

Seminar:
Statistische Methoden in der
Infektionsepidemiologie
Festlegung von Impfstrategien

Florian Hiemeyer
hieflo@stat.uni-muenchen.de

11. Dezember 2003

Zusammenfassung

Behandelt werden die Möglichkeiten zur Einbindung von Impfstrategien in das stochastische *SIR*-Modell, zur Modellierung von Heterogenität in Impfmodellen und zur Schätzung der Impfwirksamkeit (bei Individuen).

Inhaltsverzeichnis

1	Einführung	2
1.1	Impfungen in Deutschland	2
1.2	Sinn und Zweck von Impfstrategien	3
1.3	Überblick	3
2	Impfung im SIR: Grundlagen	3
2.1	Definitionen und Begriffe	3
2.2	Standard <i>SIR</i> Modell mit Impfung	4
2.3	Schätzungen von Konfidenzintervallen für v_c	5
3	Heterogenität in Impfmodellen	7
3.1	Problemstellung und Umsetzung	7
3.2	Überprüfung einer Impfstrategie	9
3.3	Simulationen von Haushaltsstrukturen	11
3.4	Auswirkungen falscher Impfstrategie-Annahmen	12
4	Schätzung der Impfwirksamkeit r	13
5	Literatur	15

1 Einführung

1.1 Impfungen in Deutschland

In Deutschland ist die **ständige Impfkommission** (STIKO) des Robert-Koch-Instituts für Impfeempfehlungen zuständig. Generell gibt es in Deutschland (anders als z. B. in den Vereinigten Staaten) allerdings keine Impfpflicht.

Aufgrund medizinischer Indikationen und zugleich in Betrachtung der drohenden direkten und indirekten Kosten eines Epidemieausbruchs, werden von der STIKO Impfeempfehlungen ausgesprochen. Diese beschreiben welche Personengruppen gegen welche Krankheiten geimpft sein sollten.

Für die folgenden Krankheiten gibt es Impfeempfehlungen:

- Cholera
- Diphtherie
- FSME (Meningoenzephalitis)
- Gelbfieber
- Hepatitis A/B
- Influenza (Grippe)
- Mumps/Masern/Röteln („MMR“)
- Pneumokokken
- Poliomyelitis (Kinderlähmung)
- Tetanus
- Tollwut
- Tuberkulose
- Typhus
- Varizellen (Windpocken/Gürtelrose)

Von den, mittlerweile in Europa ausgerotteten, Pocken und der Pest abgesehen, sind in dieser Liste also zahllose „Menschheitsgeißeln“ versammelt. Doch auch die „harmloseren“ Infektionskrankheiten wie die Grippe darf man nicht unterschätzen: Diese Krankheiten richten als Epidemien oft enormen ökonomischen Schaden durch Arbeitsausfälle und medizinische Versorgung an.

Vorwiegend ökonomische Probleme im Zusammenhang mit Impffragestellungen treten auch in der Agrarwirtschaft und der Computertechnik (dort: Impfung = „Virens scanner aktiviert“) auf. In diesem Seminarbericht wird die Betrachtung ökonomischer Aspekte jedoch außer Acht gelassen.

1.2 Sinn und Zweck von Impfstrategien

Impfstrategien in stochastischen Modellen werden hauptsächlich entwickelt, um

1. aus den Informationen einer Epidemie den mindestens zu impfenden Bevölkerungsanteil zu schätzen, ab dem ein großer Ausbruch unmöglich ist.
2. die Auswirkungen von Heterogenität in der Bevölkerung auf die Wirksamkeit der Impfung (zur Vermeidung einer Epidemie) zu analysieren.
3. die Analyse der Wirksamkeit (auf das einzelne Individuum) einer neu-entwickelten Impfung zu schätzen.

Dabei ist die Theorie der Impfstrategien relativ leicht mit dem – um Heterogenität erweiterten – *SIR*-Modell zu erklären.

1.3 Überblick

Der Vortrag beginnt daher mit der Einführung geimpfter und ungeimpfter Anfälliger in einem Multityp-*SIR*-Modell. Zentrale Aufgabe ist hierbei die Schätzung des kritischen Impfanteils v_c .

Einige Simulationen zum Thema „Impfstrategien in Haushalten“ sollen dann verdeutlichen, dass auch Bevölkerungsstrukturen mit Impfstrategien erfassbar sind. Gleichzeitig zeigen einige Simulationen, in welcher Weise falsche Annahmen vor dem Impfen sich auf den Gesamtschutz der anfälligen Bevölkerung auswirken.

Im letzten Abschnitt werden dann Möglichkeiten zum Schätzen der Wirksamkeit einer Impfung (beim einzelnen Individuum) erläutert.

2 Impfung im SIR: Grundlagen

2.1 Definitionen und Begriffe

In einer Epidemie mit *SIR* Struktur, $E_{n,m}(\lambda, I)$, sind zu Beginn der Epidemie m Infizierte und n Anfällige. Kontakte eines Anfälligen zu einem spezifischen Infizierten finden exponentialverteilt mit der Intensität λ/n statt, und eine Verteilung I entscheidet über die Länge der Infektion bei einem Infizierten.

Aus den Daten der Epidemie lässt sich dann die *Basis-Reproduktionsrate* R_0 bzw. θ schätzen. Dies ist sowohl für eine vollständig beobachtete Epidemie als auch für eine nur in der Endgröße beobachtete Epidemie möglich.

In der (großen) Bevölkerung sind einige Anfällige geimpft und andere nicht. Der Anteil der Geimpften gegenüber allen Anfälligen sei v .

Die Impfung reduziert die Übertragungsrate für Geimpfte auf $\lambda(1 - r)/n$

gegenüber λ/n für die ungeimpften Anfälligen. Damit ist r die relative Verringerung der Übertragungsrate der Krankheit, auch als *Impfwirksamkeit* bezeichnet (zunächst wird sie als bekannt vorausgesetzt). Einen besonderen Spezialfall stellt dabei die perfekte Impfung, mit 100prozentiger (und lebenslanger) Immunität bei $r = 1$ dar.

Aufgrund der aus einem Krankheitsausbruch bekannten Daten soll nun der kritische Impfanteil v_c der Anfälligen bestimmt werden, bei dessen Überschreitung die Bevölkerung vor zukünftigen, großen Ausbrüchen geschützt ist.

Diesen Zustand bezeichnet man als *Herdenimmunität*.

2.2 Standard *SIR* Modell mit Impfung

Um geimpfte und ungeimpfte Anfällige in das *SIR*-Modell einzubeziehen, benötigt man das Multityp-Modell (siehe auch den Bericht von Marietta Rottenkolber). Dabei werden zwei Gruppen modelliert ($k=2$). Die Gruppenanteile betragen $\pi_1 = 1 - v$ und $\pi_2 = v$. Dann ist π_2 der Anteil der Geimpften an der Bevölkerung. Die Übertragungsraten für die beiden Gruppen betragen:

$$\lambda_{11} = \lambda_{21} = \lambda \text{ und } \lambda_{12} = \lambda_{22} = \lambda(1 - r). \quad (1)$$

Die Impfung senkt (als Annahme) nur das Risiko infiziert zu werden, die Infektivität und Infektionsdauer im Falle einer Infektion wird für beide Gruppen als gleich angenommen, also auch $\iota_1 = \iota_2 = \iota$. Der mindestens zu impfende Anteil v_c („kritischer Impfanteil“) der Anfälligen lässt sich dann nach dem Konzept des Multityp-Modells (im Fall $k=2$) berechnen. Dazu bestimmt man zunächst die Formel für die Basis-Reproduktions-Rate als

$$R_v = \underbrace{\iota\lambda}_{\theta}(1 - v) + \underbrace{\iota\lambda(1 - r)}_{\theta'}v \quad (2)$$

Dabei sollen θ und θ' verdeutlichen, dass es sich bei der Formel um ein gewogenes Mittel der Basisreproduktionsraten der Geimpften und Ungeimpften handelt

Ziel der Impfstrategie ist natürlich, mittels der Größe von v die Bedingung $R_v \equiv \theta \leq 1$ zu erreichen. Durch Auflösen zeigt man, dass gilt:

$$v_c = \frac{1}{r} \left(1 - \frac{1}{\theta} \right) \quad (3)$$

Dann erfüllt $E_{n,m}(\lambda, I)$ die sogenannte *Herdenimmunität*, falls $v \geq v_c$.

Sollte $v_c > 1$ gelten, das heisst es wären mehr als 100% der Anfälligen zu impfen, um einen größeren Ausbruch zu verhindern, dann kann Herdenimmunität nicht erreicht werden.

Im Spezialfall einer perfekten Impfung gilt: $v_c = 1 - \frac{1}{\theta}$. Dann ist $v_c < 1$ sicher erfüllt, und v_c nähert sich der Eins nur, falls θ sehr groß ist (asymptotisch). Dies ist auf Abbildung 1 deutlich zu sehen.

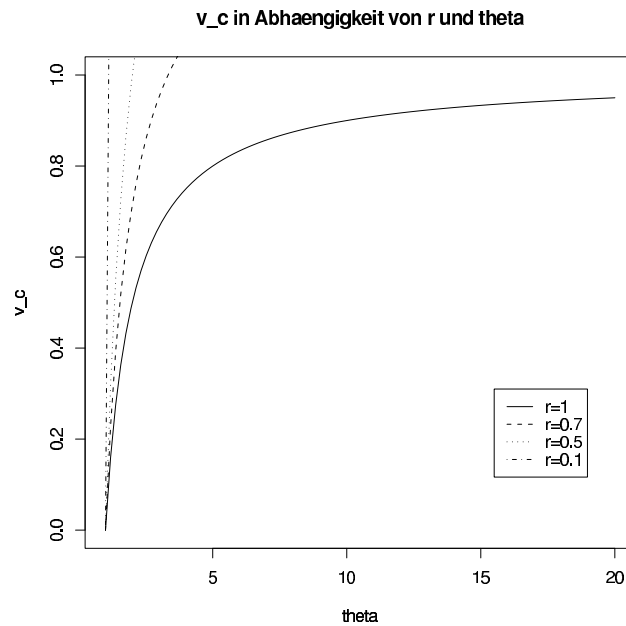


Abbildung 1: Der kritische Impfanteil v_c , als Funktion von r und θ (im Wertebereich $[0,1]$)

Für geringere Werte von r sieht die Kurve ähnlich aus, die Funktionswerte überschreiten dann aber für gewisse θ den Wert $v_c = 1$: So ist wenn $r = 0.7$ der Wert $v_c = 1$ für $\theta = 3, \bar{3}$ erreicht, bei $r = 0.5$ bereits für $\theta = 2$. Im Falle $r = 0.1$ liegt der erforderliche Impfanteil dann schon für eine sehr geringe Bruttoreproduktionsrate ($\theta = 1, \bar{1}$) bei 1, siehe Abbildung 1.

2.3 Schätzungen von Konfidenzintervallen für v_c

Um einen Schätzer und das zugehörige Konfidenzintervall für v_c zu erhalten, sind aus den Daten für eine Epidemie zunächst ein ML-Schätzer für θ und der dazugehörige Standardfehler $s.e.(\hat{\theta})$ zu bestimmen. Zu beachten sind hierbei allerdings wieder die unterschiedlichen Beobachtungsmöglichkeiten: Vollständige Beobachtung und Beobachtung der Endgröße (siehe das Referat von Christine Hiebl). Die Impfwirksamkeit r wird weiterhin als bekannt und konstant vorausgesetzt.

Aufgrund der Invarianzeigenschaft des ML-Schätzers für θ lässt sich dann auch ein Schätzer für v_c bestimmen:

$$\hat{v}_c = \frac{1}{r} \left(1 - \frac{1}{\hat{\theta}} \right) \quad (4)$$

Der Schätzer für v_c nimmt Werte auf ganz \mathbb{R} an.

Für den Standardfehler $s.e.(\hat{v}_c)$ gilt dann (Anwendung der δ -Methode):

$$s.e.(\hat{v}_c) = \frac{s.e.(\hat{\theta})}{r\hat{\theta}^2} \quad (5)$$

Für \hat{v}_c und den zugehörigen Standardfehler gelten im Falle

- vollständiger Beobachtung, mit $\hat{\theta} = \hat{\lambda}\hat{\gamma}^{-1}$:

$$\hat{v}_c = \frac{1}{r}(1 - \hat{\gamma}\hat{\lambda}^{-1}) \quad (6)$$

$$s.e.(\hat{v}_c) = \frac{1}{r\hat{\theta}} \sqrt{\frac{\bar{Z}^{-1} + (\bar{Z} + \mu)^{-1}}{n}} \quad (7)$$

- der Beobachtung der Endgröße, hierbei können zwei Fälle unterschieden werden:

- die exakte Berechnung, mit $\hat{\theta} = \frac{\frac{1}{n} + \frac{1}{n-1} + \dots + \frac{1}{n-(Z-1)}}{\bar{Z} + \mu}$:

$$\hat{v}_c = \frac{1}{r} \left(1 - \frac{\bar{Z} + \mu}{\frac{1}{n} + \frac{1}{n-1} + \dots + \frac{1}{n-(Z-1)}} \right) \quad (8)$$

- die Approximation, mit $\hat{\theta} \approx -\log(1 - \bar{Z})/(\bar{Z} + \mu)$:

$$\hat{v}_c \approx \frac{1}{r} \left(1 + \frac{(\bar{Z} + \mu)}{\log(1 - \bar{Z})} \right) \quad (9)$$

Für beide Berechnungsmethoden bei Beobachtung der Endgröße gilt:

$$s.e.(\hat{v}_c) = \frac{1}{r\sqrt{n}(\bar{Z} + \mu)\hat{\theta}^2} \sqrt{\frac{\bar{Z}}{1 - \bar{Z}} + \hat{\theta}^2(\bar{Z} + \mu)}. \quad (10)$$

Aus den Standardfehlern und der asymptotischen Normalität von \hat{v}_c läßt sich ein Konfidenzintervall für \hat{v}_c konstruieren. Dieses Intervall läßt sich am sinnvollsten als *einseitiges Konfidenzintervall für v_c* schätzen, denn eine Überschätzung des kritischen Impfteils hat weniger schwerwiegende Folgen als eine Unterschätzung.

Man schätzt also die obere Intervallgrenze (G_o) so, dass man ein einseitiges Konfidenzintervall für v_c zum Niveau α erhält:

$$G_o = \hat{v}_c + z_{1-\alpha} s.e.(\hat{v}_c), \quad KI : (-\infty; G_o].$$

Interpretation (jeweils zur Sicherheitswahrscheinlichkeit $1 - \alpha$):

- Ist $G_o > 1$, dann lässt sich Herdenimmunität nicht erreichen.
- Ist $G_o \leq 0$, dann muß niemand geimpft werden, da dann $\theta \leq 1$.
- Gilt $0 < G_o \leq 1$ lässt sich Herdenimmunität erreichen, wenn mindestens der Anteil G_o der Anfälligen geimpft ist.

Beispiel 1: Verwendet werden die Daten und Ergebnisse des Epidemieausbruchs auf „Tristan de Cunha“ (siehe Vortrag von Christine Hiebl).

Gegeben: $\hat{\theta} = 1.056$, $s.e.(\hat{\theta}) = 0.2381$, $n = 254$, $m = 1$, $Z = 39$. Hier zunächst die Ergebnisse für die exakte Berechnung von \hat{v}_c aus den Formeln (8) und (10):

- r allgemein: $\hat{v}_c = 0.05303/r$, $s.e.(\hat{v}_c) = 0.21352/r$
- $r = 1$: $\hat{v}_c = 0.05303$, $s.e.(\hat{v}_c) = 0.21352$
- $r = 0.5$: $\hat{v}_c = 0.10606$, $s.e.(\hat{v}_c) = 0.4270$

Für die Konfidenzintervallgrenze G_o gilt für $\alpha = 0.05$: $G_o = 0.35231/r$. Bereits bei $r = 1$ muss dann mit $G_o = 0.35231$ über ein Drittel der Bevölkerung geimpft werden, um mit 95%-iger Sicherheit den wahren Wert von v_c zu überdecken! Für $r = 0.5$ ist mit $G_o = 0.70462$ zum selben α dann schon fast 75% der Bevölkerung zu impfen.

Bei Verwendung der Näherungsformel für \hat{v} aus Formel (9) gilt mit $\hat{\theta} \approx 1.0585$: $\hat{v}_c \approx 0.05527/r$, $s.e.(\hat{v}_c) \approx 0.12726/r \Rightarrow$ für $n=254$ mit der Formel (10) noch schlechte Näherung des Standardfehlers!

3 Heterogenität in Impfmodellen

3.1 Problemstellung und Umsetzung

Im *SIR*-Modell mit Impfung lässt sich – neben der Impfung – natürlich auch noch weitere Struktur bzw. weitere Gruppen einführen. Dabei ist besonders auch das Haushaltsmodell interessant, dass von Marietta Rottenkolber in ihrem Seminarvortrag bereits kurz angesprochen wurde.

Um das Haushaltsmodell kombiniert mit Impfungen zu Simulationszwecken einzuführen, erweist sich bereits die Sellke-Konstruktion im Multityp-Modell als relativ kompliziert: denn die Zahl verschiedener Gruppen steigt rapide an, gleichzeitig muss auch die Veränderung der Anfälligkeit durch die Impfung implementiert sein. Zudem erschwert die reine Fixierung auf die bloßen verschiedenen Kontaktparameter der einzelnen Gruppen das Verständnis des Modells.

Deswegen habe ich, in der Programmiersprache *Java*, eine echt objektorientierte Simulationsmöglichkeit implementiert, die theoretisch beliebige

strukturelle Komplexität unter den Anfälligen (und den Infektiösen) ermöglicht. Programmiert wurde direkt nach der Definition des *SIR*-Modells (S. 11, Andersson).

Auch ermöglicht die objektorientierte Implementierung die Verwendung anderer – geeigneter – Verteilungen als die Exponentialverteilung für die Simulation der Kontakte mit Infektiösen an. So bietet sich bei erstem naiven Nachdenken zum Beispiel die Gamma-Verteilung an. Diese Idee habe ich aber nicht weiterverfolgt.

Es können nicht nur Objekte wie Haushalte oder gar Sportvereine erzeugt werden, sondern auch Attribute eingeführt werden, deren verschiedene Kombination (weiblich, Kind) etc. zu verschiedener Anfälligkeit, führen kann. Prinzipiell wäre auch die Implementierung von unterschiedlicher Infektiosität oder unterschiedlichen Infektionsdauerverteilungen möglich. Damit sind auch die Individuen (und die Strukturen in der Bevölkerung) keine unabhängig in verschiedenen Vektoren stehende drögen Zahlenreihen, sondern können auch textuell und erklärend auf dem Bildschirm ausgegeben werden. Hier ein Beispiel für ein Individuum meiner Simulation:

```
Id: 54 Inf: true, Rem: false, Vac: false, I: 2.4582713141147714,
localInt: 2.0, globInt: 1.0
```

Dieses Individuum hat die Identifikationsnummer 54, ist infiziert und noch nicht wieder genesen, nicht geimpft und hat noch eine bestimmte Zeit in der es infektiös ist. Gegenüber Anfälligen in seinem Haushalt hat es eine Infektionsintensität von 2.0 und gegenüber allen anderen eine Intensität von 1.0. Ähnlich lassen sich auch Haushalte und andere Strukturen ausgeben.

Ein großer Arbeitsanteil zur Bestimmung der individuellen Anfälligkeitsparameter λ und der Ermittlung der Infektionsdauer aus I kann dabei auf die Initialisierung der Bevölkerung verschoben werden. Damit kann der eigentliche „Epidemiealgorithmus“ des Programms verhältnismäßig knapp gehalten werden.

Der Algorithmus der Epidemie besteht darin, aus der Gesamtbevölkerung der Anfälligen zunächst den Zeitpunkt des nächsten Ansteckungs- und des nächsten Genesungsereignisses zu bestimmen, und dann aus diesen beiden schlußendlich das tatsächlich eintretende Ereignis. Dieses wird ausgeführt und das betroffene Individuum infiziert/entfernt. Anschließend erfolgt erneut wieder die Suche nach dem nächsten Ereignis.

Der Abbruch der Epidemie erfolgt – definitionsgemäß – sobald keine Infizierten mehr in der Bevölkerung vorhanden sind.

Der letzte Vorteil ist die Tatsache, dass sich mit dieser Simulation nicht nur die Endgröße, sondern auch der zeitliche Verlauf der Zahl der Infizierten und der Anfälligen bestimmt werden kann.

Der große Nachteil der objektorientierten *Java* Implementation des Haushaltsmodells ist die lange Laufzeit, die eine Simulation mit vielen Durchläufen

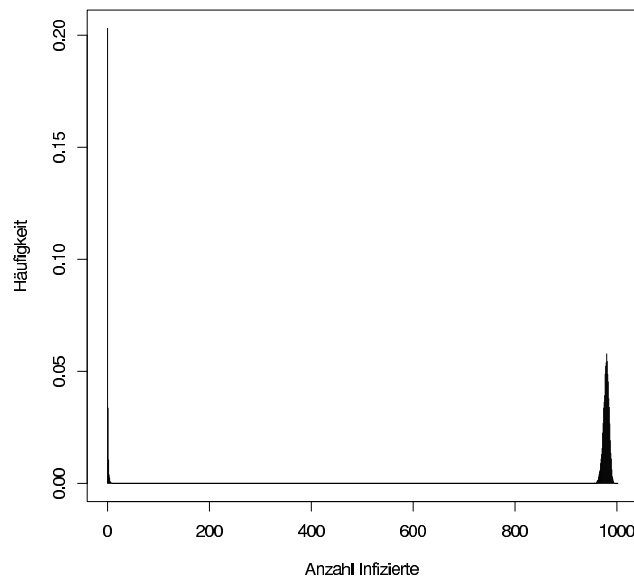


Abbildung 2: Diese Simulation mit 10000 Durchläufen zeigt eine Epidemie mit $\theta = 4.0$, $n = 1000$ und $m = 1$ ohne Impfstrategie.

erfordert: Denn bei der Ermittlung des nächsten Ereignisses muss stets wieder fast die gesamte Bevölkerung „durchblättert“ werden. Besonders viel Zeit nimmt dazu auch die erforderliche dynamische Speicherzuteilung für die Objekte in Anspruch!

3.2 Überprüfung einer Impfstrategie

Zwar wurden die Formeln für die kritische Impfgröße schon mathematisch bewiesen. Zu Anschauungszwecken ist aber auch eine Kontrolle per Simulation möglich.

Aus der Beobachtung der Endgröße einer einzelnen Simulation mit $n = 1000$ und $m = 1$ mit wahren $\theta = 14.5$ habe ich über die Formeln (8),(10) $\hat{v}_c = 0.72546/r$ und $s.e.(\hat{v}_c) = /r$ bestimmt. Mit der Annahme einer perfekten Impfung mit $r = 1$ ergibt sich dann $G_o = 0.72729$.

Um eine halbwegs runde Zahl an Geimpften zu bekommen, habe ich v_c großzügig auf 75% gesetzt.

Eine anschließende Simulation mit einer Simulationsanzahl von 1000 brachte das Ergebnis in Abbildung 3.

Dort ist zu sehen, dass die Epidemien durch die Impfung – wie gewünscht – deutlich unter Kontrolle gehalten werden. Der visuelle Eindruck wird durch die Mittelwerte der beiden Modelle bestätigt:

Die durchschnittliche Anzahl von Infizierten beträgt 729.10 (emp. Standardabw.: 426.72), wenn nicht geimpft wird. Bei Durchimpfungsgrad 75% beträgt

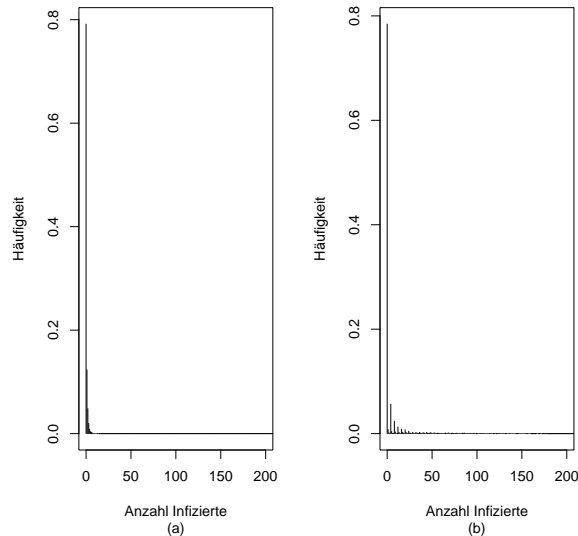


Abbildung 4: eine Epidemie mit $v_c = 0.75$, $R_0 = 4$ und (a): gleichmäßiger Verteilung der Geimpften über alle Haushalte, (b): Durchimpfung von 75% der Bevölkerung, durch Impfen kompletter Haushalte (10000 Simulationen).

die durchschnittliche Anzahl nur noch 7.6496 (emp. Standardabw.: 18.733).

3.3 Simulationen von Haushaltsstrukturen

Es wird angenommen, dass es 250 Haushalte mit jeweils 4 Mitgliedern gibt. Alle 4 Mitglieder werden als gleich anfällig gegenüber der Krankheit angesehen. Betrachtet wird eine Epidemie wie im vorherigen Kapitel: $\lambda_G = 1.1$, $\lambda_L = 6.0$, $m = 1$, $v_c = 0.75$, $r = 1$ und $n = 1000$. Dann gilt $R_0 = M_L \lambda_G \nu = 4.0$ (S. 55f, Andersson). Es werden 10000 Wiederholungen der Epidemie betrachtet.

Die beiden Abbildungen 4 (a) und (b) zeigen die Simulationsergebnisse. Dabei sind in Abbildung 4 (a) 3 von 4 Haushaltsmitgliedern in jedem Haushalt geimpft, in Bild 4 (b) sind ebenfalls 75% der Bevölkerung, aber in jedem geimpften Haushalt jeweils alle Mitglieder durchgeimpft. Der Effekt der Impfung hängt jetzt nicht mehr nur von der Anzahl der Geimpften, sondern eben auch von der Struktur der Bevölkerung ab.

Denn obwohl nun der Anteil der Geimpften in beiden Haushalten gleich ist, sind dennoch die Ergebnisse unterschiedlich:

- im Modell in dem 75% jedes Haushalts geimpft ist, beträgt die mittlere Anzahl der Infizierten: 0.3662 (emp. Standardabw.: 0.91820).
- im Modell in dem 75% der Bevölkerung, und zwar komplette Haushalte, durchgeimpft sind beträgt die mittlere Anzahl der Infizierten: 6.3315 (emp. Standardabw.: 20.662).

Am sinnvollsten (S.119f, Andersson): Gleichmäßige Impfung durch alle Haushalte.

3.4 Auswirkungen falscher Impfstrategie-Annahmen

Betrachtet wird jetzt eine Epidemie mit einer verdeckten Struktur: Frauen und Männer sind doppelt so anfällig für die Epidemie. Der Wert für den kritischen Anteil der zu Impfenen wird aber ohne Beachtung der Struktur geschätzt. Besteht nun dennoch Herdenimmunität?

Auch dieses Problem kann mit Simulationen behandelt werden. Dazu wird zunächst eine Epidemie mit $r = 1, n = 1000, m = 1, 10000$ Simulationen betrachtet. Der Anteil der Geschlechter ist jeweils 50%, die Anfälligkeit der Männer ist $\theta_m = 1$, die der Frauen $\theta_f = 3$. Damit bestimmt sich $\theta = R_0 = 2$, dann gilt (diesmal mit der deterministischen Formel): $v_c = 0.5$. Bei Impfung ohne Betrachtung der unterschiedlichen Anfälligkeit (also zufälliger Auswahl der zu impfenden Personen) erhält man für die Zahl der Infizierten:

- Mittelwert : 9.3698
- Standardabw : 24.849

Bei Impfung der Personen proportional zu ihrer Anfälligkeit - also unter Beachtung der vorhandenen Struktur - (25% der Männer, 75% der Frauen) erhält man analog:

- Mittelwert : 2.7490
- Standardabw : 7.4202

Nicht betrachtete Strukturen können also - trotz eigentlich korrektem Impfanteil für ein homogenes Modell - zu einem großen Ausbruch der Krankheit führen.

Da in der Realität wahre Heterogenität - insbesondere sozialer Art - nicht vollständig erfasst werden kann, ist immer mit Verzerrungen durch latente Beziehungen in der Bevölkerung zu rechnen.

Das bedeutet auch, dass die Ergebnisse der Impfanteilsschätzungen immer nur als *untere* Grenze für den wahren Anteil von v_c zu betrachten ist.

4 Schätzung der Impfwirksamkeit r

Bei neu eingeführten Impfungen ist die Impfwirksamkeit r , also die relative Verringerung der Infektivität von Geimpften gegenüber Ungeimpften, unter Umständen noch nicht bekannt. Auch ist es möglich, dass die Größe von r vor Einführung einer Impfstrategie an der tatsächlichen Bevölkerung überprüft werden soll. Dies ist generell mit zwei Verfahren möglich.

Klassischerweise ist es üblich eine sogenannte „gematchte“ Studie durchzuführen, bei der man zu jedem Geimpften einen passenden Nicht-Geimpften mit möglichst ähnlichen soziodemographischen Faktoren und ähnlicher Infektionsexposition betrachtet. Diese Studienpartner können zum Beispiel Geschwisterpaare sein.

Durch die aufwändige Studienstruktur liefern diese Untersuchungen sehr gut interpretierbare Ergebnisse, die allerdings mit hohen Durchführungskosten verbunden sind.

Eine Alternative zu diesen Untersuchungen bietet das *SIR*-Modell. Betrachtet wird dabei eine große Bevölkerung, von denen nur ein vernachlässigbar kleiner Teil geimpft ist, und nur eine kleine Anzahl zu Beginn der Epidemie infiziert ist (also m klein). Sei ν die Anzahl der (gesondert ausgewählten) Geimpften unter den n Anfälligen. Dann muss $\pi_v = \nu/n$ vernachlässigbar klein sein. Nach Ablauf der Epidemie ist unter den Ungeimpften ein Anteil von \bar{Z}_u und bei den Geimpften ein Anteil von \bar{Z}_v infiziert, für den Gesamtanteil der Infizierten \bar{Z} gilt also:

$$\bar{Z} = \pi_v \bar{Z}_v + (1 - \pi_v) \bar{Z}_u \approx \bar{Z}_u \quad (11)$$

Aus dem Themenkomplex des Multityp-*SIR*-Modells ist für einen großen Ausbruch bekannt, dass (\bar{Z}_u, \bar{Z}_v) asymptotisch gegen (τ_u, τ_v) konvergiert. Für die Werte (τ_u, τ_v) ist bekannt, dass:

$$1 - \tau_u = \exp\{-\iota\lambda[\pi_v\tau_v + (1 - \pi_v)\tau_u]\} \quad (12)$$

$$1 - \tau_v = \exp\{-(1 - r)\iota\lambda[\pi_v\tau_v + (1 - \pi_v)\tau_u]\} \quad (13)$$

Dann lässt sich durch Logarithmieren und Dividieren der folgende Zusammenhang herstellen:

$$1 - r = \frac{\log(1 - \tau_v)}{\log(1 - \tau_u)}. \quad (14)$$

Wenn man nun für die asymptotischen Werte τ_v und τ_u die entsprechenden Schätzer \bar{Z}_v und \bar{Z}_u einsetzt, erhält man einen Schätzer für die Impfwirksamkeit r :

$$\hat{r} = 1 - \frac{\log(1 - \bar{Z}_v)}{\log(1 - \bar{Z}_u)}. \quad (15)$$

Dabei kann \hat{r} Werte in $(-\infty, 1]$ annehmen. Eine Interpretation der Werte von \hat{r} erfolgt bei der Betrachtung des Konfidenzintervalls des Schätzers weiter unten. Dieser Schätzer benötigt die oben gemachte Annahme, dass π_v vernachlässigbar ist, noch nicht.

Beim Bestimmen eines Schätzers für $s.e.(\hat{r})$ – auf einfachem Wege – wird diese Annahme aber wichtig: Dann kann nämlich der Gesamtanteil der Infizierten als approximativ unabhängig gegenüber \bar{Z}_v betrachtet werden. Wegen dieser approximativen Unabhängigkeit ist die Wahrscheinlichkeit für einen Geimpften infiziert zu werden

$$\tau_v = 1 - \exp\{-(1-r)\nu\lambda[\pi_v\tau_v + (1-\pi_v)\tau_u]\}. \quad (16)$$

Für Z_v gilt dann:

$$Z_v \sim \text{Bin}(\nu, \tau_v). \quad (17)$$

Der Wert von Z_u wird von Z_v nicht beeinflusst, und kann daher als deterministisch gegenüber Z_v betrachtet werden. Dann sieht man aus Formel (15), dass

$$\text{Var}(\hat{r}) = \frac{\text{Var}[\log(1 - \bar{Z}_v)]}{[\log(1 - \bar{Z}_u)]^2}. \quad (18)$$

Mittels der δ -Methode, und der Varianz von Z_v lässt sich die Varianz von $\log(1 - \bar{Z}_v)$ bestimmen. Man erhält schliesslich:

$$s.e.(\hat{r}) = \frac{\sqrt{\bar{Z}_v(1 - \bar{Z}_v)}}{\sqrt{\nu}(1 - \bar{Z}_u)|\log(1 - \bar{Z}_u)|}. \quad (19)$$

Bei der Impfwirksamkeit r birgt nur eine Überschätzung des wahren Parameters die Gefahr der Unterschätzung des kritischen Impfanteils, was wiederum zur Folge haben kann, dass ein großer Ausbruch nicht unmöglich ist.

Deshalb ist es an dieser Stelle angebracht, ein einseitiges Konfidenzintervall zu schätzen. Mit der Annahme asymptotischer Normalverteilung von r gilt

$$G_u = \hat{r} - z_{1-\alpha} s.e.(\hat{r}) \quad (20)$$

und man erhält das $KI : [G_u; 1]$, dass mit Wahrscheinlichkeit $1 - \alpha$ den wahren Parameter r überdeckt.

Interpretation (jeweils zur Sicherheitswahrscheinlichkeit $1 - \alpha$):

- Ist $G_u \leq 0$, dann ist die Impfung gar nicht wirksam, die Annahme $r = 0$ ist legitim.
- Gilt $0 < G_u \leq 1$, dann überschätzt man bei der Annahme $r = G_u$ die Impfwirksamkeit nicht.

Zur Kritik an dieser Methode zur Bestimmung von \hat{r} ist allerdings zu sagen, dass die Anzahl anfänglich Infizierter m klein sein muss (woran soll man

das in der Realität erkennen?), und eine größere Epidemie der Krankheit abgewartet werden muss.

Beispiel 2: Gegeben: $\nu = 185$, $Z_v = 16$, $\bar{Z}_u = 0.42$. Dann ist $\bar{Z}_v = 16/185 \approx 0.08649$ und $\hat{r} = 1 - \frac{-0.09046}{-0.54473} = 0.83394$ und $s.e.(\hat{r}) = \frac{0.28109}{1.1240} = 0.25008$. Bei Annahme von $r = G_u = 0.42256$ unterschätzt man den wahren Wert von r mit 95%-iger Sicherheit nicht.

Kommentar zum Beispiel: r liegt sicher echt in $[0, 1]$. Deswegen ist für diesen Parameter unter Umständen eine (logit) Transformation des Schätzers und des Standardfehlers mit anschließendem Rücktransformieren sinnvoll! Problem: Vor der Durchführung der Transformation ist zu klären, ob \hat{r} ein ML-Schätzer ist!

5 Literatur

Andersson, Hakan und Britton, Tom; Stochastic epidemic models and their statistical analysis, Springer-Verlag, 2000