



# Bachelorarbeit - Simulation eines räumlichen SEIR Modells

Referent: Vjekoslav Drvar  
Betreuer: Prof. Dr. Konstantinos Panagiotou

München, 15. September 2021

Ludwig-Maximilians-Universität  
Mathematisches Institut

# Inhaltsverzeichnis

<b>1 Einführung und Motivation</b>	<b>3</b>
<b>2 Compartment-Modelle</b>	<b>4</b>
2.1 Das klassische SIR Modell mit ODE's [1] . . . . .	4
2.1.1 Satz über Eindeutigkeit der Lösung von SIR [2] . . . . .	5
2.1.2 Die Reproduktionszahl: $\sigma$ [3] . . . . .	5
2.1.3 Der Grenzwert des Anteils von S: $s_\infty$ . . . . .	6
2.1.4 Lösbarkeit des SIR-Modells [4] . . . . .	6
2.1.5 Beispiele von SIR Trajektorien . . . . .	7
2.2 Weitere Variationen von SIR-Modellierung . . . . .	8
2.2.1 SIS-Modell . . . . .	8
2.2.2 Modelle mit dem Überträger (C) Compartment . . . . .	9
2.2.3 SEIR-Modell . . . . .	9
2.3 Geometrische Modelle . . . . .	10
<b>3 'Das' Modell</b>	<b>11</b>
3.1 Vorbereitung: Verteilungen . . . . .	11
3.1.1 Diskrete Gleichverteilung (Abk. <i>Unif</i> ) [5] . . . . .	11
3.1.2 Poisson-Verteilung (Abk. <i>Poiss</i> ) [6] . . . . .	11
3.1.3 Exponentialverteilung (Abk. <i>Exp</i> ) [7] . . . . .	12
3.1.4 Pareto-Verteilung [8] . . . . .	12
3.2 Interpretation und Nachteile des SIR Modells . . . . .	14
3.3 Beschreibung des Modells . . . . .	15
3.4 Begründung der Wahl der Verteilungen . . . . .	17
3.5 Mögliche Verallgemeinerungen . . . . .	17
3.6 Vermutungen . . . . .	18
3.7 Graphische Ergebnisse [9] . . . . .	18
3.8 Schlussfolgerungen I . . . . .	22
3.9 Quantitative Ergebnisse . . . . .	22
3.9.1 Ergebnisse für 'große' $\alpha$ 's . . . . .	22
3.9.2 Ergebnisse für 'kleine' $\alpha$ 's . . . . .	26
3.10 Schlussfolgerungen II . . . . .	26
3.11 Diskussion . . . . .	27

# Abbildungsverzeichnis

1 SIR Umlauf . . . . .	5
2 Eine SIR Trajektorie mit $\beta = 3, \gamma = 4$ [10] . . . . .	7
3 Eine SIR Trajektorie mit $\beta = 5, \gamma = 1$ [10] . . . . .	8
4 SIRC Umlauf . . . . .	9
5 Veranschaulichung der Poisson-Dichte für verschiedene Parameter . . . . .	11
6 Veranschaulichung der Dichte der Exponentialverteilung für verschiedene Parameter . . . . .	12

7	Veranschaulichung der Pareto-Dichte für verschiedene Parameter . . . . .	13
8	Veranschaulichung der homogenen Struktur des SIR Modells; Das Farbsystem ist z.B. S - blau, I - rot, R - grün . . . . .	14
9	Skizze des SEIR-Mechanismus . . . . .	16
10	Skizze des räumlichen Mechanismus . . . . .	16
11	Illustration eines Torus [11]; man kann sich die räumliche Konfiguration der Population als die Fläche eines Torus vorstellen . . . . .	16
12	Simulationsergebnis für $\alpha = 4$ . . . . .	19
13	Simulationsergebnis für $\alpha = 3$ . . . . .	20
14	Simulationsergebnis für $\alpha = 2.5$ . . . . .	21
15	Simulationsergebnisse für $\alpha = 4, \lambda_{Poisson} = 1$ . . . . .	23
16	Simulationsergebnisse für $\alpha = 4, \lambda_{Poisson} = 2$ . . . . .	23
17	Simulationsergebnisse für $\alpha = 4, \lambda_{Poisson} = 3$ . . . . .	24
18	Simulationsergebnisse für $\alpha = 5, \lambda_{Poisson} = 1$ . . . . .	24
19	Simulationsergebnisse für $\alpha = 5, \lambda_{Poisson} = 2$ . . . . .	25
20	Simulationsergebnisse für $\alpha = 5, \lambda_{Poisson} = 3$ . . . . .	25

# 1 Einführung und Motivation

Im Jahr 2020 war die ganze Welt Zeuge einer Epidemie - COVID-19 Epidemie. Obwohl sehr viel im tagtäglichen Leben dadurch schlimmer worden ist, entsteht mit dieser Situation ein zusätzlicher Anreiz, diese Problematik wissenschaftlich unter die Lupe zu nehmen.

Insbesondere wurden zur Frage der Verbreitung einer Epidemie auch verschiedene mathematische Modelle entwickelt. Die wahrscheinlich bekannteste Klasse von solchen Modellen sind die sogenannten Compartment-Modelle. Sie sind ein wichtiges Werkzeug zur Analyse und Kontrolle einer Krankheit worden; nachdem man das Modell spezifiziert und die Variablen der Realität nach einstellt ist es möglich, verschiedenste Schlussfolgerungen über die epidemiologische Situation zu treffen. Außerdem, Simulationen von solchen Modellen ermöglichen es, Theorien aufzubauen und zu beurteilen, kausale Verbindungen zu untersuchen und bestimmte Fragen zu beantworten. Weiterhin, mit genug Erfahrung und Intuition ist es möglich, dass Modelle dieser Art zu legislativen Entscheidungen beibringen, insbesondere zu denen, die die Verbreitung einer Krankheit verhindern sollen. Genau dieser letzte Punkt ist von zentraler Bedeutung in dieser Arbeit.

Nämlich, wir bauen und simulieren ein geometrisches Compartment-Modell. Solche Modelle sind von großer Relevanz in diesem Bereich, da sie eine natürliche Erweiterung von den klassischen deterministischen Modellen sind. Klassische Modelle arbeiten mit Raten und da wird die Population als homogen betrachtet; d.h. je zwei Individuen sind im gleichen Verhältnis zueinander. Insbesondere, den Individuen wird kein 'Standort' zugeordnet, was in der Realität schon der Fall ist.

Die Hauptfragen lauten: Wann ist die Krankheit unter Kontrolle und wann stirbt sie aus? Wann und unter welchen Bedingungen verursacht eine kranke Person nur einen lokalen Cluster in ihrer Umgebung? Nachdem wir einige stochastische Eigenschaften spezifizieren ist es möglich, sehr elegante und sinnvolle mathematische Vermutungen über diese Frage zu machen. Wir werden sehen, dass diese Vermutungen die Realität gut widerspiegeln.

Einmal wir mit Sicherheit sagen können, dass das Modell eine legitime (aber natürlich vereinfachte) Version der epidemiologischen Situation darstellt, können wir uns fragen, wie man die neuen Erkenntnisse im echten Leben sinnvoll nutzen kann. Die wichtige Frage, die man sich zu dieser Simulation stellt ist: Machen Bewegungsrestriktionen Sinn im Kontext von Verhindern der Verbreitung einer Krankheit? Wenn ja, wie genau sollen sie definiert sein? Entstehen globale Clusters trotz ihrer manchmal? Spielt es eine Rolle wie lang die Inkubation und die Phase, in der man ansteckend ist, dauern? Ist es relevant wie viel Menschen eine Person in einem Tag im Durchschnitt trifft? Wie soll man die Wahrscheinlichkeit modellieren, dass man nach dem Treffen mit einer kranken Person selber krank wird?

Genau diese Problematik werden wir mit dieser Simulation, die durch das Kombinieren von Kenntnissen aus Informatik und Mathematik entstanden ist, untersuchen und die wesentlichen Resultate kommentieren. Zuerst präsentieren wir die Grundlagen der Theorie von Compartment-Modellen und einige Variationen davon. Danach konzentrieren wir uns auf die eigentliche Simulation: die Mechanismen der Infektionsverbreitung und die Simulationsergebnisse. Die Simulation wurde komplett in C++ geschrieben.

## 2 Compartment-Modelle

Compartment-Modelle in Epidemiologie sind vereinfachte mathematische Modelle einer Infektionskrankheit. Die Population wird in verschiedene Kategorien (Compartments) mit verschiedenen Namen zugeordnet. Im elementaren Fall sind die Compartments Anfällig (S - Susceptible), Infiziert (I - Infected) und Entfernt (R - Removed). Die Reihenfolge, in der die Compartments angegeben werden, weist oft darauf hin, wie der Umlauf zwischen den Compartments modelliert wird.

### 2.1 Das klassische SIR Modell mit ODE's [1]

Ein bekannter und relevanter Ansatz zur Modellierung der Dynamik zwischen den Compartments ist derjenige mit gewöhnlichen Differentialgleichungen. Da modelliert man die Übergangsgraten zwischen den Compartments abhängig von den Niveaus der Compartments selbst, als auch von einigen extern definierten Konstanten. Also, eine Population von  $N \in \mathbb{N}$  Individuen wird in die oben angegebenen Compartments verteilt; startend ab einem Nullzeitpunkt, die Anzahl von Individuen in dem jeweiligen Compartment wird als eine zeitabhängige Funktion definiert:  $S(t)$ ,  $I(t)$ ,  $R(t)$ . Weiter definiert man Konstanten  $\beta, \gamma \in \mathbb{R}^+$ , die die Übergangsgraten modellieren. Solch ein Modell kann dementsprechend folgende Form haben:

$$\begin{aligned}\frac{dS}{dt} &= -\frac{\beta IS}{N} \\ \frac{dI}{dt} &= \frac{\beta IS}{N} - \gamma I \\ \frac{dR}{dt} &= \gamma I\end{aligned}\tag{1}$$

Mit den Anfangsbedingungen:

$$\begin{aligned}S(0) &= S_0 \geq 0 \\ I(0) &= I_0 \geq 0 \\ R(0) &= R_0 \geq 0\end{aligned}\tag{2}$$

Also, die Rate des Übergangs von S nach I ist negativ abhängig sowohl von dem Niveau von S als auch von dem von I - je mehr es Anfällige und Infizierte gibt, desto mehr Neuinfizierte gibt es. Symmetrisch dazu, mit der gleichen Rate (hier ist sie positiv) wächst I; ein Anteil von I, abhängig von dem Niveau von I, wechselt zu R.

Es ist wichtig klarzumachen, dass in diesem Modell alle Komponenten deterministisch sind (keine Unsicherheit, keine stochastische Modellierung) und dass der geometrische Standort von einzelnen Individuen keine Rolle spielt, bzw. überhaupt nicht spezifiziert wird.



Abbildung 1: SIR Umlauf

### 2.1.1 Satz über Eindeutigkeit der Lösung von SIR [2]

*Satz:*

Das Differentialgleichungssystem im SIR Modell ist eindeutig lösbar.

*Beweis:*

Nachdem man die Gleichungen in (1) durch die konstante Populationsgröße  $N$  teilt, bekommt man:

$$\begin{aligned} \frac{ds}{dt} &= -\beta is \\ \frac{di}{dt} &= \beta is - \gamma i \end{aligned} \tag{3}$$

Mit den Anfangsbedingungen:

$$\begin{aligned} s(0) &= s_0 \geq 0 \\ i(0) &= i_0 \geq 0 \end{aligned} \tag{4}$$

mit  $r(t) = 1 - s(t) - i(t)$ , wobei  $s(t)$ ,  $i(t)$ ,  $r(t)$  die Anteile der Compartments sind. Das Dreieck  $T$  in der  $si$  Phasenebene ausgedrückt als:

$$T = \{(s, i) | s \geq 0, i \geq 0, s + i \leq 1\} \tag{5}$$

ist positiv invariant und eindeutige Lösungen existieren in  $T$  für alle positive Zeitpunkte, somit ist das Modell mathematisch und epidemiologisch wohldefiniert. [12]  $\square$

### 2.1.2 Die Reproduktionszahl: $\sigma$ [3]

Eine wichtige Maßzahl zu diesem Modell, von der die Dynamik des I-Compartments abhängt, ist die sogenannte Reproduktionszahl,  $\sigma = \frac{\beta}{\gamma}$ . Sie bezeichnet die erwartete Anzahl von Neuinfektionen von einer einzigen Infektion in einer Population, in der alle anfällig sind. Nachdem man die zweite Differentialgleichung umschreibt, bekommt man:

$$\frac{dI}{dt} = (\sigma \frac{S}{N} - 1)\gamma I \tag{6}$$

Dies impliziert folgendes:

$$\sigma * S(0) > N \implies \frac{dI}{dt}(0) > 0 \quad (7)$$

Das heißt, ein richtiger Ausbruch passiert und die Anzahl von Infizierten steigt.

Andererseits, man hat auch:

$$\sigma * S(0) < N \implies \frac{dI}{dt}(0) < 0 \quad (8)$$

Das heißt, kein richtiger Ausbruch passiert und die Anzahl von Infizierten sinkt.

Im Spezialfall, in dem (fast) alle Individuen am Anfang zu S gehören, kann man die obigen Formeln auch auf die folgende Weise schreiben:

$$\begin{aligned} \sigma > 1 &\implies \frac{dI}{dt}(0) > 0 \\ \sigma < 1 &\implies \frac{dI}{dt}(0) < 0 \end{aligned} \quad (9)$$

### 2.1.3 Der Grenzwert des Anteils von S: $s_\infty$

*Satz:* [2]

Sei  $(s(t), i(t))$  die Lösung von (3) in T. Falls  $\sigma * s_0 > 1$ , dann  $i(t) \rightarrow 0$ , als  $t \rightarrow \infty$ . Falls  $\sigma * s_0 \leq 1$ , dann nimmt  $i(t)$  sein Maximum  $i_{max} = i_0 + s_0 - \frac{1}{\sigma} - \frac{\ln(\sigma * s_0)}{\sigma}$  an und dann  $i(t) \rightarrow 0$ , als  $t \rightarrow \infty$ . Der anfällige Anteil ist die eindeutige Lösung in  $(0, \frac{1}{\sigma})$  der Gleichung:

$$i_0 + s_0 - s_\infty + \ln\left(\frac{s_\infty}{s_0}\right)/\sigma = 0 \quad (10)$$

*Beweis:* Der Beweis wird hier ausgelassen.  $\square$

Die Interpretation dazu ist ähnlich wie in 2.1.2. Falls fast alle Individuen am Anfang anfällig sind und die Reproduktionszahl größer als 1 ist, werden alle Individuen in einem Moment krank. D.h., am Ende gibt es keine anfälligen Individuen. Umgekehrt, falls die Reproduktionszahl kleiner oder gleich 1 ist, nicht alle Individuen werden infiziert und ein Anteil der Population bleibt 'für ewig' anfällig.

### 2.1.4 Lösbarkeit des SIR-Modells [4]

Im Jahr 2014 haben Harko et al. eine pseudo-analytische Lösung (enthält Integrale, die man numerisch lösen muss) zum SIR-Modell hergeleitet. Sie sieht folgendermaßen aus:

$$\begin{aligned} S(u) &= S(0)u \\ I(u) &= N - R(u) - S(u) \\ R(u) &= R(0) - \rho * \ln(u) \end{aligned} \quad (11)$$

für die Parametrisierung:

$$t = \frac{N}{\beta} \int_u^1 \frac{du^*}{u^* I(u^*)}, \quad \rho = \frac{\gamma N}{\beta} \quad (12)$$

mit den Anfangsvoraussetzungen:

$$(S(1), I(1), R(1)) = (S(0), N - R(0) - S(0), R(0)), \\ u_T < u < 1, \text{ wobei } I(u_T) = 0 \quad (13)$$

### 2.1.5 Beispiele von SIR Trajektorien

Zunächst präsentieren wir zwei SIR Trajektorien mit verschiedenen Parametern. In der ersten stirbt die Krankheit aus, d.h.  $\sigma \leq 1$ . In der zweiten werden dennoch alle Individuen infiziert, d.h.  $\sigma > 1$ .

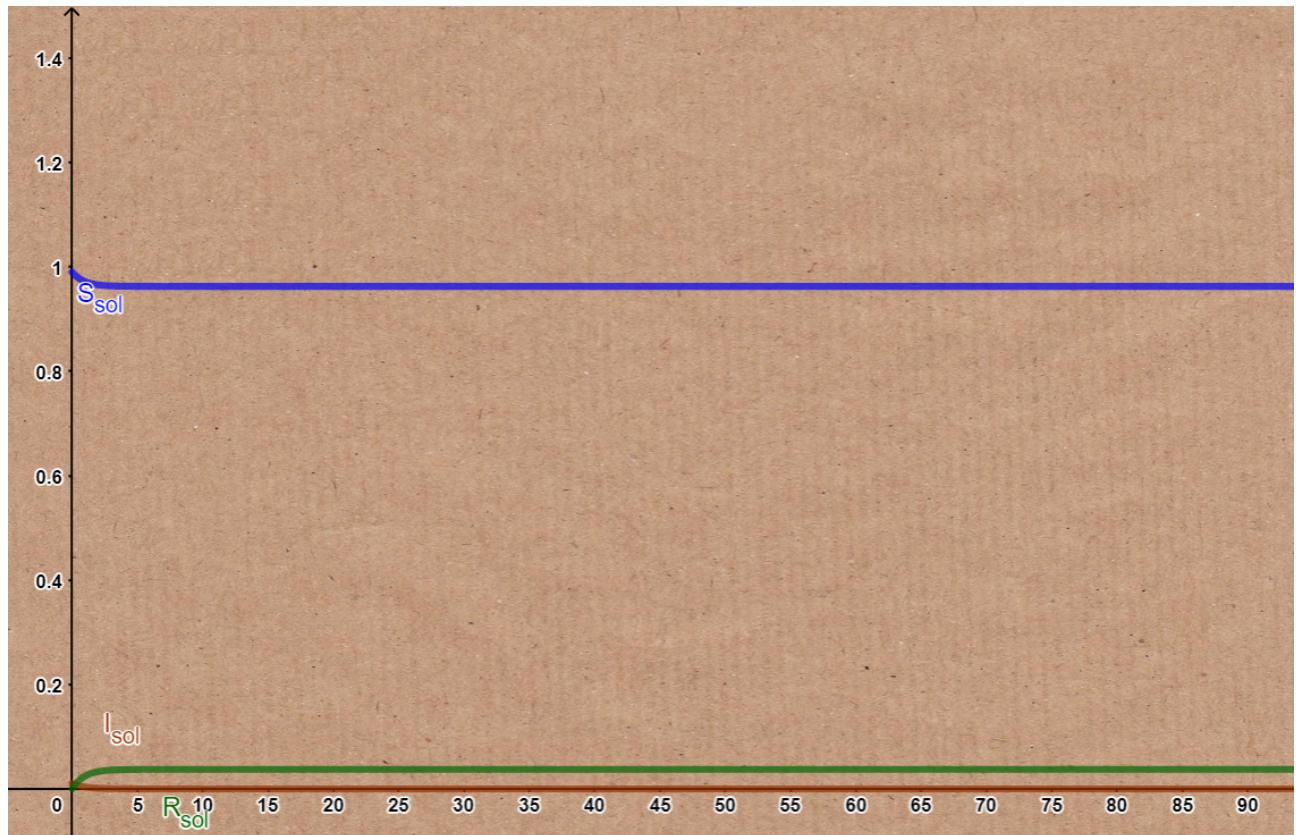


Abbildung 2: Eine SIR Trajektorie mit  $\beta = 3, \gamma = 4$  [10]

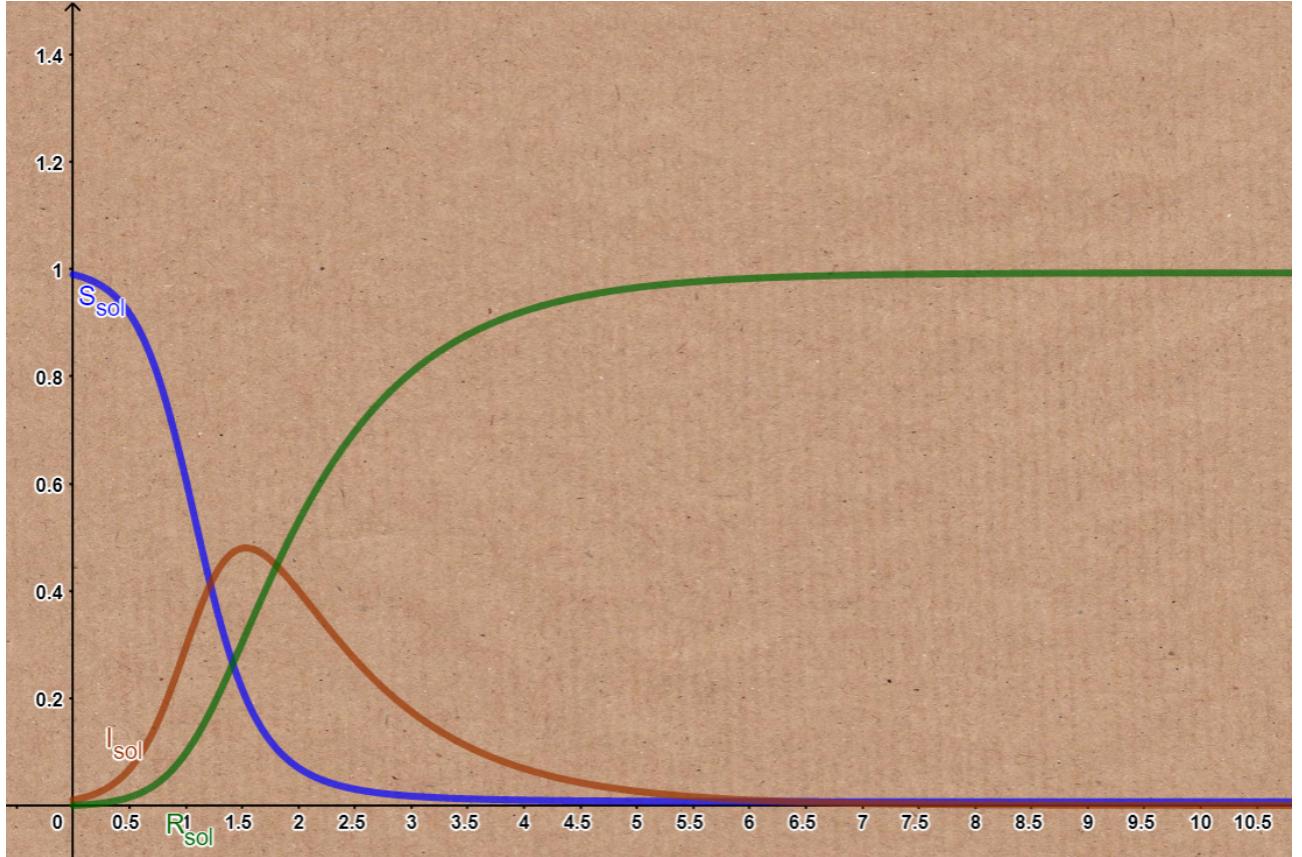


Abbildung 3: Eine SIR Trajektorie mit  $\beta = 5, \gamma = 1$  [10]

## 2.2 Weitere Variationen von SIR-Modellierung

### 2.2.1 SIS-Modell

Für einige Krankheiten entwickelt man keine langfristige Immunität. Für diesen Fall ist das R-Compartment dementsprechend übrig, also man wechselt für ewig zwischen S und I. Folgende Form bietet sich an:

$$\begin{aligned}\frac{dS}{dt} &= \frac{-\beta IS}{N} + \gamma I \\ \frac{dI}{dt} &= \frac{\beta IS}{N} - \gamma I\end{aligned}\tag{14}$$

Es ist aus dem tagtäglichen Leben leicht einzusehen, dass die SIR- und SIS-Strukturen für die Mehrheit der infektiösen Krankheiten zu viel vereinfachen. Einerseits, viele Krankheiten haben kompliziertere Phasen: Menschen entwickeln eine Immunität, einige bekommen die Krankheit wieder, Viren mutieren usw. Andererseits, wir schon erwähnt, berücksichtigen diese zwei Modelle die Population als eine homogene Menge, was auch problematisch ist.

## 2.2.2 Modelle mit dem Überträger (C) Compartiment

Einige Menschen, die eine infektiöse Krankheit haben, erholen sich nie hundertprozentig und tragen die Krankheit passiv weiter. Somit führt man eine Alternative zum R-Compartiment: das C-Compartiment (Carrier).

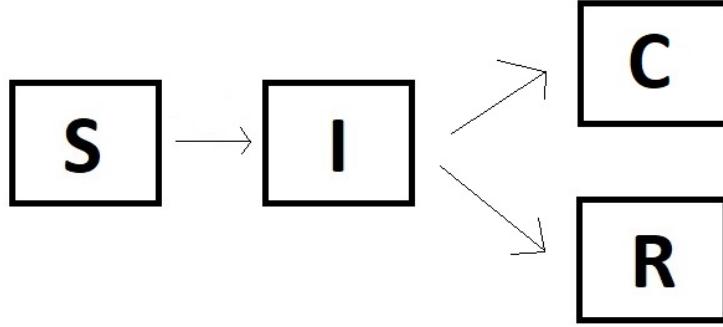


Abbildung 4: SIRC Umlauf

Solche Struktur kann man z.B. so modellieren, wobei  $\beta_1, \beta_2, \gamma_1, \gamma_2 \in \mathbb{R}^+$

$$\begin{aligned}
 \frac{dS}{dt} &= (-\beta_1 I - \beta_2 C) \frac{S}{N} \\
 \frac{dI}{dt} &= (\beta_1 I + \beta_2 C) \frac{S}{N} - (\gamma_1 + \gamma_2) I \\
 \frac{dR}{dt} &= \gamma_1 I \\
 \frac{dC}{dt} &= \gamma_2 I
 \end{aligned} \tag{15}$$

## 2.2.3 SEIR-Modell

Es ist oft der Fall, dass man, nach einer Infektion, nicht sofort für andere infektiös ist. Somit muss man ein zusätzliches Compartiment zwischen S und I einführen: E - Ausgesetzt (Exposed). Zusätzlich zu den Objekten, die in 2.1 definiert wurden, seien  $\alpha, \mu \in \mathbb{R}^+$ . Somit ist folgendes Differentialgleichungssystem möglich:

$$\begin{aligned}
 \frac{dS}{dt} &= \mu N - \mu S - \frac{\beta IS}{N} \\
 \frac{dE}{dt} &= \frac{\beta IS}{N} - (\mu + \alpha) E \\
 \frac{dI}{dt} &= \alpha E - (\gamma + \mu) I \\
 \frac{dR}{dt} &= \gamma I - \mu R
 \end{aligned} \tag{16}$$

Dabei modelliert  $\alpha$  die Dauer der E-Phase und  $\mu$  die Rate der Mortalität und Fertilität. Diese Compartment-Struktur nutzen wir später im eigentlichen Modell der Simulation.

Für diese Version wird die Reproduktionszahl folgendermaßen definiert:

$$\sigma = \frac{\alpha}{\mu + \alpha} * \frac{\beta}{\mu + \gamma} \quad (17)$$

Es ist möglich, Resultate über  $s_\infty$  für dieses Model analog zu denen für SIR zu treffen. [2]

## 2.3 Geometrische Modelle

Im Gegensatz zur oben besprochenen Klasse von Modellen, in geometrischen Modellen spielt der Standort eines Individuums schon eine Rolle in der Verbreitung der Krankheit. Hier wird üblicherweise 'die Welt' als eine (große) Matrix definiert, wobei jedes Element der Matrix eine Person darstellt. Danach spezifiziert man eine Reihe von epidemiologisch relevanten Eigenschaften die man danach an die Individuen stochastisch zuteilt. Abhängig davon, unter welchen Verteilungen diese Eigenschaften liegen, passieren, wie sich man auch vorstellen kann, verschiedene Phänomene. Genau diese Abhängigkeiten sind häufig von zentraler Wichtigkeit bei solchen Modellen, diese Arbeit ist keine Ausnahme diesbezüglich.

# 3 'Das' Modell

## 3.1 Vorbereitung: Verteilungen

### 3.1.1 Diskrete Gleichverteilung (Abk. *Unif*) [5]

Mit dieser Verteilung tritt jedes mögliche Ergebnis mit der gleichen Wahrscheinlichkeit ein. Der Grundgedanke einer Gleichverteilung ist, dass es keine Präferenz gibt. Sei  $\Omega$  eine nichtleere endliche Menge. Dann ist bei einer Gleichverteilung die Wahrscheinlichkeit  $P(A)$  eines Ereignisses  $A$  mit  $A \subseteq \Omega$  durch die folgende Formel definiert:

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\text{Anzahl der Elemente in } A}{\text{Anzahl der Elemente in } \Omega} \quad (18)$$

### 3.1.2 Poisson-Verteilung (Abk. *Poiss*) [6]

Die Poisson-Verteilung ist eine diskrete Wahrscheinlichkeitsverteilung. Sie wird durch einen reellen Parameter  $\lambda > 0$  bestimmt, der den Erwartungswert und gleichzeitig die Varianz der Verteilung beschreibt. Sie wird folgendermaßen definiert, eine Zufallsvariable  $X$  ist Poissonverteilt mit Parameter  $\lambda \in \mathbb{R}^+$ , falls:

$$P_\lambda(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad \forall k \in \mathbb{N}_0, \quad \lambda > 0 \quad (19)$$

Wie schon erwähnt, für die ersten zwei Momente gilt folgendes:

$$\mathbb{E}[X] = \lambda, \quad \text{Var}[X] = \lambda \quad (20)$$

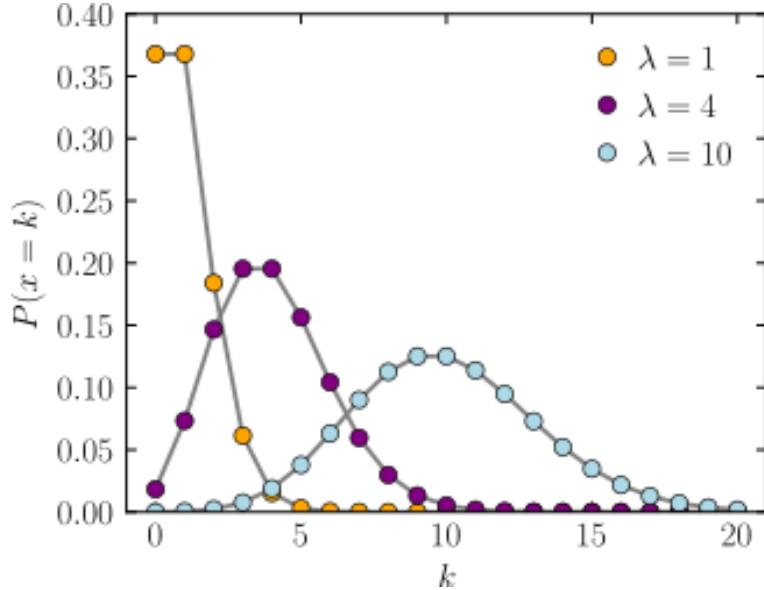


Abbildung 5: Veranschaulichung der Poisson-Dichte für verschiedene Parameter

### 3.1.3 Exponentialverteilung (Abk. *Exp*) [7]

Die Exponentialverteilung ist eine stetige Wahrscheinlichkeitsverteilung über der Menge der nicht-negativen reellen Zahlen, die durch eine Exponentialfunktion gegeben ist. Sie wird folgendermaßen definiert, eine Zufallsvariable  $Y$  ist exponentialverteilt mit Parameter  $\lambda \in \mathbb{R}^+$ , falls sie die folgende Dichte besitzt:

$$f_\lambda(y) = \mathbb{1}_{\{y>0\}} \lambda e^{-\lambda y} \quad (21)$$

Für die ersten zwei Momente gilt folgendes:

$$\mathbb{E}[Y] = \frac{1}{\lambda}, \text{Var}[Y] = \frac{1}{\lambda^2} \quad (22)$$

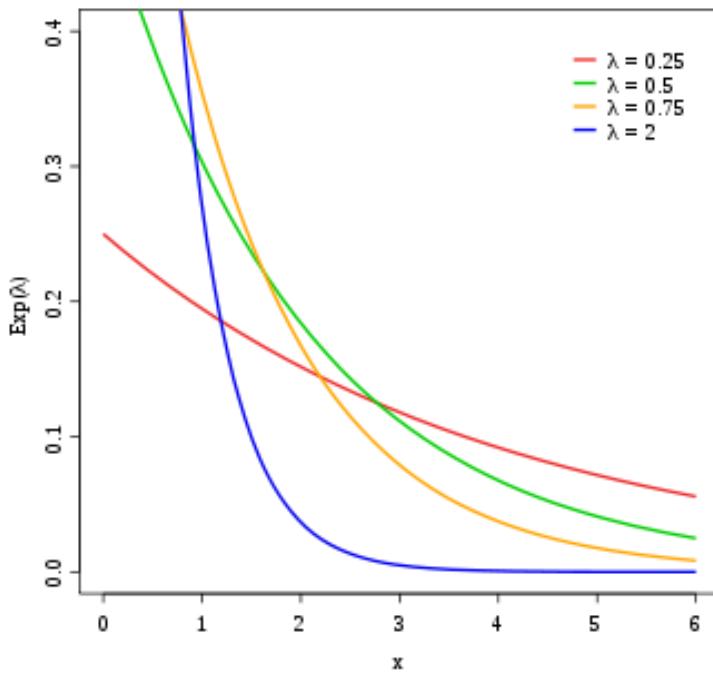


Abbildung 6: Veranschaulichung der Dichte der Exponentialverteilung für verschiedene Parameter

### 3.1.4 Pareto-Verteilung [8]

Die Pareto-Verteilung ist eine stetige Wahrscheinlichkeitsverteilung auf einem rechtsseitig unendlichen Intervall  $[x_{min}, \infty]$  mit Parametern  $\alpha > 1, x_{min} \geq 1$ . Sie ist skaleninvariant und genügt einem Potenzgesetz. Für den Spezialfall, dass  $x_{min} = 1$  (der für uns später relevant ist) wird sie folgendermaßen definiert; eine Zufallsvariable  $Z$  ist Pareto-verteilt mit Parameter  $\alpha > 1$ , falls sie die folgende Dichte besitzt:

$$f_\alpha(z) = \mathbb{1}_{\{z>1\}}(\alpha - 1) * z^{-\alpha} \quad (23)$$

*Lemma:*

Für die Dichte  $f_\alpha(z)$  gilt folgendes:

- $\alpha > 2 \iff \mathbb{E}[Z] < \infty$
- $\alpha > 3 \iff \text{Var}[Z] < \infty$

*Beweis:*

$$\begin{aligned} \mathbb{E}[Z] &= \int_1^\infty z * f(z) dz = \int_1^\infty z * (\alpha - 1) * z^{-\alpha} dz = \int_1^\infty (\alpha - 1) * z^{1-\alpha} dz = \\ &= (\alpha - 1) \int_1^\infty z^{1-\alpha} dz \implies (\mathbb{E}[Z] < \infty \iff \alpha > 2) \end{aligned} \quad (24)$$

$$\begin{aligned} \text{Var}[Z] &= \mathbb{E}[(Z - \mathbb{E}[Z])^2] = \mathbb{E}[Z^2] - \mathbb{E}[Z]^2 = \\ &= (\alpha - 1) \int_1^\infty z^{2-\alpha} dz - ((\alpha - 1) \int_1^\infty z^{1-\alpha} dz)^2 \implies (\text{Var}[Z] < \infty \iff \alpha > 3) \end{aligned} \quad (25)$$

□

Im Weiteren, für die ersten zwei Momente, sobald sie existieren, gilt folgendes:

$$\mathbb{E}[Z] = \frac{\alpha - 1}{\alpha - 2}, \text{Var}[Z] = \frac{\alpha - 1}{\alpha - 3} - \left(\frac{\alpha - 1}{\alpha - 2}\right)^2 \quad (26)$$

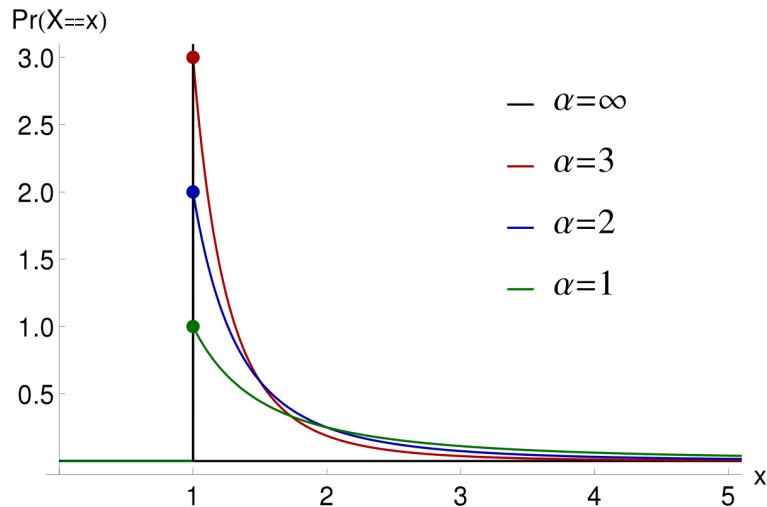


Abbildung 7: Veranschaulichung der Pareto-Dichte für verschiedene Parameter

## 3.2 Interpretation und Nachteile des SIR Modells

Das klassische SIR Modell erklärt vieles, aber bei seinem Entwurf wurden einige starke Annahmen gemacht. Es arbeitet mit Raten, die die Übergänge zwischen Compartments erklären. Obwohl diese Raten zum Teil die Interaktion zwischen Individuen aus verschiedenen Compartments gut erklären, vereinfachen sie aber die Situation in einem Aspekt.

Die Population in dem SIR wird als eine homogene Menge von Individuen betrachtet. Insbesondere, die Struktur des Modells impliziert, dass alle anfälligen Individuen in einem Zeitpunkt die gleiche Wahrscheinlichkeit haben, die Krankheit zu bekommen. D.h., es wird unter Gleichverteilung 'gewürfelt', welche anfällige Individuen in der Tat krank werden.

Wir wissen aber, dass die Akteure in einem epidemiologischen Modell selten auf die gleiche Weise mit allen anderen Akteuren interagieren, und dass dementsprechend das Verhältnis zwischen zwei Individuen schon eine Rolle in dem spielt, wie wahrscheinlich es ist, dass ein Individuum wegen der Interaktion mit dem Anderen krank wird. Hier wird insbesondere der geometrische Aspekt betont, also der Abstand zwischen je zwei Individuen.

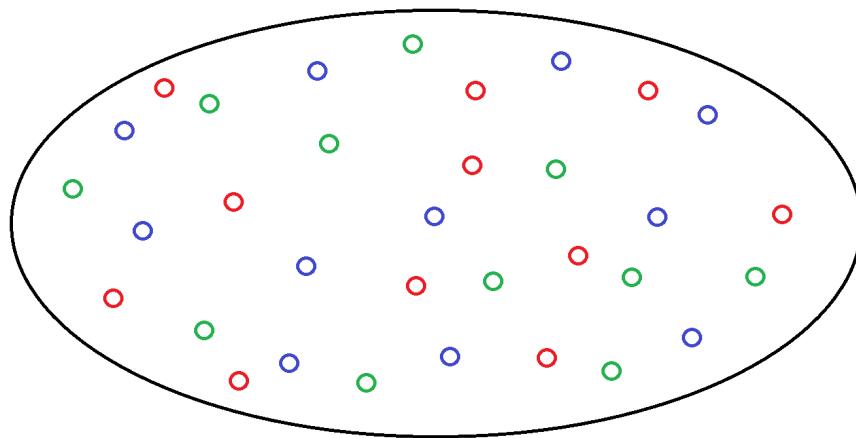


Abbildung 8: Veranschaulichung der homogenen Struktur des SIR Modells; Das Farbsystem ist z.B. S - blau, I - rot, R - grün

Dies führt uns zum eigentlichen geometrischen Modell der Simulation. Nämlich, wir definieren jedem Individuum der Population einen eindeutigen Standort. Die epidemiologisch relevanten Interaktionen werden dann auch von diesen Standorten abhängig sein.

### 3.3 Beschreibung des Modells

Die Simulation nimmt folgende Parameter als Eingabe:

$$n \in \mathbb{N}, \alpha > 1, \lambda_{meet} > 0, \lambda_{inc} > 0, \lambda_{inf} > 0, days \in \mathbb{N} \quad (27)$$

Die Population, in der wir eine Epidemie simulieren, wird als eine  $n^2$  Matrix dargestellt und wir nutzen die SEIR Struktur (2.2.3). Die Nachbarschaftsregeln sind wie auf einem Torus; die Person  $(n, n)$  hat z.B. folgende Nachbarn:  $(n - 1, n), (n, n - 1), (1, n), (n, 1)$ .

Jedem Individuum werden folgende Zufallszahlen **einmal** zugeordnet:

- $A \sim \lfloor Exp(\lambda_{inc}) \rfloor$ : die Dauer der E-Phase, einmal das Individuum infiziert wird
- $B \sim \lfloor Exp(\lambda_{inf}) \rfloor$ : die Dauer der I-Phase, einmal die E-Phase des Individuums vorbei ist

Es werden  $days$  Tage simuliert. Am Anfang der Simulation sind alle Individuen gesund außer eines: des sogenannten 'patient-zero'. In der Sprache der Compartment-Modelle heißt das, dass alle Individuen außer einem im S-Compartment sind. Den 'patient-zero' setzt man o.B.d.A. ins Zentrum der Matrix und man geht davon aus, dass er schon ansteckend ist. Also, er gehört zum I-Compartment.

Jedem Individuum wird die folgende Zufallszahl **jeden Tag** zugeordnet:

- $M \sim Poiss(\lambda_{meet})$ : Die Anzahl der anderen Individuen, die das Individuum an dem Tag besucht (Besuche interpretiert man als Interaktionen; nach einem Treffen eines S- und I-Individuums fängt für das Erste seine E-Phase an.)

Jedem Individuum werden folgende Zufallszahlen für **jeden der  $M$  Besuche** zugeordnet:

- $D \sim \lfloor Pareto(\alpha) \rfloor$ : Der Abstand zu dem zubesuchenden Individuum
- $W \sim Unif(\{0, 1, \dots, D\})$ : Der horizontale Teil der Bewegung für den Besuch
- $H \sim Unif(\{-1, 1\})$ : Die Orientierung des horizontalen Teils der Bewegung
- $V \sim Unif(\{-1, 1\})$ : Die Orientierung des vertikalen Teils der Bewegung

Somit befindet sich ein Individuum für einen bestimmten Besuch  $H * W$  nach rechts und  $V * (W - D)$  nach unten und trifft da eine Person.

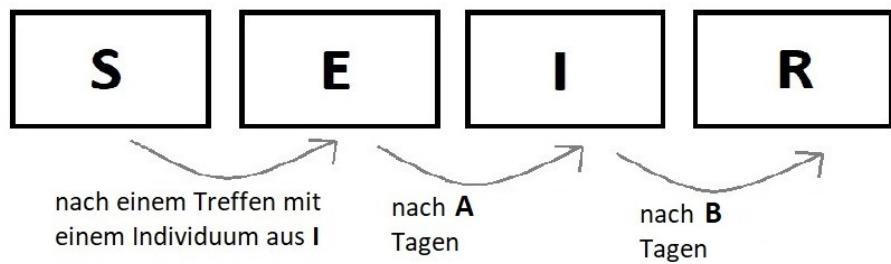
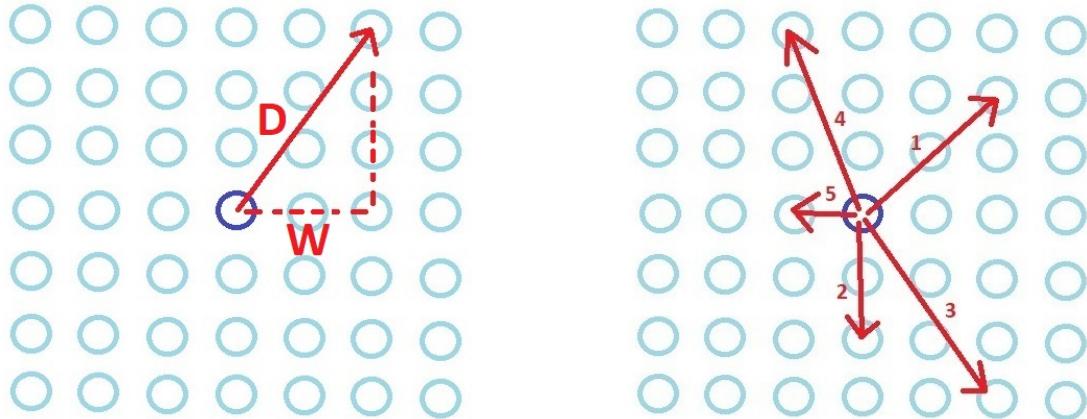


Abbildung 9: Skizze des SEIR-Mechanismus



(a) Skizze eines konkreten Besuchs;  
der Fall wenn  $D = 5, W = 2, H = 1, V = -1$

(b) Skizze aller Besuche eines Individuums an ei-  
nem Tag; der Fall wenn  $M = 5$

Abbildung 10: Skizze des räumlichen Mechanismus

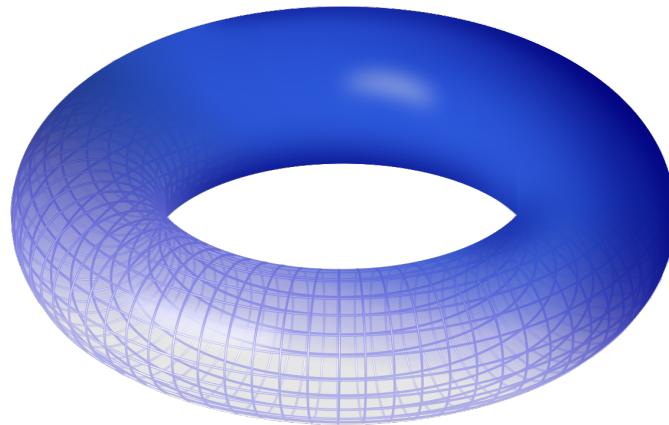


Abbildung 11: Illustration eines Torus [11]; man kann sich die räumliche Konfiguration der Population als die Fläche eines Torus vorstellen

## 3.4 Begründung der Wahl der Verteilungen

Im Modell benutzen wir vier Verteilungen:

- Gleichverteilung 3.1.1

Sie benutzen wir für Zufallszahlen, bei denen keine reine Präferenz vorliegt: für den horizontalen Anteil der Bewegung für einen Besuch und die Orientierungen der jeweiligen Bewegungskomponente

- Poisson-Verteilung 3.1.2

Sie benutzen wir für die Anzahl von Individuen, die ein Individuum in einem Tag besucht. Wenn man zufällige natürliche Zahlen modellieren muss, ist die Poisson-Verteilung oftmals eine plausible Wahl; sie hat nur einen Parameter und der Erwartungswert und die Varianz entsprechen genau diesem Parameter.

- Exponentialverteilung 3.1.3

Sie wird als Modell vorrangig bei der Beantwortung der Frage nach der Länge von zufälligen Zeitintervallen benutzt. Deswegen ist sie auch für unsere zwei Phasendauern eine gute und plausible Wahl.

- Pareto-Verteilung 3.1.4

Diese Verteilung nutzen wir um zu entscheiden, wie weit ein zu besuchendes Individuum ist. Die Hauptidee dahinter ist, große Abstände potenziert zu 'bestrafen'. Die Verteilung hat auch nur einen Parameter,  $\alpha > 1$ , durch ihn wird reguliert wie weit sich Individuen im Modell bewegen.

## 3.5 Mögliche Verallgemeinerungen

Es bieten sich natürlich einige Verallgemeinerungen, die man zusätzlich sinnvoll zum Modell einführen könnte:

- Die wahrscheinlich erste Verallgemeinerung die man erwarten kann ist, dass die Simulation nicht in diskreten Zeitpunkten funktioniert, sondern dass das Modell in stetiger Zeit gebaut wird.
- Man kann, außer einfach einer, eine breitere Klasse von Verteilungen für jeden Mechanismus in Betracht ziehen. Die Verteilungen, die oben genutzt werden, sind natürlich nicht die einzige natürliche Wahl.
- Die Annahme der sicheren und ewigen Immunität kann man auch aufheben und diese Eigenschaft auf eine elegante Weise stochastisch würfeln.
- Analog dazu, es ist möglich zu modellieren, dass das Übertragen der Krankheit nicht bei jedem Besuch erfolgt, sondern dass man stattdessen dieses Ereignis auch stochastisch definiert.

- Die 'Welt' in diesem Modell wird definiert als eine  $n * n$  Matrix,  $n \in \mathbb{N}$ , also eine Teilmenge von  $\mathbb{Z}^2$ . Man kann das Ganze diesbezüglich verallgemeinern und die 'Welt' als eine Teilmenge von  $\mathbb{Z}^m$ ,  $m \in \mathbb{N}$  definieren und alle Mechanismen entsprechend anpassen. Die Interpretation der Ergebnisse von solch einem Modell wäre aber vermutlich schwieriger.

### 3.6 Vermutungen

Uns interessiert vor allem der Einfluss zweier Variablen:

- $\alpha$ : Wie es schon in 3.1.4 gezeigt wurde, hat die dazugehörige Pareto-Verteilung genau dann eine endliche Varianz, wenn  $\alpha > 3$ . Dies kann man in unserem Modell so interpretieren, dass die Individuen kontrolliert nah zu ihrem Standort bleiben. In dem Fall soll auch die Verbreitung der Krankheit einigermaßen stabil erfolgen.
- $\lambda_{meet}$ : Diese Variable spielt im Modell der Simulation eine ähnliche Rolle wie die Reproduktionszahl in dem klassischen SIR Modell. Im Erwartungswert trifft ein Individuum genau  $\lambda_{meet}$  andere Individuen. Die Vermutung ist, dass die Krankheit aussterben soll, falls  $\lambda_{meet} \leq 1$ .

Also, wir betrachten prinzipiell die Ergebnisse davon, was passiert wenn wir 'ceteris paribus' die zwei oben erwähnten Parameter iteriert ändern (wobei wir natürlich interessante Marginalwerte berücksichtigen).

### 3.7 Graphische Ergebnisse [9]

Zunächst präsentieren wir einige Simulationsergebnisse für folgende Parametereinstellungen:  $n = 400$ ,  $\lambda_{meet} = 3$ ,  $\lambda_{inc} = 1$ ,  $\lambda_{inf} = 1$ ,  $days = 100$ . Es wird dreimal simuliert für 3 Wahlen für den 'Abstandspараметer':  $\alpha \in \{4, 3, 2.5\}$ . Das Farbsystem lautet: S - grün, E - gelb, I - rot, R - blau. Es werden nur gewählte Tage gezeigt, der Kompaktheit halber. Hier liegt der Fokus auf der Variable  $\alpha$ . Deswegen 'versichern' wir, dass die Krankheit nicht ausstirbt, indem wir ein 'sicheres'  $\lambda_{meet}$  wählen.

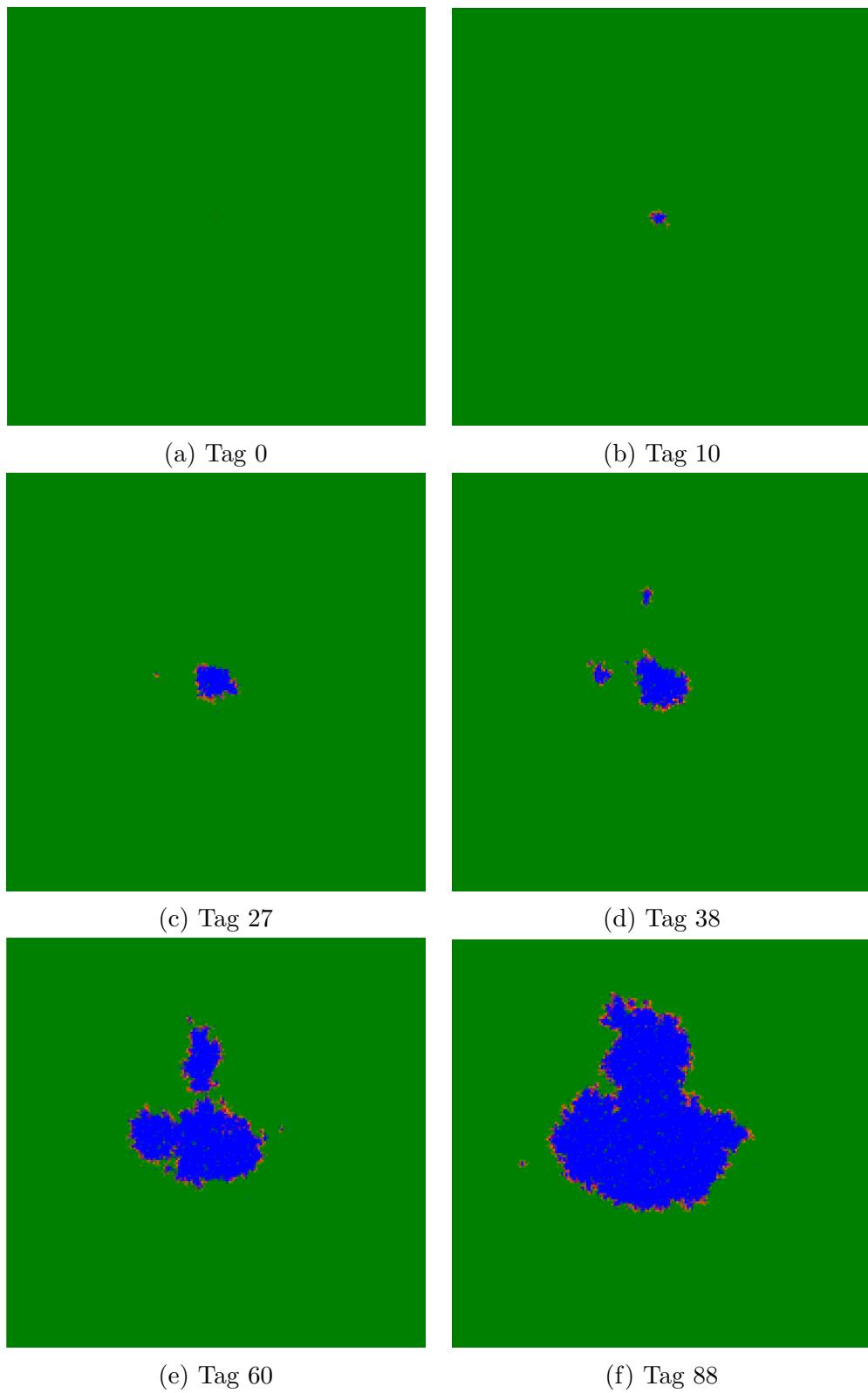
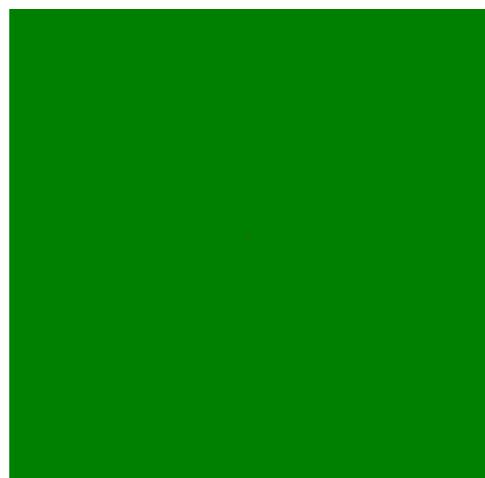
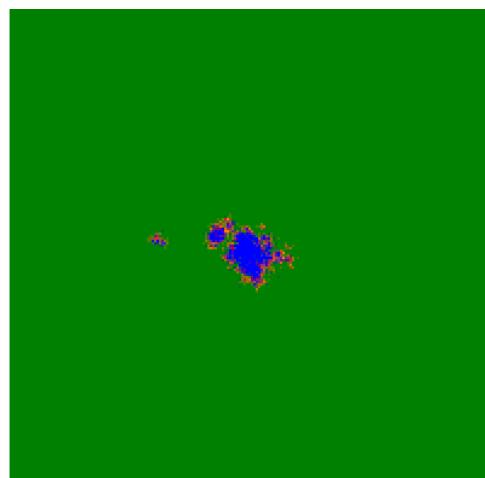


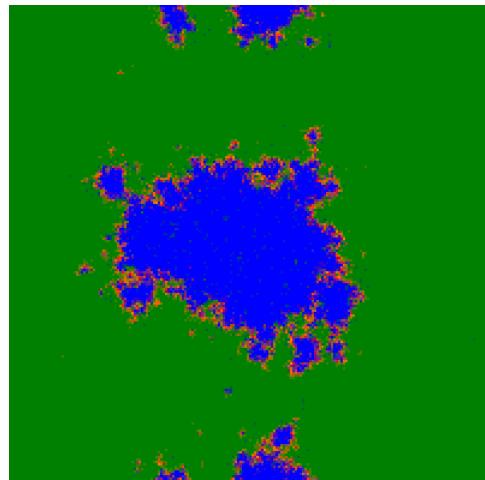
Abbildung 12: Simulationsergebnis für  $\alpha = 4$



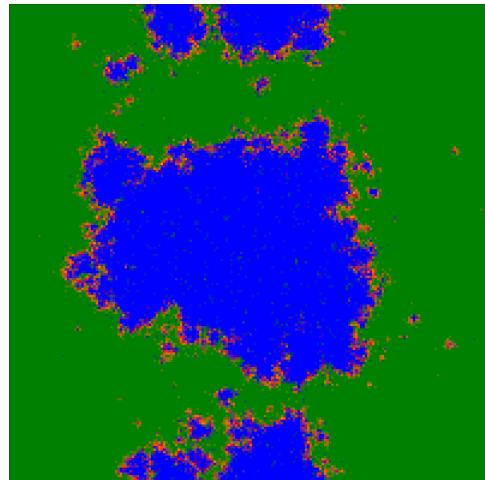
(a) Tag 0



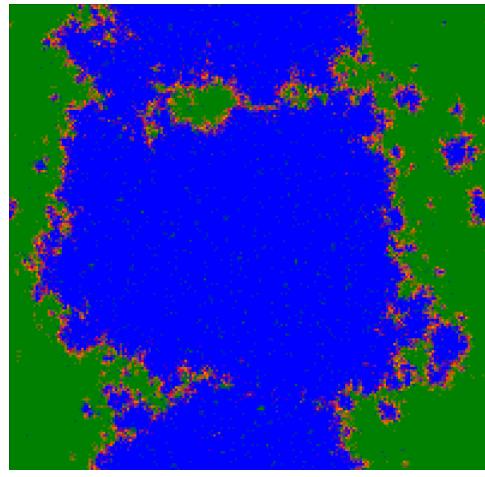
(b) Tag 21



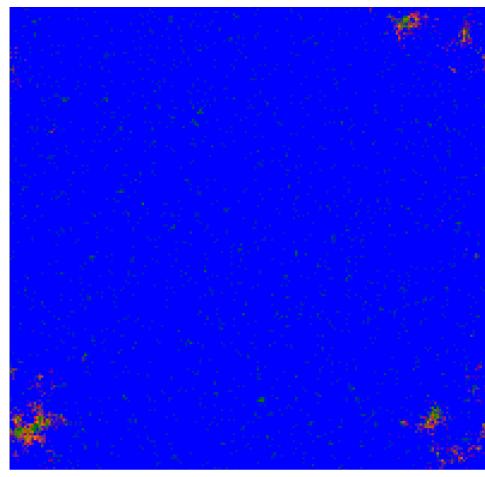
(c) Tag 50



(d) Tag 60



(e) Tag 75



(f) Tag 100

Abbildung 13: Simulationsergebnis für  $\alpha = 3$

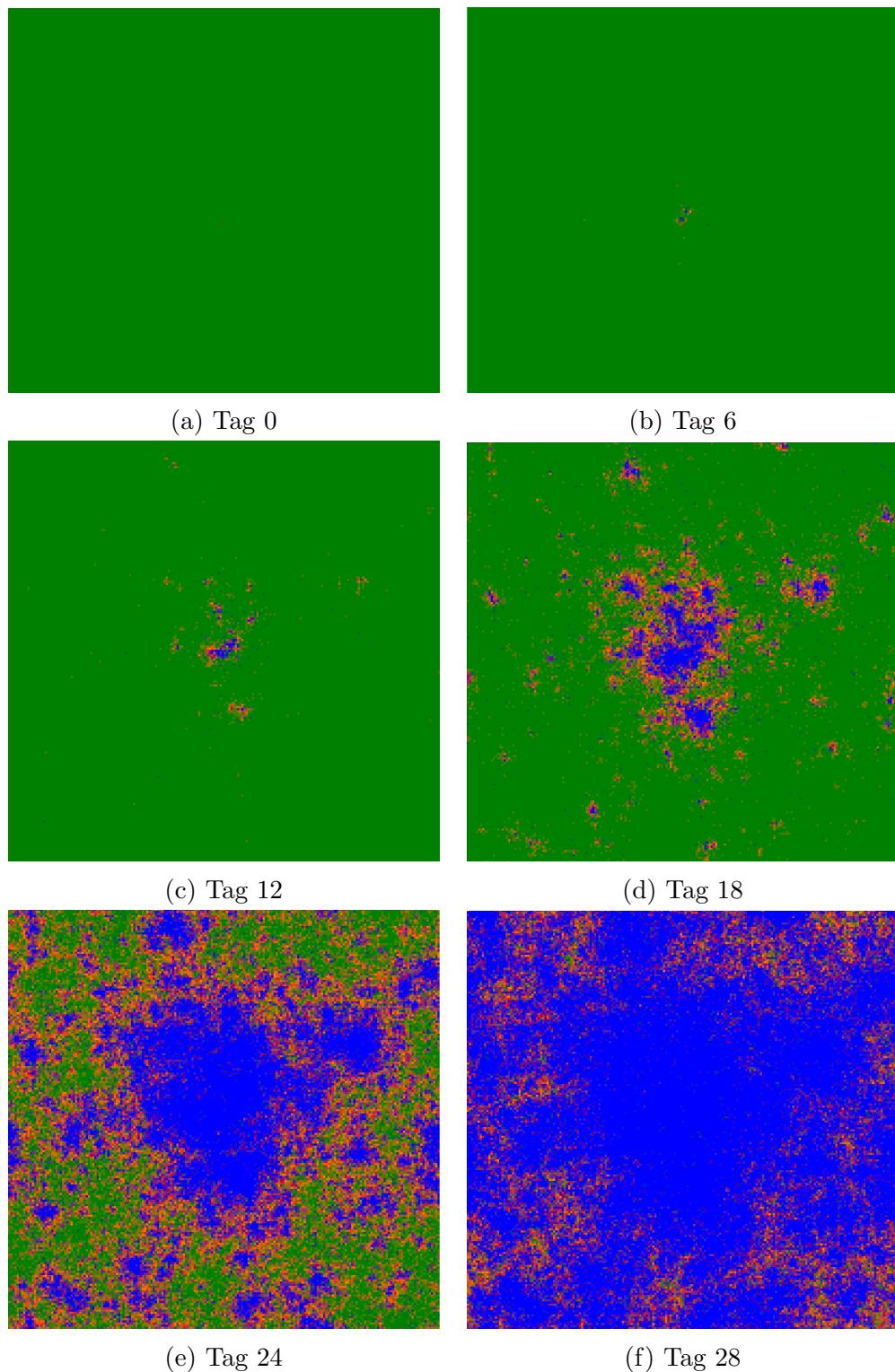


Abbildung 14: Simulationsergebnis für  $\alpha = 2.5$

## 3.8 Schlussfolgerungen I

Wir kommentieren jede Simulation separat:

- $\alpha = 4$ : Da  $\alpha$  wesentlich größer als 3 ist, sieht man, wie angekündigt, eine lokale Verbreitung der Krankheit.
- $\alpha = 3$ : In diesem marginalen Fall sieht man auch marginales Verhalten. Die Verbreitung sieht einigermaßen lokal aus, aber es entstehen auch neue Clusters.
- $\alpha = 2.5$ : Da  $\alpha$  wesentlich kleiner als 3 ist, sieht man, wie angekündigt, eine blitzschnelle globale Verbreitung der Krankheit.

Die restlichen Parameter (abgesehen von  $\lambda_{meet}$ ) beeinflussen nicht die 'Lokalität' der Verbreitung, nur die Geschwindigkeit mit der sie sich realisiert. Für diese Parameter machen wir folgende Bemerkungen:

- $n$ : Je größer  $n$  ist, desto größer ist die 'Welt' in der wir simulieren. Es ist natürlich wünschenswert, dass dieser Parameter groß ist, um die Plausibilität der Verbreitung zu erlauben. Andererseits, große  $n$ 's bringen mit sich auch große Initialisierungskosten.
- $\lambda_{inc}$ : Je länger die E-Phase ist (äq. wie je größer  $\lambda_{inc}$  ist), desto langsamer ist die Verbreitung, da es länger dauert, dass ein Individuum infektiös wird.
- $\lambda_{inf}$ : Je länger die I-Phase ist (äq. wie je größer  $\lambda_{inf}$  ist), desto schneller ist die Verbreitung, da Individuen länger infektiös bleiben.

Im nächsten Kapitel beschäftigen wir uns mehr mit dem Parameter  $\lambda_{meet}$ .

## 3.9 Quantitative Ergebnisse

Nun simulieren wir mehr als einmal, um einige asymptotische Resultate zu zeigen. Erstens beschäftigen wir uns mit 'großen'  $\alpha$ 's, bei denen die Verbreitung kontrolliert ist d.h. falls  $\alpha > 3$ . Danach betrachten wir was passiert, falls die Verbreitung flexibler ist, d.h. falls  $\alpha \leq 3$ .

### 3.9.1 Ergebnisse für 'große' $\alpha$ 's

Hier berücksichtigen wir Fälle wobei  $\alpha > 3$ . Es wird jeweils 100 mal simuliert und uns interessiert welcher Anteil der Population in irgendeinem Moment krank war. Deswegen schauen wir wer am Ende der jeweiligen Simulation zu den Compartments I oder R gehört. Es wird für folgende Parameter simuliert:

$$days = 500, n = 400, \lambda_{inc} = 1, \lambda_{inf} = 1, \alpha \in \{4, 5\}, \lambda_{meet} \in \{1, 2, 3\}, \quad (28)$$

Die Kommentare zum jeweiligen Histogramm der Simulation befinden sich unter dem jeweiligen Bild.

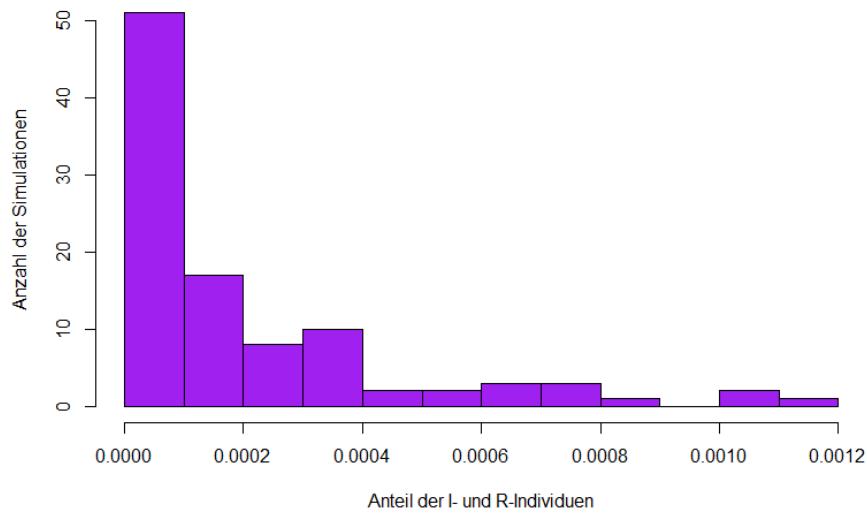


Abbildung 15: Simulationsergebnisse für  $\alpha = 4, \lambda_{Poisson} = 1$

In diesem Fall ist die Krankheit praktisch ausgestorben, fast niemand war krank.

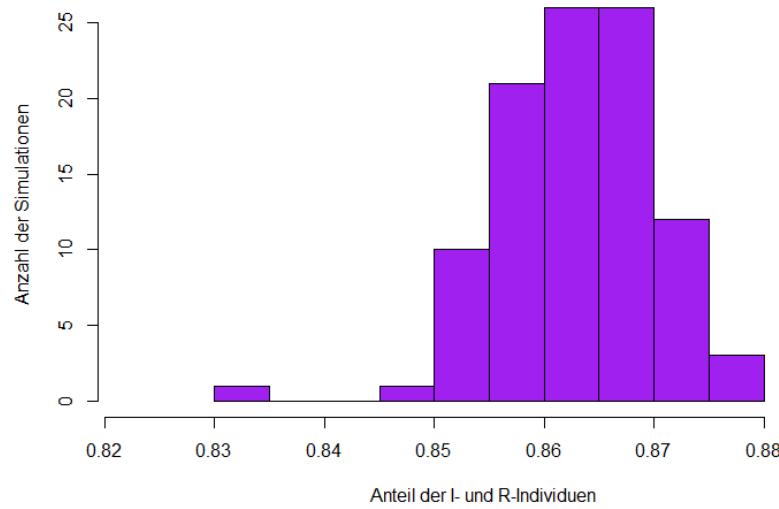


Abbildung 16: Simulationsergebnisse für  $\alpha = 4, \lambda_{Poisson} = 2$

Hier war ein wesentlicher Anteil der Population krank, vermutlich wären alle in I oder R, falls wir mehrere Tage simulieren würden.

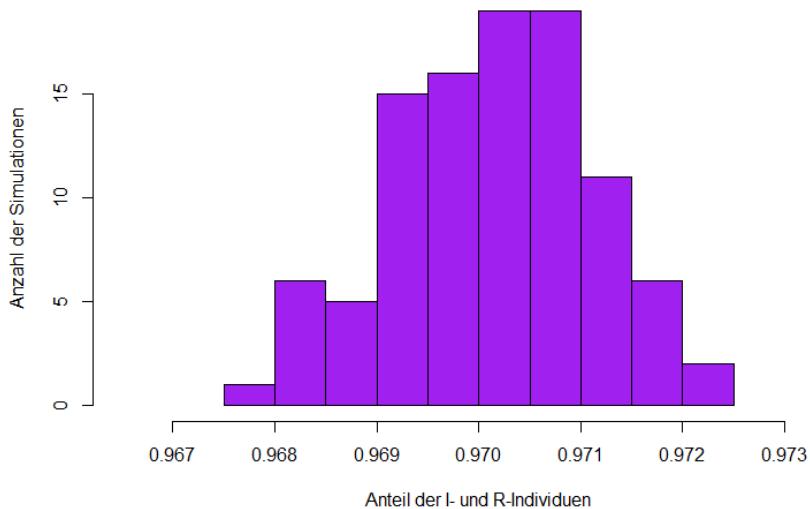


Abbildung 17: Simulationsergebnisse für  $\alpha = 4, \lambda_{Poisson} = 3$

In dieser Simulation sind praktisch alle Individuen krank worden.

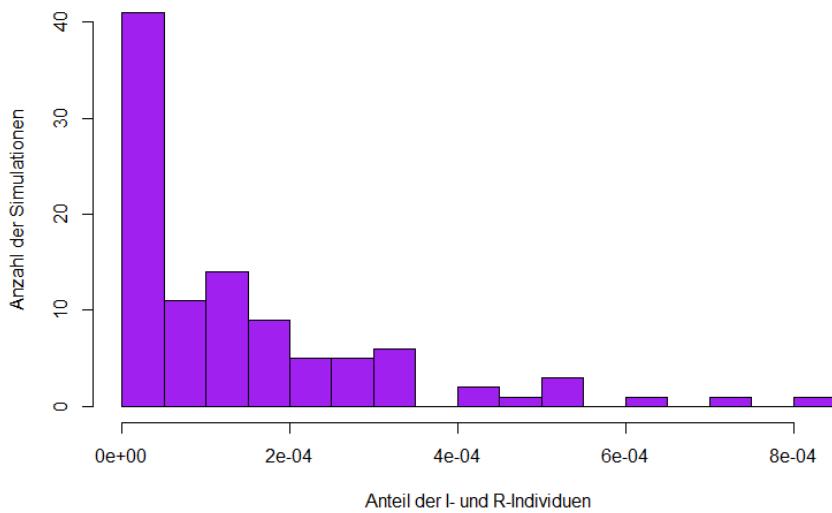


Abbildung 18: Simulationsergebnisse für  $\alpha = 5, \lambda_{Poisson} = 1$

Ähnlich wie im Fall für  $\alpha = 4, \lambda_{meet} = 1$ , ist effektiv niemand krank worden.

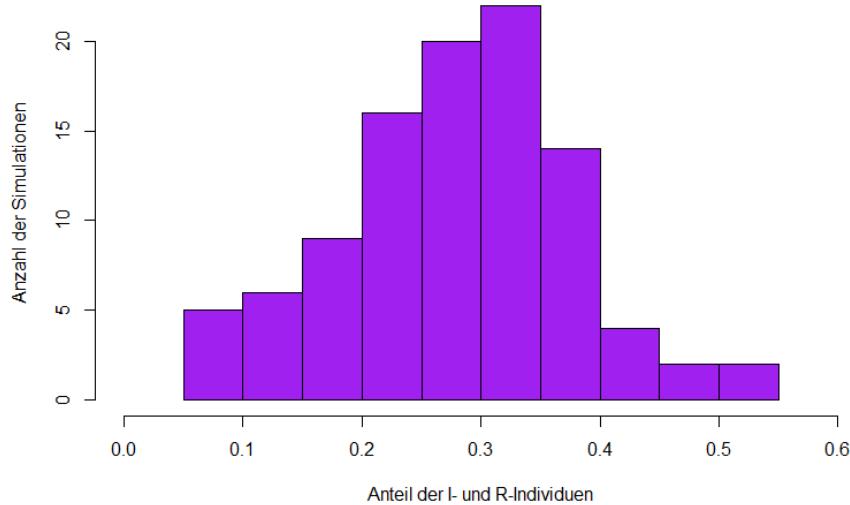


Abbildung 19: Simulationsergebnisse für  $\alpha = 5, \lambda_{Poisson} = 2$

Hier ist auch ein wesentlicher Anteil der Population krank worden, allerdings ist dieser Anteil kleiner als im Fall für  $\alpha = 4, \lambda_{meet} = 2$ . Man kann aber auch hier erwarten, dass alle Individuen früher oder später krank werden.

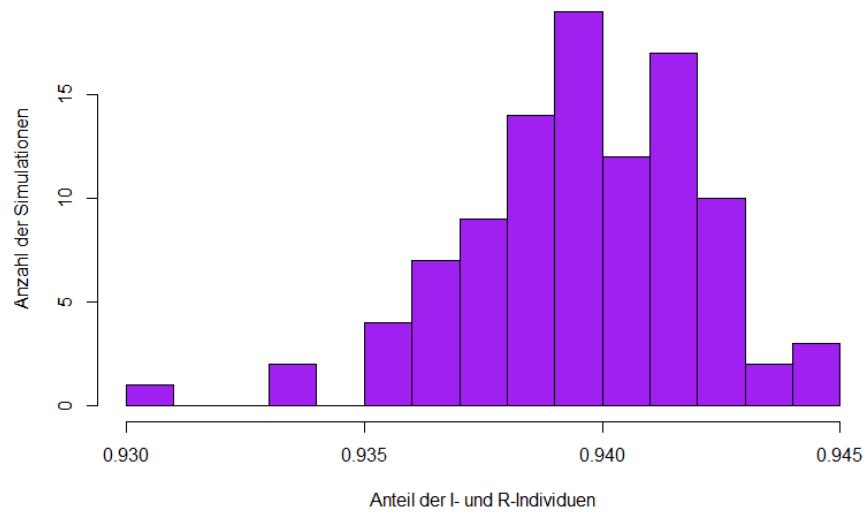


Abbildung 20: Simulationsergebnisse für  $\alpha = 5, \lambda_{Poisson} = 3$

Wieder, analog zum Fall für  $\alpha = 5, \lambda_{meet} = 3$  sind mehr oder weniger alle Individuen krank worden.

### 3.9.2 Ergebnisse für 'kleine' $\alpha$ 's

Nun simulieren wir Verbreitungen für  $\alpha < 3$ . In diesem Kapitel zeigen wir keine Histogramme, sonder eine Tabelle mit den wichtigsten statistischen Informationen. Die Größe der (I oder C) Clusters wird folgendermaßen definiert:

- ein kleiner Cluster: weniger als 5% der Population ist in I oder C
- ein mittelgroßer Cluster: zwischen 5% und 95% der Population ist in I oder C
- ein großer Cluster: mehr als 95% der Population ist in I oder C

Es wird für folgende Parameter 100 mal simuliert:

$$days = 500, n \in \{100, 400\}, \lambda_{inc} = 1, \lambda_{inf} = 1, \alpha \in \{2, 2.5\}, \lambda_{meet} \in \{1, 2\} \quad (29)$$

$n$	100				400			
$\alpha$	2.5		2		2.5		2	
$\lambda_{meet}$	1	2	1	2	1	2	1	2
<b>Kleine Clusters</b>	0.46	0.04	0.27	0.04	0.02	0.00	0.01	0.00
<b>Mittelgroße Clusters</b>	0.54	0.20	0.73	0.00	0.98	0.00	0.99	0.00
<b>Große Clusters</b>	0.00	0.76	0.00	0.96	0.00	1.00	0.00	1.00
<b>Emp. Erwartungswert</b>	0.30	0.92	0.50	0.93	0.53	0.95	0.69	0.96
<b>Emp. Standardabweichung</b>	0.28	0.19	0.31	0.19	0.08	0.00	0.07	0.00

## 3.10 Schlussfolgerungen II

In beiden Klassen von Ergebnissen sieht man einen starken Einfluss von dem Parameter  $\lambda_{meet}$ . Der Kompaktheit halber ordnen wir die Kommentare zur jeweiligen Klasse zu:

- Ergebnisse für 'große'  $\alpha$ 's 3.9.1:  
Es ist ersichtlich, dass es in diesem Fall entscheidend ist, ob  $\lambda_{meet} \leq 1$ . Der Abstandsparameter ist schon streng und verursacht eine kontrollierte Verbreitung. Wenn es also noch dazu der Fall ist, dass eine kranke Person nicht mehr als eine neue infiziert, stirbt diese Krankheit in der Regel aus.
- Ergebnisse für 'kleine'  $\alpha$ 's 3.9.2:  
Hier ist die Situation ein wenig komplizierter. Im Prinzip impliziert ein genug kleines  $\lambda_{meet}$  das Aussterben der Krankheit, allerdings kann ein genug kleines  $\alpha$  (hektische Verbreitung) diesen Effekt durchaus negieren. Auch, größere  $n$ 's bringen der 'Natürlichkeit' der Verbreitung bei.

Es ist in beiden Fällen essenziell einzusehen, dass das Zusammenspiel von den zwei besprochenen Parametern zentral ist, falls man die Verbreitung der Infektion in diesem Modell untersuchen will.

### 3.11 Diskussion

Kein Modell ist eine perfekte Widerspiegung der Realität, dies ist natürlich auch hier der Fall. Deswegen geben wir einige Limitationen an, an denen diese Simulation leidet.

- Es gibt weitere kompliziertere Zusammenhänge zwischen Individuen. Die genauen Regeln dafür, wie sich Menschen tagtäglich miteinander treffen sind fast undefinierbar, dementsprechend ist es ziemlich klar, dass wir im Sinne davon die Situation drastisch vereinfachen.
- Die für die Simulation definierte Welt ist nur pseudo-unendlich. D.h., nach  $n$  Schritten in einer Richtung befindet sich man immer noch am gleichen Ort. Obwohl dies, streng genommen, auch für unsere Welt gilt, ist diese Modellierung im epidemiologischen Kontext nicht die absolut ideale Wahl.
- Die Individuen in der Simulation werden dauerhaft immun. Da diese medizinische Problematik noch immer diskutiert wird und da es für z.B. COVID-19 schon Fälle gab, dass man zum zweiten Mal krank worden ist, könnte man diese Restriktion auch aufheben und diese Eigenschaft sinnvoll stochastisch modellieren.
- Prozesse in der Natur laufen kontinuierlich, unser Modell ist diskret. Die Zeit ist nicht diskretisiert, jede Beobachtung die wir machen oder simulieren ist nur ein zeitliches 'Snapshot' eines Objektes, welches ein Subjekt stetiger Änderung ist. Daher sind auch die Ergebnisse dieser Simulation 'cum grano salis' zu nehmen.

## Inhalt der beigelegten CD

Auf der beigelegten CD befindet sich eine digitale (PDF) Version dieser Arbeit als auch eine C++ Datei - die eigentliche Simulation.

## Literatur

- [1] W. O. Kermack and A. G. McKendrick, “Contributions to the mathematical theory of epidemics–i. 1927.” *Bulletin of mathematical biology*, vol. 53, no. 1-2, pp. 33–55, 1991.
- [2] H. W. Hethcote, “The mathematics of infectious diseases,” *SIAM review*, vol. 42, no. 4, pp. 599–653, 2000.
- [3] Wikipedia - Compartmental models in epidemiology. [Online]. Available: [en.wikipedia.org/wiki/Compartmental\\_models\\_in\\_epidemiology](https://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology)
- [4] T. Harko, F. S. Lobo, and M. Mak, “Exact analytical solutions of the susceptible-infected-recovered (sir) epidemic model and of the sir model with equal death and birth rates,” *Applied Mathematics and Computation*, vol. 236, pp. 184–194, 2014.
- [5] Wikipedia - Gleichverteilung. [Online]. Available: <https://de.wikipedia.org/wiki/Gleichverteilung>
- [6] Wikipedia - Poisson distribution. [Online]. Available: [https://en.wikipedia.org/wiki/Poisson\\_distribution](https://en.wikipedia.org/wiki/Poisson_distribution)
- [7] Wikipedia - Exponentialverteilung. [Online]. Available: <https://de.wikipedia.org/wiki/Exponentialverteilung>
- [8] Wikipedia - Pareto-Verteilung. [Online]. Available: <https://de.wikipedia.org/wiki/Pareto-Verteilung>
- [9] C. Sorg, “Verbreitung von Epidemien (Bachelorarbeit),” 2021.
- [10] Geogebra SIR Visualization. [Online]. Available: <https://www.geogebra.org/m/nbjfjtpv>
- [11] Wikipedia - Torus. [Online]. Available: <https://de.wikipedia.org/wiki/Torus>
- [12] H. W. Hethcote, “Qualitative analyses of communicable disease models,” *Mathematical biosciences*, vol. 28, no. 3-4, pp. 335–356, 1976.

## **Ehrenwörtliche Erklärung**

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Arbeit selbständig angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

---

München, 15. September 2021

Ort und Datum

---

Unterschrift (Vjekoslav Drvar)