 1.2: Andamento della durata dell'epidemia al variare di N e R

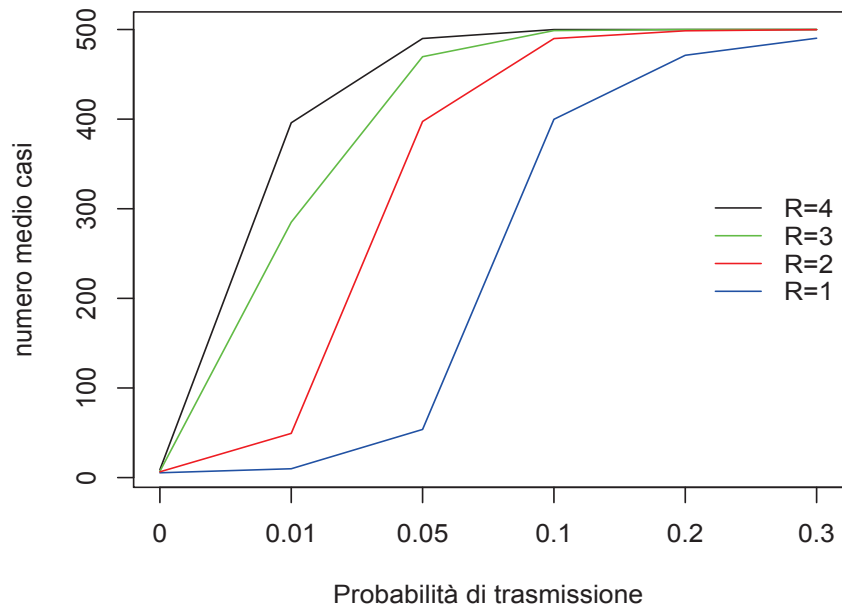


Figura 1.3: Andamento del numero medio di malati per epidemia al variare di  $p$  e  $R$



# Capitolo 2

## Estensioni del modello SIR

### Introduzione

In questo capitolo verranno esposti alcuni modelli, ottenuti modificando quello visto nel capitolo precedente. Per ognuno di questi è stato valutato l'andamento del numero medio di casi, la durata media dell'epidemia e come, entrambi, dipendano dalla probabilità di contagio, dal numero di incontri giornalieri e dalla durata dell'infezione. Come prima modifica è stato valutato un tempo di malattia  $R$  aleatorio, ottenendo quindi un modello molto simile a quello di partenza ma con un elemento di variabilità. È stato considerato quindi il modello SEIMR, in cui per indicare gli individui malati non è più considerato un unico stato, ma si distinguono tre stadi differenti del decorso dell'infezione: una prima fase in cui i soggetti sono infettati ma non contagiosi, un secondo periodo di contagiosità asintomatica e infine un terzo periodo in cui gli individui risultano essere contagiosi e manifestano i sintomi della malattia. Inoltre è stato introdotto un parametro  $\lambda$  per pesare le probabilità di incontro di individui appartenenti a stati diversi (persone che manifestano i sintomi sono presumibilmente più difficili da incontrare in quanto si assentano momentaneamente dalle loro attività). Il secondo modello valutato è il SEIMRS in cui si prevede la possibilità di riprendere la malattia, può essere, ad esempio, il caso dell'influenza o in generale di malattie non immunizzanti. In questo modello è stata introdotta la probabilità  $\gamma$  di tornare suscettibile che varia al variare delle patologie considerate. Più è alto il valore che assume più è raro che la malattia considerata conferisca l'immunità. Infine è stato valutato il modello SEIMS, dove non è contemplata la possibilità di immunità e che risulta adatto a modellare patologie che possono interessare più volte uno stesso individuo (ad esempio il raffreddore).

## 2.1 Modello con R aleatorio

Nella prima variante considerata è stato utilizzato il modello (1.11) ma con tempo di infezione aleatorio. Questo genere di assunzione risulta essere verosimile in quanto non tutti gli individui reagiscono allo stesso modo alle cure ed è possibile che abbiano tempi di guarigione diversi. La durata dell'infezione è stata modellata con una distribuzione *Gamma*  $(\mu, \sigma^2)$  con  $\mu \in \{1, 2, 3, 4\}$ , per ricondursi agli esempi fatti nell'articolo<sup>1</sup>, e  $\sigma^2 = 1$  eccetto nel caso  $\mu = 1$  per cui si ha  $\sigma^2 = 0.5$ . Fatta eccezione per la durata del periodo di infezione, nel nuovo modello valgono le assunzioni fatte in (1.11). Considerando R aleatorio, il modello non è più definito con precisione in quanto ogni individuo ha un comportamento assestante e quindi non generalizzabile. Si è proceduto allora solo tramite simulazione.

È stato valutato, quindi, il verificarsi o meno di cambiamenti nell'estensione e nella durata dell'epidemia. In Figura 2.2 si nota che l'epidemia tende a durare leggermente di più, soprattutto con  $\mu > 3$  e  $N < 6$ . Sotto la soglia dei 6 giorni, al crescere di  $\mu$ , aumenta la durata dell'epidemia; per valori maggiori di N le differenze nella durata si assottigliano, infatti indipendentemente da  $\mu$ , converge attorno ai 20-25 giorni. Caso a parte si ha con  $\mu = 1$ , per cui l'epidemia tende a durare molto poco. Alle stesse conclusioni si giunge osservando anche l'andamento della durata al variare di  $p$  e  $R$ , Figura 2.1. Rispetto al modello con R fisso, la durata dell'epidemia è in media maggiore; si raggiunge comunque il picco massimo con  $p = 0.05$ . Le differenze maggiori si hanno con  $R = 1$  e  $R = 2$ , per i quali si notano aumenti considerevoli nella durata dell'epidemia soprattutto con  $p < 0.2$ .

Osservando invece Figura 2.3 si nota che l'introduzione della variabilità nella durata dell'infezione porta a un leggero aumento nel numero di persone coinvolte.

Si può pensare quindi che, l'introduzione della variabilità nella durata dell'infezione, non comporti un aumento nel numero dei casi, bensì che questi vengano distribuiti su un periodo leggermente più lungo. Non essendo emerse particolari novità rispetto al modello di partenza, nei prossimi modelli R sarà considerato costante, lasciando inalterata la struttura di Catena di Markov.

---

<sup>1</sup>Quando in seguito si indicherà  $R = j$  si intenderà  $E(R) = j$



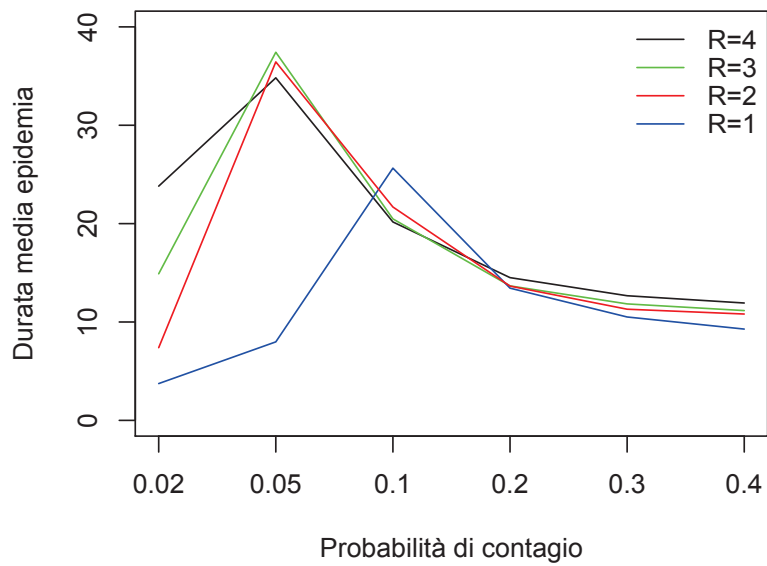


Figura 2.1: Andamento della durata dell'epidemia al variare di  $p$  e  $R$

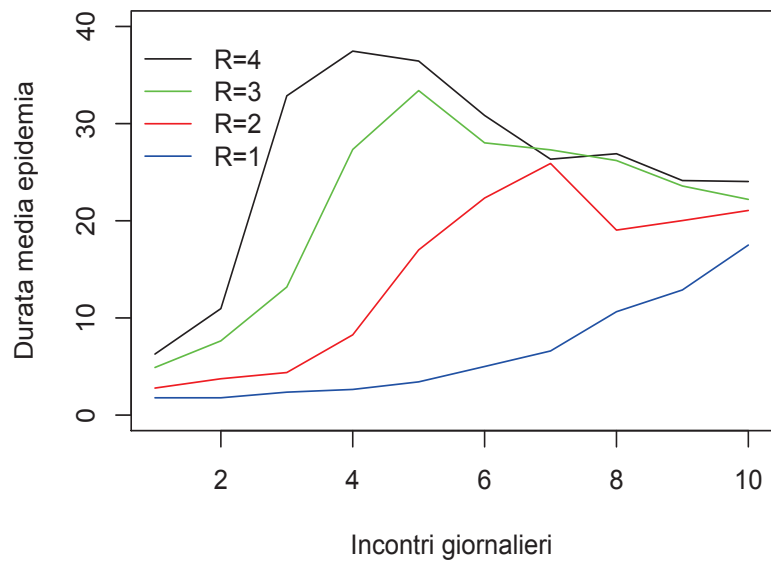


Figura 2.2: Andamento della durata dell'epidemia al variare di  $N$  e  $R$

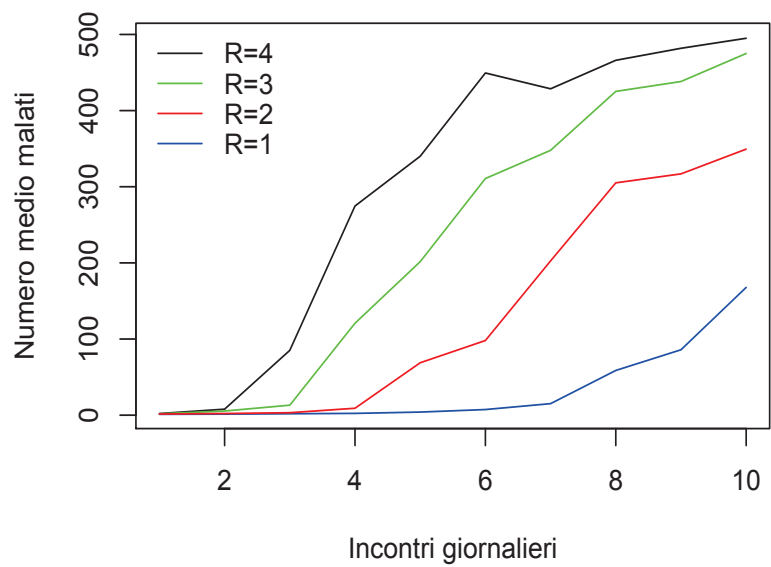


Figura 2.3: Andamento del numero medio di malati al variare di  $N$  e  $R$

## 2.2 Modello SEIMR

A partire dalla struttura del modello considerato in [1] è stato realizzato un nuovo modello strutturato in modo tale da considerare la presenza di un tempo di latenza e trattare in modo differente la condizione di malato e le probabilità di contagio date dall'incontro di individui appartenenti a stati differenti.

L'obiettivo è quindi ottenere uno strumento che si adatti a malattie più lunghe, con un decorso in parte più articolato e che possa cogliere meglio l'andamento di chi manifesta i sintomi della malattia. Le idee alla base del nuovo modello sono:

- la presenza di un tempo di latenza-incubazione: per qualche giorno dopo il contagio l'individuo oltre a non presentare i sintomi della malattia non è contagioso;
- la presenza di due fasi differenti nel periodo di contagiosità: la prima in cui i soggetti sono contagiosi asintomatici, la seconda in cui i soggetti sono contagiosi e manifestano i sintomi della malattia;
- la non equiprobabilità degli incontri: si suppone che gli individui che presentano i sintomi della malattia si assentino dalle proprie attività e si incontrino meno facilmente.

Si passa quindi da un modello *SIR* a un modello *SEIMR*, dove *S* e *R* sono gli stessi dei modelli citati in precedenza, mentre *I* è suddiviso in: *E*, stato che comprende gli individui contagiati ma non ancora nè contagiosi nè sintomatici, *I*, comprendente le persone contagiose asintomatiche ed infine *M*, in cui gli individui infetti sono contagiosi e manifestano i sintomi della malattia. Ogni individuo, una volta infettato, in modo deterministico passa dagli stati *EIM* nei quali rimane rispettivamente  $R_1$ ,  $R_2$ ,  $R_3$  giorni al termine dei quali diventa rimosso. La struttura del modello che sarà utilizzato è quella definita in (1.11) e si avrà quindi un modello markoviano a  $\{0, 1\}^{(R+1)n}$  stati, dove

$R = R_1 + R_2 + R_3$  e dove gli stati sono ordinati per rappresentare rispettivamente la condizione di sano, malato da un giorno, da due ecc, fino ad arrivare all'ultimo stato che indica la situazione di rimosso.

Data l'indipendenza dei soggetti, ancora una volta è possibile considerare la matrice che descrive l'andamento del singolo individuo che è

$$P_i = \begin{bmatrix} \alpha & \beta & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & 1 \end{bmatrix}$$

dove  $\alpha$  e  $\beta$  non sono più quelle in (1.14), bensì sono state ricalcolate considerando  $p_1, p_2, p_3$ , che sono rispettivamente le probabilità di incontrare una persona appartenente allo stato  $E, I, M$ . Al fine di ottenere  $p_3$  più bassa rispetto alle altre due, per poter considerare meno frequente l'incontro di una persona malata sintomatica, è stato introdotto un parametro  $\lambda \in (0, 1)$ . Per calcolare la probabilità di contagio al tempo  $t$ ,  $p_1$  può essere accorpata alla probabilità d'incontro di una persona sana, in quanto gli individui in  $I_1$  non sono contagiosi, si ha quindi:

$$\begin{aligned} p_3 &= \frac{\lambda y_3}{n-1} & (2.1) \\ p_2 &= \frac{y_2}{n-1-y_3} \left(1 - \frac{\lambda y_3}{n-1}\right) \\ p_s &= (1 - p_2 - p_3) = \left(1 - \frac{\lambda y_3}{n-1}\right) \left(\frac{n-1-y_2-y_3}{n-1-y_3}\right) \end{aligned}$$

con  $y_2$  e  $y_3$  il numero di persone rispettivamente in  $I$  ed  $M$  al tempo  $t$ . Si noti che se  $\lambda \rightarrow 1$  si ottengono le stesse probabilità del modello (1.11), mentre se  $\lambda \rightarrow 0$ ,  $p_3 = 0$  e si ha il caso limite nel quale non si incontra nessun malato sintomatico (può essere il caso di malattie particolarmente gravi che portano all'isolamento).

Siano  $j_{i2}, j_{i3}, j_{is} = N_i(t) - j_{i2} - j_{i3}$  rispettivamente il numero il persone, incontrate al tempo  $t$  dall'individuo  $i$ -esimo, appartenenti agli stati  $I, M$  e sane, la probabilità di incontrare questa proporzione di individui è

$$P_{j_{i2}, j_{i3}}^i(y_2, y_3, N_i(t); n) = \sum_{j_{i2}, j_{i3}=1}^{N_i(t)} \frac{N_i(t)!}{j_{i2}! j_{i3}! j_{is}!} p_2^{j_{i2}} p_3^{j_{i3}} p_s^{j_{is}} \quad (2.2)$$

Considerando l'uguaglianza (1.13) si può valutare inoltre la probabilità di contagio come

$$\begin{aligned} p_{j_{i2}+j_{i3}} &= 1 - (1-q)^{j_{i2}+j_{i3}} \\ &= 1 - ((1-q)^{j_{i2}} (1-q)^{j_{i3}}) \end{aligned} \quad (2.3)$$

Allora la probabilità di contagio  $\beta$  al tempo  $t + 1$  risulta essere

$$\begin{aligned}
\beta &= P(Y_0^i(t+1) = 1 | N_i(t) = N, X^i(t) = 1, Y_2(t) = y_2, Y_3(t) = y_3) \\
&= \sum_{j_{i2}, j_{i3}=1}^N p_{j_{i2}+j_{i3}} \frac{N!}{j_{i2}! j_{i3}! j_{is}!} p_2^{j_{i2}} p_3^{j_{i3}} p_s^{j_{is}} \\
&= \sum_{j_{i2}, j_{i3}=1}^N (1 - (1-q)^{j_{i2}} (1-q)^{j_{i3}}) \frac{N!}{j_{i2}! j_{i3}! j_{is}!} p_2^{j_{i2}} p_3^{j_{i3}} p_s^{j_{is}} \\
&= \sum_{j_{i2}, j_{i3}=1}^N \frac{N!}{j_{i2}! j_{i3}! j_{is}!} p_2^{j_{i2}} p_3^{j_{i3}} p_s^{j_{is}} - \sum_{j_{i2}, j_{i3}=1}^N \frac{N!}{j_{i2}! j_{i3}! j_{is}!} (p_2(1-q))^{j_{i2}} (p_3(1-q))^{j_{i3}} p_s^{j_{is}} \\
&= 1 - \sum_{j_{i2}, j_{i3}=1}^N \frac{N!}{j_{i2}! j_{i3}! j_{is}!} (p_2(1-q))^{j_{i2}} (p_3(1-q))^{j_{i3}} p_s^{j_{is}}
\end{aligned} \tag{2.4}$$

in cui si può riconoscere il polinomio di Newton per cui, sostituendo i valori ottenuti nella (2.1) si ottiene che la probabilità che l'individuo  $i$ -esimo si infetti al tempo  $t + 1$  è pari a

$$\begin{aligned}
\beta &= 1 - (p_2(1-q) + p_3(1-q) + (1-p_2-p_3))^{N(t)} \\
&= 1 - (1 - p_2q - p_3q)^{N(t)} \\
&= 1 - \left( 1 - \frac{y_2}{n-1} \frac{n-1-\lambda y_3}{n-1-y_3} q - \frac{\lambda y_3}{n-1} q \right)^{N(t)} \\
&= 1 - \left( 1 - \frac{q}{n-1} \left( \frac{y_2(n-1-\lambda y_3)}{n-1-y_3} + \lambda y_3 \right) \right)^{N(t)}
\end{aligned} \tag{2.5}$$

La probabilità di sfuggire al contagio sarà di conseguenza

$$\alpha = 1 - \beta = \left( 1 - \frac{q}{n-1} \left( \frac{y_2(n-1-\lambda y_3)}{n-1-y_3} + \lambda y_3 \right) \right)^{N(t)} \tag{2.6}$$

Si noti che assumendo diverse le probabilità di contagio derivanti dall'incontro di persone appartenenti a diversi stati, la formula (2.5) diventa

$$\beta = 1 - \left( 1 - \frac{1}{n-1} \left( \frac{y_2(n-1-\lambda y_3)}{n-1-y_3} q_{c2} + \lambda y_3 q_{c3} \right) \right)^{N(t)} \tag{2.7}$$

con  $q_{c2}$  e  $q_{c3}$  le probabilità di contagio entrando in contatto con un individuo rispettivamente delle classi I e M.

Utilizzando il modello descritto in precedenza, tramite simulazione, sono state fatte alcune analisi esplorative per valutare le differenze rispetto al modello di partenza. Inizialmente sono stati valutati l'andamento del numero medio di malati e la durata dell'epidemia al variare del numero di incontri giornalieri e della probabilità di contagio. Sono stati considerati valori simili a quelli utilizzati in [1] proprio per poter verificare le presenza o meno di differenze. Dai risultati ottenuti si nota che, all'aumentare di  $R$ , aumenta il numero medio di coinvolti nell'epidemia fino ad arrivare quasi all'intera popolazione quando  $R=6$ . Rispetto al modello descritto nell'articolo si nota però un ritardo, infatti l'andamento che in questo caso si ottiene con  $R=4$ , nell'articolo era stato ottenuto simile con  $R=3$ . Si nota infine che il numero di coinvolti nell'epidemia tende comunque ad essere leggermente inferiore rispetto al caso di partenza.

Anche la durata dell'epidemia dipende sia da  $R$  sia dal numero di incontri giornalieri. Infatti da Figura 2.4 si nota che la durata raggiunge una punta massima e poi converge attorno ai 40 giorni, indipendentemente da  $R$  che invece influisce su quando avviene il picco (al crescere di  $R$  aumenta il numero di incontri giornalieri necessari)

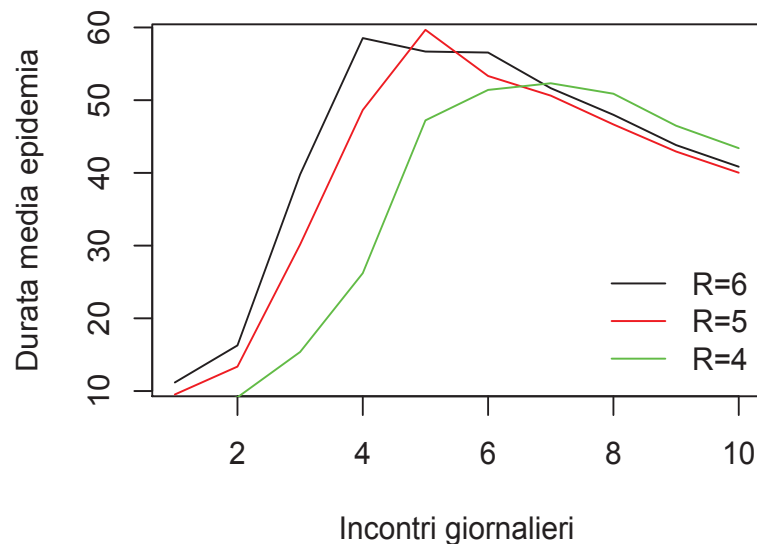


Figura 2.4: Andamento medio della durata dell'epidemia al variare di  $N$  e  $R$ .

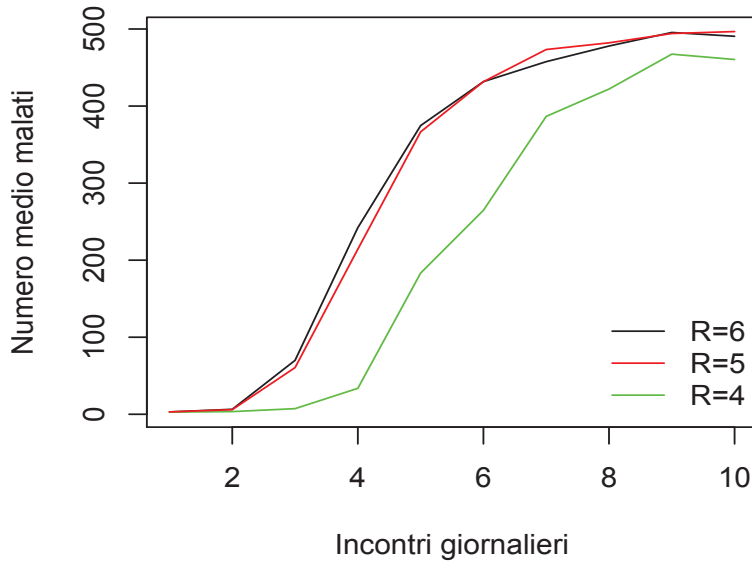


Figura 2.5: Andamento numero medio malati al variare di  $N$  e  $R$

Quando  $N < 6$  si riscontra un andamento differente al variare di  $R$ , mentre con un numero maggiore di incontri giornalieri la durata dell'epidemia tende a convergere al valore suddetto. La differenza principale rispetto al modello di partenza è che la convergenza avviene ad un valore maggiore (40 giorni invece di circa 20).

Anche valutando diverse probabilità di contagio si nota che, mediamente, la durata dell'epidemia è compresa tra i 35 e i 50 giorni. Siano  $q_{c2}$  e  $q_{c3}$  rispettivamente le probabilità di contagio entrando in contatto con persone delle classi  $I$  e  $M$ . I valori di  $q_{c2}$  utilizzati sono quelli indicati in Figura 2.6 mentre  $q_{c3} = q_{c2}/2$ . Fissato  $R$ , non emergono differenze rilevanti al variare di queste due probabilità; fissata una coppia  $(q_{c2}, q_{c3})$ , si nota invece che tendenzialmente si ha durata leggermente minore con un valore più basso di  $R$ .

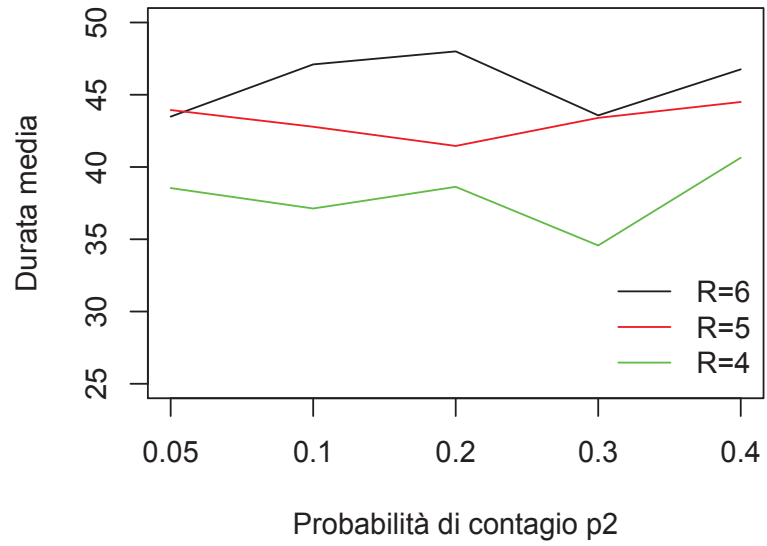


Figura 2.6: Andamento medio della durata dell'epidemia al variare di  $q_{c2}$ ,  $q_{c3}$  e R



Una delle novità inserite nel modello riguarda però la distinzione tra lo stato di contagioso asintomatico e quello sintomatico. È ipotizzabile, come già accennato in precedenza, che le persone che non manifestano i sintomi pur essendo contagiose siano più facilmente incontrabili rispetto a chi li manifesta. Questa distinzione è stata ottenuta introducendo il parametro  $\lambda$ ; è stato valutato quindi come cambia il numero medio di malati al variare di questo parametro.

Da Figura 2.7 si nota come, a pari condizioni di altri valori<sup>2</sup>, all'aumentare di  $\lambda$ , tenda ad aumentare il numero medio di contagiati, con una percentuale di popolazione infettata compresa tra il 73%, con  $\lambda = 0.01$  e il 93%, con  $\lambda = 0.9$ . Considerando invece condizioni simili<sup>3</sup> ma con  $q_{c2} > q_{c3}$ , sono state ottenute meno differenze al variare di  $\lambda$  come si può notare in Tabella 2.1.

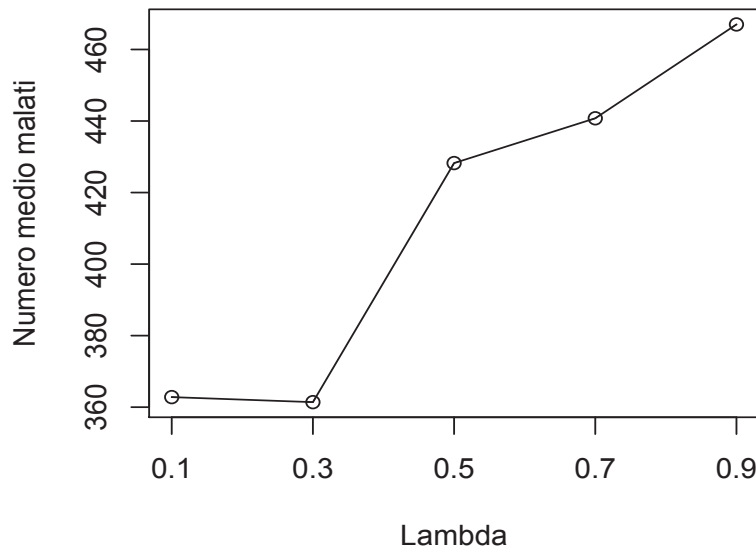


Figura 2.7: Andamento numero medio contagiati al variare di lambda

<sup>2</sup>Sono stati considerati  $q_{c2} = q_{c3} = 0.1$ ,  $R_2 = 3$ ,  $R_3 = 2$ .

<sup>3</sup> $q_{c2} = 0.1$ ,  $q_{c3} = 0.05$ ,  $R_2 = R_3 = 3$ , il periodo di infezione è leggermente più lungo ma compensa un valore di  $q_{c3}$  più basso. Sono state fatte queste scelte per rendere i due risultati confrontabili in termini di numero medio di casi

0.1	0.3	0.5	0.7	0.9
73,21	76.34	75.38	86.52	86.27

Tabella 2.1: Percentuale malati a fine epidemia al variare di  $\lambda$

Ipotizzando quindi una struttura di incontri più realistica, si nota come non venga infettata l'intera popolazione. Infatti, considerando valori medi di  $\lambda$  (tipo 0.4, 0.5), che non portano quindi a nessun caso limite, si nota che la percentuale di popolazione interessata dalla malattia oscilla tra il 75-80%. Per arrivare ad avere la quasi totalità della popolazione coinvolta bisogna considerare  $\lambda = 0.9$  e quindi andare a rientrare nel caso limite in cui gli incontri di persone appartenenti ai due stadi della malattia sono equiprobabili.

Per valutare gli effetti di  $\lambda$ , sono state considerate due situazioni differenti:

- $q_{c3} < q_{c2}$  e  $R3$  relativamente lungo: può essere il caso in cui una cura impiega più tempo ad agire, in questo periodo l'individuo è contagioso ma meno rispetto alla fase asintomatica;
- $q_{c3} = q_{c2}$  e  $R3$  molto corto: può essere il caso di malattie con cure che in breve tempo inibiscono la capacità di contagio dell'individuo.

Effettuando qualche prova è stato notato che il numero medio di coinvolti nell'epidemia sembra dipendere maggiormente dalla probabilità di contagio nella fase sintomatica piuttosto che dalla lunghezza di questa fase.

Infatti, considerando  $\lambda = 0.4$ , valore per cui nè si sconfinava in uno dei casi limiti nè si considerano gli incontri equiprobabili, con  $q_{c3} = 0.02$  e  $R3$  abbastanza lungo si ottengono i valori in Tabella 2.2

R3=3	R3=4	R3=5	R3=6
70.68	73.30	72.12	78.11

Tabella 2.2: Percentuale malati a fine epidemia al variare di  $R_3$

Considerando invece probabilità di contagio più alte ( $q_{c2} = q_{c3} = 0.1$ ) ma periodo di contagiosità sintomatica limitato, si ottengono percentuali di persone coinvolte nell'epidemia più alte, come da Tabella 2.3.

R3=1	R3=2	R3=3
77,61	80.80	85.61

Tabella 2.3: Percentuale malati a fine epidemia al variare di  $R_3$ , con  $q_{c2} = q_{c3}$

Confrontando i valori ottenuti si nota che, quando la probabilità di contagio nella fase sintomatica è bassa anche se il periodo di contagiosità è lungo, la percentuale di coinvolti nell'epidemia è minore rispetto al caso in cui  $R_3$  è breve ma la probabilità di contagio è maggiore. Quindi si può ritenere che risulti pesare maggiormente, sull'estensione della malattia, la probabilità di contagio nella fase sintomatica piuttosto che la durata di questo periodo.

Volendo ipotizzare possibili provvedimenti per limitare la diffusione dell'epidemia, potrebbe forse essere più utile un tentativo di riduzione della probabilità di contagio nella fase sintomatica piuttosto che rispetto a una riduzione di  $R_3$ .

## 2.3 Modello SEIMRS

Il modello descritto in precedenza può essere adattato a malattie che non necessariamente conferiscono l'immunità.

La struttura del modello rimane all'incirca la stessa; ogni individuo può ammalarsi con probabilità  $\beta$ . Nel caso in cui avvenga il contagio l'individuo trascorrerà rispettivamente  $R1$   $R2$   $R3$  giorni negli stati E, I, M. Finita la fase di contagiosità sintomatica ogni individuo può diventare immune con probabilità  $1 - \gamma$  o tornare suscettibile con probabilità  $\gamma$ .

Il modello di riferimento è sempre il (1.11) con  $\{0, 1\}^{(R+1)n}$  stati con  $R = R1 + R2 + R3$ . La matrice che descrive l'andamento dell' $i$ -esimo individuo è

$$P_i = \begin{bmatrix} \alpha & \beta & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \gamma & \dots & \dots & \dots & \dots & \dots & 0 & 1 - \gamma \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & 1 \end{bmatrix}$$

Le probabilità  $\alpha$  e  $\beta$  sono quelle definite in (2.5),  $\gamma$  dipende dalla malattia considerata. Come prima analisi esplorativa è stato valutato l'andamento del numero di coinvolti nell'epidemia al variare di  $\gamma$ .

Si è notato che, fissati  $R1, R2, R3, q_{c2}$  e  $q_{c3}$ , all'aumentare della probabilità di tornare suscettibili diminuisce il numero di immuni<sup>4</sup>.

0.05	0.1	0.3	0.5	0.6	0.7	0.8	0.9
89.4	86	78	63.9	56	46.9	33.6	20

Tabella 2.4: Percentuale di immuni al variare di  $\gamma$

Un elemento interessante emerso dalle simulazioni è che nonostante la presenza di uno stato assorbente (rimossi), non tutta la popolazione viene assorbita, bensì si suddivide tra rimossi e suscettibili. Questo avviene perchè l'andamento dell'epidemia è descrivibile con una catena di Markov nel momento in cui sia presente almeno una persona contagiosa; quando questo non avviene l'epidemia termina e, in base al valore di  $\gamma$ , gli  $n$  individui considerati si dividono tra sani e rimossi.

<sup>4</sup>I valori utilizzati sono  $q_{c2} = q_{c3} = 0.1$  e  $R1=1, R2=4$  e  $R3=2$

È stato osservato quindi che la percentuale media di immuni-suscettibili subisce notevoli variazioni al variare di  $\gamma$  ma anche nei casi limite non si arriva mai ad avere tutta la popolazione in un unico stato. Infatti anche quando la probabilità di tornare suscettibili è molto bassa una parte, seppur piccola, della popolazione non viene assorbita nella classe dei rimossi.

0.05	0.1	0.3	0.5	0.6	0.7	0.8	0.9
7.6	9.6	19.8	32.8	41.6	50.8	64.4	79.4

Tabella 2.5: Percentuale di suscettibili a fine epidemia al variare di  $\gamma$

0.05	0.1	0.3	0.5	0.6	0.7	0.8	0.9
725	782	844	989	1112	1244	1356	1527

Tabella 2.6: Numero medio di casi a fine epidemia al variare di  $\gamma$  e con  $n=800$

La durata dell'epidemia non sembra essere molto influenzata dal valore di  $\gamma$ , si assesta attorno ai 55-60 giorni. Gli unici casi in cui la durata risulta essere maggiore, sono quando  $\gamma$  assume valori elevati e quindi molto frequentemente gli individui, una volta finita la malattia, tornano ad essere suscettibili e il bacino di persone che può ammalarsi è molto ampio.

Osservando l'andamento dei malati si nota inoltre che all'aumentare di  $\gamma$  aumentano i casi di contagio ma il picco massimo di casi avviene all'incirca nello stesso periodo. L'andamento delle infezioni non cambia molto al variare di  $\gamma$ , anche se quando questa assume valori molto alti emergono due picchi di uguale entità a breve distanza l'uno dall'altro. Quindi il fatto che una malattia conferisca l'immunità o meno influisce molto sul numero di casi e solo in parte sulla durata e sulle dinamiche di diffusione. Da figura Figura2.9 si nota inoltre che, all'aumentare della probabilità di contagio, il numero di infettati aumenta notevolmente (si quintupla passando da 344 a 2248). Facendo variare R2 e R3 si nota che, come nei modelli precedenti, il numero di infezioni aumenta all'aumentare della durata del periodo di infezione. Infine è stata valutata la dipendenza del numero di casi e della durata dal numero di incontri giornalieri. Fissata  $\gamma = 0.5$  e utilizzati i valori precedenti si ottengono i valori in Tabella2.8 Si nota che all'aumentare di N i casi aumentano in modo consistente. Da Tabella2.7 si nota che anche la durata dipende da N, infatti all'aumentare del numero di incontri aumenta la durata.

1	2	3	4	5	6
11.32	28.88	53.76	60.72	65.02	67.40

Tabella 2.7: Durata medie dell'epidemia al variare di N

1	2	3	4	5	6
1.34	17.04	454,10	790.96	1021.98	1046.34

Tabella 2.8: Numero medio di casi al variare di N

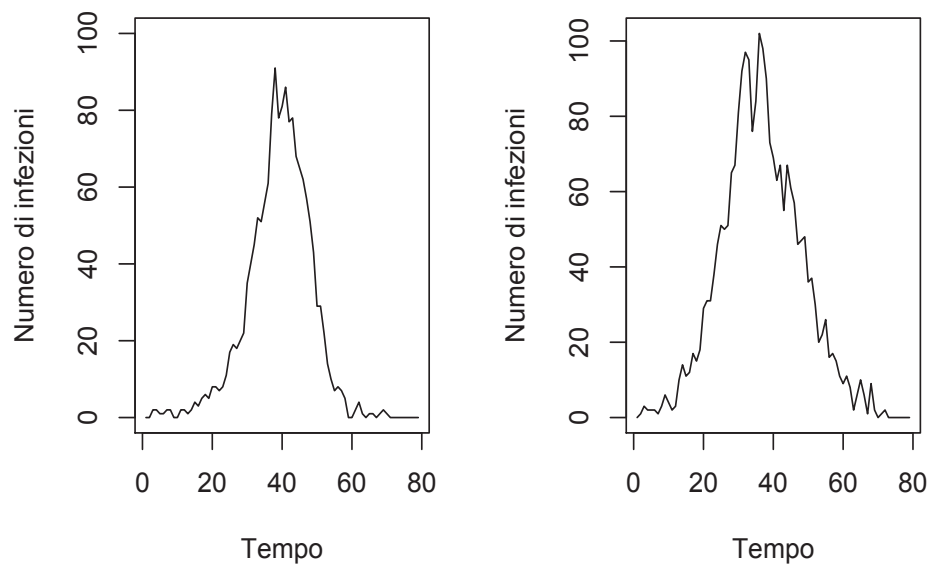


Figura 2.8: Andamento numero medio di infezioni con  $\gamma = 0.05$  e  $\gamma = 0.9$

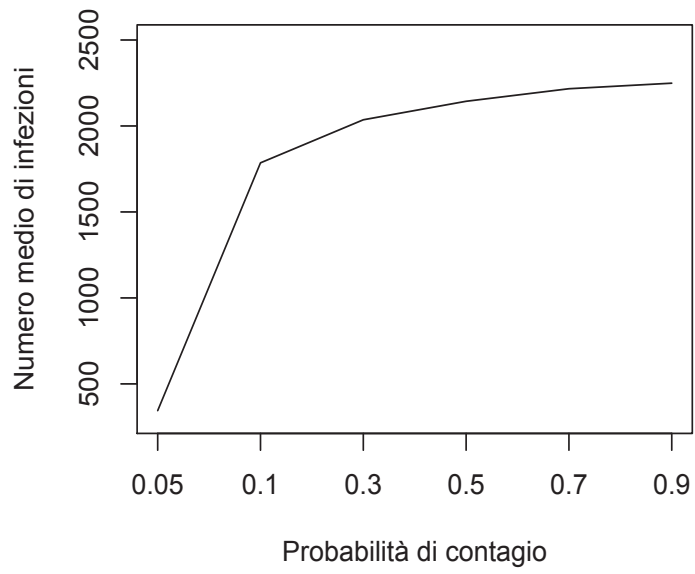


Figura 2.9: Andamento numero medio di infezioni al variare di  $q$

## 2.4 Modello SEIMS

In quest'ultima variante sono state considerate quelle patologie che non conferiscono alcun tipo di immunità per cui, una volta esaurito il periodo di malattia, ogni individuo torna ad essere suscettibile. Può essere il caso di malattie non gravi, come il raffreddore, che possono interessare una stessa persona molte volte. Anche in questo caso la struttura del modello rimane all'incirca la stessa. Ogni individuo può ammalarsi con probabilità  $\beta$ . Nel caso in cui avvenga il contagio l'individuo trascorrerà rispettivamente  $R1$   $R2$   $R3$  giorni negli stati E, I, M. A differenza di quanto succedeva nei modelli precedenti, alla fine del periodo di contagiosità sintomatica ogni individuo torna suscettibile; i periodi  $R1$  e  $R2$  sono ipotizzabili relativamente brevi, basti pensare proprio al raffreddore che ha un periodo di incubazione molto breve e quasi subito si manifestano i sintomi. Anche in questo caso il modello di riferimento è (1.11) e la matrice che descrive l'andamento dell' $i$ -esimo individuo è:

$$P_i = \begin{bmatrix} \alpha & \beta & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & 1 \\ 1 & \dots & \dots & \dots & \dots & \dots & 0 & 0 \end{bmatrix}$$

con  $\alpha$  e  $\beta$  le probabilità definite in (2.5).

Fissati gli altri valori, è stato valutato l'andamento del numero medio di casi, dei suscettibili e della durata dell'epidemia al variare della probabilità di contagio nella fase sintomatica<sup>5</sup>. Dalle prime esplorazioni si è osservato subito che, a causa dell'assenza di immunità, i casi di infezione aumentano in modo considerevole rispetto alle tipologie di malattie viste in precedenza e anche la durata dell'epidemia è molto più alta. Per come è strutturata la malattia, se in ogni tempo ci fosse sempre un nuovo contagiato, l'epidemia non finirebbe mai. Si può notare che se la  $q_{c3} < 0.5$ , l'epidemia tende a finire, infatti raggiunto un certo tempo  $t$  non si verificano più contagi e tutti tornano ad essere suscettibili. Quando la probabilità di contagio nella fase sintomatica supera il 50%, l'epidemia non si esaurisce, bensì il numero di sani tende a convergere ad un valore che dipende da  $q_{c3}$ , quindi costantemente ci sono nuovi casi che mantengono attiva l'epidemia, come da Tabella2.9.

<sup>5</sup>Sono stati considerati  $R1 = R2 = 1$ ,  $q_{c2} = 0.05$ ,  $R3 = 4$  e  $n = 800$



0.1	0.3	0.5	0.7	0.8	0.9
100	100	82.5	67	61.7	58.4

Tabella 2.9: Percentuale di suscettibili dopo 150 giorni dall'inizio dell'epidemia al variare di  $q_{c3}$

0.1	0.3	0.5	0.7	0.8	0.9
695	3022	7833	11835	13370	14734

Tabella 2.10: Numero di casi dopo 150 giorni dall'inizio dell'epidemia

Il numero medio di casi aumenta notevolmente all'aumentare della probabilità di contagio, fino ad arrivare a valori molto alti. Al di là della probabilità di contagio, l'andamento della malattia è comunque molto simile: c'è un unico picco considerevole di malati, raggiunto all'incirca allo stesso tempo, dopo il quale si ha la situazione descritta in precedenza. Come già accennato, con  $q_{c3} < 0.5$  l'epidemia finisce e il numero di nuove infezioni è pari a zero; negli altri casi il numero di nuovi contagiati oscilla in un intervallo, abbastanza ristretto, comprendente valori maggiori al crescere della probabilità di contagio, come da Figura 2.10.

Anche la durata del periodo di contagiosità sintomatica influisce sul numero di casi. Infatti si nota che all'aumentare di  $R_3$  aumentano i casi di contagio. Le differenze però, come si può vedere in Tabella 2.11, sono meno significative rispetto a quelle dovute a  $q_{c3}$

R3=2	R3=3	R3=4	R3=5
2282	2643	2942	3288
6877	7629	7763	7829
11498	11421	11816	11856
14486	14711	14756	14776

Tabella 2.11: Numero di casi dopo 150 giorni dall'inizio dell'epidemia al variare di  $R_3$  e con  $q_{c3} \in \{0.3, 0.5, 0.7, 0.9\}$

Dai risultati ottenuti si può ritenere che il modello sia adatto a descrivere patologie con grande diffusione, con un periodo di infezione relativamente breve e diverso tipo di contagiosità.

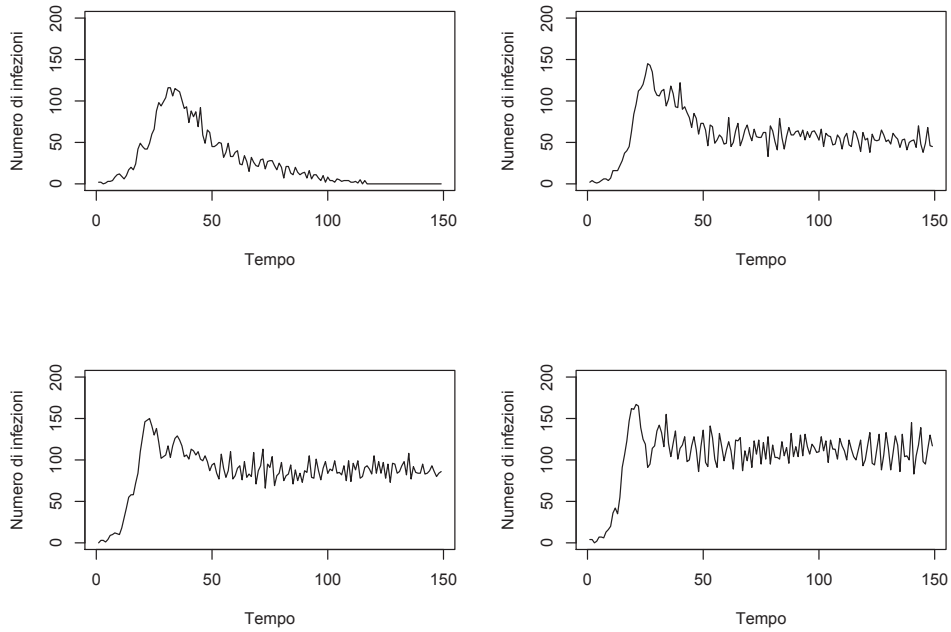


Figura 2.10: Andamento infezioni con:  $q_{c3} = 0.3, 0.5, 0.7, 0.9$

# Capitolo 3

## Il modello SIR a popolazione aggregata

### 3.1 Introduzione

Tuckwell e Williams in [1] propongono e utilizzano il modello, descritto nel Capitolo 1, in cui la popolazione è considerata individuo per individuo. Questa scelta è dettata da una maggiore semplicità e gestibilità delle matrici di transizione. A riscontro di un guadagno in termini di minore difficoltà di implementazione si ha un aumento notevole dello spazio degli stati della catena che però, simulando per individuo, non comporta particolari problemi a livello di implementazione.

Gli autori nell'articolo citano la possibilità di considerare un modello in cui la popolazione, in ogni tempo, venga considerata aggregata ma questa via non viene approfondita anche perchè risulta essere molto più impegnativa a livello computazionale e non solo.

In questo capitolo si cercherà di percorrere questa direzione. Inizialmente verrà implementato un modello a popolazione aggregata per gestire una malattia con durata di infezione di un giorno. Verranno stimati i tempi medi di assorbimento, che in realtà altro non sono che la durata dell'epidemia considerando diversi stati di partenza. Verranno valutati, inoltre, la durata massima dell'epidemia, il momento in cui si verifica, a quale stato di partenza è associata e i legami con la numerosità della popolazione.

Infine questo modello verrà allargato per gestire una malattia con periodo di infezione di due giorni. Anche in questo caso verrà stimata la durata e ne verranno valutati i picchi massimi al variare della probabilità di contagio e della numerosità della popolazione. Per entrambi i modelli ci sono notevoli restrizioni per quanto riguarda quest'ultimo aspetto infatti, per problemi

computazionali, non è stato possibile considerare una popolazione molto elevata. La dimensione massima della popolazione che è stata presa in considerazione è, nel primo caso  $n = 100$  mentre nel secondo non è stato possibile superare le 50 unità.

## 3.2 Struttura del modello e calcolo della durata media

Tutti i modelli visti in precedenza e le simulazioni sono basati su una struttura per individuo. Una via alternativa può essere data dall'utilizzo di una apparentemente semplice catena di Markov  $V(t)$  in cui il numero di stati è

$$\binom{n + R + 1}{n}$$

che risulta essere decisamente inferiore rispetto a quello del modello per individui in cui se ne hanno

$$2^{(R+1)n}$$

Gli stati di  $V(t)$  sono vettori di  $k$  elementi (con  $k = R + 1$ ): il primo indica quante persone sono suscettibili, il secondo quante persone sono malate da un giorno ecc...

Per cui la struttura del singolo stato è:

$$S = (X(t), Y_0(t), \dots, Y_R(t)) \quad (3.1)$$

dove  $X(t)$  indica il numero di persone che al tempo  $t$  sono suscettibili,  $Y_0(t)$  le persone che sono state infettate in  $t$  e in generale  $Y_l(t)$  indica il numero di individui che in  $t$  sono infettati da  $l$  unità di tempo. Per il conteggio delle persone presenti in ogni classe si sfrutta quindi una struttura simile a quella del modello per individuo, ma che in questo caso in ogni tempo si valuta complessivamente il totale delle persone presenti.

La probabilità che l'individuo  $i$  non si ammali incontrando  $N_i(t)$  è data da

$$\alpha = 1 - \left(1 - \frac{py}{n-1}\right)^{N_i(t)} \quad (3.2)$$

Il modello risulta però molto più difficile da gestire, infatti la matrice di transizione di questo processo di Markov, che bisognerebbe utilizzare in sede di simulazione, risulta avere dimensioni considerevoli. Basti pensare che con

solamente  $n = 4$  e considerando un solo giorno di malattia la catena avrebbe 15 stati. È evidente che quindi all'aumentare della popolazione e della durata del tempo di infezione lo spazio degli stati diventi poco gestibile e impegnativo a livello di simulazione (già con una popolazione di 20 individui e una malattia di 3 giorni si ottengono 4845 stati).

La matrice  $P$  ha il numero di stati descritto in precedenza e preso un generico stato  $(i_1, i_0, i_1, \dots, i_{k-1})$  le possibili transizioni sono:

- $(i_1, i_2, \dots, i_k) \rightarrow (i_1 - j, j, i_2, \dots, i_{k-1}), \text{con } 0 \leq j \leq i_1, \text{max}(i_2, \dots, i_{k-1} > 0)$
- $(i_1, 0, 0, \dots, 0) \rightarrow (i_1, 0, 0, \dots, 0) \quad \text{per } 0 \leq i \leq n$

Il primo caso raggruppa tutte le possibili transizioni in cui avviene un determinato numero di contagi ( $j$ ) e in cui gli individui già infettati scalano in avanti di un giorno; le probabilità di transizione sono determinate da (3.2). Il secondo descrive la situazione che si verifica nel momento in cui ci sono ancora dei suscettibili ma non c'è più nessun malato e quindi nessuno che possa contagiare. La catena pertanto entra in uno stato assorbente, per cui rimane dov'è con probabilità pari ad 1. Se anche  $i_1 = 0$ , allora non vi sono più suscettibili, per tanto tutta la popolazione considerata è stata coinvolta nell'epidemia; si ha quindi lo stato assorbente di fine epidemia. Il vantaggio di questo tipo di modellizzazione è dato dal fatto che, avendo tutta la popolazione aggregata, si può provare a valutare l'andamento di  $P^n$ , capire quindi come tenda a distribuirsi la popolazione con il passare del tempo e valutare il tempo medio di assorbimento, ossia la durata dell'epidemia al variare dello stato di partenza.

Siano  $m$  il numero di stati,  $T = \{1, 2, 3, \dots, s\}$  l'insieme degli stati transienti,  $R = \{s + 1, s + 2, \dots, m\}$  l'insieme degli stati ricorrenti e  $Q$  la matrice  $s \times s$  estratta da  $P$ , allora  $k = (k_1, k_2, \dots, k_s)$  vettore dei tempi di assorbimento in  $R$  è ottenuto come

$$k = (I - Q)^{-1}(1, 1, \dots, 1)^T \quad (3.3)$$

È stato implementato il modello per gestire una malattia della durata di un giorno (dopo il quale l'individuo rientra nella categoria dei rimossi) e con la probabilità di contagio definita precedentemente. È stato testato inizialmente su una piccola popolazione,  $n = 5$ , con  $p = 0.1$  e con un solo incontro al giorno. Siano  $X(t), Y(t)$  rispettivamente il numero di sani e malati al tempo  $t$ , gli stati ammissibili sono allora 21 coppie di valori.

Gli stati assorbenti sono quelli con zero malati, per cui non sono presenti contagiosi che possano infettare i suscettibili oppure tutti sono stati contagiati per cui non si hanno né malati né suscettibili. A partire dalla matrice

$P$ , che non è stata riportata in quanto composta quasi tutta da zeri, sono stati calcolati i tempi medi di assorbimento e sono stati confrontati con la durata media dell'epidemia ottenuta con il modello (1.11) con stessi valori di  $R, p$  e  $n$ .

Applicando la (3.3) ai dati sono stati ottenuti i tempi medi di assorbimento in Tabella3.1, ai quali sono stati accostati quelli ottenuti utilizzando il modello per individuo.

Confrontandoli si nota che sono molto simili, il che significa che i due modelli portano alle stesse considerazioni. I tempi medi ottenuti portano a concludere che, con questi particolari  $n, R, p$ , l'epidemia termina molto velocemente. È stato valutato quindi come varia il tempo medio di assorbimento al variare di  $p$  ottenendo i risultati in Tabella3.2. Si nota che, all'aumentare della probabilità di contagio, aumenta il tempo medio di assorbimento nelle classi ricorrenti e quindi anche la durata dell'epidemia. Questo risultato risulta più chiaro valutando anche il numero di casi. Infatti all'aumentare di  $p$ , cresce il numero di contagiati, si allunga il periodo di tempo in cui sono presenti persone malate, con conseguente aumento della durata dell'epidemia.

L'incremento di  $k$  all'aumentare di  $p$  non è costante per tutti gli stati; l'aumento massimo si ha quando si considera (4,1) come stato di partenza, ed è pari al 50%, mentre l'incremento minimo si ha partendo dallo stato (1,1) dove si raggiunge solo l'11,3%.

*Osservazione:* Nelle tabelle riferite all'esempio con  $n = 5$  sono stati riportati solo una parte degli stati. Infatti ci sono altri ipotetici stati transienti, della tipologia  $(0,i)$  con  $i \geq n$ , che non sono stati riportati perchè d'interesse irrilevante, in quanto da questi non si sviluppa l'epidemia e quindi in un tempo unitario portano il processo allo stato  $(0,0)$ .

stato di partenza	$k$	$k^*$
4,1	1.08	1.09
3,2	1.11	1.14
3,1	1.06	1.08
2,3	1.10	1.11
2,2	1.07	1.09
2,1	1.04	1.06
1,4	1.08	1.09
1,3	1.06	1.06
1,2	1.04	1.04
1,1	1.02	1.03

Tabella 3.1: Durata epidemia con modello a popolazione aggregata e per individuo

stato di partenza	$k_{0.3}$	$k_{0.5}$	$k_{0.7}$	$k_{0.9}$
4,1	1.26	1.45	1.67	1.89
3,2	1.35	1.58	1.80	1.99
3,1	1.19	1.32	1.46	1.61
2,3	1.35	1.55	1.73	1.87
2,2	1.23	1.39	1.53	1.67
2,1	1.12	1.20	1.29	1.38
1,4	1.24	1.40	1.56	1.72
1,3	1.18	1.30	1.42	1.54
1,2	1.12	1.20	1.28	1.36
1,1	1.06	1.10	1.14	1.18

Tabella 3.2: Tempi medi di assorbimento al variare della probabilità di contagio

n	$k_{0.1}$	$k_{0.3}$	$k_{0.5}$	$k_{0.7}$	$k_{0.9}$
10	1.23	1.63	1.98	2.31	2.63
15	1.33	1.84	2.26	2.67	3.08
25	1.50	2.11	2.63	3.18	3.76
50	1.78	2.49	3.19	3.98	4.89
70	1.91	2.68	3.48	4.41	5.53
100	2.04	2.89	3.81	4.90	6.28

Tabella 3.3: Massima durata media dell'epidemia al variare di  $n$  e  $p$

È stato valutato anche l'andamento della durata massima dell'epidemia, elemento molto utile soprattutto quando la popolazione è elevata. Questo valore, però, non si ottiene sempre con gli stessi stati di partenza: spesso lo stato che la determina è  $(\lfloor n/2 \rfloor + 1, \lfloor n/2 \rfloor - 1)$ , come nell'esempio citato in precedenza. Infatti da Tabella3.1 si nota che la durata massima è stata raggiunta proprio in  $(3,2)$ . È stata valutata quindi  $P^n$  in questo stato per capire come tende a suddividersi la popolazione quando viene raggiunto lo stato di convergenza.

Si nota che con una probabilità di oltre l'88% l'epidemia termina senza altre infezioni (assorbimento nello stato  $(3,0)$ ), invece con il 10% della probabilità si infetta ancora un individuo e l'epidemia termina con 3 casi (assorbimento in  $(2,0)$ ). È stato valutato anche lo stato  $(2,3)$  in quando porta a una durata dell'epidemia molto simile a quella che si ottiene con  $(3,2)$ . Anche in questo caso la probabilità che non si verificano nuovi casi e che l'epidemia si concluda con i due malati iniziali, è molto alta (circa l'88%) mentre la probabilità che si ammali un terzo individuo è circa l'11%.

Il metodo implementato permette, pur con problemi computazionali, di valutare la durata massima dell'epidemia anche per una popolazione numericamente abbastanza elevata. Da Tabella3.3 si nota che, sia all'aumentare della numerosità della popolazione, sia al crescere dell'infettività della malattia, aumenta la durata massima raggiunta.

La probabilità di contagio, però, ha maggior peso, rispetto a  $n$ , sull'incremento della durata. Tuttavia, per valori alti di  $p$ , anche  $n$  pesa maggiormente. L'entità dell'aumento non è quindi costante, ma cresce al crescere della numerosità della popolazione. Valutando infatti  $p = 0.1$  e  $p = 0.9$  si passa da un incremento del 113,8% con  $n = 10$ , al 208,8 % con  $n = 100$

Il modello è stato ampliato per valutare una malattia della lunghezza di due giorni durante i quali gli individui infetti hanno uguale probabilità di contagiare. Il numero di stati della catena è molto elevato già con una popolazione abbastanza ridotta e questo ne limita le applicazioni. I tempi medi tendono



ad essere superiori, infatti considerando lo stesso caso utilizzato in precedenza<sup>1</sup> e valutando stati di partenza simili, si nota che l'epidemia ha maggiore durata. Ad esempio partendo dallo stato (3,1), e considerando il caso con un periodo di malattia di un solo giorno, il tempo di assorbimento era 1.06 (come da Tabella3.1), in questo caso partendo dallo stato (3,1,0) sono necessari 2,23 giorni, come da Tabella3.5. Anche in questo caso è stata valutata  $P^n$  per lo stato che porta alla durata maggiore dell'epidemia, ossia (3,2,0), e per quello che porta ad una durata molto simile, ossia (2,3,0). Partendo dai due stati, l'andamento ottenuto è molto simile. Infatti, considerando (3,2,0) come stato di partenza c'è il 73% della probabilità che l'epidemia termini con nessun altro nuovo infettato (assorbimento in (3,0,0)), il 21% che si infetti un'altra persona (assorbimento in (2,0,0)) e il 5% che si ammaliano tutti. Partendo invece da (2,3,0) con una probabilità del 73% non si ammala nessun altro, con il 23% della probabilità si ammala un altro individuo e con il 3,3% l'epidemia si conclude perchè si ammaliano tutti.

A differenza del caso della malattia con durata giornaliera, ci sono probabilità, seppur non elevate, che anche con un tasso di contagiosità della malattia abbastanza basso, si ammaliano tutti. È stato valutato inoltre il tempo massimo di assorbimento al variare della numerosità della popolazione e della probabilità di contagio<sup>2</sup> e si è notato che all'aumentare di  $n$  aumenta il tempo massimo di assorbimento raggiunto.

---

<sup>1</sup> $n = 5, p = 0.1$  e un solo incontro al giorno

<sup>2</sup>Non è stato però possibile utilizzare  $n > 30$  per problemi di calcolo

All'incremento della probabilità di contagio i tempi aumentano notevolmente e, indipendentemente da  $n$ , con  $p = 0.9$  il tempo ottenuto è circa 2.5 volte quello che si ha con  $p = 0.1$ ; anche alla crescita di  $n$  corrisponde una crescita della durata massima raggiunta.

Confrontando dove è possibile i risultati dei due modelli si nota che, a parità di numerosità della popolazione, quando la durata dell'infezione si prolunga per più tempo, l'epidemia raggiunge picchi di durata maggiori, in entrambi i casi la dipendenza dalla probabilità di contagio è molto più consistente rispetto alla dipendenza da  $n$ .

$n$	$k_{0.1}$	$k_{0.3}$	$k_{0.5}$	$k_{0.7}$	$k_{0.9}$
10	2,74	3.86	4.85	5.65	6.13
15	2.99	4.34	5.63	6.75	7.25
25	3,36	5.02	6.79	8.39	9.21
28	3.50	5.25	7.50	9.17	10.02

Tabella 3.4: Massimo tempo medio di assorbimento al variare di  $n$  e  $p$

stato di partenza	$k$
4,1,0	2,32
4,0,1	1.21
3,2,0	2.44
3,1,1	2.30
3,1,0	2.23
3,0,2	1.30
3,0,1	1.15
2,3,0	2.42
2,2,1	2.33
2,2,0	2.29
2,1,2	2.24
2,1,1	2.19
2,1,0	2.15
2,0,3	1.29
2,0,2	1.20
2,0,1	1.10
1,4,0	2.28
1,3,1	2.23
1,3,0	2.21
1,2,2	2.19
1,2,1	2.16
1,2,0	2.14
1,1,3	2.14
1,1,2	2.12
1,1,1	2.09
1,1,0	2.07
1,0,4	1.20
1,0,3	1.15
1,0,2	1.10
1,0,1	1.05

Tabella 3.5: Durata epidemia



# Capitolo 4

## Applicazioni

Sulla base dei modelli esposti in precedenza sono state effettuate alcune applicazioni a patologie note. Sono state considerate due diverse malattie ad ognuna delle quali è stato applicato il modello che meglio poteva rappresentarne le peculiarità. Le due malattie considerate sono la varicella e la scarlattina e per entrambe è stato considerato un modello SEIMR, in quanto malattie immunizzanti. In entrambi i casi è stato valutato l'andamento del numero medio di contagiati, dei malati sintomatici e la durata dell'epidemia, al variare del numero medio di incontri e nel caso della scarlattina anche al variare della probabilità di contagio.

### 4.1 Varicella

La varicella è una malattia infettiva esantematica e si manifesta con un'eruzione cutanea di vescicole e bolle.

È causata dal virus varicella-zoster, un virus a DNA della famiglia degli Herpes virus. Il contagio avviene per via aerea mediante goccioline di saliva nell'aria o tramite contatto diretto con le lesioni da varicella o zoster. Nonostante sia una malattia immunizzante, il virus rimane latente nell'individuo per tutta la vita e può risvegliarsi, anche dopo decenni, dando luogo all'Herpes Zoster.

Un individuo che viene contagiato mantiene in corpo il virus latente per 15-20 giorni durante i quali non è né contagioso né manifesta i sintomi (E). Questa fase è seguita da un periodo di 4-5 giorni in cui l'individuo è altamente contagioso ma non manifesta i sintomi (I); si ha quindi l'ultima fase di contagiosità sintomatica che dura 4-6 giorni (M) in cui l'individuo ha una contagiosità ridotta del 65-70% rispetto alla fase precedente.

Il modello SEIMR sembra essere molto adatto a gestire questa situazione in

quanto la patologia presenta, ben distinte, proprio tutte le fasi del modello. Per provare a descrivere questa patologia è stato utilizzato quindi un modello SEIMR (1.11) con:  $R1 = 16, R2 = 4, R3 = 5, p_{c2} = 0.65$  e  $p_{c3} = 0.18$ . È stato valutato l'andamento della durata e del numero medio di malati totali e sintomatici. Gli incontri giornalieri sono stati modellati con delle distribuzioni uniformi di media  $\mu \in \{2, 5, 7, 10\}$  e  $\sigma^2 = 1.3$ .

Da Figura 4.1 si nota che sono presenti più picchi di malati sintomatici ciò dipende dal numero di contatti giornalieri. Quando  $\mu < 10$  sono identificabili tre picchi di malattia di cui due più consistenti. Quando il numero di contatti aumenta tende ad esserci un unico picco importante accompagnato da due cicli minori. Da notare come passando da  $\mu = 5$  a  $\mu = 7$  si inverte l'ordine delle due punte massime di malati. L'aumentare di  $\mu$  comporta anche l'anticipazione della punta massima di malattia (con  $\mu = 10$  avviene al cinquantesimo giorno). Anche l'andamento totale dei malati conferma l'andamento descritto in precedenza. Il numero di incontri giornalieri influenza quindi molto l'andamento dell'epidemia e ne dilata-accorcia i tempi.

Indipendentemente da  $\mu$ , eccetto quando il numero medio di contatti è pari a uno, nell'arco dei 90 giorni si ammala l'intera popolazione. In realtà anche nel caso con  $\mu = 1$  si ammala l'intera popolazione, serve solamente un periodo di tempo maggiore.

Questo risultato è in linea con la patologia considerata in quanto, molto spesso, anche un solo caso di varicella comporta innumerevoli contagi: basti pensare alle epidemie tra i bambini. Il numero medio di incontri, come già accennato, influisce sulla durata dell'epidemia, infatti dai risultati in Tabella 4.1 si nota che all'aumentare di  $\mu$  diminuisce la durata dell'epidemia, in quanto essendo gli incontri più frequenti e la malattia molto contagiosa, tutti si ammalano velocemente.

2	5	7	10	20
115	74.3	70.2	68.2	55.3

Tabella 4.1: Durata media dell'epidemia al variare di  $\mu$

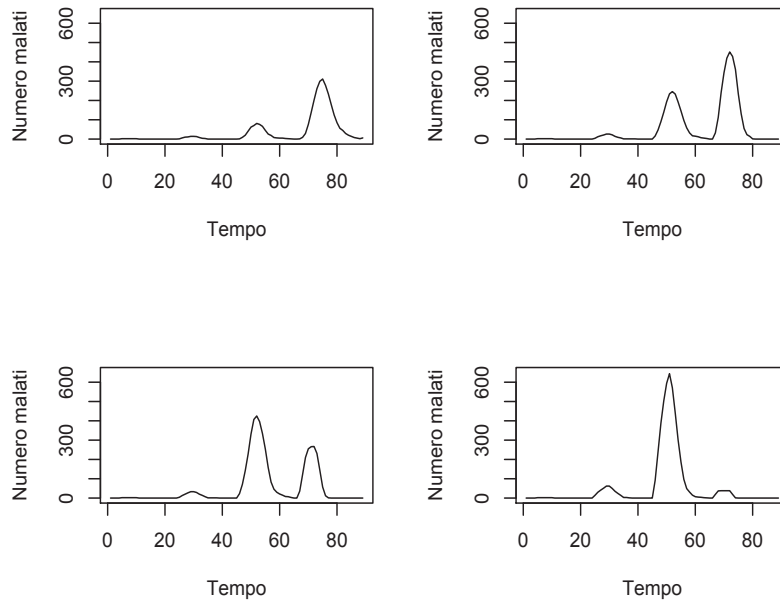


Figura 4.1: Andamento malati sintomatici con  $\mu = 2, 5, 7, 10$

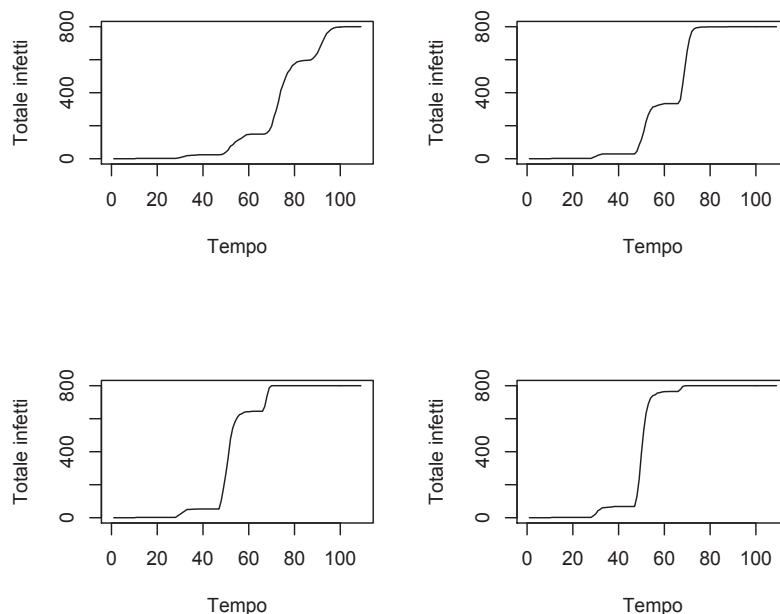


Figura 4.2: Andamento malati con  $\mu = 2, 5, 7, 10$

I risultati esposti fino ad ora sono stati ottenuti considerando due malati iniziali ( $y_2 = 2$ ). Provando a considerare una proporzione di malati iniziali più alta si è notato che, indipendentemente da  $y_2$ , la malattia mantiene l'andamento a due o più picchi. Al variare di  $y_2$  cambiano le caratteristiche dei cicli della malattia. La tendenza comune è che maggiore è il numero di malati iniziali, minore è il tempo necessario per incontrare i picchi di contagiati sintomatici. Infatti i cicli di malattia, sia quando sono tre, sia quando sono due sono sempre più anticipati.

Se si considera invece l'evoluzione dell'andamento dei malati, fissato un numero medio d'incontri e al crescere di  $y_2$ , si nota che: quando  $\mu = 1$ , si mantengono i tre picchi di malati con pochi cambiamenti, con  $\mu = 5$  si mantengono le due ondate di contagiati ma diventano più equilibrate (coinvolgono circa lo stesso numero di malati) con una tendenza all'inversione d'ordine di importanza. Con  $\mu = 7$  si ha una situazione simile a quella precedente e infine con  $\mu = 10$  si mantengono due picchi di importanza differente ma si invertiti nel tempo.



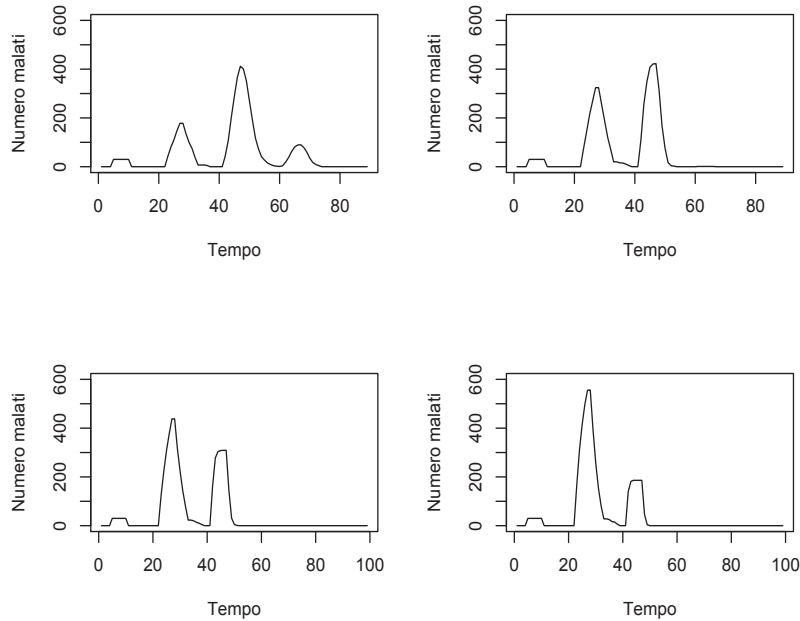


Figura 4.3: Andamento malati sintomatici con  $\mu = 2, 5, 7, 10$ , e 30 malati iniziali

I risultati ottenuti sono abbastanza simili all'andamento reale della varicella, il SEIMR sembra quindi adatto a descrivere la varicella proprio per la struttura di questa malattia che rispetta esattamente tutti gli stati del modello. Con un modello SIR, probabilmente, non si riuscirebbe a descrivere altrettanto bene l'andamento di questa patologia, in quanto non si coglierebbe la differenza tra la fase sintomatica e quella non, che in questo caso hanno durata e probabilità di contagio degli individui molto diversa.

La grande diffusione di questa patologia è dovuta all'alta contagiosità nei giorni immediatamente precedenti alla comparsa dei sintomi, per cui se si volesse ipotizzare un qualche intervento per ridurre la diffusione della malattia, un tentativo volto a ridurre  $R_3$  risulterebbe poco utile, potrebbe risultare più efficiente allora un vaccino che ridurrebbe la diffusione della malattia.

## 4.2 Scarlattina

La scarlattina è una malattia infettiva causata dal batterio *Streptococco beta-emolitico eritrogenico*. Come la varicella è una malattia esantematica ed è quindi caratterizzata da eruzioni cutanee nonché dalla caratteristica colorazione della lingua prima bianca e poi lampona. È una malattia molto diffusa tra i bambini, i casi tra le persone adulte sono rari e per questo i risultati che si otterranno andranno interpretati come se l'intera popolazione fosse composta solo da bambini.

Il contagio avviene entrando a contatto con goccioline di saliva o attraverso oggetti (per esempio, giocattoli, libri, indumenti, posate, bicchieri), sui quali lo streptococco sopravvive a lungo. La malattia ha un periodo di incubazione di 1-3 giorni, in cui l'individuo è contagioso ma non manifesta alcun sintomo. In seguito si manifestano i primi sintomi, mal di testa e febbre alta, dopo 1-4 giorni compaiono i classici sintomi della malattia. Una volta incominciata la cura antibiotica dopo 24-48 ore il malato non risulta più contagioso. Per descrivere l'andamento dell'epidemia è applicabile un modello SEIMR in cui la classe I comprenderà sia i giorni in cui l'individuo non manifesta i sintomi ed è contagioso sia quelli in cui l'individuo manifesta i primi sintomi che però non sono direttamente riconducibili alla patologia in oggetto, per cui non inizia ancora la cura specifica. Per valutare la probabilità di contagio è stato considerato il tasso di riproduttività della malattia ( $R_0$ ) ed è stato comparato con quello della varicella.

La scarlattina ha  $R_0 \simeq 4$ , mentre la varicella ha  $R_0 \simeq 11$  e probabilità di contagio di circa il 65%, per cui la probabilità di contagio della scarlattina risulta all'incirca 0.20. Il modello utilizzato è sempre il (1.11) e gli incontri sono stati modellati con delle distribuzioni uniformi di media  $\mu \in \{2, 5, 7, 10\}$  e  $\sigma^2 = 1.3$ .

Se si valuta l'andamento dei malati sintomatici al variare di  $\mu$  si nota che, a differenza della varicella, c'è un unico picco di malattia dopo il quale l'epidemia termina. In Figura 4.4 è riportato come esempio l'andamento con  $\mu = 2$ . Il numero medio di incontri quindi non influenza la struttura dell'andamento dell'epidemia ma ne modifica le dinamiche temporali, infatti all'aumentare di  $\mu$  il picco massimo di malati è anticipato. Il numero di incontri condiziona quindi anche la durata della malattia che decresce al crescere di  $\mu$ , come da Tabella 4.2. Infine da  $\mu$  dipende anche il numero di casi, anche se con  $\mu > 5$  le differenze sono minime.

2	5	7	10
67	50	45	30

Tabella 4.2: Durata media dell'epidemia al variare di  $\mu$

2	5	7	10
230	480	497	499

Tabella 4.3: Numero medio di casi al variare di  $\mu$

Essendo il valore utilizzato per  $p$  abbastanza approssimativo, sono stati considerati anche altri valori nell'intervallo  $p \pm 0.05$  ed è stato valutato come possa variare il numero medio di coinvolti al variare di  $p$  e del numero di incontri giornalieri.

Le differenze maggiori al variare di  $p$  si riscontrano quando gli incontri sono abbastanza limitati, infatti con  $\mu > 3$  tendono ad ammalarsi tutti, indipendentemente da  $p$ . Testando quindi altri valori della probabilità di contagio non emergono andamenti differenti.

Possiamo ipotizzare allora che, se considerassimo un'intera popolazione di bambini, nell'arco di tre mesi la maggior parte di questi si ammalerebbe.

La scarlattina, a differenza della varicella, non manifesta un andamento particolare; questo potrebbe voler dire che i valori utilizzati per le stime sono troppo approssimativi oppure che il modello non risulta perfettamente adatto a questa tipologia di malattia. Ad esempio la fase I in cui l'individuo infettato non manifesta i sintomi non è così ben definita come nel caso precedente: non è una vera fase asintomatica ma comprende un periodo in cui i sintomi sono già presenti anche se non sono quelli tipici della patologia e quindi anche la stima della lunghezza delle fasi ne risente. Potrebbero quindi essere necessari o altri studi di approfondimento sulla malattia, per capire se l'andamento ottenuto è corretto, o provare ad applicare un modello leggermente differente.

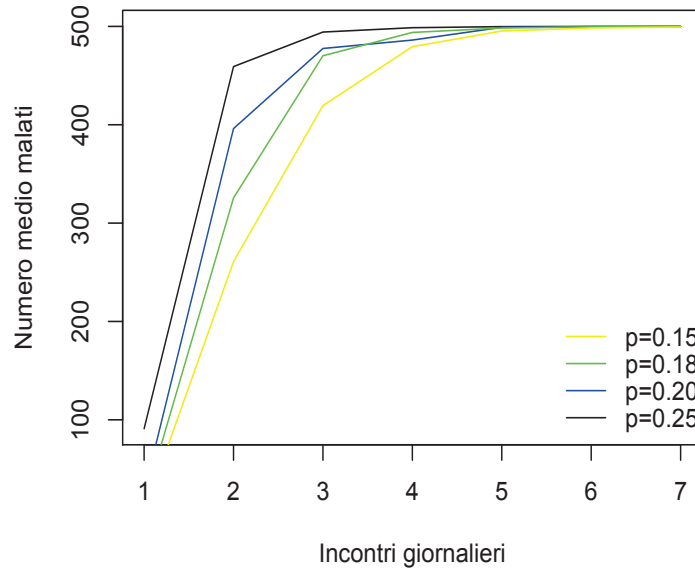


Figura 4.4: Andamento numero medio malati al variare del numero d'incontri giornalieri

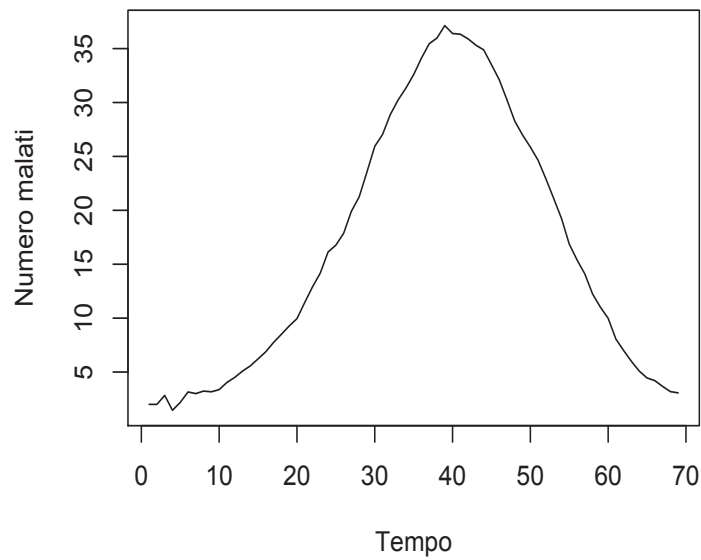


Figura 4.5: Andamento medio dei malati sintomatici con  $\mu = 2$

## Conclusioni

I modelli elaborati, come era logico attendersi, forniscono risultati in parte in linea con quelli del modello SIR originario, ma in parte anche innovativi. L'introduzione della differenziazione degli stadi della malattia permette una valutazione migliore del decorso di alcune patologie e la descrizione dell'andamento di altre più complesse. Nel modello SEIMR ciò che sembra influenzare maggiormente la diffusione dell'epidemia è la contagiosità nel periodo sintomatico, piuttosto che la lunghezza della fase (che comunque condiziona anch'essa l'estensione della patologia). Anche la struttura che permette di valutare gli incontri non equiprobabili, sembra essere utile al fine di descrivere in modo più realistico l'epidemia e porta a confermare la tendenza per cui la probabilità di contagio possa avere un peso maggiore. Il modello SEIMR si adatta bene a malattie che, come la varicella, presentano distintamente le fasi di latenza, contagiosità asintomatica e il periodo con sintomi e possibilità di infettare. Permette di evidenziare andamenti che un modello SIR normale non riuscirebbe a cogliere. Potrebbe essere quindi uno strumento più accurato e preciso per predire lo sviluppo di alcuni tipi di epidemia.

Gli altri modelli studiati, possono essere adatti per modellare patologie di massa e per studiarne la diffusione. Anche in questi modelli, la probabilità di contagio nella fase sintomatica pesa molto sul numero di casi.

Nei SEIMS questa probabilità influenza, inoltre, la possibile fine o meno dell'epidemia. Infatti, quanto tale probabilità risulta essere alta, tendenzialmente c'è un continuo ricambio di malati e l'epidemia tende a perdurare.

Nei SEIMRS invece, considerando malattie solo in alcuni casi immunizzanti, sul numero di casi incide molto la probabilità di diventare immuni o meno.

Per quanto riguarda l'approccio a popolazione aggregata, sarebbe un buon modello per poter valutare, in ogni tempo, l'andamento della popolazione vista nel suo insieme. Purtroppo però, presenta alcuni limiti computazionali. Sarebbe necessario o un maggiore supporto informatico o ipotizzare una diversa struttura in cui gli stati non aumentino notevolmente all'aumentare della durata del periodo di infezione e della popolazione. Così come è formulato ora, il modello potrebbe essere adatto a gestire fenomeni molto rari in cui anche la popolazione potenzialmente coinvolta sia molto limitata. Si potrebbe pensare, in alternativa, a cambiare l'unità temporale che nei vari modelli è stata considerata pari al giorno. In questo modo l'arco temporale considerato sarebbe più ampio e gli stati da valutare non aumenterebbero. Andrebbe però valutata la perdita in precisione.



# Appendice A

## Esempio di codice R utilizzato

### Modello SEIMR per la varicella

```
#Modello SEIMR per la varicella

# La funzione permette di valutare la durata dell'epidemia
# il numero di individui presenti in ogni tempo e in ogni
# stato della malattia

ammalatore.mod<-function(n,t,a,b)
{
  inco<-matrix(sample(a:b,n*t,replace=TRUE),nrow=n,ncol=t)
  nstati <- 4
  lambda <- 0.3
  R1 <- 16
  R2 <- 4
  R3 <- 5

  #numero di persone per stato a t=0
  sum.i1 <- num.i1 <- 0
  sum.i2 <- num.i2 <- 2
  sum.i3 <- num.i3 <- 0

  finiti<- 0
  sani<- n - (num.i1 + num.i2+ num.i3)
  durata=0
```

```

#definizione matrice persona-tempo in t=0
mat.dat<- array(0,dim=c(n,t,nstati))
mat.dat[, ,1]<-1 #[, ,1]

selezionati.i1<-sample(1:n,sum.i1)
mat.dat[selezionati.i1,1:t,1]<- 0
mat.dat[selezionati.i1,1,2]<- 1

selezionati.i2<-sample(setdiff(1:n,selezionati.i1),sum.i2)
mat.dat[selezionati.i2,1:t,1]<- 0
mat.dat[selezionati.i2,1,3]<- 1

ris.finali <- matrix(NA,nrow=t-1,ncol=5)
r1 <- r2 <- r3 <- array(0,dim=n)

# implementazione della matrice persona-tempo
for (j in 1:(t-1))
{
#p di incontro delle persone dei vari stati
p3 <- (lambda*num.i3)/(n)
p2 <- (num.i2/(n))*((n-(lambda*num.i3))/(n-num.i3))
p1 <- 1 - (p2 + p3)

#vettori contenenti indici delle persone nello stato i2, i3
indici.i2<-which(mat.dat[,j,3]==1)
indici.i3<-which(mat.dat[,j,4]==1)
for (i in 1:n)
{
if(mat.dat[i,j,1]==1)
{
if(r3[i]!=R3)
{
incontri<-rmultinom(1,inco[i,j],prob=c(p1,p2,p3))
ris2=0
ris3=0

if((incontri[2]>0) & (length(indici.i2)>incontri[2]))
{

vet2<-rbinom(incontri[2],1,0.65)
ris2<-sum(vet2)

```



```

}

if((incontri[3]>0) & (length(indici.i3)>incontri[3]))
{

vet3<-rbinom(incontri[3],1,0.18)
  ris3<-sum(vet3)
}

if((ris2>0)|(ris3>0))
{
mat.dat[i,j+1,2] <- 1
mat.dat[i,(j+1):t,1]<- 0
sum.i1=sum.i1+1
sani=sani-1

}
}
else
{
mat.dat[i,j+1,1]=mat.dat[i,j,1]
}
}
#parte della funzione che gestisce la parte determenistica.
#Per ogni stato si controlla da quanto tempo l'individuo
# vi permane,se haraggiunto il numero di giorni prefissato
#scala allo stato successivo.

if(mat.dat[i,j,2] == 1)
{
if(r1[i]==R1)
{
mat.dat[i,j+1,3] <- 1
mat.dat[i,j+1,2]<- 0
sum.i2=sum.i2+1
sum.i1=sum.i1-1
}
}
else
{
mat.dat[i,j+1,2] <- mat.dat[i,j,2]
r1[i] <- r1[i]+1

```

```

}
}

if(mat.dat[i,j,3] == 1)
{
if(r2[i]==R2)
{
mat.dat[i,j+1,4] <- 1
mat.dat[i,j+1,3]<- 0
sum.i3=sum.i3+1
sum.i2=sum.i2-1
}
else
{
mat.dat[i,j+1,3] <- mat.dat[i,j,3]
r2[i] <- r2[i]+1
}
}

if(mat.dat[i,j,4] == 1)
{
if(r3[i]==R3)
{
mat.dat[i,j+1,1] <- 1
mat.dat[i,j+1,4]<- 0
sum.i3=sum.i3-1
finiti=finiti+1
}
else
{
mat.dat[i,j+1,4] <- mat.dat[i,j,4]
r3[i] <- r3[i]+1
}
}

}

#calcolo numero di persone per stato
num.s <- length(which(mat.dat[,j,1]==1))

```

```

num.i1 <- length(which(mat.dat[,j,2]==1))
num.i2 <- length(which(mat.dat[,j,3]==1))
num.i3 <- length(which(mat.dat[,j,4]==1))
num.si1 <- num.s+num.i1

mal <- sum.i1+sum.i2+sum.i3
#definizione matrice dei risultati
ris.finali[j,1] <- sani
ris.finali[j,2] <- sum.i1
ris.finali[j,3] <- sum.i2
ris.finali[j,4] <- sum.i3
ris.finali[j,5] <- finiti

if((sum.i1+sum.i2+sum.i3)!=0)
{
durata=durata+1
}
}
list(risultati=ris.finali,durata=durata)
#ris.finali[,4]
}

#### Simulazione andamento malati contagiosi
#bisogna modificare la funzione precedente ed avere in output solo la quarta
#colonna 4 della matrice ris.finali

Nsim=100
ris<- matrix(NA,nrow=Nsim,ncol=79)

for (i in 1:Nsim)

{
ris[i,]<- ammalatore.mod(800,80,1,3)
}

esem.13<-applay(ris,2,mean)
esem.37<-applay(ris,2,mean)

```

```

#Funzione per costruire la matrice di transizione per il modello
#con popolazione aggregata per malattie della durata di u
#n giorno riempimento <- function(pop,n=1,p=0.1)
{
nstati=2
dim <- choose(pop+nstati,pop)
transizione <- matrix(0,ncol=dim,nrow=dim)

#comandi per la descrizione alfanumerica degli stati

legenda <- paste(pop,"0",sep="")
out <- matrix(c(pop,0),ncol=2)
colnames(out) <- c("sani","malati")
#definizione degli stati e la relativa etichetta
for(i in 1:pop)
{
k <- i:0
legenda <- c(legenda,paste(pop-i,"",k,sep=""))
temp <- matrix(c(rep(pop-i,length(k)),k),ncol=2,byrow=FALSE)
out <- rbind(out,temp)
}
colnames(transizione) <- rownames(transizione) <- legenda

#funzione in cui si calcolano le probabilita di transizione

probante <- function(out,pos.partenza,pos.arrivo,pop,n,p)
{
partenze <- out[pos.partenza,]
arrivi <- out[pos.arrivo,]

sp <- partenze[1]
mp <- partenze[2]
sa <- arrivi[1]
ma <- arrivi[2]

choose(sp,sa)* ((1-p*mp/pop)^n)^sa * (1-(1-p*mp/pop)^n)^ma
}
probante.vec <- Vectorize(probante,vectorize.args="pos.arrivo")

# si differenziano i vari casi e si associa ad ogni stato la
#probabilita di transizione

```

```

for(i in 1:nrow(out))
{
if((out[i,1] > 0) & (out[i,2] > 0))
{
temp <- out[i,]
k <- temp[1]:0
arrivi <- paste(k,",",max(k)-k,sep="") #definisco stati di arrivo
posizioni <- tapply(arrivi,1:length(arrivi),function(x) which(legenda==x))

transizione[i,posizioni] <- probante.vec(out,i,posizioni,pop,n,p)
}

if(out[i,1]==0)
{
transizione[i,nrow(out)] <- 1
}

if(out[i,2]==0)
{
transizione[i,i] <- 1
}
}

transizione
}

#calcolo tempo medio di assorbimento con popolazione=5
a <- riempimento(5)
num<-ncol(a)
#codice per l'identificazione della posizione degli stati transienti
# e dei ricorrenti
tenere<-array(1,dim=num)
ass<- 0
for(i in 1:num)
{
if(a[i,i]==1)
{
ass=ass+1
tenere[i]=0
}
}

```

```

}
trans=num-ass
#costruzione della matrice degli stati transienti
s<- a
indice<-which(tenere==0)
s<-s[-c(indice),-c(indice)]

#calcolo del tempo di assorbimento
m2<- ncol(s)
i<- diag(1,nrow=m2,ncol=m2) #quante sono le righe di s
l<-solve(i-s)
m<- ncol(l)
uni<- matrix(1,nrow=m,ncol=1)
tempi<- l\%*\%uni

```

# Bibliografia

- [1] H.C. Tuckwell, R.J. Williams *Some properties of a simple stochastic epidemic model of SIR type*(2006)
- [2] G. Gaeta, *Modelli matematici in Biologia.* (2007) Springer
- [3] M. E. Halloran, I. M. Longini, Jr. Claudio, J. Struchiner, *Design and Analysis of Vaccine Studies*, (2009), SPIN Springer
- [4] A.M. Lilienfeld *Epidemiology of infectious and non-infectious disease:some comparisons* (1973), American Journal of Epidemiology
- [5] K.O. Kermack, A.G. McKendrick, *A Contribution to the Mathematical Theory of Epidemics*(1927)
- [6] J.A. Jacquez *A Note on Chain-Binomial Models of Epidemic Spread: What is Wrong with the Reed-Frost Formulation?*, (1987)
- [7] <http://www.unc.edu/rls/s940/samsidisdyntut.pdf>
- [8] [www.science.unitn.it/anal1/biomat/note/epidemomog.pdf](http://www.science.unitn.it/anal1/biomat/note/epidemomog.pdf)
- [9] [http://www.stats.ox.ac.uk/data/assets/file/0013/3352/infection\\_rates.pdf](http://www.stats.ox.ac.uk/data/assets/file/0013/3352/infection_rates.pdf)
- [10] [http://en.wikipedia.org/wiki/Compartmental\\_models\\_in\\_epidemiology](http://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology)
- [11] <http://it.wikipedia.org/wiki/Varicella>
- [12] <http://www.epicentro.iss.it/problemi/varicella/varicella.asp>
- [13] <http://www.salute.gov.it/malattieInfettive>
- [14] <http://it.wikipedia.org/wiki/Scarlattina>