

Sabrina SCHERIAU

**Vergleich univariater Methoden zur
Erkennung von Änderungen der
wöchentlichen Anzahl an gemeldeten
Campylobacter-Infektionen**

MASTERARBEIT

zur Erlangung des akademischen Grades einer Diplom-Ingenieurin

Masterstudium Operations Research und Statistik



Technische Universität Graz

Betreuer:

Univ.-Prof. Dipl.-Ing. Dr.techn. Ernst STADLOBER

Institut für Statistik

Graz, im Juli 2012

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am
.....
(Unterschrift)

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

.....
date
.....
(signature)

Kurzfassung

Das Monitoring von Krankheitsfällen hat in den letzten Jahren immer mehr an Bedeutung gewonnen. Während es in anderen europäischen Ländern wie England, Deutschland und den Niederlanden bereits unterstützende Computersysteme zur Ausbruchsfriherkennung gibt, ist dies in Österreich gerade in Entstehung. Das Ziel dieser Masterarbeit ist es daher, verschiedene univariate *change-point detection* Methoden theoretisch zu beschreiben und deren Anwendbarkeit für gemeldete Campylobacter-Infektionen in Österreich zu untersuchen. Beschrieben werden (i) ein Algorithmus der statistischen Prozesskontrolle mit verschiedenen Adaptierungen, (ii) ein Bayes-basierter Ansatz sowie (iii) die eigens für diese Thematik entwickelte Methode von Farrington et al. (1996). Diese Algorithmen werden in einer Monte-Carlo-Simulationsstudie anhand gängiger Performance-Maße evaluiert. Die praktische Umsetzbarkeit der vorgestellten Methoden wird anhand von Daten aus dem Bundesland Steiermark, welche die Anzahl der wöchentlich gemeldeten Campylobacter-Infektionen in den Jahren 2007-2011 beinhalten, aufgezeigt und deren Tragfähigkeit bewertet. In Zukunft soll diese wöchentliche Analyse eine wichtige Hilfe für die Experten der Nationalen Referenzzentrale für Campylobacter darstellen. Die Masterarbeit entstand in Kooperation mit der Österreichischen Agentur für Gesundheit und Ernährungssicherheit (AGES) unter der Betreuung von Mag. Dr. Hans-Peter Stüger.

Abstract

The monitoring of infectious diseases has gained in importance over the past years. Whereas countries like England, Germany and the Netherlands already use automated outbreak-detection systems for different infectious diseases, the development of such systems is a current issue in Austria. The goal of this master thesis is on the one hand a detailed description of different univariate change-point detection methods and on the other hand the application of these methods to campylobacter-infections in Austria. Therefore a CUSUM algorithm, which originates from statistical process control, a Bayesian approach and the Farrington algorithm (Farrington et al., 1996) are analysed. Besides the theoretical description of these methods a Monte-Carlo simulation study is carried out for comparison with respect to common performance criteria. The usability of the proposed methods is shown and evaluated on the basis of data from Styria including weekly appearances of notified campylobacteriosis cases in the years 2007 to 2011. It is planned that the weekly analyses may support the experts of the National Reference Laboratory for Campylobacter (NRLC) in detecting outbreaks as early as possible. This master thesis has been written in cooperation with the Austrian Agency for Health and Food Safety (AGES).

Danksagung

An dieser Stelle möchte ich mich beim Team der Abteilung DSR-Statistik der AGES bedanken. Besonderer Dank gilt hier vor allem Herrn Dipl.-Ing. Dr. Johannes Hofrichter und Herrn Mag. Dr. Hans-Peter Stüger, die die Durchführung der Masterarbeit ermöglicht haben.

Herzlich bedanken möchte ich mich auch bei Univ.-Prof. Dipl.-Ing. Dr.techn. Ernst Stadlober für die Unterstützung und die Betreuung bei der Erstellung dieser Masterarbeit.

Inhaltsverzeichnis

Abkürzungsverzeichnis	ix
Tabellenverzeichnis	xi
Abbildungsverzeichnis	xiii
1 Einleitung	1
2 On-line Detektions-Algorithmen	3
2.1 CUSUM - Eine Methode der statistischen Prozesskontrolle	4
2.1.1 Allgemeine Einführung in die CUSUM-Methode	4
2.1.2 CUSUM für Mitglieder der einparametrischen Exponentialfamilie	6
2.1.3 Unbekannte Parameter θ_0 und θ_1	9
2.1.4 Mögliche Adaptierungen bei zeitabhängigem Mittelwert	12
2.2 Ein Algorithmus basierend auf Bayes-Inferenz	15
2.2.1 Kurzer Überblick über Grundlagen der Bayes-Theorie	16
2.2.2 Der Algorithmus	18
2.3 Der Farrington-Algorithmus	21
2.4 Detektions-Algorithmen in R: Das <i>'surveillance'</i> package	25
2.4.1 Allgemeine Beschreibung der Datenstruktur	26
2.4.2 CUSUM-Detektoren	27
2.4.3 Bayes-basierter Algorithmus	28
2.4.4 Farrington-Algorithmus	30
3 Optimalitätskriterien und Performance	33
3.1 Die ARL und das minimale ARL_1 -Kriterium	33
3.1.1 CUSUM und ARL-Kriterien	34
3.2 Die <i>false alarm probability</i>	39
3.3 Sensitivität und Spezifität	40
4 Simulationsstudie	41
4.1 Grundmodelle für den <i>in-control</i> Mittelwert	42

Inhaltsverzeichnis

4.2	CUSUM - Vergleich möglicher Adaptierungen	44
4.2.1	Normalverteilungsapproximation vs. CUSUM mit zeitabhängigen Parametern	44
4.2.2	Regression Charts (LR- und GLR-Detektoren)	50
4.3	Bayesscher-Algorithmus	57
4.4	Farrington-Algorithmus	60
4.5	Fazit	66
5	Detektion lebensmittelbedingter Krankheitsausbrüche	69
5.1	Gesetzliche Regelungen und Definitionen	69
5.2	Zoonosenerreger Campylobacter	70
5.3	Datensituation in Österreich	71
5.3.1	Datenaufbereitung	72
5.3.2	Eine erste deskriptive Analyse	73
5.3.3	Schlussfolgerungen für die praktische Betrachtung	77
5.4	Anwendung der Algorithmen	78
5.4.1	CUSUM - Poisson- und NegBin-Regression Charts	78
5.4.2	Bayes-basierter Ansatz	86
5.4.3	Farrington Algorithmus	92
5.5	Grafische Analyse für die Bezirke der Steiermark	97
6	Resümees	101
6.1	Bewertung der Methoden aufgrund der praktischen Anwendung	101
6.2	Bewertung der Realisierbarkeit eines Monitoring-Systems in Österreich	102
	Literaturverzeichnis	105
	Anhang	109

Abkürzungsverzeichnis

AGES Österreichische Agentur für Gesundheit und Ernährungssicherheit

BMG Bundesministerium für Gesundheit

C. coli *Campylobacter coli*

C. jejuni *Campylobacter jejuni*

EFSA European Food Safety Authority

EMS Epidemiologisches Meldesystem für Infektionskrankheiten

NRZC Nationale Referenzzentrale für Campylobacter

ZoonG Bundesgesetz zur Überwachung von Zoonosen und Zoonoseerregern

Tabellenverzeichnis

4.1	Sensitivität und Spezifität für Rossi bzw. Rogerson	50
4.2	Threshold-Werte für LR-Detektor	51
4.3	Threshold-Werte für GLR-Detektor	52
4.4	ARL für LR- und GLR-Detektor für Modelle A1-A3	53
4.5	ARL für LR- und GLR-Detektor für Modelle B1-B3	53
4.6	Sensitivität und Spezifität für LR- und GLR-Detektor für Modelle A1-A3	54
4.7	Sensitivität bzw. Spezifität (verwendeter Threshold h) für LR- und GLR- Detektor für Modelle B1-B3	54
4.8	Threshold-Werte für LR-Detektor bei additiven Zuwachs von $2 \cdot sd$. . .	55
4.9	Threshold-Werte für GLR-Detektor bei additiven Zuwachs von $2 \cdot sd$. .	55
4.10	ARL für LR- und GLR-Detektor für Modelle A1-A3 bei additiven Zu- wachs von $2 \cdot sd$	55
4.11	ARL für LR- und GLR-Detektor für Modelle B1-B3 bei additiven Zu- wachs von $2 \cdot sd$	56
4.12	Sensitivität und Spezifität für LR- und GLR-Detektor für Modelle A1- A3 bei additiven Zuwachs von $2 \cdot sd$	56
4.13	Sensitivität bzw. Spezifität (verwendeter Threshold h) für LR- und GLR- Detektor für Modelle B1-B3 für Parametershift $2 \cdot sd$	57
4.14	Parametereinstellungen bei Anwendung des Algorithmus basierend auf Bayes-Inferenz	58
4.15	MC-Schätzer für ARL_0 bzw. ARL_1 bei Bayes für Modelle A1-A3	59
4.16	MC-Schätzer für ARL_0 bzw. ARL_1 bei Bayes für Modelle B1-B3	59
4.17	Schätzer für die Fehlalarm-Wahrscheinlichkeit bei Bayes im ersten Jahr	59
4.18	Schätzer für die Detektions-Wahrscheinlichkeit bei Bayes in 3 Wochen .	59
4.19	Schätzwerte für Sensitivität bzw. Spezifität bei Bayes für Modelle A1-A3	60
4.20	Schätzwerte für Sensitivität bzw. Spezifität bei Bayes für Modelle B1-B3	60
4.21	Verwendete Parametereinstellung bei Anwendung des Farrington-Algo- rithmus	61
4.22	MC-Schätzwerte für ARL_0 bzw. ARL_1 bei Farrington für Modelle A1-A3	62
4.23	MC-Schätzwerte für ARL_0 bzw. ARL_1 bei Farrington für Modelle B1-B3	62
4.24	Schätzer für Sensitivität bzw. Spezifität bei Farrington für Modelle A1-A3	65
4.25	Schätzer für Sensitivität bzw. Spezifität bei Farrington für Modelle B1-B3	65

Tabellenverzeichnis

5.1	Einteilung des Jahres 2007 und des Anfangs des Jahres 2008 in Kalenderwochen	73
5.2	Anzahl der angezeigten Alarme bei Analyse der <i>C. jejuni</i> Fälle mittels GLR-Detektor und unterschiedlichen h -Werten	82
5.3	Anzahl der angezeigten Alarme bei Analyse der <i>Campylobacter coli</i> (<i>C. coli</i>) Fälle mittels GLR-Detektor und unterschiedlichen h -Werten	86
5.4	Anzahl der angezeigten Alarme bei Analyse der <i>Campylobacter jejuni</i> (<i>C. jejuni</i>) Fälle mittels Bayes-Algorithmus	87
5.5	Anzahl der angezeigten Alarme bei Analyse der <i>C. coli</i> Fälle mittels Bayes-Algorithmus	89
5.6	Anzahl der angezeigten Alarme bei Analyse der <i>C. jejuni</i> Fälle mittels Farrington-Algorithmus	92
5.7	Anzahl der angezeigten Alarme bei Analyse der <i>C. coli</i> Fälle mittels Farrington-Algorithmus	94

Abbildungsverzeichnis

2.1	Referenzbeobachtungen Bayes-basierter Algorithmus	20
2.2	Histogramm Posterior-prädiktive Verteilung	21
2.3	Menge der Referenzbeobachtungen Farrington-Algorithmus	22
2.4	Obere Schranken des Bayes-basierten Algorithmus für unterschiedliche Wahl von α	29
2.5	Obere Schranken des Farrington-Algorithmus für unterschiedliche Transformationen.	31
3.1	Histogramm der simulierten <i>run length</i> für Rogerson-CUSUM-Detektor	38
4.1	Modelle Situation A	43
4.2	Modelle Situation B	44
4.3	ARL0 sowie ARL1 für Rossi- bzw. Rogerson-CUSUM	47
4.4	P_{fa} und P_{fa}^1 für Rossi bzw. Rogerson	48
4.5	PFA in Abhängigkeit von verschiedenen h -Werten für GLR-Detektor für Modelle A2 und B3.	52
4.6	Monte-Carlo-Schätzer der <i>false alarm probability</i> bei Anwendung des Farrington-Algorithmus	63
4.7	Monte-Carlo-Schätzer der Detektionswahrscheinlichkeit bei Anwendung des Farrington-Algorithmus	64
5.1	Gemeldete Campylobacteriose-Erkrankungen in Österreich bzw. der Steiermark 2007-2011	74
5.2	Relative Anteile der bestimmten Campylobacter-Spezies an der Anzahl der gemeldeten Campylobacteriose-Erkrankungen in der Steiermark. . .	75
5.3	Zeitreihe der gemeldeten Campylobacter-Erkrankungen steiermarkweit, die auf den Campylobacter-Erreger <i>C. coli</i> zurückzuführen sind, für 2007-2011.	76
5.4	Zeitreihe der gemeldeten Campylobacter-Erkrankungen steiermarkweit, verursacht durch den Campylobacter-Erreger <i>C. jejuni</i> in den Jahren 2007-2011.	76
5.5	Jahresmittelwerte und -varianzen der <i>C. jejuni</i> bzw. <i>C. coli</i> Infektionen in der Steiermark	77

Abbildungsverzeichnis

5.6	Modellselektion zur Analyse der <i>C. jejuni</i> Infektionen mittels GLR-Detektor	80
5.7	Ergebnis GLR-Detektor für <i>C. jejuni</i> Infektionen (steiermarkweit) . . .	81
5.8	Vorhersage für <i>C. jejuni</i> Infektionen (steiermarkweit) mit bzw. ohne Neuanpassung des Modells	82
5.9	Modellselektion zur Analyse der <i>C. jejuni</i> Infektionen (steiermarkweit) mittels GLR-Detektor	84
5.10	Ergebnisse GLR-Detektor für <i>C. coli</i> Infektionen (steiermarkweit) I . .	85
5.11	Ergebnisse GLR-Detektor für <i>C. coli</i> Infektionen (steiermarkweit) II . .	85
5.12	Ergebnisse Bayes-Methode für <i>C. jejuni</i> Infektionen (steiermarkweit) .	90
5.13	Ergebnisse Bayes-Methode für <i>C. coli</i> Infektionen (steiermarkweit) . . .	91
5.14	Ergebnisse Farrington-Algorithmus für <i>C. jejuni</i> Infektionen (steiermarkweit)	95
5.15	Ergebnisse Farrington-Algorithmus für <i>C. coli</i> Infektionen (steiermarkweit)	96
5.16	Anzahl der gemeldeten <i>C. jejuni</i> Erkrankungen in den ersten 12 Wochen im Jahr 2011 in der Steiermark nach Bezirk.	98
5.17	Anzahl der gemeldeten <i>C. jejuni</i> Erkrankungen in den ersten 12 Wochen im Jahr 2011 in der Steiermark nach Bezirk und je 100 000 EW.	98
5.18	Anzahl der gemeldeten <i>C. jejuni</i> Erkrankungen in KW 22-25 im Jahr 2011 in der Steiermark nach Bezirk.	98
5.19	Anzahl der gemeldeten <i>C. jejuni</i> Erkrankungen in KW 22-25 im Jahr 2011 in der Steiermark nach Bezirk und je 100 000 EW.	98
5.20	Anzahl der gemeldeten <i>C. coli</i> Erkrankungen in KW 29-35 im Jahr 2011 in der Steiermark nach Bezirk.	99
5.21	Anzahl der gemeldeten <i>C. coli</i> Erkrankungen in KW 29-35 im Jahr 2011 in der Steiermark nach Bezirk und je 100 000 EW.	99
5.22	Anzahl der gemeldeten <i>C. coli</i> Erkrankungen im Jahr 2011 in der Steiermark nach Bezirk.	99
5.23	Anzahl der gemeldeten <i>C. coli</i> Erkrankungen im Jahr 2011 in der Steiermark nach Bezirk und je 100 000 EW.	99
5.24	Anzahl der gemeldeten <i>C. jejuni</i> Erkrankungen in KW 22-25 im Jahr 2011 in der Steiermark nach Region und je 100 000 EW.	100
5.25	Anzahl der gemeldeten <i>C. coli</i> Erkrankungen im Jahr 2011 in der Steiermark nach Bezirk und je 100 000 EW.	100

1 Einleitung

Mit der laufenden Zunahme an *Campylobacter*-Infektionen in Österreich hat auch das Interesse an unterstützenden Monitoring-Systemen zur Ausbruchsfrüherkennung zugenommen. So ist insbesondere die Österreichische Agentur für Gesundheit und Ernährungssicherheit (AGES) daran interessiert, Computersysteme, wie sie bereits in Ländern wie England, Deutschland und den Niederlanden Anwendung finden, zu realisieren.

Die vorliegende Arbeit verfolgt drei Ziele:

1. Die Beschreibung und den Vergleich von verschiedenen univariaten *change-point detection* Methoden.
2. Das Finden einer geeigneten Methode zur Ausbruchsfrüherkennung von *Campylobacter*-Ausbrüchen in Österreich.
3. Das Bewerten der Realisierbarkeit eines Monitoring-Systems. Dieses soll durch laufende Bewertung von wöchentlichen Fallzahlen bei der Erkennung von Krankheitsausbrüchen unterstützen.

Die Kapitel 2 und 4 widmen sich dem ersten Ziel. Beschrieben wird einerseits die CUSUM-Methode, die ihren Ursprung in der statistischen Prozesskontrolle hat, und einige Adaptierungen, um auf die spezielle Struktur der Zeitreihen von Krankheitsfällen einzugehen. Des Weiteren werden ein Bayes-basierter Ansatz (Riebler 2004) und der Algorithmus von Farrington et al. (1996) theoretisch aufgearbeitet. Anhand von sechs Modellen, die verschiedene Krankheiten (hinsichtlich Saisonalität und Anzahl der wöchentlichen Fallzahlen) imitieren, werden mit Monte-Carlo-Simulation gängige Performance-Maße, die in Kapitel 3 vorgestellt werden, ermittelt. Anhand der Ergebnisse werden die Methoden miteinander verglichen. Insbesondere werden ihre Eigenschaften im Zusammenhang mit den diskutierten Modellen genauer untersucht und gegenübergestellt.

Im Anwendungsteil der Arbeit (Kapitel 5) werden Daten von Krankheitsfällen lebensmittelbedingter Krankheiten, die von der Nationalen Referenzzentrale für *Campylobacter* (NRZC) der AGES aufgezeichnet wurden, mit Hilfe der vorgestellten Methoden analysiert. Dabei handelt es sich um Fälle von *Campylobacter*-Infektionen, die in den Jahren 2007-2011 gemeldet wurden. Mit 5000-6000 Fällen pro Jahr gehen die meisten in Österreich

1 Einleitung

gemeldeten lebensmittelbedingten Infektionen auf Campylobacter-Erreger zurück. Eine Unterteilung nach Campylobacter-Spezies (*C. jejuni* und *C. coli*) ist hinsichtlich Ausbruchserkennung notwendig. Da diese Spezifizierung noch nicht flächendeckend durchgeführt wird, werden in Kapitel 5 nur Ergebnisse der Steiermark diskutiert, da hier laut NRZC von vollständigen Daten ausgegangen werden kann. Zusätzlich zu der univariaten Analyse mittels der vorgestellten Methoden wird mit Bezirkskarten-Plots versucht, auch einen örtlichen Zusammenhang in die Ausbruchserkennung miteinzubeziehen.

Die Realisierbarkeit eines automatisierten Monitoring-Systems wird schlussendlich in Kapitel 6 diskutiert.

Alle rechnergestützten Auswertungen und Analysen in der vorliegenden Arbeit wurden mittels der Statistiksoftware R (R Development Core Team 2012) durchgeführt.

2 On-line Detektions-Algorithmen

Ziel dieses Kapitels ist es, verschiedene Algorithmen zur *on-line* Detektion von Änderungen in einer sequentiellen Datenreihe vorzustellen und bzgl. ihrer Eigenschaften zu analysieren. Die Methoden sollen insbesondere hinsichtlich der Anwendung auf diskrete Zeitreihen, die Anzahlen von Krankheitsfällen wiedergeben, untersucht werden. Nach einer allgemeinen Betrachtung soll deshalb vor allem auf Algorithmuseigenschaften hinsichtlich der Modellierung durch Poisson- und Negativ-Binomial-Verteilung eingegangen werden.

On-line Algorithmen haben das Ziel, Änderungen in einer Datenreihe schnellstmöglich zu erkennen. Dazu wird mittels vorliegender Beobachtungen der Vergangenheit eine Stoppzeit bestimmt. Die Stoppzeit ist jener Zeitpunkt, an dem eine ausreichende Vermutung für eine Änderung vorliegt (vgl. Sakar und Meeker 1998).

$\{X_k\}_{k \geq 1}$ sei eine Zeitreihe von diskreten Zufallsvariablen mit Wahrscheinlichkeitsfunktion $f(\cdot | \theta)$. Mithilfe der nachfolgenden Algorithmen sollen erhebliche Änderungen in Form eines Zuwachses des Verteilungsparameters θ von einem *in-control* Parameter θ_0 zu einem *out-of-control* Parameter θ_1 detektiert werden. Eine wichtige Rolle spielt hierbei vor allem der unbekannte zu schätzende *change-point* τ . Nach Höhle und Mazick (2010) spricht man hier im Sinne der statistischen Prozesskontrolle (SPK) davon, dass der Prozess vor dem Zeitpunkt τ *in-control* und ab dem Zeitpunkt τ *out-of-control* ist. Mittels eines Detektionsverfahrens wird für jeden Zeitpunkt $t \geq 1$ der wahre (unbekannte) Status $d(t)$ des Prozesses mittels der verfügbaren Information $x_t = \{x_k | k \leq t\}$ geschätzt. Dazu wird eine Statistik $S(x_t)$ berechnet und der Schätzer

$$\hat{d}(t) = \begin{cases} 0, & \text{für } S(x_t) \leq h \text{ (in-control)} \\ 1, & \text{für } S(x_t) > h \text{ (out-of-control)} \end{cases}$$

ermittelt, wobei h ein gewählter Threshold ist. Des Weiteren ergibt sich die Zufallsvariable $T_A = \min\{t \geq 1 | S(x_t) > h\}$ als vorhin erwähnte Stoppzeit. Sie beschreibt den Zeitpunkt des ersten *out-of-control* Alarms (vgl. Höhle und Mazick 2010).

2.1 CUSUM - Eine Methode der statistischen Prozesskontrolle

Die SPK hat ihren Ursprung in der industriellen Produktüberwachung mit dem Hauptziel, Prozessänderungen rasch zu erkennen. Bei rechtzeitiger Erkennung kann unter Umständen mittels Korrekturmaßnahmen eine Vielzahl an fehlerhaften Produkten verhindert werden (Montgomery 2009).

Die Kontrollkarte (Control-Chart) ist das wichtigste Instrument der SPK. Diese geht zurück auf Shewart, der den Shewart-Control-Chart während der 1920er Jahre entwickelte (Winkel und Zhang 2007). Control-Charts helfen bei der Entscheidung, ob eine Änderung im Prozess so beachtlich ist, dass sie das Ergebnis tatsächlich beeinflusst (Benneyan et al. 2003).

Das nachfolgende Kapitel beschreibt einen Nachfolger des Shewart-Control-Charts und geht auf seine Anwendung in der Überwachung von Krankheitsausbrüchen ein. Bei der vorgestellten Methode beruht die eingangs erwähnte Statistik $S(x_t)$ auf dem Likelihood-Quotienten

$$\frac{f(\cdot | \theta_1)}{f(\cdot | \theta_0)}$$

(Basseville und Nikiforov 1993). Im Hinblick auf die Analyse von Häufigkeiten von Krankheitsfällen soll in dieser Arbeit der einseitige CUSUM-Algorithmus vorgestellt werden. D. h. es wird davon ausgegangen, dass es keine negativen Beobachtungswerte gibt und dass nur positive Änderungen detektiert werden sollen (vgl. Höhle 2006).

2.1.1 Allgemeine Einführung in die CUSUM-Methode

Der CUSUM-Algorithmus kann auf verschiedene Weisen hergeleitet werden. Die nachfolgende Darstellung orientiert sich an jener, die den Algorithmus als Wiederholungen eines *Sequential Probability Ratio Test (SPRT)* definiert (vgl. Basseville und Nikiforov 1993; Hawkins und Olwell 1998).

Der Sequential Probability Ratio Test

Ausgehend von den vorliegenden unabhängigen Beobachtungen x_1, \dots, x_n soll die Hypothese $H_0 : \theta = \theta_0$ gegen $H_1 : \theta = \theta_1$ getestet werden. Seien $f(\cdot | \theta_i)$ die zu Hypothese i korrespondierenden Wahrscheinlichkeits- bzw. Dichtefunktionen. Die Teststatistik S

2.1 CUSUM - Eine Methode der statistischen Prozesskontrolle

des SPRT ist gegeben durch die Summe der Logarithmen von Likelihood-Quotienten

$$S = \sum_{i=1}^n S_i, \quad \text{mit} \quad S_i = \log \left(\frac{f(x_i | \theta_1)}{f(x_i | \theta_0)} \right).$$

Für gewählte Schranken ϵ und h hat man die drei Entscheidungsmöglichkeiten

- i. H_0 wird akzeptiert, falls $S \leq -\epsilon$,
- ii. H_1 wird akzeptiert, falls $S \geq h$,
- iii. gilt $-\epsilon < S < h$, wird mit der nächsten Beobachtung x_{n+1} fortgesetzt

(vgl. Hawkins und Olwell 1998).

CUSUM und der SPRT

Page (1954) definiert den CUSUM-Algorithmus als Wiederholung von *Sequential Probability Ratio Tests* mit der Idee, den SPRT so oft zu wiederholen, bis der Wert der Teststatistik $S \geq h$ wird. Der erste Zeitpunkt, an dem H_1 akzeptiert wird, wird dann als *alarm time* t_A bestimmt. Als optimalen Wert für ϵ definierte Page den Wert 0 (Basseville und Nikiforov 1993).

Für die Annahme, dass θ_0 und θ_1 bekannt sind, folgt aus diesen Überlegungen die rekursive Darstellung für die kumulierte Summe

$$C_k = \max \left(0, C_{k-1} + \log \left(\frac{f(x_k | \theta_1)}{f(x_k | \theta_0)} \right) \right) = \max(0, C_{k-1} + S_k) \quad \text{mit} \quad C_0 = 0. \quad (2.1)$$

Die Alarmzeit ist eine Zufallsvariable und ist definiert durch $T_A = \min\{k \mid C_k \geq h\}$ (Basseville und Nikiforov 1993).

Der entscheidende Unterschied zwischen dem CUSUM-Detektor und dem SPRT liegt darin, dass die Nullhypothese im Falle des CUSUM-Algorithmus niemals akzeptiert wird. Bevor die Teststatistik einen negativen Wert annehmen kann, wird die Summe mittels der obigen Vorschrift (2.1) auf 0 gesetzt (Hawkins und Olwell 1998). Andererseits erhöht sich bei einem Hinweis gegen die Nullhypothese (= der wahre Parameter entspricht dem *in-control* Parameter θ_0) die kumulierte Summe um den Wert S_k , für den in diesem Fall $S_k > 0$ gilt (vgl. Höhle 2010).

Da die rekursive Darstellung lediglich für bekannte Parameterwerte θ_0 und θ_1 anwendbar ist, wird häufig die zu (2.1) äquivalente Darstellung

$$C_k = \max_{1 \leq k \leq t} \left(\sum_{i=k}^t \log \left(\frac{f(x_i | \theta_1)}{f(x_i | \theta_0)} \right) \right) \quad (2.2)$$

2 On-line Detektions-Algorithmen

betrachtet (Höhle 2010). Diese kann einfach für unbekannte Parameter θ_0 und θ_1 adaptiert werden, indem die unbekannt Parameter durch entsprechende Schätzungen ersetzt werden.

FIR (Fast Initial Response) CUSUMs

Um die Performance des CUSUM-Algorithmus hinsichtlich der durchschnittlichen Laufzeit bis zum ersten Alarm (ARL) zu verbessern, schlugen Lucas und Crosier (1982) vor, der kumulierten Summe einen „Vorsprung“ zu geben. D. h. anstatt mit $C_0 = 0$ zu starten, wird ein Wert $V > 0$ gewählt, sodass $C_0 = V$ gilt. Lucas und Crosier (1982) haben gezeigt, dass damit Änderungen schneller detektiert werden, falls bereits zu Beginn der Anwendung des CUSUM-Algorithmus ein Shift im Parameter vorliegt. Ist die zu betrachtende Zeitreihe zum Zeitpunkt 0 im *in-control* Zustand, führt die FIR CUSUM allerdings zu einer etwas langsameren Detektion. Der im Artikel vorgeschlagene Wert $V = \frac{h}{2}$ hat sich als Standard für diese Adaption durchgesetzt (vgl. Hawkins 1992).

2.1.2 CUSUM für Mitglieder der einparametrischen Exponentialfamilie

Definition 2.1 (Einparametrische Exponentialfamilie). Eine Verteilung mit Verteilungsparameter θ gehört zur einparametrischen Exponentialfamilie, falls deren Dichte- bzw. Wahrscheinlichkeitsfunktion dargestellt werden kann als

$$\begin{aligned} f(x | \theta) &= u(x)c(\theta) \exp(w(\theta)t(x)) \\ &= \exp(w(\theta)t(x) + \log(u(x)) + \log(c(\theta))), \end{aligned}$$

wobei $u(x) \geq 0$, $t(x)$ reellwertige Funktionen, die von der Beobachtung x abhängen und $c(\theta)$, $w(\theta)$ reellwertige Funktionen in θ (Casella und Berger 2002).

Für Mitglieder der einparametrischen Exponentialfamilie ergibt sich die spezielle Form des Log-Likelihood-Quotienten als

$$\begin{aligned} S_k &= \log(f(x_k | \theta_1)) - \log(f(x_k | \theta_0)) \\ &= w(\theta_1)t(x_k) - w(\theta_0)t(x_k) + \log u(\theta_1) - \log u(\theta_0) \\ &= t(x_k) [w(\theta_1) - w(\theta_0)] + v(\theta_1) - v(\theta_0). \end{aligned} \tag{2.3}$$

Da positive Änderungen im Parameter θ ($\theta_1 > \theta_0$) detektiert werden sollen, gilt $w(\theta_1) - w(\theta_0) > 0$. Dies führt ausgehend von der rekursiven Darstellung (2.1) mit

$$Y_k := t(X_k), \quad \kappa := -\frac{v(\theta_1) - v(\theta_0)}{w(\theta_1) - w(\theta_0)} \quad \text{und} \quad D_{k-1} := \frac{C_{k-1}}{w(\theta_1) - w(\theta_0)}$$

2.1 CUSUM - Eine Methode der statistischen Prozesskontrolle

zum CUSUM-Detektor für Mitglieder der einparametrischen Exponentialfamilie

$$D_0 = 0, \quad D_k = \max(0, D_{k-1} + Y_k - \kappa), \quad (2.4)$$

wobei der Zeitpunkt des Alarms T_A gegeben ist durch

$$T_A = \min \{k \geq 1 \mid D_k \geq H\} \quad \text{mit} \quad H = \frac{h}{w(\theta_1) - w(\theta_0)} \quad (2.5)$$

(Van Dobben de Bruyn 1968).

Poisson-CUSUM

Die Poisson-Verteilung findet häufig Anwendung bei der Modellierung des Eintrittes eines zufälligen Ereignisses je Einheit. So wird zum Beispiel die Anzahl der Defekte X , die pro Produktionseinheit bei einem Produkt auftreten, mittels Poisson-Verteilung modelliert (vgl. Montgomery 2009).

Definition 2.2 (Poisson-Verteilung). Ist $X \sim Poi(\mu)$, dann gilt $E(X) = Var(X) = \mu$ sowie

$$\begin{aligned} f(x \mid \mu) &= \frac{e^{-\mu} \mu^x}{x!}, \quad x \geq 0 \\ &= \exp(x \log(\mu) + \log(1/x!) - \mu) \end{aligned}$$

(vgl. Casella und Berger 2002).

Definition 2.3 (Poisson-CUSUM-Detektor). Sei $X_k \sim Poi(\mu)$, dann wird eine positive Änderung im Mittelwert von μ_0 zu $\mu_1 > \mu_0$ mittels (2.4), der dazugehörigen Alarmzeit (2.5) und den Parametern

$$Y_k = X_k, \quad w(\mu) = \log(\mu) \quad \text{und} \quad \kappa = \frac{\mu_1 - \mu_0}{\log(\mu_1) - \log(\mu_0)}$$

detektiert (vgl. Hawkins und Olwell 1998).

Negativ-Binomial-CUSUM

Eine mögliche Definition der Negativ-Binomial-Verteilung beruht auf der Modellierung der Zufallsvariablen $X = \text{Anzahl der Misserfolge bis zum Eintritt des } r\text{-ten Erfolges}$, wobei r fest ist.

2 On-line Detektions-Algorithmen

Definition 2.4. $X \sim \text{NegBin}(p, r)$, falls

$$f(x|r, p) = \binom{r+x-1}{x} p^r (1-p)^x, \quad x \geq 0$$

gilt. Weiters ist $E(X) = \mu = \frac{r(1-p)}{p}$ und $\text{Var}(X) = \frac{1}{p}\mu$ (Hawkins und Olwell 1998).

Mittels der Parametrisierung $p = c/(1+c)$ ($c > 0$) für ganzzahlige nichtnegative Zufallsvariablen folgt die alternative Darstellung der Wahrscheinlichkeitsfunktion als

$$\begin{aligned} f(x|r, c) &= \frac{(r+x-1)!}{x!(r-1)!} \left(\frac{c}{1+c}\right)^r \left(1 - \frac{c}{1+c}\right)^x \\ &= \frac{\Gamma(r+x)}{x!\Gamma(r)} \frac{c^r}{(1+c)^{r+x}}, \quad x \geq 0. \end{aligned} \quad (2.6)$$

Für diese ergeben sich die ersten beiden Momente $E(X) = \mu = \frac{r}{c}$ und $\text{Var}(X) = \mu(1+1/c)$. Für festes r zählt auch die Negativ-Binomial-Verteilung zur einparametrischen Exponentialfamilie mit Parameter c (Hawkins und Olwell 1998).

Definition 2.5 (Negativ-Binomial-CUSUM-Detektor). Sei $X_k \sim \text{NegBin}(r, c)$ mit r fest. Das CUSUM-Schema (2.4) mit

$$Y_k = X_k, \quad w(c) = -\log(1+c), \quad v(c) = r \log\left(\frac{c}{1+c}\right) \quad \text{und} \quad \kappa = \frac{r \log\left(\frac{c_0(1+c_1)}{c_1(1+c_0)}\right)}{\log\left(\frac{1+c_0}{1+c_1}\right)}$$

detektiert einen Parameter-Shift von c_0 zu $c_1 > c_0$ (vgl. Watkins et al. 2008).

Da für diese Parametrisierung $\mu = \frac{r}{c}$ gilt, korrespondiert die Detektion einer positiven Änderung des Parameters c zur Detektion einer negativen Änderung im Mittelwert der zu überwachenden Zeitreihe.

Hinsichtlich der Analyse von Krankheitsfällen wird jedoch die Erkennung eines erhöhten Mittelwerts von Interesse sein, weshalb in weiterer Folge der Detektor

$$D_k = \min(0, D_{k-1} + X_k - \kappa), \quad T_A = \min(k \geq 1 \mid D_k < -H) \quad (2.7)$$

gemeint ist, falls vom Negativ-Binomial-CUSUM die Rede ist (Hawkins und Olwell 1998).

Bemerkung 2.1 (Zusammenhang NegBin- und Poi-Verteilung). Im Allgemeinen sind die Tails der NegBin-Verteilung stärker als jene der Poisson-Verteilung. Häufig wird auch die reparametrisierte Form der Wahrscheinlichkeitsfunktion betrachtet, die sich durch Parametrisierung der Parameter r und p in Abhängigkeit von μ und σ^2 ergibt.

2.1 CUSUM - Eine Methode der statistischen Prozesskontrolle

Es gilt dann $p = r/(\mu + r)$, $r = \mu^2/(\sigma^2 - \mu)$. Die reparametrisierte Wahrscheinlichkeitsfunktion lautet

$$\begin{aligned} f(x | r, p) &= f\left(x | r, \frac{r}{r + \mu}\right) \\ &= \binom{r + x - 1}{x} \left(\frac{r}{r + \mu}\right)^r \left(1 - \frac{r}{r + \mu}\right)^x \\ &= \frac{\Gamma(r + x)}{\Gamma(x + 1)\Gamma(r)} \left(\frac{1}{1 + \frac{1}{r}\mu}\right)^r \left(\frac{\frac{1}{r}\mu}{1 + \frac{1}{r}\mu}\right)^x. \end{aligned} \quad (2.8)$$

Die Varianz dieser reparametrisierten Form ist dann

$$\text{Var}(X) = \sigma^2 = \mu + \frac{1}{r}\mu.$$

In diesem Fall kann $1/r$ als Dispersionsparameter interpretiert werden. Für $r \rightarrow \infty$ folgt daraus die Poisson-Verteilung mit Parameter μ (vgl. Hawkins und Olwell 1998).

Wahl der Parameter κ und h . Der Referenzwert κ wird vollständig durch die Werte des *in-control* Parameters θ_0 und des *out-of-control* Parameters θ_1 bestimmt.

Beispiel 2.1. Soll zum Beispiel ein Shift im Mittelwert der zu untersuchenden Zeitreihe $X_i \sim \text{Poi}(\mu)$ von $\mu_0 = 4$ zu $\mu_1 = 7$ detektiert werden, folgt daraus $\kappa = 5.36$.

Die Wahl des Thresholds h orientiert sich allerdings an der Laufzeit, welche im Wesentlichen die Performance des Algorithmus bestimmt. Genauer gesagt, ist hier die *in-control* Laufzeit ARL_0 von Interesse. Diese ist definiert als die zufällige Anzahl der Beobachtungen, die benötigt wird, um nach dem Start der Überwachung den Wert h bzw. H zu über- oder unterschreiten (vgl. Paul et al. 2007). Kapitel 3 enthält detailliertere Ausführungen zur Wahl von h und zur Performance-Analyse.

Bemerkung 2.2 (CUSUM-Methoden zur Gesundheitsüberwachung). Kumulierte Summen eignen sich aufgrund ihrer Sensitivität gegenüber kleineren kontinuierlichen Änderungen im Mittelwert der gemeldeten Krankheitsfälle auch für länger anhaltende Epidemien, wie etwa Influenzaerkrankungen. Jedoch können aus diesem Grund auch Unregelmäßigkeiten bzw. bereits kleinere Änderungen im Meldesystem die Robustheit der CUSUMs beeinflussen (vgl. Unkel et al. 2012).

2.1.3 Unbekannte Parameter θ_0 und θ_1

Der CUSUM beruht, wie viele andere Algorithmen zur Detektion von Änderungen in einer Zeitreihe, auf der Annahme, dass die zu untersuchenden Parameter vor und

2 On-line Detektions-Algorithmen

nach der zu detektierenden Änderung bekannt sind. Da dies vor allem im Rahmen der Gesundheitsüberwachung kaum zutrifft, ist die übliche Vorgehensweise, wie unter anderem in Höhle und Mazick (2010) beschrieben, den Parameter θ_0 aus einem Datensatz von historischen Beobachtungen zu schätzen und den *out-of-control* Parameter θ_1 als Funktion von θ_0 darzustellen.

Bei dem zur Schätzung von θ_0 herangezogenen Datensatz sollten die Beobachtungen tatsächlich aus der *in-control* Verteilung stammen, da Ausbrüche in diesen Beobachtungen (Ausreißer) die Schätzung des *in-control* Parameters verfälschen würden. Des Weiteren ist zu beachten, dass der CUSUM-Detektor und die in Kapitel 3 folgende Performance-Analyse keinerlei Unsicherheiten, die im Zuge der Parameterschätzung entstehen, miteinbeziehen.

GLR-Ansatz für unbekanntem *out-of-control* Parameter θ_1

Als Alternative zur Bestimmung von θ_1 beschreibt Lorden (1971) den *Generalized Likelihood Ratio* (GLR) Algorithmus als Verallgemeinerung der Methode von Page (1954). Dabei wird die Verteilungsfamilie $F(\cdot | \theta_1)$ durch eine Familie $\{F(\cdot | \theta), \theta \in \Theta\}$ mit unbekanntem Parameter θ ersetzt.

In weiterer Folge wird anstelle der Summe der Log-Likelihood-Quotienten das Supremum

$$\sup_{\theta \in \Theta} \sum_{i=k}^t \log \left(\frac{f(x_i | \theta)}{f(x_i | \theta_0)} \right) \quad (2.9)$$

betrachtet.

Daraus ergibt sich für den CUSUM-Algorithmus als modifizierte Alarmzeit

$$T_A = \min \left\{ k \mid \max_{1 \leq k \leq t} \sup_{\theta \in \Theta} \sum_{i=k}^t \log \left(\frac{f(x_i | \theta)}{f(x_i | \theta_0)} \right) \geq h \right\} \quad (2.10)$$

(siehe auch Basseville und Nikiforov 1993).

Zur Ermittlung dieser Alarmzeit ist für jeden möglichen Änderungszeitpunkt $k \in [1, t]$ eine Maximierung des Log-Likelihood-Quotienten über alle $\theta \in \Theta$ notwendig, was für große t sehr rechenaufwendig wird (Höhle 2006).

Lai und Shan (1999) erläutern eine Vereinfachung dieser Maximierung im Falle der Normalverteilung. Höhle (2006) sowie Höhle und Paul (2008) geben unter bestimmten Voraussetzungen eine effiziente Evaluierung für Daten aus der Poisson- sowie der

2.1 CUSUM - Eine Methode der statistischen Prozesskontrolle

Negativ-Binomial-Verteilung an. Vorausgesetzt wird hier die Charakterisierung des *out-of-control* Parameters θ_1 durch eine multiplikative Änderung im *in-control* Mittelwert θ_0 um den Wert $\exp(\eta)$.

Der Parameter η ist dabei im Allgemeinen unbekannt und beschreibt den additiven Zuwachs auf der Logarithmus-Skala. Der *out-of-control* Mittelwert wird demnach durch

$$\mu_1 = \mu_0 \cdot \exp(\eta) \quad (2.11)$$

modelliert.

Unter der bereits erwähnten Parametrisierung (2.8) der *NegBin*(p, r)-Verteilung mit $p = r/(r + \mu)$ ergibt sich für Beobachtungen x_1, \dots, x_t , die aus einer solchen Verteilung stammen, für den Log-Likelihood-Quotienten

$$\begin{aligned} \log \left(\frac{f(x_i | \mu)}{f(x_i | \mu_0)} \right) &= r \left(-\log \left(1 + \frac{1}{r} \mu \right) + \log \left(1 + \frac{1}{r} \mu_0 \right) \right) \\ &\quad + x_i \left(\log(\mu) - \log \left(1 + \frac{1}{r} \mu \right) - \log(\mu_0) + \log \left(1 + \frac{1}{r} \mu_0 \right) \right). \end{aligned}$$

Ersetzt man den Parameter μ durch die Charakterisierung $\mu = \mu_0 \exp(\eta)$, folgt

$$\begin{aligned} \log \left(\frac{f(x_i | \mu)}{f(x_i | \mu_0)} \right) &= r \log \left(\frac{1 + \frac{1}{r} \mu_0}{1 + \frac{1}{r} \mu_0 \exp(\eta)} \right) \\ &\quad + x_i \left(\log(\mu_0) + \eta - \log(\mu_0) + \log \left(\frac{1 + \frac{1}{r} \mu_0}{1 + \frac{1}{r} \mu_0 \exp(\eta)} \right) \right) \\ &= x_i \eta + (x_i + r) \log \left(\frac{1 + \frac{1}{r} \mu_0}{1 + \frac{1}{r} \mu_0 \exp(\eta)} \right). \end{aligned}$$

Eine ML-Schätzung des unbekanntes *out-of-control* Mittelwerts ist hier somit äquivalent zur ML-Schätzung des unbekanntes Zuwachses η . Dafür ist die Summe der Log-Likelihood-Funktionen der Beobachtungen x_k, \dots, x_t bzgl. η zu optimieren. Höhle und Paul (2008) berechnen hierfür die ersten beiden Ableitungen (nach Lawless 1987) und lösen

$$\frac{\partial}{\partial \eta} \sum_{i=k}^t \log f(x_i | \eta) = 0$$

numerisch, woraus man $\hat{\eta}_{k,t}$ als *Maximum Likelihood Estimator* (MLE) für η in Abhängigkeit von den Beobachtungen x_k, \dots, x_t erhält. Als Startwert schlagen Höhle und Paul (2008) den Wert $\hat{\eta}_{k+1,t}$ vor.

Da nur positive Änderungen von Interesse sind, ist noch $\eta \geq 0$ zu gewährleisten, was durch Wahl von $\hat{\eta}_{k,t}^+ = \max(0, \hat{\eta}_{k,t})$ erfolgt.

2 On-line Detektions-Algorithmen

In weiterer Folge ist

$$\sup_{\mu \in \Theta} \sum_{i=k}^t \log \frac{f(x_i | \mu)}{f(x_i | \mu_0)} = \hat{\eta}_{k,t}^+ \sum_{i=k}^t x_i + \sum_{i=k}^t \left((x_i + r) \log \left(\frac{1 + \frac{1}{r} \mu_0}{1 + \frac{1}{r} \mu_0 \exp(\hat{\eta}_{k,t}^+)} \right) \right) \quad (2.12)$$

rekursiv zu ermitteln, um schließlich die Stoppzeit (2.10) zu erhalten.

Für $X_i \sim Poi(\mu_0)$ gilt die exakte Lösung

$$\hat{\eta}_{k,t}^+ = \log \left(\sum_{i=k}^t x_i \right) - \log \left(\sum_{i=k}^t \mu_0 \right).$$

Des Weiteren vereinfacht sich das Supremum (2.12) wegen $\frac{1}{r} \rightarrow 0$ im Poisson-Fall zu

$$\hat{\eta}_{k,t}^+ \sum_{i=k}^t x_i + (1 - \exp(\hat{\eta}_{k,t}^+)) \sum_{i=k}^t \mu_0$$

(vgl. Höhle und Paul 2008).

2.1.4 Mögliche Adaptierungen bei zeitabhängigem Mittelwert

Gerade bei Zeitreihen, welche die Anzahl an Krankheitsfällen widerspiegeln, ist häufig ein saisonaler Trend zu erwarten. Tatsächlich variiert der *in-control* Mittelwert oftmals über die Zeit aufgrund kurz- bzw. langfristiger Trends (vgl. Höhle et al. 2007). Die nachfolgenden drei Methoden von Rossi et al. (1999), Rogerson und Yamada (2004) und Höhle und Paul (2008) widmen sich dieser Sachlage.

Normalverteilungsapproximation

Sei X_k Poisson-verteilt mit zeitabhängigem *in-control* Mittelwert $\mu_{0,k}$. Rossi et al. (1999) regten eine Transformation der Zufallsvariable X_k zur Normalverteilung an. Um in weiterer Folge die Anwendung einer Normalverteilungs-CUSUM zur Erkennung von Änderungen im Erwartungswert auf die transformierten Daten zu ermöglichen. Die Transformationsvorschrift ergibt sich dabei als Kombination der beiden allgemeinen Ansätze

- $X \stackrel{as}{\sim} N(\mu, \sqrt{\mu}) \quad \Rightarrow \quad Z_1 = \frac{X - \mu}{\sqrt{\mu}} \stackrel{as}{\sim} N(0, 1)$

2.1 CUSUM - Eine Methode der statistischen Prozesskontrolle

- Wurzel-Transformation der Poisson-Verteilung

$$\sqrt{X} \stackrel{appr.}{\sim} N\left(\sqrt{\mu}, \frac{1}{2}\right) \Rightarrow Z_2 = 2\left(\sqrt{X} - \sqrt{\mu}\right) \stackrel{as}{\sim} N(0, 1)$$

und lautet

$$Y_k = \frac{Z_{1,k} + Z_{2,k}}{2} = \frac{X_k - 3\mu_{0,k} + 2\sqrt{\mu_{0,k}X_k}}{2\sqrt{\mu_{0,k}}} \stackrel{appr.}{\sim} N(0, 1). \quad (2.13)$$

Die Zufallsvariable Y_k in (2.13) ist approximativ standardnormalverteilt, denn aufgrund der hohen Korrelation der Zufallsvariablen $Z_{1,k}$ und $Z_{2,k}$ gilt

$$Var(Y_k) = \frac{1}{4}(Var(Z_{1,k}) + Var(Z_{2,k}) + 2Cov(Z_{1,k}, Z_{2,k})) \approx \frac{1}{2} + \frac{1}{2}Corr(Z_{1,k}, Z_{2,k}) \approx 1.$$

Wie bereits erwähnt, wird dann der CUSUM-Detektor (2.4) mit $\kappa = \frac{\mu_0 + \mu_1}{2}$ auf Y_k angewendet.

Bemerkung 2.3. Insbesondere für geringe Anzahlen (z. B. für $\mu < 2$) sind bei dieser Transformation Abweichungen zur Performance einer Normalverteilungs-CUSUM möglich. Eine Abwandlung, um diese Schwachstelle auszugleichen, wird in Rossi et al. (2010) vorgeschlagen. Die Konstante κ wird dabei durch den zeitabhängigen Wert

$$\kappa_k = \frac{\left(\sqrt{\frac{\mu_{1,k}}{\mu_{0,k}}} + 1\right)^2 - 4}{2\sqrt{1/\mu_{0,k}}}$$

ersetzt. Daraus resultiert der Detektor

$$D_k = \max(0, D_{k-1} + Y_k - \kappa_k). \quad (2.14)$$

CUSUM mit zeitabhängigen Parametern

Rogerson und Yamada (2004) schlagen vor, die CUSUM-Methode mit zeitabhängigen Parameterwerten für κ und den oberen Grenzwert H zu verwenden, woraus sich der CUSUM-Detektor

$$D_k = \max\left(0, D_{k-1} + \frac{H}{h_k}(x_k - \kappa_k)\right) \quad (2.15)$$

mit Alarmzeit

$$T_A = \min\{k \geq 1 \mid D_k \geq H\} \quad (2.16)$$

2 On-line Detektions-Algorithmen

ergibt. Die Wahl von H erfolgt äquivalent zur Wahl des Thresholds für einen Detektor mit Poisson-Parameter $\mu_0 = \sum_{k=1}^t \mu_{0k}$, konstantem Wert κ und festgelegter ARL_0 . Der zeitabhängige obere Grenzwert h_k wird festgelegt, in dem zu jedem Zeitpunkt k der Threshold für einen CUSUM-Detektor mit bestimmter ARL_0 und mit Referenzwert

$$\kappa_k = \frac{\mu_{1k} - \mu_{0k}}{\log(\mu_{1k}) - \log(\mu_{0k})} \quad (2.17)$$

fixiert wird.

Poisson und Negative-Binomial Regression Charts

Eine andere Möglichkeit, mit dieser Situation umzugehen, bieten Höhle (2006) sowie Höhle und Paul (2008) unter Verwendung eines generalisierten linearen Modells für den *in-control* Mittelwert, in dem sowohl eine saisonale als auch eine lineare Trendkomponente eingebunden werden kann. Weiters wird wiederum der *out-of-control* Zustand in Form eines additiven Zuwachses des Mittelwerts auf der Logarithmus-Skala definiert, d. h.

$$\mu_{1,k} = \mu_{0,k} \exp(\eta). \quad (2.18)$$

Der zeitabhängige *in-control* Mittelwert wird durch das Log-lineare Modell

$$\log(\mu_{0,k}) = \beta_0 + \beta_1 k + \sum_{i=1}^I (\gamma_i \sin(\omega i k) + \delta_i \cos(\omega i k)) \quad (2.19)$$

beschrieben. Im obigen Modell ist I die Anzahl der harmonischen Schwingungen, $\omega = \frac{2\pi}{T}$ und T die betrachtete Periode, z. B. $T = 52$ Wochen. Der Summenterm beschreibt somit den saisonalen Trend. Soll dieser nicht berücksichtigt werden, ist $I = 0$ zu setzen. Ist andererseits kein linearer Trend in den Daten zu erkennen, gilt $\beta_1 = 0$.

Beispiel 2.2. Für $Y_k \sim Poi$, $T = 52$ und $I = 1$ liefert der R-Befehl

```
glm(y ~ 1 + k + sin(2*pi*52) + cos(2*pi*52),  
     family = poisson, data=dat)
```

das gewünschte Modell für den *in-control* Mittelwert. Für $Y_k \sim NegBin$ kann die Funktion `glm.nb` aufgerufen werden.

```
glm.nb(y ~ 1 + k + sin(2*pi*52) + cos(2*pi*52), link=log)
```

2.2 Ein Algorithmus basierend auf Bayes-Inferenz

Zusätzlich zu den Parameterschätzern $\hat{\beta}_i, i = 0, 1, \hat{\gamma}_i$ und $\hat{\delta}_i$ liefert der R-Output einen Schätzer $\hat{\theta}$, woraus sich die Schätzung für den Dispersionsparameter $\frac{1}{r} = \frac{1}{\theta}$ der Negativ-Binomial-Verteilung ergibt.

Aus der Kenntnis des zu detektierenden Shifts von η folgt sofort, dass der *out-of-control* Mittelwert μ_{1k} für alle Zeitpunkte $k = 1, \dots, t$, bekannt ist. In diesem Fall kann die Alarmzeit des CUSUM-Detektors, in Abhängigkeit von den bekannten zeitabhängigen *in-control* und *out-of-control* Mittelwerten,

$$T_A = \min \left\{ k \mid \max_{1 \leq k \leq t} \sum_{i=k}^t \log \left(\frac{f(x_i \mid \mu_{1k})}{f(x_i \mid \mu_{0k})} \right) \geq h \right\} \quad (2.20)$$

als Alarmzeit verwendet werden.

Der Vorteil dieser Methode von Höhle und Paul (2008) liegt vor allem in der Adaptionsmöglichkeit, falls η und somit der *out-of-control* Mittelwert unbekannt ist. Mittels der Überlegungen aus dem GLR-Ansatz ist auch bei zeitabhängigen Mittelwerten eine effiziente Berechnung des MLEs $\hat{\eta}$ möglich. Hier findet dann die Alarmzeit

$$T_A = \min \left\{ k \mid \max_{1 \leq k \leq t} \sup_{\eta} \sum_{i=k}^t \log \left(\frac{f(x_i \mid \mu_{0k} \exp(\eta))}{f(x_i \mid \mu_{0k})} \right) \geq h \right\} \quad (2.21)$$

Anwendung.

2.2 Ein Algorithmus basierend auf Bayes-Inferenz

Der in Riebler (2004) beschriebene Algorithmus beruht auf der Bayes-Theorie und bietet einen alternativen Ansatz zur Ausbruchserkennung (vgl. auch Höhle und Riebler (2005) sowie Höhle (2007)).

Seien $X_k \stackrel{iid}{\sim} Poi(\mu)$ verteilt mit unbekanntem Parameter μ . Für vorliegende Beobachtungen x_1, \dots, x_t soll mittels Bayes-Inferenz über den unbekanntem Parameter μ eine Prognose für die $(t + 1)$ -ste Beobachtung X_{t+1} ermittelt werden. Der Vergleich dieser Vorhersage mit der tatsächlichen Beobachtung x_{t+1} soll dann zu einer Entscheidung für oder gegen einen Alarm führen.

2.2.1 Kurzer Überblick über Grundlagen der Bayes-Theorie

Im folgenden Abschnitt werden einige grundlegende Begriffe der Bayes-Inferenz angegeben. Die nachfolgenden Erläuterungen sowie die Notation orientieren sich an Held (2008).

Sei X eine Zufallsvariable mit Dichte $f(x | \theta)$. Anhand einer Beobachtung von $X = x$ soll Inferenz über den unbekanntem (möglicherweise vektorwertigen) Parameter θ der „wahren“ Verteilung von X betrieben werden. Die Idee der Bayes-Inferenz ist, den zu untersuchenden Verteilungsparameter θ als Zufallsvariable zu definieren. Unter Annahme einer Priori-Verteilung $f(\theta)$ und der vorliegenden Beobachtung kann in weiterer Folge die Posteriori-Verteilung $f(\theta | x)$ mithilfe des Satz von Bayes

$$f(\theta | x) = \frac{f(x | \theta)f(\theta)}{\sum_{\theta \in \Theta} f(x | \theta)f(\theta)} \quad (2.22)$$

ermittelt werden. Der Nenner in (2.22) stellt lediglich eine Proportionalitätskonstante dar, die gewährleistet, dass $f(\theta | x)$ eine Dichtefunktion (Integral über den Gesamtbereich ist Eins) darstellt.

Wahl der Priori-Verteilung

Eine Möglichkeit für die Wahl der Priori-Verteilung ist die Wahl von $f(\theta)$ so, dass $f(\theta | x)$ und $f(\theta)$ zur selben Verteilungsklasse gehören. In diesem Fall handelt es sich um eine konjugierte Priori-Verteilung.

Beispiel 2.3 (Konjugierte Priori der Poisson-Verteilung). Sei $X \sim Poi(\mu)$. Wird als Priori-Verteilung für den Parameter μ die $Gamma(\alpha, \beta)$ -Verteilung gewählt, folgt für die Posteriori-Verteilung

$$\begin{aligned} f(\mu | x) &\propto f(x | \mu)f(\mu) \\ &\propto \mu^{\alpha-1} e^{-\beta\mu} e^{-\mu} \mu^x \\ &\propto \mu^{(\alpha+x)-1} e^{-\mu(\beta+1)}, \end{aligned}$$

was dem Kern einer $Gamma(\alpha + x, \beta + 1)$ -Verteilung entspricht.

Liegen Beobachtungen x_1, \dots, x_t vor, so entspricht die Posteriori-Verteilung der $Gamma(\alpha + \sum_{k=1}^t x_k, \beta + t)$ -Verteilung.

Um den Einfluss der Priori-Verteilung zu verringern, wird oftmals auch eine „vage“ Verteilung (zum Beispiel mit großer Varianz) gewählt. Dies kann zu einer uneigentlichen

2.2 Ein Algorithmus basierend auf Bayes-Inferenz

Priori-Verteilung führen, für die im diskreten Fall

$$\sum_{\theta \in \Theta} f(\theta) = \infty$$

gilt (vgl. Held 2008).

Eine weitere mögliche Priori-Verteilung ist invariant bezüglich eindeutigen Transformationen von θ und wird durch Jeffreys' Priori festgelegt. Sie ist definiert durch

$$f(\theta) \propto \sqrt{E(I(\theta))},$$

wobei $I(\theta) = -\frac{\partial^2}{\partial \theta^2} \log(f(x | \theta))$ die Fisher-Information ist.

Beispiel 2.4 (Jeffreys' Priori für die Poisson-Verteilung). Sei $X \sim Poi(\mu)$. Die Fisher-Information

$$\begin{aligned} I(\mu) &= -\frac{\partial^2}{\partial \mu^2} \left(x \log(\mu) + \log\left(\frac{1}{x!}\right) - \mu \right) \\ &= \frac{1}{\mu^2} x \end{aligned}$$

führt zur Dichtefunktion von Jeffreys' Priori

$$f(\mu) \propto \sqrt{\frac{1}{\mu} E(X)} = \frac{1}{\sqrt{\mu}}.$$

Daraus folgt, dass die *Gamma* $(\frac{1}{2}, 0)$ -Verteilung Jeffreys' Priori-Verteilung für den Poisson-Parameter μ ist.

Vorhersagen in der Bayes-Inferenz

$(X_k)_{k=1}^t$ habe Dichte $f(\cdot | \theta)$. Gesucht ist eine Vorhersage für X_{t+1} . Es gelte die Unabhängigkeit von (X_1, \dots, X_t) und $X_{t+1} | \theta$. Die posteriori-prädiktive Verteilung ist dann gegeben durch

$$\begin{aligned} f(x_{t+1} | x_1, \dots, x_t) &= \int f(x_{t+1}, \theta | x_1, \dots, x_t) d\theta \\ &= \int f(x_{t+1} | \theta) f(\theta | x_1, \dots, x_t) d\theta. \end{aligned}$$

Die posteriori-prädiktive Verteilung berücksichtigt im Gegensatz zur einfachen Plug-in-Vorhersage $f(x_{t+1} | \hat{\theta})$ bereits Unsicherheiten in der Parameterschätzung.

Beispiel 2.5 (Posteriori-Prädiktive Verteilung für Poisson-Parameter mit konjugierter Priori). Wie in Beispiel 2.3 gelte

$$\begin{aligned} X_{t+1} | \mu &\sim Poi(\mu) \\ \mu | x_1, \dots, x_t &\sim Gamma(\tilde{\alpha}, \tilde{\beta}) \end{aligned}$$

2 On-line Detektions-Algorithmen

mit $\tilde{\alpha} = \alpha + \sum_{k=1}^t x_k$ und $\tilde{\beta} = \beta + t$. Die posteriori-prädiktive Verteilung lässt sich in diesem Fall exakt berechnen und es gilt

$$\begin{aligned} f(x_{t+1} | x_1, \dots, x_t) &= \int_0^\infty f(x_{t+1} | \mu) f(\mu | x_1, \dots, x_t) d\mu \\ &= \int_0^\infty \frac{e^{-\mu} \mu^{x_{t+1}}}{x_{t+1}!} \cdot \frac{\tilde{\beta}^{\tilde{\alpha}}}{\Gamma(\tilde{\alpha})} \mu^{\tilde{\alpha}-1} e^{-\mu \tilde{\beta}} d\mu \\ &= \frac{\tilde{\beta}^{\tilde{\alpha}}}{x_{t+1}! \Gamma(\tilde{\alpha})} \cdot \int_0^\infty \exp\left(-\mu (\tilde{\beta} + 1)\right) \mu^{(\tilde{\alpha}+x_{t+1})-1} d\mu \\ &= \frac{\tilde{\beta}^{\tilde{\alpha}}}{x_{t+1}! \Gamma(\tilde{\alpha})} (\tilde{\beta} + 1)^{-(\tilde{\alpha}+x_{t+1})} \Gamma(\tilde{\alpha} + x_{t+1}). \end{aligned}$$

Durch einfache Umformung dieses Ergebnisses folgt die Dichtefunktion der Poisson-Gamma-Verteilung $PoiGa(\tilde{\alpha}, \tilde{\beta}, 1)$, d.h.

$$f(x_{t+1} | x_1, \dots, x_t) = \left(\frac{\tilde{\beta}}{1 + \tilde{\beta}}\right)^{\tilde{\alpha}} \frac{1}{\Gamma(\tilde{\alpha})} \frac{1}{x_{t+1}!} \Gamma(\tilde{\alpha} + x_{t+1}) \left(\frac{1}{\tilde{\beta} + 1}\right)^{x_{t+1}}.$$

2.2.2 Der Algorithmus

Seien $X_k \stackrel{iid}{\sim} Poi(\mu)$. Wie bereits erwähnt, soll die $(t + 1)$ -ste Beobachtung mittels Bayes-Inferenz prognostiziert werden. Dazu wird für den unbekannt Parameter μ die zur Poisson-Verteilung konjugierte Priori-Verteilung $Gamma(\alpha, \beta)$ -Verteilung gewählt. Aus Beispiel 2.3 folgt für die Posteriori-Verteilung

$$\mu | x_1, \dots, x_t \sim Gamma\left(\alpha + \sum_{k=1}^t x_k, \beta + t\right). \quad (2.23)$$

Die posteriori-prädiktive Verteilung entspricht, wie in Beispiel 2.5 gezeigt, der Poisson-Gamma-Verteilung mit Parametern $\tilde{\alpha} = \alpha + \sum_{k=1}^t x_k$, $\tilde{\beta} = \beta + t$ und 1. Für $\tilde{\alpha} \in \mathbb{N}$ gilt $\Gamma(\tilde{\alpha}) = (\tilde{\alpha} - 1)!$. Daraus folgt mit

$$\frac{\Gamma(\tilde{\alpha} + x_{t+1})}{\Gamma(\tilde{\alpha}) x_{t+1}!} = \binom{\tilde{\alpha} + x_{t+1} - 1}{\tilde{\alpha} - 1}$$

unmittelbar aus der $PoiGa(\tilde{\alpha}, \tilde{\beta}, 1)$ -Wahrscheinlichkeitsfunktion die Wahrscheinlichkeitsfunktion der $NegBin\left(\tilde{\alpha}, \tilde{\beta}/(\tilde{\beta} + 1)\right)$ -Verteilung. Die Poisson-Gamma-Verteilung ist somit „eine Verallgemeinerung der Negativ-Binomial-Verteilung auf reelle α “ (Held 2008).

Daraus folgt

$$X_{t+1} \mid x_1, \dots, x_t \sim \text{NegBin} \left(\alpha + \sum_{k=1}^t x_k, \frac{\beta + t}{\beta + t + 1} \right) \quad (2.24)$$

(vgl. Höhle und Riebler 2005).

Um „möglichst wenig“ Vorwissen in die Priori-Verteilung einzubringen, kann Jeffreys' Priori, die in diesem Fall der $\text{Gamma}(\frac{1}{2}, 0)$ -Verteilung entspricht, gewählt werden (Held 2008), d. h. in den vorangegangenen Formeln ist $\alpha = \frac{1}{2}$ und $\beta = 0$.

Saisonale Schwankungen

Für das Monitoring von Krankheitsfällen ist die Annahme $X_k \stackrel{iid}{\sim} \text{Poi}$ für alle $k = 1, \dots, t$ nicht realistisch. Da Ausbrüche von Krankheiten meist eine saisonale Abhängigkeit aufweisen, sollten für die oben erläuterte Prognose nicht alle verfügbaren Beobachtungen herangezogen werden. Deshalb wird der Prognose eine Referenzmenge an Beobachtungen zu Grunde gelegt.

Zuallererst wird eine Zeiteinheit festgelegt. In dieser Arbeit wird als Zeiteinheit eine Woche (=7 Tage) vorgegeben. Des Weiteren wird das aktuelle Jahr als Jahr 0 und die aktuelle Woche als Woche s definiert. Die Referenzmenge soll aus jenen Wochen von vergangenen Jahren bestehen, welche die Woche s als zentrale Woche einbetten, und im aktuellen Jahr aus entsprechenden Wochen vor der Woche s bestehen.

Wie in Höhle (2007) wird aufgrund des vorhandenen Datenmaterials die Anzahl b von Vorjahren, die Anzahl w an Wochen um die aktuelle Woche s sowie die Anzahl der Wochen w_0 im Jahr 0 vor Woche s bestimmt. Diese werden dann zur Bildung der Referenzmenge R herangezogen. Sei dazu X_{s_0} die Beobachtung der Woche s im Jahr 0, dann ist

$$R = R(b, w, w_0) = \left(\bigcup_{j=1}^b \bigcup_{i=-w}^w x_{(s+i)_j} \right) \cup \bigcup_{i=-w_0}^{-1} x_{(s+i)_0}. \quad (2.25)$$

Beispiel 2.6. Im aktuellen Jahr soll die Anzahl der Krankheitsfälle in Woche $s = 19$ beurteilt werden. Dazu stehen Beobachtungen der vergangenen $b = 3$ Jahre zur Verfügung. Weiters sind auch die Krankheitsfälle des aktuellen Jahres der ersten 18 Wochen bekannt. Zur Bildung der Menge der Referenzbeobachtungen sollen jeweils vier Wochen vor und nach der 19. Woche in den vorangegangenen b Jahren miteinbezogen werden, d. h. $w = 4$. Vom aktuellen Jahr sollen ebenfalls vier Wochen vor der aktuellen Woche s in die Referenzmenge inkludiert werden, daraus folgt $w_0 = 4$. In Abbildung 2.1 sind die zur Verfügung stehenden

2 On-line Detektions-Algorithmen

Wochen in Form von Balken zu sehen. Der letzte rot-markierte Balken entspricht der Woche 19 im aktuellen Jahr, für die eine Prognose getroffen werden soll. Die restlichen farbigen Balken bilden die Menge der Referenzbeobachtungen. Schraffierte Balken heben die Woche s im jeweiligen vorangegangenen Jahr hervor. Insgesamt könnten $3 \cdot 52 + 18 = 174$ Beobachtungen herangezogen werden. Tatsächlich basiert die Entscheidung des Algorithmus jedoch nur auf $|R| = b(2w + 1) + w_0 = 27 + 4 = 31$ Referenzbeobachtungen, um Auswirkungen von saisonalen Schwankungen zu eliminieren.

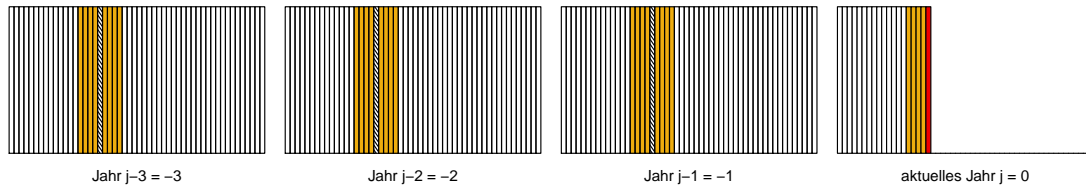


Abbildung 2.1: Referenzbeobachtungen des Bayes-basierten Algorithmus zu Beispiel 2.6.

Für die Bayes-Prognose gilt somit $R = \{y_1, \dots, y_n\}$, $Y_{n+1} = X_{s_0}$ sowie

$$Y_{n+1} \mid y_1, \dots, y_n \sim \text{NegBin} \left(\frac{1}{2} + \sum_{y_k \in R} y_k, \frac{|R|}{|R| + 1} \right).$$

Für einen gewählten Quantilparameter α wird mittels des kleinsten Wertes x_α , für den $P(Y_{n+1} \leq x_\alpha \mid y_1, \dots, y_n) \geq 1 - \alpha$ gilt, die Alarmentscheidung für die aktuelle Woche getroffen. Ein Alarm wird ausgegeben, falls die realisierte Beobachtung $y_{n+1} = x_{s_0} \geq x_\alpha$ ist (Höhle 2007; Höhle und Riebler 2005).

Beispiel 2.7 (Fortsetzung Beispiel 2.6). Angenommen die vorliegenden Beobachtungen x_1, \dots, x_{174} stammen aus einer $Poi(\mu)$ -Verteilung. Für die Menge der Referenzbeobachtungen gilt

$$\begin{aligned} R &= \{x_{(15)_{-3}}, \dots, x_{(23)_{-3}}, x_{(15)_{-2}}, \dots, x_{(23)_{-2}}, x_{(15)_{-1}}, \dots, x_{(23)_{-1}}, x_{(15)_0}, \dots, x_{(18)_0}\} \\ &= \{y_1, \dots, y_{31}\}. \end{aligned}$$

Die posteriori-prädiktive Verteilung für $(Y_{32} \mid y_1, \dots, y_{31}) = (X_{(19)_0} \mid R)$ entspricht daher der $\text{NegBin} \left(0.5 + \sum_{y_k \in R} y_k, 31/32 \right)$ -Verteilung.

Für $\alpha = 0.05$ ist $x_{0.05} = 5$, sodass $P(Y_{32} \leq x_{0.05} \mid y_1, \dots, y_{31}) \geq 0.95$. Werden in der 19. Woche des aktuellen Jahres mehr als fünf Krankheitsfälle gemeldet, ist der Grenzwert überschritten und ein Alarm wird angezeigt (siehe Abbildung 2.2).

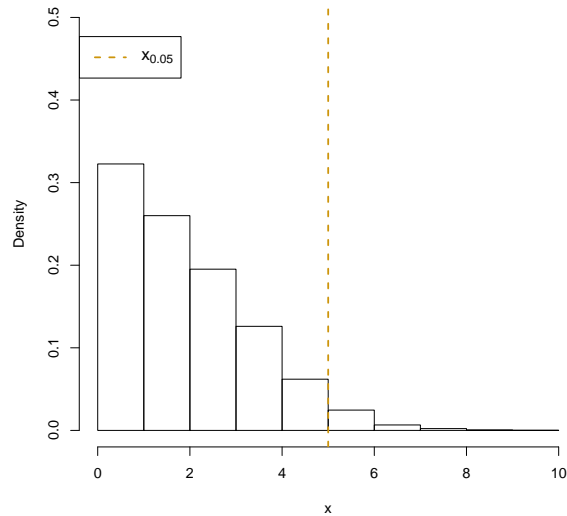


Abbildung 2.2: Histogramm der posterior-prädiktiven Verteilung zu Beispiel 2.7.

2.3 Der Farrington-Algorithmus

Der Algorithmus von Farrington et al. (1996) wurde speziell für die Detektion von Krankheitsausbrüchen entwickelt. Er ist ein *on-line* Algorithmus, der auf die Erkennung von auffällig hohen Anzahlen von Krankheitsfällen, die pro Woche berichtet werden, fokussiert ist.

Mittels eines auf Überdispersion, saisonale und langfristige Trends angepassten log-linearen Modells sollen die Anzahl X_i der Woche t_i und ein oberes Limit geschätzt werden. Auch hier wird bei Überschreitung des gewählten Limits ein Alarm (entspricht einem Krankheitsausbruch) gemeldet.

Wie bei der Methode basierend auf dem Bayes-Ansatz werden auch bei der Methode von Farrington Referenzwerte herangezogen, um so saisonale Schwankungen zu eliminieren. Dazu werden nur Daten von vergleichbaren Perioden in vergangenen Jahren in die Schätzung miteinbezogen. Sei s die aktuelle Woche im Jahr $j = 0$ und X_s die zu vorhersagende Anzahl. Dann werden Beobachtungen der Wochen $[s-w, s+w]$ der Jahre $[j-b, j-1]$ für die Berechnungen herangezogen. Insgesamt fließen somit $n = b(2w+1)$ Wochen als Referenzbeobachtungen $R_{farrington} = \{x_1, \dots, x_n\}$ in die Modellanpassung und somit in die Vorhersage ein.

2 On-line Detektions-Algorithmen

Beispiel 2.8. Wie im Beispiel zur Referenzmengen-Bildung des Bayes-basierten Algorithmus soll auch hier die Anzahl der Krankheitsfälle in Woche $s = 19$ beurteilt werden. Wiederum stehen Beobachtungen der vergangenen $b = 3$ Jahre zur Verfügung. Weiters gelte $w = 4$. Beobachtungen des aktuellen Jahres werden hier nicht berücksichtigt. Die Anzahl der Referenzbeobachtungen entspricht somit $|R_{farrington}| = b(2w + 1) = 27$ gegenüber den 31 Referenzbeobachtungen in Beispiel 2.6. In Abbildung 2.3 definiert der rote Balken im Jahr 0 die aktuelle Woche, für die die Anzahl an Krankheitsfällen prognostiziert werden soll. Die restlichen farbigen Balken bilden die Menge der Referenzbeobachtungen. Schraffierte Balken heben dabei die Woche s im jeweiligen vorangegangenen Jahr hervor.

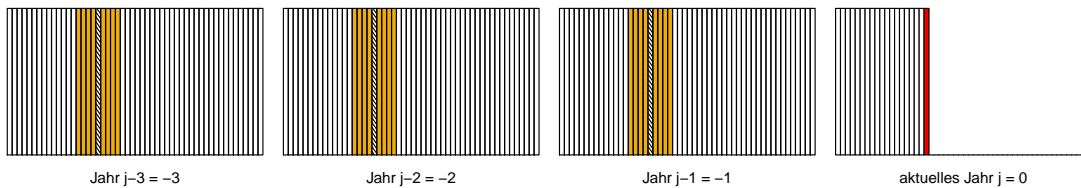


Abbildung 2.3: Referenzbeobachtungen des Farrington-Algorithmus zu Beispiel 2.8.

Der Algorithmus von Farrington et al. (1996) beruht auf neun Schritten, die im folgenden kurz erläutert werden.

Schritt 1. Fitte das Ausgangsmodell

Für die Referenzwerte wird ein Poisson-GLM mit Überdispersion und einem log-link angepasst. D. h. die Anzahl der gemeldeten Krankheitsfälle in Woche i ist eine Zufallsvariable mit $E(X_i) = \mu_i$ und $Var(X_i) = \phi V(\mu_i)$, wobei die Varianzfunktion $V(\mu_i) = \mu_i$ bekannt sei. Da diese Varianzannahme zu keinem Mitglied der Exponentialfamilie korrespondiert, wird anstelle der Likelihood-Funktion die sogenannte Quasi-Likelihood-Funktion zur Parameterschätzung herangezogen (vgl. Friedl 2000).

Für das Modell

$$\log(\mu_i) = \alpha + \beta t_i \quad (2.26)$$

erhält man in einem ersten Schritt den Schätzer für den Dispersionsparameter ϕ (mit Gewichten ω_i , die vorerst alle gleich 1 sind) als

$$\hat{\phi} = \max \left(\frac{1}{n-p} \sum_{i=1}^n \omega_i \frac{(x_i - \hat{\mu}_i)^2}{\hat{\mu}_i}, 1 \right) \quad (2.27)$$

und den Schätzer für μ_i als

$$\hat{\mu}_i = \exp(\hat{\alpha} + \hat{\beta} t_i). \quad (2.28)$$

Bemerkung 2.4. Für die konkrete Anwendung ist insbesondere Überdispersion der Daten von Interesse. Da bei mehrfacher Neugewichtung (siehe Schritte 3-5) die Schätzung $\hat{\phi}$ sehr klein werden kann, wird mittels des Maximums in (2.27) Unterdispersion ausgeschlossen.

Schritt 2. Berechnung der Gewichte ω_i

Um den Einfluss von Krankheitsausbrüchen in den herangezogenen Referenzwochen zu mindern, wird in Farrington et al. (1996) vorgeschlagen, die Referenzbeobachtungen bei der Berechnung des Grenzwertes neuerlich zu gewichten. Für Poisson-verteilte Daten wird hierzu die Gewichtsfunktion

$$\omega_i = \begin{cases} \frac{\gamma}{s_i^2}, & s_i < 1 \\ \gamma, & s_i \geq 1 \end{cases}$$

gewählt, wobei

$$s_i = \frac{3 \left(x_i^{2/3} - \hat{\mu}_i^{2/3} \right)}{2\hat{\mu}_i^{1/6} \sqrt{\hat{\phi}(1 - h_{ii})}}$$

die Residuen sind, welche sich aus den Schätzungen von Schritt 1 ergeben.

Die Werte h_{ii} sind die Diagonalelemente der Hat-Matrix des Modells und γ ist eine vorgegebene Konstante, sodass sich die Gewichte auf n summieren.

Da Beobachtungen mit hohen Residuen eine geringere Gewichtung erhalten, wird der Einfluss von Krankheitsausbrüchen in den Referenzbeobachtungen auf die Berechnung des Thresholds reduziert.

Bemerkung 2.5. Sind die Daten Poisson-verteilt, d. h. gilt $\phi = 1$, so sind s_i die standardisierten Anscombe-Residuen (siehe Davison und Snell 1991).

Schritte 3 - 5.

Nach der Berechnung der Gewichte wird das Modell neuerlich angepasst, wobei eine revidierte Schätzung für ϕ ermittelt und eine Neuskalierung des Modells durchgeführt wird.

Schritt 6. Prüfe Signifikanz des Trends

Eine weitere Verbesserung kann durch genauere Betrachtung des linearen Trends im Modell erzielt werden. So berücksichtigen Farrington et al. (1996) diesen Trend im Modell lediglich, wenn Daten von mindestens drei Jahren vorliegen, der Trend signifikant

2 On-line Detektions-Algorithmen

zum Niveau $\alpha = 0.05$ ist und - um eine unrealistische Extrapolation zu vermeiden - $\hat{\mu}_0 \leq \max(x_i \mid i = 1, \dots, n)$.

Schritt 7. Wiederholung der Schritte 1 - 6

Schritt 8. Berechnung der oberen Grenze

Der Threshold h_0 wird als obere Grenze des $(1 - 2\alpha)\%$ -Prädiktionsintervalls für X_0 (die Beobachtung der aktuellen Woche)

$$\hat{\mu}_0 \pm z_{1-\alpha} \sqrt{\text{Var}(X_0 - \hat{\mu}_0)}$$

definiert, wobei $z_{1-\alpha}$ das $(1 - \alpha)$ -Quantil der $N(0, 1)$ -Verteilung ist. Ist die tatsächlich gemeldete Anzahl $x_0 > h_0 = \hat{\mu}_0 + z_{1-\alpha} \sqrt{\text{Var}(X_0 - \hat{\mu}_0)}$, so wird für diesen Zeitpunkt ein Ausbruch gemeldet.

Korrektur der Schiefe. Vor allem für geringe Anzahlen ist die Berechnung des Prädiktionsintervalls unter Verwendung der Normalverteilungsquantile aufgrund der Schiefe der Poisson-Verteilung (diese entspricht $1/\sqrt{\mu}$) nicht adäquat. Als korrigierende Transformation wird in Farrington et al. (1996) $g(X) = X^{\frac{2}{3}}$ vorgeschlagen. Mittels der Taylor-Approximation von $g(X_0)$ um μ_0

$$X_0^{\frac{2}{3}} \approx \mu_0^{\frac{2}{3}} + (X_0 - \mu_0) \cdot \frac{2}{3} \cdot \mu_0^{-\frac{1}{3}} + \dots$$

erhält man

$$\text{Var}\left(X_0^{\frac{2}{3}}\right) \approx \frac{4}{9} \phi \mu_0^{\frac{1}{3}}, \quad \text{Var}\left(\hat{\mu}_0^{\frac{2}{3}}\right) \approx \frac{4}{9} \mu_0^{-\frac{2}{3}} \text{Var}(\hat{\mu}_0)$$

sowie

$$\text{Var}\left(X_0^{\frac{2}{3}} - \hat{\mu}_0^{\frac{2}{3}}\right) \approx \frac{4}{9} \left(\phi + \frac{\text{Var}(\hat{\mu}_0)}{\mu_0} \right) \mu_0^{\frac{1}{3}}.$$

Nach Rücktransformation folgt die obere Grenze des Prädiktionsintervalls für X_0

$$h_0^{(2/3)} = \hat{\mu}_0 \left(1 + \frac{2}{3} z_{1-\alpha} \sqrt{\frac{\hat{\phi} \hat{\mu}_0 + \text{Var}(\hat{\mu}_0)}{\hat{\mu}_0^2}} \right)^{\frac{3}{2}},$$

welche nun an Stelle von h_0 als Threshold dient.

Varianzstabilisierende Transformation. Ein alternatives Prädiktionsintervall erhält man, wenn man die Daten nach der varianzstabilisierenden Wurzel-Transformation transformiert. Für $g(X) = \sqrt{X}$ ist im Fall $X \sim Poi(\mu)$ die Varianz dann nicht mehr vom Erwartungswert abhängig, denn es gilt $Var(\sqrt{X}) \approx \frac{1}{4}$. Diese Überlegungen führen zu

$$Var(\sqrt{X_0}) \approx \frac{\phi}{4}, \quad Var(\sqrt{\hat{\mu}_0}) \approx \frac{Var(\hat{\mu}_0)}{4\mu_0}, \quad Var(\sqrt{X_0} - \sqrt{\hat{\mu}_0}) = \frac{1}{4} \left(\phi + \frac{Var(\hat{\mu}_0)}{\mu_0} \right)$$

und somit in weiterer Folge zu

$$h_0^{(1/2)} = \hat{\mu}_0 \left(1 + \frac{1}{2} z_{1-\alpha} \sqrt{\frac{\hat{\phi}\hat{\mu}_0 + Var(\hat{\mu}_0)}{\hat{\mu}_0^2}} \right)^2.$$

In allen drei Fällen wird die aktuelle Beobachtung x_0 als Ausbruch markiert, falls sie größer als der angegebene Threshold ist.

9. Output

Schlussendlich wird noch ein „Überschreitungs-Score“

$$\mathbb{X} = \begin{cases} 0, & \text{falls } \leq 5 \text{ Beobachtungen in den letzten 4 Wochen} \\ \frac{x_0 - \hat{\mu}_0}{h_0 - \hat{\mu}_0}, & \text{falls } > 5 \text{ Beobachtungen in den letzten 4 Wochen} \end{cases}$$

berechnet (mit h_0 abhängig davon, welche Transformation in Schritt 8 durchgeführt wurde). Um zu verhindern, dass sporadische Fälle von Krankheiten mit ohnehin geringen Anzahlen zu einem Alarm führen, wird nur für Beobachtungen mit $\mathbb{X} > 1$ ein Alarm ausgegeben.

2.4 Detektions-Algorithmen in R: Das 'surveillance' package

R ist eine unter <http://cran.r-project.org/> frei erhältliche Open-Source-Statistik-Software, die der von John Chambers et al. der Bell Laboratories entwickelten Programmiersprache und -umgebung S sehr ähnlich ist. Mittels verschiedenster *packages* ist die Standardimplementation laufend erweiterbar (R Development Core Team 2012).

Auswertungen mittels Methoden der klassischen statistischen Prozesskontrolle zur Erkennung von Prozessänderungen in stetigen sowie auch in diskreten Daten, welche u.

2 On-line Detektions-Algorithmen

a. in Montgomery (2009) ausführlich beschrieben werden, sind u. a. mit den beiden *packages* 'qAnalyst' von Andrea Spano und 'qcc' von Luca Scrucca möglich.

Ein weiteres *package*, das Tools der statistischen Prozesskontrolle umfasst, ist das von Sven Knoth entwickelte 'spc' *package*, das im Wesentlichen Funktionen zur Evaluierung von Control-Charts enthält. Das 'surveillance' *package*, welches von Höhle et al. entwickelt und 2010 veröffentlicht wurde, beinhaltet statistische Methoden zum Monitoring von diskreten Zeitreihen mit dem Ziel, Änderungen in Datenreihen, die aus der Gesundheitsüberwachung stammen, zu detektieren (vgl. Höhle et al. 2012).

Das *package* 'surveillance' umfasst alle Implementationen der im vorangegangenen Kapitel erläuterten Methoden. Es wurde daher für die in Kapitel 5 dargestellte Analyse von in Österreich gemeldeten Campylobacteriose-Erkrankungen ausgewählt. Im Folgenden soll kurz auf Funktionen des *package*, die für diese Analyse von Bedeutung sind, eingegangen werden. Für eine umfassendere Beschreibung siehe Höhle (2007) sowie Höhle et al. (2012).

2.4.1 Allgemeine Beschreibung der Datenstruktur

Für die Anwendung der Methoden sind zwei Datenstrukturen von Bedeutung. Einerseits die S3-Klasse `disProg` für univariate Zeitreihen sowie die S4-Klasse `sts` (*surveillance time series*), die für die Analyse multivariater Zeitreihen herangezogen werden kann. Beide Datenstrukturen verbinden die Beobachtungen mit der angegebenen Zeitperiode - `freq = 52` steht hier für 52 Wochen. Weiters kann ein Startjahr sowie ein Startmonat mit `start=c(2001,1)` und der Status der Zeitreihe mittels `state` (*in-control* $\equiv 0$, *out-of-control* $\equiv 1$) definiert werden.

```
> set.seed(123)
> t <- 1:208
> x <- rpois(length(t), mu.a1)
> dP <- create.disProg(week=t, observed=x,
                      start=c(2001,1), state=rep(0,length(t)), freq=52); str(dP)
List of 8
 $ week          : int [1:208] 1 2 3 4 5 6 7 8 9 10 ...
 $ observed      : num [1:208, 1] 0 1 0 2 2 0 1 2 1 0 ...
 ..- attr(*, "dimnames")=List of 2
 .. ..$ : NULL
 .. ..$ : chr "observed"
 $ state        : num [1:208, 1] 0 0 0 0 0 0 0 0 0 0 ...
 ..- attr(*, "dimnames")=List of 2
 .. ..$ : NULL
 .. ..$ : chr "observed"
 $ start        : num [1:2] 2001 1
 $ freq         : num 52
```

2.4 Detektions-Algorithmen in R: Das 'surveillance' package

```
$ neighbourhood : NULL
$ populationFrac: num [1:208, 1] 1 1 1 1 1 1 1 1 1 1 ...
$ epochAsDate   : logi FALSE
- attr(*, "class")= chr "disProg"
```

Mit `disProg2sts` kann jedes Objekt der Klasse `disProg` zu einem Objekt der Klasse `sts` umgewandelt werden. Die `sts` Klasse bildet eine Erweiterung für Zeitreihen, die gleichzeitig erfasst wurden. So können in dieser Klasse zum Beispiel gemeldete Krankheitsfälle verschiedener Bundesländer aufgezeichnet werden. Mithilfe der zusätzlichen Slots `map` und `neighbourhood` können auch räumliche Beziehungen in den Zeitreihen erfasst werden.

2.4.2 CUSUM-Detektoren

Der einseitige CUSUM-Detektor ist mittels der Funktion `algo.cusum` anwendbar.

- Mit `range` werden die zu untersuchenden Wochen definiert.
- `k` und `h` korrespondieren mit den Parametern κ und h aus Kapitel 2.1. Diese können auch mit den Funktionen `findH`, `findK` und `find.kh` bestimmt werden.
- `alpha` legt den Parameter der reparametrisierten Negativ-Binomial-Verteilung fest, sodass $Var = \mu + \alpha \cdot \mu^2$.
- Mittels `trans="rossi"` wird die Normalverteilungsapproximation nach Rossi et al. (1999) durchgeführt.

Die Adaption von Rogerson und Yamada (2004) für Poisson-Zufallsvariablen verwendet anstelle konstanter Werte für κ und h zeitabhängige Parameter. Die zeitabhängigen Werte h_k werden bei bekanntem *in-control* Mittelwert mit der Funktion `hValues` ermittelt. Der angepasste CUSUM-Detektor ist über die Funktion `algo.rogerson` aufrufbar.

Der GLR-Ansatz für den unbekanntem *out-of-control* Parameter θ_1 wird mit zwei Methoden durchgeführt: `algo.glrpois` für Poisson- und `algo.glrnb` für Negativ-Binomial-Zufallsvariablen. Wird `mu0` nicht spezifiziert, dann wird der *in-control* Mittelwert durch ein generalisiertes lineares Modell für Poisson- bzw. NegBin-Zufallsvariablen geschätzt.

- Mit `mu0 = list(S = 1, trend = TRUE, refit = FALSE)` kann die Anzahl der zu inkludierenden harmonischen Schwingungen bestimmt sowie ein Trend berücksichtigt werden. Des Weiteren kann mit `refit = TRUE` angegeben werden, dass das GLM nach Ausgabe eines Alarms neu angepasst wird. In diesem Fall werden

2 On-line Detektions-Algorithmen

alle Beobachtungen bis zum Zeitpunkt des letzten Alarms zur Anpassung des neuen Modells herangezogen (vgl. Wimmer und Höhle 2008).

- `c.ARL` bestimmt den Threshold h , wie in (2.10).
- Für `theta=NULL` wird der MLE $\hat{\eta}_{k,t}$ für den unbekannt additiven Zuwachs auf der Logarithmus-Skala η rekursiv berechnet. Wird die zu detektierende Änderung bereits vor Beginn der On-line Detektion festgelegt - z. B. durch Expertenwissen - wird für `theta` dieser definierte Wert übergeben. In diesem Fall handelt es sich dann um einen gewöhnlichen Likelihood-Ratio-Detektor.
- Die Ermittlung des Maximums über alle Zeitpunkte $k \in \{1, \dots, t\}$ in (2.20) bzw. in (2.21) ist für eine hohe Anzahl an vorliegenden Beobachtungen t sehr aufwendig. In diesem Fall kann es sinnvoll sein, die Window-limited-GLR-Methode zu bemühen. Diese wurde erstmals in Willsky und Jones (1976) vorgestellt. Wie in Höhle und Paul (2008) erläutert, wird hier lediglich über $k \in \{t-M, \dots, t-\tilde{M}+1\}$ maximiert. Dabei beschreibt M die Anzahl der Beobachtungen, die aus der Vergangenheit berücksichtigt werden sollen. \tilde{M} ist die Anzahl der Beobachtungen, um eine zuverlässige Schätzung zu erhalten, wobei hier $\tilde{M} = 1$ ausreichend ist (vgl. Höhle und Paul 2008). Allgemein gilt $1 \leq \tilde{M} < M$. Sollen zu jedem Zeitpunkt alle vorherigen zur Verfügung stehenden Beobachtungen in die Schätzung eingehen, kann dies durch $M = -1$ erreicht werden.
- Als Output erhält man
 - i. die Zeitpunkte, in denen das Supremum der Summe der Log-Likelihoodquotienten den gewählten Threshold überschritten hat und
 - ii. je nach Wahl entweder den Wert dieser Suprema zum jeweiligen Zeitpunkt (`ret= "value"`) oder die notwendige Anzahl an Krankheitsfällen für die Überschreitung des Thresholds (`ret="cases"`).

2.4.3 Bayes-basierter Algorithmus

Der Befehl `algo.bayes` ermöglicht die Analyse einer univariaten Zeitreihe von gemeldeten Krankheitsfällen. Da der in Kapitel 2.2 erläuterte Algorithmus auf einer Menge von Referenzwerten basiert, ist anzugeben, für wieviele Vorjahre Daten vorliegen und wieviele Wochen um die aktuelle Woche berücksichtigt werden sollen.

Mit den Einstellungen `b=3`, `w=4` wird zum Beispiel definiert, dass die Alarmentscheidung der aktuellen Woche auf Beobachtungen der letzten drei Jahre beruht. Um saisonale Schwankungen zu eliminieren, werden in den vorangegangenen Jahren lediglich vier

2.4 Detektions-Algorithmen in R: Das 'surveillance' package

Wochen vor bzw. nach der aktuellen Woche betrachtet. Bei Festlegung der Parameter \mathbf{b} und \mathbf{w} ist vor allem darauf zu achten, dass aus der Vergangenheit tatsächlich zumindest $\text{freq} \cdot b + w$ Beobachtungen vorliegen.

Durch `actY = TRUE` fließen im aktuellen Jahr zusätzlich die letzten vier Wochen vor der aktuellen Woche in die Menge der Referenzbeobachtungen ein.

Der zu wählende Quantilparameter α kann durch `alpha` festgelegt werden. Abbildung 2.4 zeigt anhand einer simulierten Zeitreihe von Krankheitsfällen exemplarisch die Auswirkung der unterschiedlichen Wahl von α . Analysiert wurde dafür die wöchentliche Anzahl an gemeldeten Krankheitsfällen für die Jahre 2001 - 2004. Die erste Bayes-Prognose wurde somit für die erste Woche im Jahr 2001 berechnet. Für diese und alle weiteren Prognosen wurde die Referenzmenge aus Beobachtungen der letzten $b = 3$ Jahre und $w = 4$ Wochen gebildet, wobei zusätzlich `actY = TRUE` gesetzt wurde. Die Analyse wurde für drei verschiedene, in der Praxis gängige Werte für α ermittelt. Zu sehen sind die dafür jeweils errechneten oberen Schranken, bei deren Überschreiten das System einen Krankheitsausbruch meldet.

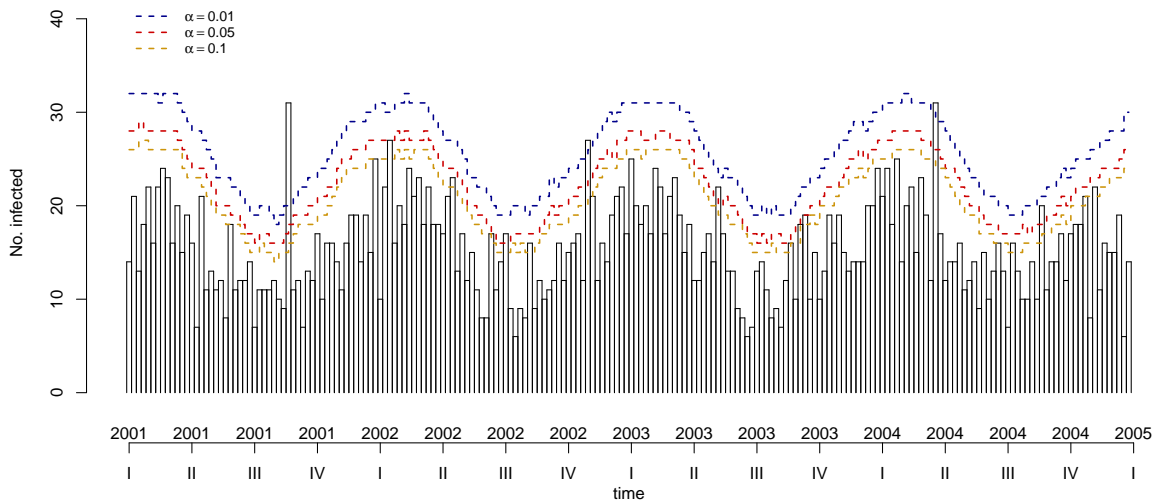


Abbildung 2.4: Obere Schranken des Bayes-basierten Algorithmus für unterschiedliche Wahl von α für Parametereinstellung $\mathbf{b} = 3$, $\mathbf{w} = 4$, `actY = TRUE`.

Die Wahl $\alpha = 0.01$ garantiert, dass $P(Y_{n+1} \leq x_{0.01} | y_1, \dots, y_n) \geq 0.99$ gilt. Hingegen wird für $\alpha = 0.10$ lediglich garantiert, dass diese Wahrscheinlichkeit mindestens 90% beträgt. Dementsprechend ist die obere Schranke für $\alpha = 0.01$ durchwegs wesentlich höher als für $\alpha = 0.05$ bzw. $\alpha = 0.1$. Dies hat zur Folge, dass für die Wahl von $\alpha = 0.10$

die häufigste Anzahl an Alarmen (in diesen drei Varianten) erreicht wird. Demnach ist auch die Wahrscheinlichkeit für einen Fehlalarm am höchsten, wobei gleichzeitig die Wahrscheinlichkeit, alle tatsächlichen Krankheitsausbrüche zu detektieren, gegenüber den anderen Varianten steigt.

2.4.4 Farrington-Algorithmus

Der Befehl `algo.farrington` führt die in Kapitel 2.3 beschriebenen Schritte zur Analyse einer univariaten Zeitreihe aus.

- Da auch die Methode von Farrington et al. (1996) auf einer Menge von Referenzbeobachtungen basiert, sind auch hier die Parameter `b`, `w` festzulegen. Allerdings fließen hier keinerlei Beobachtungen aus dem aktuellen Jahr in die Referenzmenge ein.
- Mit `reweight = TRUE` kann angegeben werden, dass eine Neugewichtung der Referenzbeobachtung nach einer ersten Schätzung erfolgt (siehe Schritte 2–5 in Kapitel 2.3).
- `trend = TRUE` legt fest, dass ein linearer Trend bei der Schätzung des *in-control* Mittelwerts berücksichtigt werden soll. Dieser wird aber nur dann tatsächlich berücksichtigt, falls
 - i. Daten von mindestens drei Jahren zur Schätzung vorliegen,
 - ii. der Trendparameter signifikant zum Niveau von 5% ist und
 - iii. $\hat{\mu}_0 \leq \max\{x_i \mid i = 1, \dots, n\}$ gilt (vgl. Farrington et al. 1996).
- Die Berechnung des oberen Grenzwertes wird mit den Parametern `alpha` und `powertrans` spezifiziert. Die $\frac{2}{3}$ -Potenz-Transformation zur Korrektur der Schiefe wird mit `powertrans = "2/3"`, die varianzstabilisierende Wurzel-Transformation mit `powertrans = "1/2"` durchgeführt. Abbildung 2.5 zeigt die unterschiedlichen oberen Schranken, die aus der Wahl verschiedener Transformationen resultieren, anhand einer simulierten Zeitreihe. Analysiert wurde wiederum die wöchentliche Anzahl an gemeldeten Krankheitsfällen für die Jahre 2001-2004. Die Referenzmenge wurde aus Beobachtungen der letzten $b = 3$ Jahre und $w = 4$ Wochen gebildet. Das Signifikanzniveau wurde mit `alpha = 0.10` festgelegt. Für jeden Zeitpunkt ist somit der Threshold h_0 zu sehen. Weiters sind die Threshold-Werte $h_0^{(1/2)}$ (rot) sowie $h_0^{(2/3)}$ (dunkelgelb) dargestellt. In diesem Fall ist der stärkste Unterschied der Schranken ohne (dunkelblaue Linie) bzw. mit Wurzel-Transformation (rote Linie) zu erkennen. Die Schranke h_0 (ohne Transformation) wird hier am häufigsten überschritten.

2.4 Detektions-Algorithmen in R: Das 'surveillance' package

Der berechnete Threshold ist die obere Grenze des jeweiligen $(1 - 2\alpha)\%$ -Prädiktionsintervalls. Das heißt, für `alpha = 0.05` erhält man ein einseitiges Signifikanzniveau von $0.05/2 = 0.025$. Vergleicht man für α wiederum die Werte 0.01, 0.05 und 0.10, erhält man die oberen Schranken der korrespondierenden 99.5%-, 97.5%- und 95%-Prädiktionsintervalle. Dementsprechend sind die oberen Schranken für $\alpha = 0.01$ durchwegs höher als jene für $\alpha = 0.05$ bzw. $\alpha = 0.10$. Wiederum erhält man für $\alpha = 0.10$ die höchste Wahrscheinlichkeit für einen Fehlalarm.

- Der Score \mathbb{X} in Schritt 9 des Algorithmus wird nach Farrington et al. (1996) auf 0 gesetzt, falls es weniger als 5 Beobachtungen in den letzten 4 Wochen gab. In der Implementation des *package* 'surveillance' korrespondiert dies mit der Einstellung `limit54 = c(5,4)`.

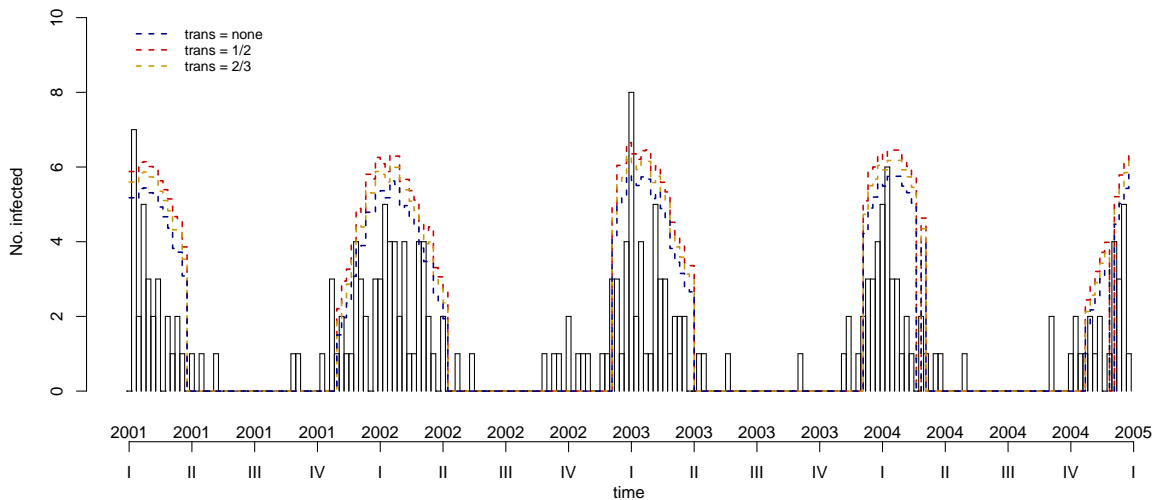


Abbildung 2.5: Obere Schranken des Farrington-Algorithmus (mit $\mathbf{b} = 3$, $\mathbf{w} = 4$) für unterschiedliche Transformationen.

3 Optimalitätskriterien und Performance

Ein optimaler Detektionsalgorithmus soll einerseits Veränderungen möglichst schnell detektieren und andererseits die Anzahl der *false alarms* (Fehlalarme) möglichst gering halten (Frisén 2007b). Seit Entwicklung der Theorie der statistischen Qualitätskontrolle haben sich verschiedene Optimalitätskriterien etabliert. Es wurden immer wieder Methoden vorgestellt, welche bzgl. eines Kriteriums optimal sind (vgl. Friséen 2003). In Theorie und Praxis reicht ein einziges Optimalitätskriterium meist jedoch nicht aus, um einen Algorithmus bewerten zu können. Einen Überblick über verschiedene Performance-Maße und ihre Anwendung auf verschiedene Methoden bieten Basseville und Nikiforov (1993), Friséen (2003) und Friséen (2007b), nach denen sich auch die Diskussion im folgenden Kapitel orientiert.

In jedem Fall ist die Performance einer Methode allerdings vom unbekanntem Änderungszeitpunkt τ und somit von entsprechenden Annahmen abhängig. So ist es in erster Linie wichtig zu unterscheiden, ob τ ein unbekannter Parameter oder eine Zufallsvariable mit Wahrscheinlichkeiten $\pi(t) = P(\tau = t)$ ist (vgl. Friséen 2003).

Eine zentrale Rolle für die Performance-Analyse spielt die *run length* (RL , *Laufzeit*), d. h. die Anzahl der Beobachtungen vom Beginn der Kontrolle bis zum ersten Alarm (vgl. Hawkins und Olwell 1998) sowie die daraus abgeleiteten Zufallsvariablen (vgl. Höhle und Mazick 2010)

- $T_A | \tau = \infty \dots$ *in-control* Laufzeit (Anzahl der Zeiteinheiten vor dem ersten *false alarm*.)
- $T_A | \tau = 1 \dots$ *out-of-control* Laufzeit (Anzahl der benötigten Zeiteinheiten, um eine bereits eingetretene Änderung zu detektieren.).

3.1 Die ARL und das minimale ARL_1 -Kriterium

Sei τ ein nicht-zufälliger unbekannter Parameter. Für diese Annahme wird häufig der Erwartungswert der *in-control* bzw. der *out-of-control* Laufzeit als Performance-Maß

3 Optimalitätskriterien und Performance

herangezogen.

Definition 3.1.

$$ARL_0 = E(T_A | \tau = \infty) \quad (3.1)$$

heißt durchschnittliche *in-control* Laufzeit.

$$ARL_1 = E(T_A | \tau = 1) \quad (3.2)$$

heißt durchschnittliche *out-of-control* Laufzeit.

Bemerkung 3.1. Obwohl diese Maße sehr häufig in der Literatur diskutiert werden und auch für die Praxis von großer Bedeutung sind, ist darauf zu achten, dass die ARL_1 lediglich eine Änderung zu Beginn der Überwachung berücksichtigt.

Nach dem in der statistischen Prozesskontrolle vielfach diskutierten Optimalitätskriterium „Minimale ARL_1 “ minimiert ein optimaler *on-line* Detektionsalgorithmus die ARL_1 für eine festgelegte ARL_0 . Der beachtliche Vorteil dieses Kriteriums ist, dass keinerlei Annahmen über die Verteilung des Änderungszeitpunkts τ erforderlich sind.

Unter Verwendung dieses Optimalitätskriteriums wird das Hauptaugenmerk auf frühe Änderungen gelegt. Dieses Kriterium sollte lediglich als deskriptives Optimalitätsmaß dienen, da Methoden, die alle Beobachtungen gleich gewichten, nach Frisé (2003) dieses Optimalitätskriterium erfüllen. Dies widerspricht allerdings der Idee einiger Methoden, welche die letzten Beobachtungen stärker gewichten als frühere (vgl. auch Frisé 2007a).

3.1.1 CUSUM und ARL-Kriterien

Für den CUSUM-Algorithmus werden als Optimalitätsmaße sehr häufig ARL_0 und ARL_1 herangezogen. Vor allem für die Wahl des Parameters κ und des Limits h ist die durchschnittliche Laufzeit ARL entscheidend. So wird durch die Wahl dieser Parameter eine ARL impliziert. Andererseits wird durch Wahl von κ und der ARL das Detektionslimit h festgelegt. Natürlich ist die Wahl wesentlich von der Verteilung der zu untersuchenden Zeitreihe abhängig, weshalb für diskrete Verteilungen wie der Poisson- und der Negativ-Binomial-Verteilung auch die zur Verfügung stehenden Parameterwerte beschränkt sind. D. h., wird eine ARL gewählt, so kann nur ein Wert h angegeben werden, sodass die korrespondierende ARL zumindest der festgelegten entspricht (vgl. Hawkins und Olwell 1998).

3.1 Die ARL und das minimale ARL_1 -Kriterium

Wie Moustakides (1986) gezeigt hat, ist der CUSUM-Algorithmus für bekannte *in-* und *out-of-control* Verteilung mit unbekanntem Änderungszeitpunkt τ optimal in dem Sinne, dass er unter allen Algorithmen mit derselben durchschnittlichen *in-control* Laufzeit ARL_0 die geringste erwartete *out-of-control* Laufzeit ARL_1 aufweist.

Wahl des Parameters h

Für die Festlegung des Thresholds h müssen im Wesentlichen entweder die ARL_0 und κ oder aber eine gewünschte ARL_0 und ARL_1 bestimmt werden.

Beispiel 3.1. Für $X \sim Poi(10)$ soll ein additiver Zuwachs im Mittelwert von $2 \cdot \sqrt{10}$ (entspricht einem Zuwachs der doppelten Standardabweichung) erkannt werden. Dabei soll eine durchschnittliche *in-control* Laufzeit von $ARL_0 = 500$ erreicht werden. In R kann der gesuchte Threshold h folgend ermittelt werden

```
> findH(ARL0=500, theta0=10, s = 2, rel.tol = 0.03, roundK = TRUE,
+       distr = "poisson", digits = 1, FIR = FALSE)
      theta0      h      k      ARL      rel.tol
10.00000000  9.40000000 12.90000000 546.44640939 0.09289282
```

Theta0 ist hierbei der *in-control* Mittelwert μ_0 . Der *out-of-control* Mittelwert wird bestimmt durch

$$\mu_1 = \mu_0 + s\sqrt{\mu_0} = 10 + 2 \cdot \sqrt{10} = 16.32.$$

Der Wert k im R-Output ist der Referenzwert

$$\kappa = \frac{\mu_1 - \mu_0}{\log(\mu_1) - \log(\mu_0)} = \frac{10 - 16.32}{\log(10) - \log(16.32)}.$$

Die tatsächlich erreichte $ARL_0 = 546.44$, für $h = 9.4$ und $\kappa = 12.9$ ist etwas höher als der geforderte Wert von 500.

Ermittlung der ARL

Es gibt im Allgemeinen drei Möglichkeiten, um die ARL für CUSUM-Methoden zu ermitteln:

- i. Lösung von Integralgleichungen,
- ii. Diskrete Approximation der Integralgleichungen mittels Markov-Ketten-Ansatzes,
- iii. Monte-Carlo-Simulation.

3 Optimalitätskriterien und Performance

Methoden 1 und 2 sind zum Beispiel für den GLR-Ansatz für unbekannte *in-* und *out-of-control* Parameter nicht realisierbar. Auch im Vergleich mit anderen Algorithmen gestaltet sich die Anwendung dieser Methoden als schwierig, da diese im Allgemeinen nur für CUSUM-Algorithmen gültig sind. Um annähernd äquivalente Voraussetzungen zu schaffen, bietet sich somit die Monte-Carlo-Simulation für einen praktischen Vergleich an.

Um trotzdem eine Idee des Markov-Ketten-Ansatzes zu vermitteln, soll in dieser Arbeit lediglich kurz auf den von Brook und Evans (1972) publizierten Ansatz eingegangen werden. Die nachfolgenden Ausführungen orientieren sich ebenfalls an Hawkins (1992). In dieser Publikation findet sich ein FORTRAN-Code zur Idee von Brook und Evans (1972) mit einigen Verfeinerungen.

Nach Definition der CUSUM C_n erfüllt diese die Markov-Eigenschaft

$$P(C_k = c_k \mid C_0 = c_0, \dots, C_{k-1} = c_{k-1}) = P(C_k = c_k \mid C_{k-1} = c_{k-1}).$$

Sind κ und h rationale Zahlen mit gemeinsamen Nenner N , so ist die CUSUM zur Analyse von diskreten Zufallsvariablen ein diskreter Markov-Prozess mit einem diskreten Zustandsraum der Nh Zustände $\mathcal{Z} = \{0, \dots, M\}$. Es sei $K = \lfloor \kappa N + 0.5 \rfloor$, $M = \lfloor hN - 0.5 \rfloor$. Die kumulierte Summe C_k kann lediglich Werte annehmen, die ganzzahlige Vielfache von $\frac{1}{N}$ sind, d.h.

$$\begin{cases} C_k \in \text{Zustand } E_0, & \text{falls } C_k = 0 \\ C_k \in \text{Zustand } E_i, & \text{falls } C_k = i/N, \quad i = 1, \dots, M-1 \\ C_k \in \text{Zustand } E_M, & \text{falls } C_k > h \end{cases}$$

mit $(M+1) \times (M+1)$ Übergangsmatrix \mathbf{P} und Übergangswahrscheinlichkeiten

$$p_{ij} = P(C_k \in E_j \mid C_{k-1} \in E_i).$$

Sei F die Verteilungsfunktion der zu untersuchenden Zufallsvariable X , so können die Übergangswahrscheinlichkeiten wie folgt berechnet werden

$$\begin{aligned} p_{00} &= F\left(\frac{K}{N}\right) \\ p_{i0} &= F\left(\frac{K-i}{N}\right) \quad i = 1, \dots, M \\ p_{0j} &= F\left(\frac{K+j}{N}\right) - F\left(\frac{K+j-1}{N}\right) \quad j = 1, \dots, M \\ p_{ij} &= F\left(\frac{K+j-i}{N}\right) - F\left(\frac{(K+j-i)-1}{N}\right) \quad i, j = 1, \dots, M. \end{aligned}$$

3.1 Die ARL und das minimale ARL_1 -Kriterium

Die gesuchte Größe ARL ist in diesem Kontext die durchschnittliche Anzahl an Schritten zwischen „Besuchen“ des Zustands E_M . Die ARL_0 ist die Anzahl an Schritten, die benötigt wird, um zum ersten Mal in den Zustand E_M zu gelangen (vgl. auch Hawkins und Olwell 1998). Die Lösung der Gleichung

$$(I - \mathbf{P})\lambda = \mathbf{1} \quad (3.3)$$

enthält die gewünschte ARL sowohl für die Standard-CUSUM als auch für die FIR-CUSUM. Denn λ_i ist gleich der durchschnittlichen Laufzeit, wenn C_0 in Zustand i ist, d. h. λ_0 ist die ARL für $C_0 = 0$, die durchschnittliche Laufzeit für einen *fast initial response* Detektor mit $V = \frac{h}{2}$ ist hier $\lambda_{(M+1)/2+1}$.

Die Funktion `ar1Cusum` des R *package* 'surveillance' führt die oben erläuterten Schritte durch. Zusätzlich kann eine *winsorization* Konstante W zur Robustifizierung gegenüber Ausreißern festgelegt werden. Beobachtungen, die größer als diese Konstante sind, werden dann in der Berechnung der kumulierten Summe durch W ersetzt. Das heißt W ist der maximal mögliche Wert, den eine einzelne Beobachtung zur CUSUM beitragen kann. Gilt $W < \kappa + h$, wird somit verhindert, dass eine einzelne Beobachtung zu einem Alarm führt. Dies reduziert Fehlalarme, falls nur dauerhafte Änderungen in der Zeitreihe erkannt werden sollen. W sollte des Weiteren nicht zu gering sein, damit genügend Beobachtungen mit ihrem tatsächlichen Wert in die CUSUM einfließen (Hawkins 1992). Der Parameter `theta` entspricht dem Verteilungsparameter der Verteilung von X . Im nachfolgenden Beispiel wird also die ARL für $X \sim Poi(3)$ mit $\kappa = 3$ und $h = 10$ bestimmt.

Beispiel 3.2. Bestimmt werden soll die *average run length* eines Poisson-CUSUM-Detektors. Für die zu untersuchenden Zufallsvariablen gelte $X_i \sim Poi(3)$ mit $\kappa = 3$ und $h = 10$.

```
ar1Cusum(h=10, k=3, theta=3, distr="poisson", W=NULL, digits=1)
$ARL
[1] 45.13
$FIR.ARL
[1] 33.75844
```

Zusätzlich zur ARL des CUSUM-Detektors mit $C_0 = 0$ wird die ARL eines FIR-CUSUM-Detektors mit $C_0 = \frac{h}{2} = 5$ ausgegeben.

Bemerkung 3.2. Wie bereits erwähnt, ist die exakte Bestimmung der ARL nur für ausgewählte Methoden möglich. Im Hinblick auf den Vergleich aller Methoden, die in dieser Arbeit diskutiert werden sollen, wird in weiterer Folge deshalb das Hauptaugenmerk auf Monte-Carlo-Simulation zur Ermittlung der ARL gelegt.

Bemerkung 3.3 (Kritik an ARL). Wird die ARL als Performance-Maß zur Analyse eines Algorithmus herangezogen, sollte man sich bewusst machen, dass der Erwartungswert die Verteilung der Laufzeit - aufgrund ihrer Schiefe - nicht optimal widerspiegelt (vgl. Hawkins und Olwell 1998).

3 Optimalitätskriterien und Performance

Daher kann es auch sinnvoll sein, die gesamte Verteilung der Laufzeit zu untersuchen oder aber die median run length (MRL) anstelle der ARLs zu verwenden, wie etwa in Gan (1993).

Des Weiteren ergibt sich bei Schätzung der ARL die Frage, wie mit nicht beobachteten Laufzeiten umgegangen wird. Diese können auftreten, wenn eine begrenzte Zeitperiode analysiert werden soll, da der Erwartungswert $E(T_A | \tau = \infty)$ sehr groß werden kann (siehe auch Beispiel 3.3).

Beispiel 3.3. Mittels Monte-Carlo-Simulation soll die ARL_0 eines CUSUM-Detektors mit der Adaption nach Rogerson und Yamada (2004) geschätzt werden. Dazu wurde für einen Zeitraum von 80 Jahren (entspricht 4160 Wochen) 1000-mal eine Zeitreihe mit saisonalem Trend und *in-control* Zustand simuliert (entspricht Situation A1 in Kapitel 4.1). Anhand jeder Replikation wurde jene Woche vermerkt, in der der erste Fehlalarm berichtet wurde. Bei zwei Replikationen (2 %) konnte diese Woche nicht bestimmt werden, da sie außerhalb des untersuchten Zeitraums lag (> 4160). Solche Fälle wurden als NA-Werte behandelt und bei der Mittelwerts- bzw. Median-Berechnung ignoriert. Die restlichen 998 Werte ergaben den MC-Schätzer für die ARL_0 .

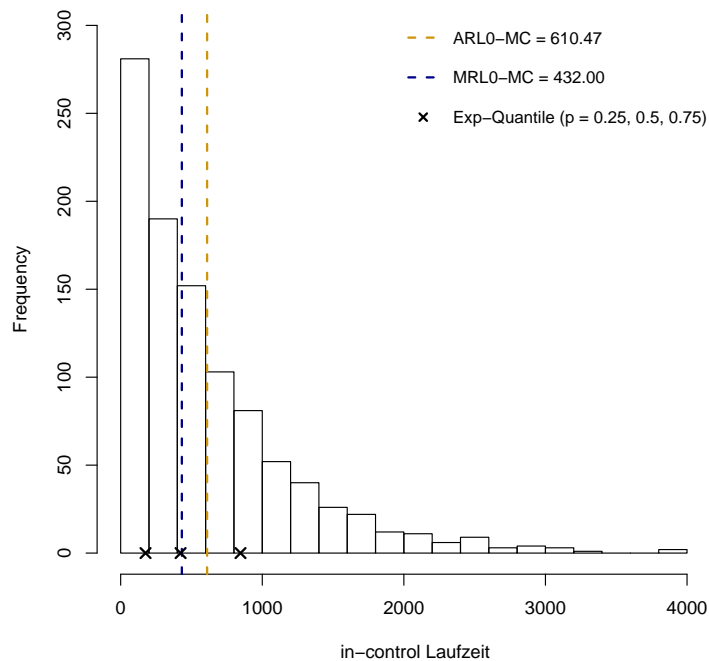


Abbildung 3.1: Histogramm von 1000 simulierten Laufzeiten bei Anwendung des Rogerson-CUSUM-Detektors.

Abbildung 3.1 zeigt den resultierenden Monte-Carlo-Schätzer für die *in-control* Laufzeit ARL_0 (entspricht dem Mittelwert) sowie den Median der durch Simulation erhaltenen Laufzeiten (entspricht dem MC-Schätzer für die *in-control median run length* MRL_0). Das heißt, nach durchschnittlich 610 Wochen ist hier mit dem ersten Fehlalarm zu rechnen. Betrachtet man den MC-Schätzer \widehat{MRL}_0 , würde man von einem ersten Fehlalarm in der 432-ten Woche ausgehen.

Das Histogramm in Abbildung 3.1 lässt weiters vermuten, dass die Laufzeit einer $Exp(\lambda)$ -Verteilung unterliegt. Dies ist tatsächlich ein asymptotisches Ergebnis, das in Lai (1995) gezeigt wird (siehe dazu auch Höhle und Paul 2008). Der MLE für den Parameter λ der Exponentialverteilung ist hier gerade $1/\widehat{ARL}_0$. In Abbildung 3.1 sind auch die Werte, der aus diesen Überlegungen folgenden Quantilschätzer (für $x_p = F^{-1}(p)$)

$$\hat{x}_p = -\widehat{ARL}_0 \cdot \log(1 - p)$$

für $p = 0.25, 0.5, 0.75$ ersichtlich.

3.2 Die false alarm probability

Definition 3.2. Die Wahrscheinlichkeit für zumindest einen Fehlalarm im Beobachtungszeitraum $[0, t]$ wird durch

$$P_{fa}(t) = P(T_A \leq t \mid \tau = \infty) \quad (3.4)$$

definiert und wird auch als *false alarm probability* oder Fehlalarm-Wahrscheinlichkeit bezeichnet.

Die *false alarm probability* ist ebenfalls ein Performance-Maß, das sich den Fehlalarmen widmet. Sie ist aber vor allem für einen kurzen Beobachtungszeitraum oftmals praktikabler als die ARL, da die Erwartung $E(T_A \mid \tau = \infty)$ sehr groß werden kann. Unabhängig von der Wahl von t kann so das Problem von nichtbeobachteten Realisierungen umgangen werden.

Häufig wird t durch die gesamte Anzahl an verfügbaren Beobachtungen bestimmt. So kann im Falle des CUSUM-Algorithmus die *false alarm probability* für die zu untersuchende Zeitreihe mithilfe des Thresholds h kontrolliert werden, sofern die Beobachtungen x_1, \dots, x_t aus der *in-control* Verteilung stammen (vgl. Höhle und Mazick 2010).

Bemerkung 3.4. Alternativ kann auch die Detektions-Wahrscheinlichkeit

$$P_{fa}^1(t) = P(T_A \leq t \mid \tau = 1) \quad (3.5)$$

3 Optimalitätskriterien und Performance

analysiert werden (Höhle und Mazick 2010). Sie findet Anwendung, falls zum Beispiel die Detektionswahrscheinlichkeit eines Ausbruchs innerhalb der ersten drei Wochen nach Ausbruch, d. h. $P(T_A \leq 3 \mid \tau = 1)$, untersucht werden soll.

3.3 Sensitivität und Spezifität

Zwei weitere Optimalitätskriterien, die im medizinischen Kontext immer wieder herangezogen werden, sind die *Sensitivität* und die *Spezifität*. Die Bezeichnungen stammen aus der Beurteilung von diagnostischen Tests. Die Sensitivität beschreibt dabei den Anteil *korrekt positiv* getesteter Personen an der Gesamtanzahl der Erkrankten. Sie gibt demnach die Wahrscheinlichkeit an, mit der ein diagnostischer Test eine erkrankte Person auch tatsächlich als krank identifiziert (DNeM 2011).

Die Spezifität hingegen gibt den Anteil der *korrekt negativ* getesteten Personen an der Gesamtanzahl der Erkrankten an.

Im Kontext der on-line Detektion von Krankheitsausbrüchen ist die Ermittlung dieser Performance-Maße nur möglich, wenn von Fachexperten tatsächliche Krankheitsausbrüche festgestellt werden konnten und diese in Form einer zusätzlichen Variable in den zu untersuchenden Datensatz eingebunden werden. Diese Variable beschreibt dann den tatsächlichen Status $d(t)$ für jede Zeiteinheit.

Ein Ausbruch gilt als korrekt detektiert, falls $\hat{d}(t) = d(t) = 1$ gilt (werden auch als *true positive* bezeichnet). Analog wird für den Fall $\hat{d}(t) = d(t) = 0$ richtigerweise kein Alarm berichtet (auch bekannt als *true negative*). Mit *false positive* werden Fälle bezeichnet, in denen zwar ein Alarm ausgegeben wurde, jedoch kein tatsächlicher Krankheitsausbruch stattfand. Analog dazu ist der Begriff *false negative* zu interpretieren.

Definition 3.3. Die Sensitivität ist demnach gegeben durch

$$Sens = \frac{\text{true positive}}{\text{false negative} + \text{true positive}} \quad , \quad (3.6)$$

wobei die Spezifität hingegen als

$$Spec = \frac{\text{true negative}}{\text{false positive} + \text{true negative}} \quad (3.7)$$

definiert wird (vgl. Höhle et al. 2012).

4 Simulationsstudie

Die Performance-Analyse anhand realer Daten gestaltet sich oftmals schwierig, da z. B. kleinere Krankheitsausbrüche eventuell unbemerkt bleiben und somit nicht in den Daten vermerkt werden. Die Kontrolle, ob tatsächliche Ausbrüche vom Algorithmus erkannt wurden, wird somit verfälscht. Zu einer ersten Beurteilung der Algorithmen bietet sich deshalb die Simulation von verschiedenen Krankheitsverläufen an.

In der vorliegenden Arbeit wurden sechs verschiedene Modelle gewählt, die sich in linearem und saisonalem Trend sowie generell in der Anzahl bzgl. der gemeldeten Krankheitsfälle pro Woche unterscheiden.

Die Beobachtungen der zu untersuchenden Zeitreihe $\{X_k\}_{k=1}^t$ wurden wie folgt simuliert. Zuerst wurde der *in-control* Mittelwert durch

$$\log(\mu_{0,k}) = \beta_0 + \beta_1 k + Season_k$$

modelliert, wobei $\beta_1 k$ den linearen Trend und $Season_k$ den saisonalen Trend beschreibt.

Die saisonale Komponente $Season_k$ kann dabei beschrieben werden durch

$$Season_k = \sum_{i=1}^I (\gamma_i \sin(\omega i k) + \delta_i \cos(\omega i k)). \quad (4.1)$$

Mit $a_i = \sqrt{\gamma_i^2 + \delta_i^2}$ und $\phi_i = \tan^{-1}(\delta_i/\gamma_i)$ ist (4.1) äquivalent zu

$$Season_k = \sum_{i=1}^I a_i \sin(\omega k + \phi_i). \quad (4.2)$$

Eine Zeitreihe ohne Krankheitsausbrüche kann durch die Annahme $X_k \sim Poi(\mu_{0,k})$ bzw. $X_k \sim NegBin(r/(r + \mu_{0,k}), r)$ gewonnen werden.

Die Funktion `sim.pointSource` des R-packages 'surveillance' ermöglicht die Simulation von sogenannten *point source outbreaks* mittels Beschreibung des Status $d(k)$ durch

4 Simulationsstudie

eine homogene Markov-Kette (Höhle 2007). Bei Ausbrüchen dieser Art wird davon ausgegangen, dass es eine bestimmte Quelle (wie z. B. ein gemeinsames Abendessen) gibt. Es wird weiters angenommen, dass keine Mensch-zu-Mensch-Übertragung stattfindet und die Quelle nach einer gewissen Zeit versiegt (BMG 2010). Der Status $d(k)$ beschreibt, wie in der Einleitung des Kapitels 2 bereits erläutert, den Zustand der zu überwachenden Zeitreihe zum Zeitpunkt k , d. h.

$$d(k) = \begin{cases} 0, & \text{falls kein Ausbruch} \\ 1, & \text{falls ein Ausbruch stattfindet.} \end{cases}$$

Die Übergangsmatrix

$$P = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix}$$

beschreibt die homogene Markov-Kette. Die Wahrscheinlichkeit für einen Ausbruch zum Zeitpunkt $k+1$, wenn es zum Zeitpunkt k keinen Ausbruch gegeben hat, wird durch $1-p$ festgelegt. Die Wahrscheinlichkeit für einen Ausbruch zum Zeitpunkt $k+1$, wenn bereits zum Zeitpunkt k ein Ausbruch vorlag, ist q . Das heißt, dass die Länge eines Ausbruchs mittels q kontrolliert werden kann (Höhle 2007).

Die Zeitreihe $\{X_k\}_{k=1}^t$ wird mittels Poisson-Verteilung simuliert, wobei sich der Mittelwert der Zeitreihe bei Auftreten eines Ausbruchs um eine zu wählende Konstante $const$ erhöht, d. h.

$$X_k \sim Poi(\mu_{0,k} + const \cdot d(k)).$$

Die zu untersuchenden Algorithmen wurden, soweit möglich, optimal zur Detektion eines Parameter-Shifts der Größe 2 angepasst. Wie diese Anpassung genau erfolgte, wird im jeweiligen Unterkapitel detaillierter erläutert.

4.1 Grundmodelle für den in-control Mittelwert

Betrachtet werden zwei Ausgangssituationen. In Situation A werden die Anzahlen von wenig auftretenden Krankheitsfällen (siehe Abb. 4.1) und in Situation B werden die Anzahlen von häufig auftretenden Krankheitsfällen (Abb. 4.2) simuliert.

Situation A

A1 Es liegt *kein linearer Trend* vor, $\beta_0 = -0.8$ und

$$\begin{aligned} Season_k^{A1} &= 0.3 \sin(\omega k) + 0.3 \cos(\omega k) \\ &= \sqrt{2 \cdot 0.3^2} \sin(\omega k + \tan^{-1}(1)). \end{aligned}$$

A2 Es liegt *kein linearer Trend* vor, $\beta_0 = -0.8$ und

$$\begin{aligned} Season_k^{A2} &= 1.3 \sin(\omega k) + 1.3 \cos(\omega k) \\ &= \sqrt{2 \cdot 1.3^2} \sin(\omega k + \tan^{-1}(1)). \end{aligned}$$

A3 Es liegt ein *schwach abfallender linearer Trend* vor mit $\beta_1 = -0.0006$, $\beta_0 = 0.5$,

$$\begin{aligned} Season_k^{A3} &= 0.8 \sin(\omega k) + 0.7 \cos(\omega k) \\ &= \sqrt{0.8^2 + 0.7^2} \sin\left(\omega k + \tan^{-1}\left(\frac{0.7}{0.8}\right)\right). \end{aligned}$$

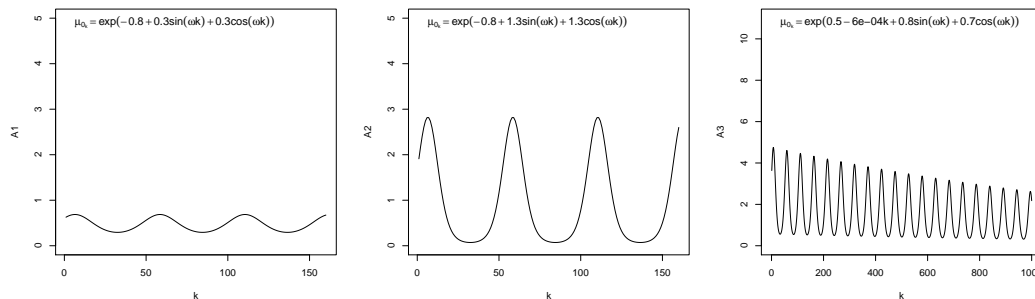


Abbildung 4.1: Modelle A1, A2 und A3

Situation B

B1 Es liegt *kein linearer Trend* vor, $\beta_0 = 2.7$ und

$$\begin{aligned} Season_k^{B1} &= 0.3 \sin(\omega k) + 0.2 \cos(\omega k) \\ &= \sqrt{0.3^2 + 0.2^2} \sin\left(\omega k + \tan^{-1}\left(\frac{0.2}{0.3}\right)\right). \end{aligned}$$

4 Simulationsstudie

B2 Es liegt *kein linearer Trend* vor, $\beta_0 = 2.7$ und

$$\begin{aligned} \text{Season}_k^{B2} &= 0.7 \sin(\omega k) + 0.7 \cos(\omega k) \\ &= \sqrt{2 \cdot 0.7^2} \sin(\omega k + \tan^{-1}(1)). \end{aligned}$$

B3 Es liegt ein *schwach abfallender linearer Trend* vor mit $\beta_1 = -0.0007$, $\beta_0 = 4.7$,

$$\begin{aligned} \text{Season}_k^{B3} &= 0.1 \sin(\omega k) + 0.1 \cos(\omega k) \\ &= \sqrt{2 \cdot 0.1^2} \sin(\omega k + \tan^{-1}(1)). \end{aligned}$$

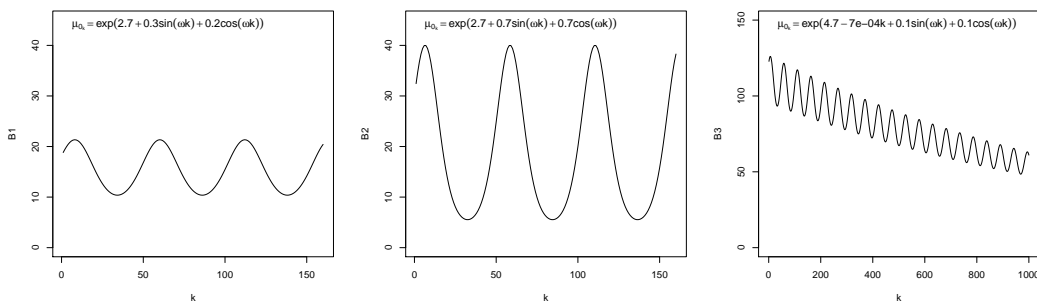


Abbildung 4.2: Modelle B1, B2 und B3

4.2 CUSUM - Vergleich möglicher Adaptierungen

Der Standard-CUSUM-Detektor wird der speziellen Struktur der Zeitreihe von gemeldeten Krankheitsfällen vor allem durch saisonale Schwankungen nicht gerecht (vgl. Rossi et al. 2010). Deshalb sollen im folgenden Kapitel die vorgestellten Adaptionen genauer untersucht werden, die auf diese spezielle Struktur eingehen.

4.2.1 Normalverteilungsapproximation vs. CUSUM mit zeitabhängigen Parametern

Festlegen der Parameter

Bei der Adaptierung von Rossi et al. (1999) werden die Beobachtungen transformiert und danach der CUSUM-Detektor für die Standard-Normalverteilung angewendet. Der

4.2 CUSUM - Vergleich möglicher Adaptierungen

in-control Mittelwert entspricht nach der Transformation dementsprechend dem Wert 0. Soll ein Shift von 2 detektiert werden, gilt $\mu_1 = \mu_0 + 2sd(\mu_0) = 2$ und somit ist $\kappa = (\mu_0 + \mu_1)/2 = 1$. Mithilfe der Funktion `xcusum.crit` des Package 'spc' kann der optimale Wert für h ermittelt werden. Es ist lediglich die gewünschte ARL_0 festzulegen. Diese wird hier, wie auch in der Literatur (siehe z.B. Hawkins und Olwell 1998; Rossi et al. 2010), auf 500 gesetzt. Mit

```
> xcusum.crit(1, 500, mu0 = 0, hs = 0, sided = "one", r = 30)
h = 2.32324
```

folgt somit der optimale Wert $h = 2.32$.

Die Adaptierung von Rogerson und Yamada (2004) hingegen beruht auf zeitabhängigen Threshold-Werten h_k , welche man mittels der Funktion `hValues` erhält. Hierbei wird die Funktion `findH` für jeden Zeitpunkt ausgewertet.

Beispiel 4.1 (Zeitabhängige Threshold-Werte für Rogerson-CUSUM). Für das Modell A2 werden unten die ersten zehn Werte als R-Output der Funktion `hValues` gezeigt. Wiederum wird die gewünschte $ARL_0 = 500$ angegeben. Weiters wird der zu detektierende Parametershift mit $s = 2$ spezifiziert. Mit `distr = "poisson"` und `theta0 = mu.a2` wird die *in-control* Verteilung, die hier der $Poi(\mu_{A2})$ -Verteilung entspricht, festgelegt. Zusätzlich zu den errechneten h Werten werden auch die jeweiligen Werte für κ (k), die tatsächlich erreichte ARL_0 sowie die erreichte relative Toleranz (Wert für das Abbruchkriterium) angegeben (siehe Höhle et al. 2012).

```
> hValues(theta0 = mu.a2, ARL0 = 500, rel.tol = 0.02, s = 2,
          roundK = TRUE, digits = 1, distr = "poisson",
          FIR = FALSE)$hValues[1:10,]
      theta0    h    k    ARL    rel.tol
[1,] 1.910222 4.9 3.1 500.8565 0.001712949
[2,] 2.166981 5.4 3.4 498.8066 0.002386852
[3,] 2.402494 5.7 3.7 582.9549 0.165909739
[4,] 2.599276 6.0 3.9 490.5270 0.018945944
[5,] 2.741111 5.8 4.2 551.9661 0.103932284
[6,] 2.815454 5.8 4.3 546.8057 0.093611451
[7,] 2.815454 5.8 4.3 546.8057 0.093611451
[8,] 2.741111 5.8 4.2 551.9661 0.103932284
[9,] 2.599276 6.0 3.9 490.5270 0.018945944
[10,] 2.402494 5.7 3.7 582.9549 0.165909739
```

Analyse der durchschnittlichen Laufzeit

In einem ersten Schritt soll analysiert werden, inwiefern die Vorgabe $ARL_0 = 500$ erfüllt wird. Dazu wurde für alle der im vorigen Kapitel vorgestellten Grundmodelle eine *in-control* Zeitreihe für einen Zeitraum von 80 Jahren (4160 Wochen) simuliert, die einer $Poi(\mu_{0,k})$ -Verteilung folgt. Diese Zeitreihe wurde mit dem jeweiligen Detektor analysiert und die Woche mit dem ersten Alarm wurde aufgezeichnet. Der nachfolgende Auszug aus dem R-Code soll exemplarisch die Ermittlung der *run length* für eine Replikation verdeutlichen.

```
t <- 1:4160
arl <- NA
x <- rpois(length(t), mu0)
dP <- create.disProg(week=t, observed=x, state=rep(0,length(t)),
                    freq=52)
rossi <- algo.cusum(dP, control = list(range = 1:length(t), k = 1,
                                     h = 2.32, trans = "rossi", alpha=alph, m = mu0))
ifelse(min(which(rossi$alarm=="1")) <= length(t),
      arl <- min(which(rossi$alarm=="1")), arl <- NA)
```

Anhand von 1000 Replikationen wurde der Monte-Carlo-Schätzer für die *in-control* Laufzeit durch Mittelwertbildung erhalten. Falls im gesamten Beobachtungszeitraum kein Alarm auftrat, wurde diese Replikation als NA-Wert nicht in die Schätzung mitbezogen.

Analog zur MC-Schätzung für die ARL_0 kann die MC-Schätzung für die *out-of-control* Laufzeit ARL_1 ermittelt werden. Der entscheidende Unterschied liegt in der Simulation der zu untersuchenden Zeitreihe. Dieser liegt bei der Berechnung der ARL_1 der *out-of-control* Mittelwert μ_1 zugrunde, das heißt, die Zeitreihe folgt einer $Poi(\mu_{0,k} + 2\sqrt{\mu_{0,k}})$ -Verteilung.

Das Ergebnis der Monte-Carlo-Simulation verdeutlicht die Schwachstelle der CUSUM-Adaption nach Rossi et al. (1999). Aufgrund der Schiefe der Poisson-Verteilung ist vor allem für geringe Anzahlen die Normalverteilungs-Approximation nicht adäquat. Daraus ergibt sich auch die deutlich geringere durchschnittliche *in-control* Laufzeit ARL_0 für die Modelle A1-A3 im Gegensatz zum geforderten Wert von 500. Im Gegensatz dazu wird für diese Modelle die gewünschte $ARL_0 = 500$ leicht übererfüllt, falls die Adaption nach Rogerson und Yamada (2004) angewendet wird. Für die Modelle B1-B3 ist der Unterschied der MC-Schätzwerte der beiden Adaptionen deutlich geringer. In allen drei Fällen wird für beide Adaptionen nahezu ein Wert von 500 erreicht.

4.2 CUSUM - Vergleich möglicher Adaptierungen

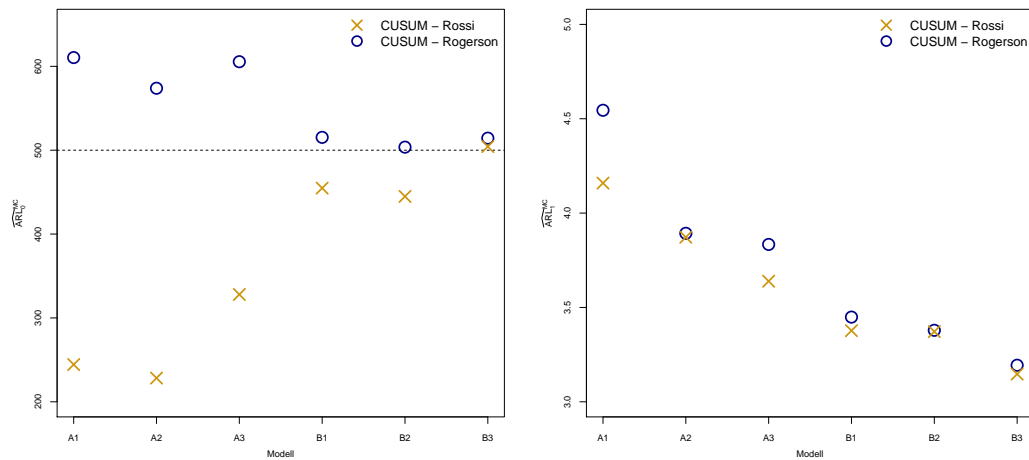


Abbildung 4.3: Zu sehen sind die MC-Schätzwerte für die durchschnittliche *in-control* Laufzeit (linkes Bild) sowie für die durchschnittliche *out-of-control* Laufzeit (rechtes Bild) der beiden CUSUM-Adaptionen nach Rossi bzw. Rogerson.

Die durchschnittliche *out-of-control* Laufzeit gibt an, nach wievielen Wochen der erste Alarm berichtet wird, falls die Zeitreihe bereits zu Beginn der Überwachung aus der *out-of-control* Verteilung stammt. Idealerweise sollte die ARL_1 möglichst gering sein, was oft allerdings nur auf Kosten einer geringeren ARL_0 erreicht werden kann. Der CUSUM-Algorithmus bietet durch die Wahl von h und κ gute Auswahlmöglichkeiten in Hinblick auf gewünschte ARL_0 - bzw. ARL_1 -Werte. Im vorliegenden Setting wurde lediglich die zu erreichende ARL_0 festgelegt. Die daraus resultierenden geschätzten ARL_1 -Werte sind in Abbildung 4.3 in der rechten Grafik dargestellt. Für Modell A1 wird beispielsweise durchschnittlich nach 4.2 Wochen (Rossi) bzw. nach 4.5 Wochen (Rogerson) der erste Alarm berichtet. Für größere Anzahlen verringert sich die geschätzte durchschnittliche Anzahl der Wochen bis zum ersten Alarm deutlich. Der lineare Trend in Modell A3 bzw. B3 hat hier keinen großen Einfluss, da bereits nach wenigen Wochen der erste Alarm ausgegeben wird. Das heißt jener Zeitraum, in dem die Auswirkungen des Trends deutlich werden, liefern keinen Beitrag zu diesem Ergebnis.

False Alarm Probability

Zur Ermittlung der *false alarm* Wahrscheinlichkeit P_{fa} wurden die ersten 52 Wochen (entspricht dem ersten Jahr) betrachtet. Geschätzt wurde somit die Wahrscheinlichkeit für einen Fehlalarm innerhalb des Beobachtungszeitraums von einem Jahr. Des Weiteren wurde die Detektions-Wahrscheinlichkeit $P_{fa}^1(3)$ geschätzt. In Abbildung 4.4 sind

4 Simulationsstudie

wiederum die Monte-Carlo-Schätzungen für $P(T_A \leq 52 \mid \tau = \infty)$ auf der linken bzw. für $P(T_A \leq 1 \mid \tau = 1)$ auf der rechten Seite zu sehen.

Wie nach Betrachtung der geschätzten durchschnittlichen *in-control* Laufzeit erwartet, ist die Wahrscheinlichkeit für einen Fehlalarm im ersten Jahr für Modelle A1-A3 bei Anwendung der Adaption nach Rossi et al. (1999) deutlich höher als bei Anwendung der Adaption nach Rogerson und Yamada (2004). Die größte Differenz ergibt sich bei Modell A2. Hier beträgt die geschätzte Wahrscheinlichkeit für einen Fehlalarm 22% bei Anwendung des Rossi-CUSUM Detektors bzw. 13% bei Anwendung des Rogerson-CUSUM Detektors. Im Gegensatz dazu sind die Schätzwerte beider Adaptionen für Modelle B1-B3 nahezu ident und in etwa 10%.

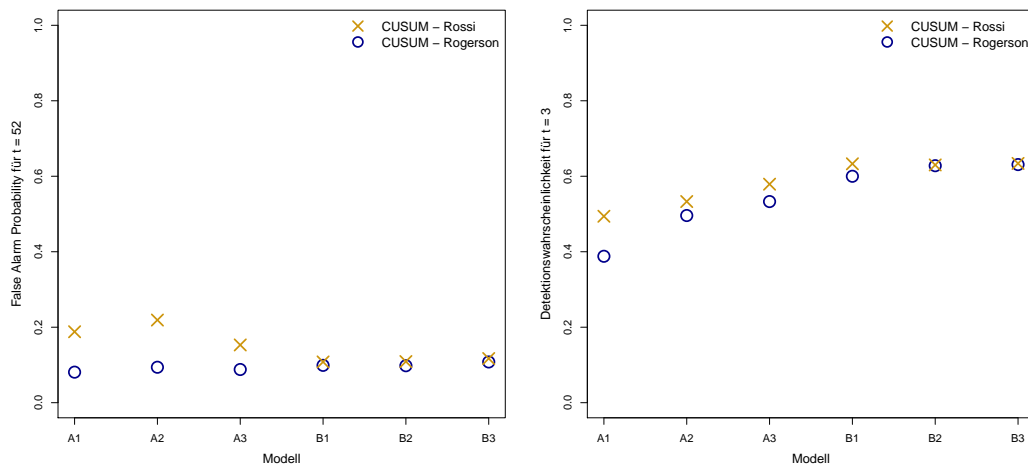


Abbildung 4.4: MC-Schätzwerte für die Fehlalarm-Wahrscheinlichkeiten $\hat{P}_{fa}(T_A \leq 52 \mid \tau = \infty)$ bzw. für die Detektions-Wahrscheinlichkeiten $P_{fa}^1(T_A \leq 3 \mid \tau = 1)$ (rechts) der beiden CUSUM-Adaptionen nach Rossi bzw. Rogerson.

Dementsprechend überraschen auch die MC-Schätzwerte für die Detektionswahrscheinlichkeit einer Änderung innerhalb der darauffolgenden 3 Wochen nicht. So sind diese mit 50% (Modell A1), 53% (Modell A2) bzw. 58% (Modell A3) bei der Rossi-Adaption deutlich höher als jene der Rogerson-Adaption (39%, 50% bzw. 53%). Ein Parametershift der Größe $\mu_0 + 2 \cdot \sqrt{\mu_0}$ zu Beginn der Überwachung wird bei den Modellen B1-B3 mit einer Wahrscheinlichkeit von etwa 60% innerhalb der ersten drei Wochen detektiert. Dieses Ergebnis kann mit beiden Adaptionen nahezu gleichermaßen erreicht werden.

Sensitivität und Spezifität

Mit der Funktion `sim.pointSource` wurde zur Ermittlung der Sensitivität und Spezifität wiederum eine Zeitreihe über den Zeitraum von 4160 Wochen simuliert. Diese enthält nun aber *point source* Ausbrüche, die der Markov-Kette

$$P = \begin{pmatrix} 0.975 & 0.025 \\ 0.950 & 0.050 \end{pmatrix}$$

folgen.

Anhand von 100 Replikationen wurde die durchschnittliche Sensitivität bzw. Spezifität mit dem Befehl `algo.quality` ermittelt. Die nachfolgenden Auszüge des R-Codes sollen die Vorgehensweise verdeutlichen.

```
set.seed(123)
dP <- sim.pointSource(p = 0.975, r = 0.05, length = 4160,
  frequency = 1, K = log(2), A = sqrt(2*0.7^2),
  alpha = 2.7, beta = 0, phi = atan(1)/omega)
rossi.b2 <- algo.cusum(dP, control = list(range = 1:4160,
  k = 1, h = 2.32, trans = "rossi", alpha = 0,
  m = mu.b2))
algo.quality(rossi.b2)
  TP FP  TN   FN Sens      Spec      dist      mlag
[1,] 62 119 3935 44 0.5849057 0.9706463 0.4161309 7.60396
```

Mit `sim.pointSource` wird eine Zeitreihe der Länge 4160 für Modell B2 simuliert. Die Zeitreihe wird mittels Rossi-CUSUM-Detektor untersucht. Die Funktion `algo.quality` liefert unter anderem die Anzahl der richtig detektierten (TP) sowie der fälschlicherweise detektierten (FP) Ausbrüche. Interessant sind hier vor allem die Spalten `Sens` und `Spec`, die die in Kapitel 3.3 definierten Performance-Maße angeben.

In Tabelle 4.1 sind die Schätzwerte, die durch Mittelwertbildung der 100 Replikationen ermittelt wurden, d. h. $\hat{E}(Sens) = \frac{1}{100} \sum_{i=1}^{100} Sens_i$, zu finden. Beide Adaptionen haben durchgängig für alle Modelle eine sehr hohe Spezifität von mindestens 92.2%. Hingegen ist die Sensitivität vor allem für Modelle A1-A3 mit maximal 4.78% sehr gering. Die Adaption nach Rossi et al. (1999) hat dabei eine durchwegs höhere Sensitivität.

		A1	A2	A3	B1	B2	B3
Sensitivität (in %)	Rossi	3.27	4.78	4.03	62.21	61.80	76.48
	Rogerson	1.64	2.82	2.32	59.09	59.06	74.60
Spezifität (in %)	Rossi	99.37	99.28	99.34	97.37	96.59	93.52
	Rogerson	99.51	99.47	99.50	96.64	95.73	92.19

Tabelle 4.1: Sensitivität und Spezifität für Rossi und Rogerson

4.2.2 Regression Charts (LR- und GLR-Detektoren)

Die Adaptierung von Höhle (2006) sowie Höhle und Paul (2008), beruhend auf generalisierten linearen Modellen, ist im Gegensatz zu den vorangegangenen Adaptierungen auch zur Schätzung des unbekanntem Parametershifts zum *out-of-control* Mittelwert anwendbar. Hierbei spricht man, wie bereits in Kapitel 2.1.3 erläutert, vom GLR-Detektor. Ist der additive Zuwachs auf der Logarithmus-Skala $\exp(\eta)$ bereits bekannt bzw. wird er vor Anwendung der Algorithmen festgelegt, handelt es sich um einen LR-Detektor. Dies verringert die Laufzeit natürlich wesentlich, da die rekursive Ermittlung des Schätzers $\hat{\eta}$ entfällt.

Entscheidend für die Performance dieser CUSUM-Adaptierungen ist wiederum die Wahl des Thresholds h . Mittels Kombination von Monte-Carlo-Schätzung und numerischer Methoden, wie zum Beispiel dem Sekantenverfahren, ist es auch hier möglich einen Threshold zu ermitteln, der zu einer festgelegten ARL_0 führt (vgl. Höhle und Paul 2008). Wie aber unter anderem in Lai und Shan (1999) sowie Lai (1995) diskutiert wird, liefert die *false alarm probability* eine Entscheidungshilfe, die wesentlich weniger rechnerintensiv ist. So kann der Threshold h bestimmt werden, indem jener Wert gewählt wird, bei dem der Monte-Carlo-Schätzer für $P_{fa} = (T_A \leq t \mid \tau = \infty)$ mindestens einem festgelegten Wert (z.B. 0.15) entspricht (vgl. Höhle und Mazick 2010).

Wahl des Parameters h für den LR-Detektor

Detektiert werden sollen Änderungen der Form

$$\mu_{1,k} = 2 \cdot \mu_{0,k}.$$

Mittels $n = 1000$ Replikationen wurde die folgende Monte-Carlo-Schätzung ermittelt: $\hat{P}_{fa}(52) = P(T_A \leq 52 \mid \tau = \infty)$ und $\hat{P}_{fa}^1(3) = P(T_A \leq 3 \mid \tau = 1)$. Die geschätzte *false*

4.2 CUSUM - Vergleich möglicher Adaptierungen

alarm probability in einem Jahr (entspricht 52 Wochen) sollte zumindest 15% betragen. Des Weiteren soll darauf geachtet werden, dass die Detektions-Wahrscheinlichkeit eines Krankheitsausbruches innerhalb von drei Wochen nach Eintritt der Änderung nicht zu gering ist (z. B. $> 30\%$).

Für Modell A1 erhält man mit der Wahl $h^{A1} = 3.4$ die gewünschte *false alarm probability* $\hat{P}_{fa} < 0.15$, jedoch ist $\hat{P}_{fa}^1 = 0.014$ sehr gering. Wird mehr Wert auf die rasche Detektion gelegt, empfiehlt es sich, hier deshalb den Wert $h_1^{A1} = 1$ zu wählen, womit $\hat{P}_{fa}^1 = 0.31$ erreicht wird.

Für Modell A2 ergibt sich mit obigem Kriterium wiederum $h^{A2} = 3.4$. Der Wert \hat{P}_{fa}^1 ist mit 0.21 aber wesentlich höher als bei Modell A1. Wird $P(T_A \leq 3 \mid \tau = 1)$ als Auswahlkriterium gewählt, ergibt sich $h_1^{A2} = 2.5$ und somit $\hat{P}_{fa}^1 = 0.36$.

Die Wahl für Modell A3, das einen linearen Trend beinhaltet, gestaltet sich schwieriger, da die Wahl von h in diesem Fall von den beobachteten 52 Wochen abhängt. Hier soll zur Wahl von h das letzte Jahr betrachtet werden. Daraus ergibt sich für dieses Modell $h^{A3} = 2.8$ ($\hat{P}_{fa} = 0.009$), mit $h_1^{A3} = 0.7$ erhält man $\hat{P}_{fa}^1 = 0.35$.

Für die Modelle B1 und B2 erhält man mit den Werten $h^{B1} = 3.4$ bzw. $h^{B2} = 3.7$, die zu $\hat{P}_{fa} \leq 0.15$ führen, bereits eine Wahrscheinlichkeit von $\hat{P}_{fa}^1 > 0.90$.

Bei Modell B3 ist analog zu Modell A3 der lineare Trend in die Überlegungen mit-einzubeziehen. Außerdem kommt es bei diesem speziellen Modell dazu, dass die Log-Likelihood-Quotienten der ersten Jahre größtenteils negativ sind. Negative Werte werden an dieser Stelle aber durch 0 ersetzt. Für die Ermittlung der Alarmzeit heißt dies insbesondere, dass in diesen Fällen der Wert des Log-Likelihood-Quotienten ($= 0$) nicht größer als h wird. Das heißt es wird unter diesen Umständen kein Alarm angezeigt. Unabhängig vom gewählten Threshold kommt es somit im ersten Jahr zur *false alarm probability* $\hat{P}_{fa} = 0$. Zur Festlegung des Thresholds h wird, wie bei Modell A3, auch hier das letzte Jahr der vorliegenden Zeitreihe analysiert. Daraus folgt für $h^{B3} = 4.3$.

In Tabelle 4.2 wird die Wahl der Threshold-Werte für die verschiedenen Modelle und Kriterien zusammenfassend dargestellt.

Kriterium	A1	A2	A3	B1	B2	B3
$\hat{P}(T_A \leq 52 \mid \tau = \infty) \leq 0.15$	3.4	3.4	2.8	3.4	3.7	4.3
$\hat{P}(T_A \leq 3 \mid \tau = 1) \geq 0.30$	1.0	2.5	0.7	3.4	3.7	4.3

Tabelle 4.2: Threshold-Werte für LR-Detektor

Wahl des Parameters h für den GLR-Detektor

Im Gegensatz zum LR-Detektor wird beim GLR-Detektor, wie bereits erwähnt, der zu detektierende Parameter-Shift nicht festgelegt, sondern während der Analyse der Zeitreihe geschätzt. Die Wahl der Threshold-Werte erfolgt jedoch analog zur Wahl beim LR-Detektor. Je nach gewähltem Kriterium ergeben sich demnach die Threshold-Werte der Tabelle 4.3.

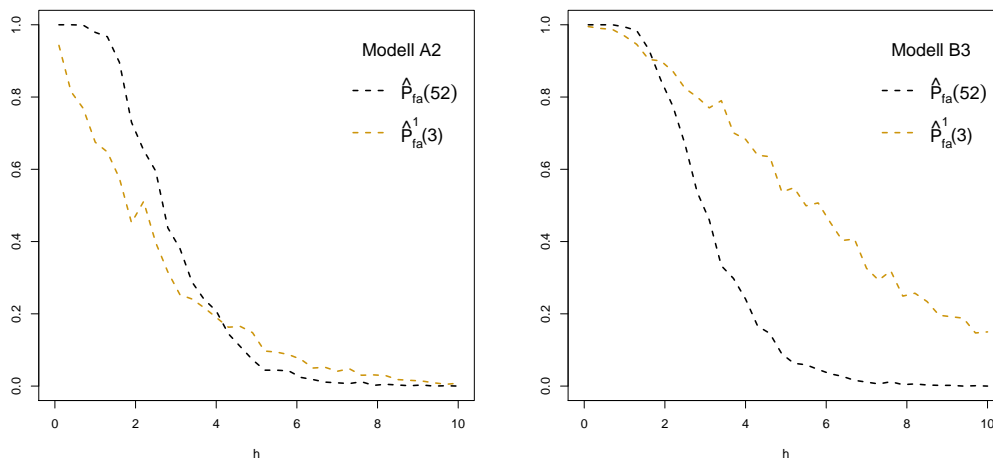


Abbildung 4.5: PFA in Abhängigkeit von verschiedenen h -Werten für GLR-Detektor für Modelle A2 und B3.

Kriterium	A1	A2	A3	B1	B2	B3
$\hat{P}(T_A \leq 52 \mid \tau = \infty) \leq 0.15$	4.3	4.3	2.8	4.6	4.9	4.9
$\hat{P}(T_A \leq 3 \mid \tau = 1) \geq 0.30$	1.6	2.8	0.7	4.6	4.9	4.9

Tabelle 4.3: Threshold-Werte für GLR-Detektor

Durchschnittliche Laufzeit

Die durchschnittliche *in-control* Laufzeit ist für den LR-Detektor, bei dem der zu detektierende Shift als bekannt vorausgesetzt und deshalb exakt angegeben wird, etwas länger als für den GLR-Detektor. Die Bedeutung der Wahl von h wird durch die deutlichen Unterschiede der geschätzten ARL_0 für die jeweiligen Modelle deutlich. Wird mit $h = 3.4$ im Modell A1 mit dem LR-Detektor durchschnittlich nach 376 Wochen

4.2 CUSUM - Vergleich möglicher Adaptierungen

der erste Fehllalarm ausgegeben, sind es bei $h = 1$ nur mehr etwa 19 Wochen. Etwas geringere Differenzen ergeben sich bei den Modellen A2 und A3. Das Verhalten des GLR-Detektors ist diesbezüglich ähnlich.

Betrachtet man die durchschnittliche *out-of-control* Laufzeit für Modelle B1-B3, fällt auf, dass bei allen 1000 durchgeführten Replikationen bereits in der ersten Woche ein Alarm berichtet wird. Dies ist darauf zurückzuführen, dass bei einer derartig hohen Anzahl von berichteten Fällen pro Woche die Änderung dermaßen stark ausfällt, dass die gewählten h Werte bereits zu Beginn überschritten werden.

	A1			A2			A3		
	h	\widehat{ARL}_0	\widehat{ARL}_1	h	\widehat{ARL}_0	\widehat{ARL}_1	h	\widehat{ARL}_0	\widehat{ARL}_1
LR-Detektor	3.40	376.01	13.77	3.40	335.10	4.66	2.80	108.74	2.70
	1.00	18.49	4.69	2.50	113.56	3.71	0.70	10.05	1.61
GLR-Detektor	4.30	347.17	15.13	4.30	301.15	5.32	2.80	72.30	2.59
	1.60	22.66	4.88	2.80	82.61	3.80	0.70	8.40	1.58

Tabelle 4.4: ARL für LR- und GLR-Detektor für Modelle A1-A3

	B1			B2			B3		
	h	\widehat{ARL}_0	\widehat{ARL}_1	h	\widehat{ARL}_0	\widehat{ARL}_1	h	\widehat{ARL}_0	\widehat{ARL}_1
LR-Detektor	3.40	284.79	1.23	3.70	385.88	1.05	4.30	2428.88	1.00
GLR-Detektor	4.60	287.40	1.27	4.90	380.75	1.04	4.90	353.76	1.00

Tabelle 4.5: ARL für LR- und GLR-Detektor für Modelle B1-B3

Sensitivität und Spezifität

Die Spezifität unterscheidet sich nur gering bzgl. Detektor, gewähltem Threshold, und analysiertem Modell. Mit Werten, die durchwegs über 90% liegen, ist die Spezifität hier mehr als akzeptabel. Die Sensitivität für Modelle A1-A3 hingegen ist sehr gering, was bedeutet, dass die Rate der richtig positiv erkannten teilweise sogar weit unter 32% liegt.

Für höhere Anzahlen an Krankheitsfällen, wie sie in den Modellen B1-B3 auftreten, fällt die Sensitivität deutlich höher aus. Für Modell B3 erreicht der LR-Detektor sogar eine Sensitivität von nahezu 77%.

4 Simulationsstudie

Um auch die Sensitivität für die Modelle A1-A3 zu erhöhen, könnte der einzige frei wählende Parameter, der Threshold h , verringert werden.

		A1		A2		A3	
Sensitivität (in %)	LR-Detektor	1.15	(3.4)	3.76	(3.4)	4.70	(2.8)
	GLR-Detektor	16.26	(1.0)	7.25	(2.5)	26.89	(0.7)
Spezifität (in %)	LR-Detektor	2.06	(4.3)	3.91	(4.3)	8.55	(2.8)
	GLR-Detektor	14.41	(1.6)	8.87	(2.8)	31.68	(0.7)
Sensitivität (in %)	LR-Detektor	99.70	(3.4)	99.66	(3.4)	99.38	(2.8)
	GLR-Detektor	94.77	(1.0)	99.11	(2.5)	91.86	(0.7)
Spezifität (in %)	LR-Detektor	99.69	(4.3)	99.66	(4.3)	98.70	(2.8)
	GLR-Detektor	95.82	(1.6)	98.75	(2.8)	89.09	(0.7)

Tabelle 4.6: Sensitivität bzw. Spezifität (verwendeter Threshold h) für LR- und GLR-Detektor für Modelle A1-A3

		B1		B2		B3	
Sensitivität (in %)	LR-Detektor	72.72	(3.4)	65.87	(3.7)	76.75	(4.3)
	GLR-Detektor	64.00	(4.6)	62.37	(4.9)	75.78	(4.9)
Spezifität (in %)	LR-Detektor	99.63	(3.4)	99.70	(3.7)	99.88	(4.3)
	GLR-Detektor	99.60	(4.6)	99.70	(4.9)	99.71	(4.9)

Tabelle 4.7: Sensitivität und Spezifität für LR- und GLR-Detektor für Modelle B1-B3

Detektion eines additiven Zuwachses der zweifachen Standardabweichung

Geht man von Poisson-verteilten Beobachtungen aus, entspricht ein additiver Zuwachs im Mittelwert um das zweifache der Standardabweichung gerade

$$\mu_{1,k} = \mu_{0,k} + 2\sqrt{\mu_{0,k}}.$$

Beruhet die *out-of-control* Zeitreihe nun auf diesem zeitabhängigen Mittelwert, ergeben sich, aufgrund der unterschiedlichen Detektionswahrscheinlichkeiten, neue Threshold-Werte hinsichtlich des Kriteriums $P(T_A \leq 3 \mid \tau = 1) \geq 0.30$. Den Tabellen 4.8 und 4.9 sind die 'neuen' Threshold Werte für den LR- bzw. GLR-Detektor zu entnehmen. Jene

4.2 CUSUM - Vergleich möglicher Adaptierungen

Threshold-Werte, die sich von denen in der obigen Ausführungen unterscheiden, sind fett gedruckt.

Hinsichtlich LR-Detektor liefern $h_1^{B1} = 3.4$ und $h_1^{B3} = 4.3$ eine geschätzte Detektionswahrscheinlichkeit von 37% und 39% und erfüllen deshalb beide Kriterien.

Beim GLR-Detektor ist bzgl. des Kriteriums $P(T_A \leq 3 \mid \tau = 1) \geq 0.30$ lediglich für Modell A1 ein geringerer Threshold zu wählen. Mit $h_1^{A1} = 4.0$ kann dann eine geschätzte Detektionswahrscheinlichkeit von 33% erreicht werden. Die Schätzwerte der Detektionswahrscheinlichkeit für Modelle A2-B3 liegen zwischen 31% und 37%.

Kriterium	A1	A2	A3	B1	B2	B3
$\hat{P}(T_A \leq 52 \mid \tau = \infty) \leq 0.15$	3.4	3.4	2.8	3.4	3.7	4.3
$\hat{P}(T_A \leq 3 \mid \tau = 1) \geq 0.30$	2.5	3.4	1.9	3.4	2.2	4.3

Tabelle 4.8: Threshold-Werte für LR-Detektor bei additiven Zuwachs von $2 \cdot sd$

Kriterium	A1	A2	A3	B1	B2	B3
$\hat{P}(T_A \leq 52 \mid \tau = \infty) \leq 0.15$	4.3	4.3	2.8	4.6	4.9	4.9
$\hat{P}(T_A \leq 3 \mid \tau = 1) \geq 0.30$	4.0	4.3	2.8	4.6	4.9	4.9

Tabelle 4.9: Threshold-Werte für GLR-Detektor bei additiven Zuwachs von $2 \cdot sd$

Für die besprochenen Performance-Maße ergeben sich daraus die folgenden Änderungen.

	A1			A2			A3		
	h	\widehat{ARL}_0	\widehat{ARL}_1	h	\widehat{ARL}_0	\widehat{ARL}_1	h	\widehat{ARL}_0	\widehat{ARL}_1
LR-Detektor	3.4	376.01	4.42	3.4	335.10	3.40	2.8	108.74	2.69
	2.5	137.29	3.73				1.9	38.47	2.05
GLR-Detektor	4.3	347.17	4.20	4.3	301.15	3.77	2.8	72.30	2.57
	4.0	229.37	3.91						

Tabelle 4.10: ARL für LR- und GLR-Detektor für Modelle A1-A3 bei additiven Zuwachs von $2 \cdot sd$

4 Simulationsstudie

	B1			B2			B3		
	h	\widehat{ARL}_0	\widehat{ARL}_1	h	\widehat{ARL}_0	\widehat{ARL}_1	h	\widehat{ARL}_0	\widehat{ARL}_1
LR-Detektor	3.4	284.79	4.39	3.7	385.88	8.76	4.3	2428.88	210.61
				2.2	103.70	6.25			
GLR-Detektor	4.6	287.40	3.21	4.9	380.75	3.35	4.9	353.76	3.15

Tabelle 4.11: ARL für LR- und GLR-Detektor für Modelle B1-B3 bei additiven Zuwachs von $2 \cdot sd$

Die Schätzwerte der durchschnittlichen *in-control* Laufzeiten sind für h -Werte nach Kriterium $P(T_A \leq 52 \mid \tau = \infty) \leq 0.15$ äquivalent zu denen im ersten Teil dieses Kapitels. Die *out-of-control* Laufzeiten sind aufgrund des unterschiedlichen Shifts im Mittelwert nicht direkt vergleichbar. Entsprechend der zugrundeliegenden Theorie ist auch an diesen Ergebnissen zu erkennen, dass ein geringerer Threshold zu einer geringeren ARL_1 führt. Hinsichtlich LR- und GLR-Detektor ergibt sich nur für Modell B3 ein deutlicher Unterschied. Sowohl die geschätzte *in-control* Laufzeit als auch die *out-of-control* Laufzeit ist beim LR-Detektor wesentlich höher. So wird ein additiver Zuwachs in Form der doppelten Standardabweichung mit dem LR-Detektor durchschnittlich erst nach 211 Wochen erkannt. Mit dem GLR-Detektor ist in diesem Fall nach durchschnittlich 3 Wochen ein Alarm zu erwarten. Eine mögliche Erklärung hierfür könnte die Tatsache sein, dass beim GLR-Detektor keine Spezifizierung des zu detektierenden Shifts vorzunehmen ist. Der LR-Detektor hingegen wurde für die Detektion eines Shifts von $\mu_{1,k} = 2\mu_{0,k}$ angepasst.

Bezüglich Sensitivität und Spezifität lassen sich keine deutlichen Unterschiede zwischen den Detektoren feststellen.

		A1		A2		A3	
Sensitivität (in %)	LR-Detektor	1.15	(3.4)	3.76	(3.4)	4.70	(2.8)
		3.08	(2.5)			9.62	(1.9)
	GLR-Detektor	2.06	(4.3)	3.91	(4.3)	8.55	(2.8)
		2.85	(4.0)				
Spezifität (in %)	LR-Detektor	99.70	(3.4)	99.66	(3.4)	99.38	(2.8)
		99.24	(2.5)			98.32	(1.9)
	GLR-Detektor	99.69	(4.3)	99.66	(4.3)	98.70	(2.8)
		99.52	(4.0)				

Tabelle 4.12: Sensitivität bzw. Spezifität (verwendeter Threshold h) für LR- und GLR-Detektor für Modelle A1-A3 bei additiven Zuwachs von $2 \cdot sd$

		B1		B2		B3	
Sensitivität (in %)	LR-Detektor	72.72	(3.4)	65.87	(3.7)	76.75	(4.3)
	GLR-Detektor	64.00	(4.6)	62.37	(4.9)	75.78	(4.9)
Spezifität (in %)	LR-Detektor	99.63	(3.4)	99.70	(3.7)	99.88	(4.3)
	GLR-Detektor	99.60	(4.6)	99.70	(4.9)	99.71	(4.9)

Tabelle 4.13: Sensitivität und Spezifität für LR- und GLR-Detektor für Modelle B1-B3 für Parametershift $2 \cdot sd$

4.3 Bayesscher-Algorithmus

Für den in Kapitel 2.2 erläuterten Algorithmus ist im Gegensatz zu den CUSUM-Methoden vom Anwender kein Threshold festzulegen. Da diese Methode referenzwertbasiert ist, ist allerdings die Anzahl der Jahre bzw. Wochen, die in die Referenzmenge einfließen sollen, im Vorhinein zu bestimmen. Im Allgemeinen werden dabei alle verfügbaren Jahre miteinbezogen. Bei der Simulation wurde insgesamt ein Zeitraum von 80 Jahren untersucht. Für den Algorithmus basierend auf Bayes-Inferenz und dem Algorithmus von Farrington wurde ein Teil dieses Zeitraums zur Bildung der Referenzmenge herangezogen. Somit konnte nicht der gesamte Zeitraum von 4160 Wochen untersucht werden.

Es wurden einerseits $b = 4$ Jahre und $w = 3$ Wochen sowie andererseits $b = 15$ Jahre und $w = 4$ Wochen zur Bildung der Referenzmenge herangezogen. Im ersten Setting kann mit der Überwachung in der 212-ten Woche, im zweiten erst in der 785-ten Woche begonnen werden. Ein direkter Vergleich der Ergebnisse für $b = 4$ bzw. $b = 15$ kann deshalb vor allem bei Modellen mit linearem Trend (A3 und B3) nicht angestellt werden.

Die Grenze, die bei Überschreitung zur Ausgabe eines Alarms führt, beruht, wie bereits erläutert, auf einem gewählten Quantilparameter α . Um den Einfluss dieses Quantilparameters zu verdeutlichen, wurden vier verschiedene Parametereinstellungen gewählt (siehe Tabelle 4.14).

4 Simulationsstudie

Bezeichnung	b	w	α
B1-1	4	3	0.05
B1-2	4	3	0.10
B2-1	15	4	0.05
B2-2	15	4	0.10

Tabelle 4.14: Parametereinstellungen bei Anwendung des Algorithmus basierend auf Bayes-Inferenz

Der *in-control* Mittelwert wurde je nach betrachteten Modell berechnet. Für Modell A1 zum Beispiel gilt

$$\mu_0^{A1} = -0.8 + 0.3 \cdot \sin(\omega k) + 0.3 \cdot \cos(\omega k)$$

sowie

$$\mu_1^{A1} = \mu_0^{A1} + 2 \cdot \sqrt{\mu_0^{A1}}.$$

Durchschnittliche Laufzeit

Für die Modelle A1, A2, B1 und B2 ergeben sich nur geringe Unterschiede zwischen den geschätzten erwarteten *in-control* bzw. *out-of-control* Laufzeiten hinsichtlich der gewählten Jahre b (siehe Tabelle 4.15 bzw. 4.16). Hierbei ist jedoch darauf hinzuweisen, dass der Beobachtungszeitraum nicht übereinstimmt. Durch die Berechnung der Modelle A1-A2, B1-B2 (wiederkehrende saisonale Schwankungen, kein linearer Trend), wiederholt sich allerdings die Grundstruktur über die gesamten 80 Jahre. Die Wahl von α beeinflusst die *in-control* Laufzeit merklich.

Mit eineinhalb bis zwei Wochen ist die geschätzte ARL_1 durchwegs sehr gering. Durch den linearen Trend wird für $b = 15$ bei Modell B3 ein deutlich höheres Niveau für den *in-control* Parameter geschätzt. Dadurch sind auch die sehr hohen Schätzwerte für ARL_0 und ARL_1 zu erklären. Nicht ganz so deutlich wirkt sich dies bei Modell A3 aus.

Die *in-control* Laufzeit ist für alle Modelle, mit Ausnahme von Modell B3, sehr gering. Das spiegelt sich auch in den MC-Schätzungen der Fehlalarm-Wahrscheinlichkeit im ersten Jahr wider. Nur für Modell B3 können hier Werte unter 1 erreicht werden (siehe Tabelle 4.17).

Entsprechend der geringen geschätzten ARL_1 ist, wie in Tabelle 4.18 zu sehen, die Detektionswahrscheinlichkeit durchwegs sehr hoch ($\geq 68\%$).

b	w	α	A1		A2		A3	
			\widehat{ARL}_0	\widehat{ARL}_1	\widehat{ARL}_0	\widehat{ARL}_1	\widehat{ARL}_0	\widehat{ARL}_1
4	3	0.05	7.93	1.62	9.07	1.55	12.65	1.92
4	3	0.10	5.59	1.43	6.24	1.38	7.60	1.43
15	4	0.05	7.54	1.56	9.58	1.61	26.76	2.26
15	4	0.10	5.96	1.55	5.59	1.42	13.18	1.74

Tabelle 4.15: MC-Schätzer für ARL_0 bzw. ARL_1 bei Bayes für Modelle A1-A3

b	w	α	B1		B2		B3	
			\widehat{ARL}_0	\widehat{ARL}_1	\widehat{ARL}_0	\widehat{ARL}_1	\widehat{ARL}_0	\widehat{ARL}_1
4	3	0.05	17.69	1.50	17.57	1.47	129.76	3.25
4	3	0.10	6.82	1.33	6.29	1.31	50.83	2.4
15	4	0.05	15.90	1.62	12.17	1.35	1852.09	144.56
15	4	0.10	8.25	1.32	7.71	1.19	1255.12	48.01

Tabelle 4.16: MC-Schätzer für ARL_0 bzw. ARL_1 bei Bayes für Modelle B1-B3

False alarm Probability

Modell	B1-1	B1-2	B2-1	B2-2
A1	1.00	1.00	1.00	1.00
A2	1.00	1.00	1.00	1.00
A3	0.99	1.00	0.92	1.00
B1	0.97	1.00	0.97	1.00
B2	0.96	1.00	0.96	1.00
B3	0.34	0.66	0.00	0.00

Tabelle 4.17: Schätzer für die Fehlalarm-Wahrscheinlichkeit bei Bayes im ersten Jahr

Modell	B1-1	B1-2	B2-1	B2-2
A1	0.68	0.78	0.75	0.76
A2	0.72	0.85	0.75	0.85
A3	0.87	0.95	0.96	0.99
B1	1.00	1.00	1.00	1.00
B2	1.00	1.00	1.00	1.00
B3	1.00	1.00	1.00	1.00

Tabelle 4.18: Schätzer für die Detektions-Wahrscheinlichkeit bei Bayes in 3 Wochen

Auch hinsichtlich Sensitivität und Spezifität zeigen die Schätzwerte keine nennenswerten Unterschiede bzgl. der unterschiedlichen Wahl von b . Die Wahl $\alpha = 0.10$ führt

4 Simulationsstudie

gegebenüber der Wahl $\alpha = 0.05$ zu einer höheren Sensitivität sowie zu einer geringeren Spezifität bei den Modellen A1-A3 (siehe Tabelle 4.19).

Für Parametereinstellung B1-2 wird für die Modelle B1-B3 sowohl eine Sensitivität als auch Spezifität von etwa 90% erreicht (vgl. Tabelle 4.20).

b	w	α	A1		A2		A3	
			\widehat{Sens}	\widehat{Spec}	\widehat{Sens}	\widehat{Spec}	\widehat{Sens}	\widehat{Spec}
4	3	0.05	31.62	86.24	35.64	84.66	31.99	86.68
4	3	0.10	44.76	76.67	49.08	73.69	44.63	77.57
15	4	0.05	32.82	86.24	32.39	89.26	23.58	91.45
15	4	0.10	46.48	75.84	51.93	71.00	38.53	81.00

Tabelle 4.19: Schätzwerte für Sensitivität bzw. Spezifität bei Bayes für Modelle A1-A3

b	w	α	B1		B2		B3	
			\widehat{Sens}	\widehat{Spec}	\widehat{Sens}	\widehat{Spec}	\widehat{Sens}	\widehat{Spec}
4	3	0.05	92.38	94.71	86.45	94.97	90.09	98.05
4	3	0.10	95.86	89.56	91.42	89.98	93.57	95.64
15	4	0.05	92.74	94.85	86.15	95.27	71.01	99.83
15	4	0.10	95.82	89.68	91.22	90.43	79.68	99.49

Tabelle 4.20: Schätzwerte für Sensitivität bzw. Spezifität bei Bayes für Modelle B1-B3

4.4 Farrington-Algorithmus

Die Methode nach Farrington et al. (1996) ist mit sechs zu wählenden Parametern sehr flexibel. Um die Auswirkungen verschiedener Einstellungen zu untersuchen, wurden die Monte-Carlo-Schätzer für die bereits behandelten Performance-Maße aufgrund von neun verschiedenen Kalibrierungen ermittelt. Die konkreten Einstellungen für die jeweilige Kalibrierung sind der Tabelle 4.21 zu entnehmen. Auch hier ist die Gruppe der Parametereinstellungen 1-, sowie die Gruppe der Parametereinstellungen 2-, differenziert zu betrachten.

Analog zum vorangegangenen Abschnitt wurden die Monte-Carlo-Schätzer durch Anwendung der Funktion `algo.farrington` ermittelt.

Parametereinstellung					
Bezeichnung	b	w	α	powertrans	trend
1-1	4	3	0.10	"2/3"	FALSE
1-2	4	3	0.10	"1/2"	FALSE
1-3	4	3	0.10	none	FALSE
2-1	15	4	0.10	"2/3"	TRUE
2-2	15	4	0.05	"2/3"	TRUE
2-3	15	4	0.10	"1/2"	TRUE
2-4	15	4	0.05	"1/2"	TRUE
2-5	15	4	0.10	none	TRUE
2-6	15	4	0.05	none	TRUE

Tabelle 4.21: Verwendete Parametereinstellung bei Anwendung des Farrington-Algorithmus

Wie bereits erläutert sorgt ein $\alpha = 0.10$ für ein einseitiges 95%-iges Prädiktionsintervall. Dementsprechend lieferte $\alpha = 0.05$ ein etwas breiteres 97.5%-iges Prädiktionsintervall.

Durchschnittliche Laufzeit

Betrachtet man die Ergebnisse aller Kalibrierungen, bei denen $b = 4$, $w = 3$, gewählt wurde, ist zu erkennen, dass die mit Abstand längste *in-control* Laufzeit für die Modelle A1-A3 mit der Wurzeltransformation erreicht werden kann (vgl. Tabelle 4.22). Dabei erhöht sich zwar ebenfalls die *out-of-control* Laufzeit, jedoch ist der Unterschied hier wesentlich geringer. Die kürzeste geschätzte *in-control* bzw. *out-of-control* Laufzeit wird ohne Transformation erreicht.

Für $b = 15$, $w = 4$ wird deutlich, dass die Wahl von $\alpha = 0.05$ durch das breitere Prädiktionsintervall zu einer höheren Laufzeit führt (sowohl für Modelle A1-A3 als auch für Modelle B1-B3).

Die deutlichsten Unterschiede ergeben sich hinsichtlich gewählter Transformation hier bei den Modellen A1-A3. So liegt zum Beispiel für Modell A2 die geschätzte *in-control* Laufzeit für $\alpha = 0.10$ bei 36.75 Wochen mit der „2/3“-Transformation, bei 51.13 Wochen mit der Wurzeltransformation bzw. bei 19.38 Wochen, falls keine Transformation

4 Simulationsstudie

durchgeführt wird.

Kalibrierung	A1		A2		A3	
	\widehat{ARL}_0	\widehat{ARL}_1	\widehat{ARL}_0	\widehat{ARL}_1	\widehat{ARL}_0	\widehat{ARL}_1
1-1	74.04	2.41	27.47	1.84	24.47	1.96
1-2	107.00	2.74	38.11	1.91	31.36	2.08
1-3	57.70	2.18	18.18	1.52	15.84	1.82
2-1	73.11	2.56	36.75	1.80	27.19	1.74
2-2	123.48	2.78	62.35	2.02	50.10	2.17
2-3	125.47	2.78	51.13	1.82	37.80	1.85
2-4	161.66	3.34	108.27	2.34	79.13	2.60
2-5	50.69	2.21	19.38	1.51	19.36	1.62
2-6	59.46	2.28	36.09	1.80	27.21	1.79

Tabelle 4.22: MC-Schätzwerte für ARL_0 bzw. ARL_1 bei Farrington für Modelle A1-A3

Kalibrierung	B1		B2		B3	
	\widehat{ARL}_0	\widehat{ARL}_1	\widehat{ARL}_0	\widehat{ARL}_1	\widehat{ARL}_0	\widehat{ARL}_1
1-1	19.29	1.56	20.67	1.51	145.97	3.56
1-2	21.31	1.59	23.35	1.54	153.33	3.67
1-3	14.61	1.49	17.52	1.46	133.50	3.40
2-1	15.75	1.58	20.23	1.33	15.00	1.72
2-2	31.29	1.85	49.21	1.58	31.40	1.97
2-3	17.90	1.67	23.52	1.33	15.34	1.73
2-4	34.19	1.94	68.54	1.60	32.31	2.02
2-5	13.16	1.56	14.42	1.26	13.83	1.69
2-6	23.82	1.79	33.02	1.45	25.32	1.90

Tabelle 4.23: MC-Schätzwerte für ARL_0 bzw. ARL_1 bei Farrington für Modelle B1-B3

Für die Modelle B1-B3 ergeben sich ähnliche Ergebnisse, wie Tabelle 4.23 entnommen werden kann. Aufgrund der hohen wöchentlichen Fallzahlen und dem saisonalen Trend sind die Laufzeiten jedoch durchwegs geringer als bei den Modellen A1-A3.

Für Modell B3 ergeben sich für $b = 4$, $w = 3$ ungewöhnlich hohe Schätzwerte, verglichen mit denen für Modelle B1 und B2. Dies ist vermutlich auf den negativen linearen Trend

zurückzuführen, der durch die gewählten Parametereinstellungen nicht berücksichtigt wird.

False alarm Probability

Die geringste Fehlalarm-Wahrscheinlichkeit wird für Modelle A1-A3 mit der Parametereinstellung 1-2 bzw. 2-4 erreicht. Die Ergebnisse für Modelle B1-B3 zeigen hingegen ein differenzierteres Bild. So ist (wie es sich auch in den Schätzungen der durchschnittlichen Laufzeit gezeigt hat) die Fehlalarm-Wahrscheinlichkeit für Modell B3 und die Parametereinstellungen 1-1, 1-2 und 1-3 sehr gering. Für $b = 15$, $w = 4$ wird die geringste Fehlalarm-Wahrscheinlichkeit mit der Wurzeltransformation und $\alpha = 0.05$ (Kalibrierung 2-4) erreicht.

Auch für die Modelle B1 und B2 liefert die Einstellung 2-4 das beste Ergebnis hinsichtlich der *false alarm probability* (siehe Abb. 4.6).

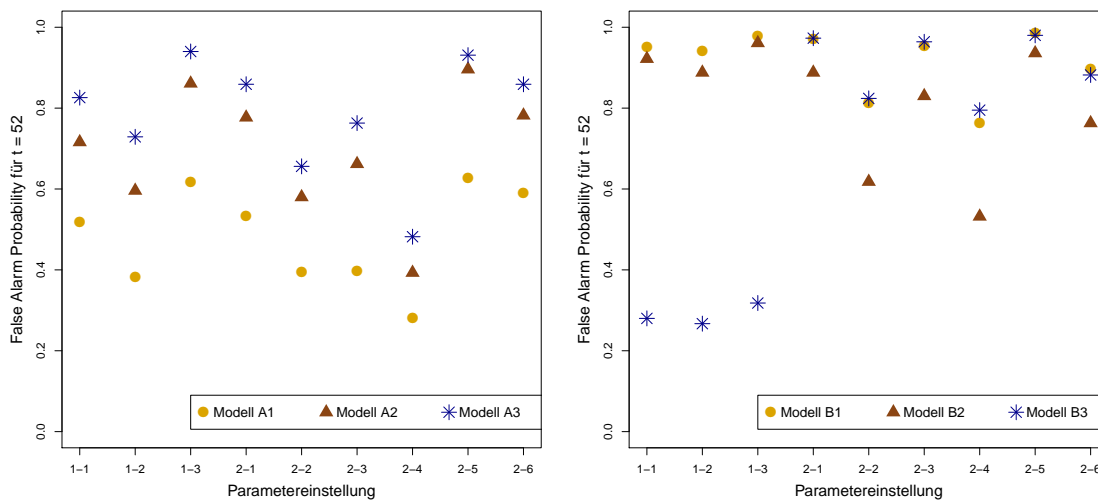


Abbildung 4.6: Zu sehen ist der Monte-Carlo-Schätzer für $P(T_A \leq 52 \mid \tau = \infty)$ bei Anwendung des Farrington-Algorithmus für simulierte Zeitreihen, die den Modellen A1-A3 bzw. B1-B3 folgen.

Die höchste Detektions-Wahrscheinlichkeit (Abb. 4.7) wird hingegen mit der Kalibrierung 1-3 bzw. 2-5 für alle Modelle erreicht. Auffallend ist hierbei, dass die Detektions-Wahrscheinlichkeit für Modell B2 und Modell B1 durchwegs nahezu 100% beträgt.

Die Auswirkung des fehlenden Trendparameters zeigt sich auch bei den Kalibrierungen

4 Simulationsstudie

1-1, 1-2 sowie 1-3 für Modell B3 in Form einer sehr geringen Detektions-Wahrscheinlichkeit (etwa 60%).

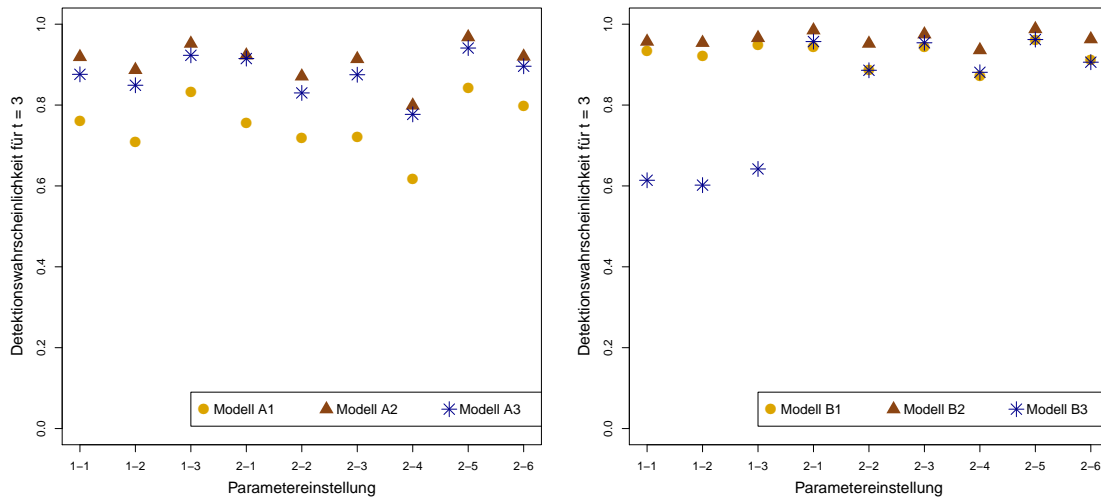


Abbildung 4.7: Zu sehen ist der Monte-Carlo-Schätzer für $P(T_A \leq 3 \mid \tau = 1)$ bei Anwendung des Farrington-Algorithmus für simulierte Zeitreihen, die den Modellen A1-A3 bzw. B1-B3 folgen.

Aufgrund der Ergebnisse ist zu vermuten, dass bei den untersuchten Modellen die Wurzeltransformation eher für eine geringe *false alarm probability* spricht, während die Detektions-Wahrscheinlichkeit eher erhöht wird, falls keine Transformation durchgeführt wird. Ersteres wird vor allem bei Betrachtung der geschätzten Fehlalarm-Wahrscheinlichkeiten mit Kalibrierung 2-4 (mit Wurzeltransformation) bzw. mit Kalibrierung 2-6 (ohne Transformation) deutlich.

Sensitivität und Spezifität

Auch für den Algorithmus von Farrington et al. (1996) gibt es keine Parametereinstellung, die eine zufriedenstellende Sensitivität der Modelle A1-A3 liefert. Wobei die Ergebnisse jedoch vermuten lassen, dass $\alpha = 0.10$ die Sensitivität gegenüber $\alpha = 0.05$ erhöht. Des Weiteren resultieren höhere Schätzwerte für den Fall, dass keine Transformation durchgeführt wurde (siehe Tabelle 4.25). Die Spezifität ist mit durchwegs zumindest 97% sehr hoch.

Ganz anders verhalten sich die Schätzungen für die Modelle B1-B3. Hier liegt auch die erreichte Sensitivität für alle betrachteten Parametereinstellungen bei zumindest

4.4 Farrington-Algorithmus

70%. Die geringsten Werte werden dabei für Modell B2 mit Kalibrierung 2-4 erreicht. Für alle Modelle (B1-B3) liefert die Einstellung 1-3 bzw. 2-5 die höchste Sensitivität. Vergleicht man die Ergebnisse lediglich hinsichtlich des Transformationstyps, so liefert die Wurzeltransformation die höchsten Schätzwerte für die Spezifität.

Kalibrierung	A1		A2		A3	
	\widehat{Sens}	\widehat{Spec}	\widehat{Sens}	\widehat{Spec}	\widehat{Sens}	\widehat{Spec}
1-1	6.37	98.47	14.91	97.52	10.60	98.23
1-2	5.00	99.01	12.57	98.23	8.73	98.79
1-3	7.89	97.77	18.20	96.18	13.01	97.25
2-1	6.37	98.43	15.42	97.33	8.75	98.26
2-2	4.96	98.99	11.66	98.50	6.66	99.00
2-3	5.03	98.98	13.29	98.02	7.26	98.78
2-4	3.81	99.35	9.32	99.11	4.92	99.42
2-5	8.03	97.66	18.73	95.89	10.53	97.40
2-6	7.18	98.06	15.60	97.27	9.12	98.10

Tabelle 4.24: Schätzer für Sensitivität bzw. Spezifität bei Farrington für Modelle A1-A3

Kalibrierung	B1		B2		B3	
	\widehat{Sens}	\widehat{Spec}	\widehat{Sens}	\widehat{Spec}	\widehat{Sens}	\widehat{Spec}
1-1	91.33	95.03	83.36	96.27	89.32	98.37
1-2	90.46	95.62	81.57	96.80	88.32	98.58
1-3	92.77	93.84	85.87	95.19	90.88	97.91
2-1	92.17	94.78	81.37	97.01	92.29	95.12
2-2	87.67	97.36	73.83	98.64	88.44	97.58
2-3	91.28	95.37	79.78	97.46	91.50	95.65
2-4	86.29	97.82	70.90	98.94	87.01	97.99
2-5	93.44	93.60	84.53	96.04	93.67	94.05
2-6	89.79	96.37	78.67	97.97	90.58	96.70

Tabelle 4.25: Schätzer für Sensitivität bzw. Spezifität bei Farrington für Modelle B1-B3

4.5 Fazit

Ist der zu detektierende Shift bekannt, kann mit der Adaption nach Rogerson und Yamada (2004) eine festgelegte ARL_0 sehr gut erreicht werden. Dies gilt auch für Zeitreihen mit geringen Anzahlen. Für Modelle A1-A3 wird mit der Adaption nach Rossi et al. (1999) die festgelegte Laufzeit deutlich unterschritten. Dementsprechend sind die Fehlalarm-Wahrscheinlichkeiten und die Sensitivität höher als bei der Adaption nach Rogerson und Yamada (2004).

In der praktischen Anwendung ist es oft schwierig, den zu detektierenden Shift allgemein zu bestimmen und eine vernünftige ARL_0 zu wählen. Bei ungeeigneter Wahl können sich starke Abweichungen hinsichtlich der Performance dieser Algorithmen ergeben, die man weder quantifizieren noch kontrollieren kann.

Der GLR-Detektor, der Farrington-Algorithmus und der Bayes-basierte Ansatz haben diesbezüglich den Vorteil, dass die Spezifizierung des Shifts nicht notwendig ist.

Für Zeitreihen mit geringen Anzahlen (Modelle A1-A3) liefert der Bayes-Algorithmus die mit Abstand geringsten ARL_0 Schätzwerte. Der GLR-Detektor hingegen liefert die höchsten (siehe Tabelle 4.10). Dementsprechend ist die Fehlalarm-Wahrscheinlichkeit für die Bayes-basierte Methode am höchsten. Die höchste Detektionswahrscheinlichkeit liefert der Algorithmus von Farrington et al. (1996). Bezüglich Sensitivität kann nur mit dem Bayesschen Algorithmus ein zufriedenstellendes Ergebnis für Modelle A1-A3 erreicht werden.

Sind Fehlalarme weitestgehend zu vermeiden, z. B. aufgrund von hohen Kosten, die ein Fehlalarm nach sich ziehen könnte, so kann für Modelle mit geringen Anzahlen der GLR-Detektor empfohlen werden. Für den Fall, dass Fehlalarme zugunsten einer höheren Sensitivität akzeptiert werden können, liefert der Algorithmus basierend auf Bayes-Inferenz die besten Ergebnisse in obiger Simulationsstudie.

Die Analyse von Zeitreihen mit hohen Anzahlen (Modelle B1-B3) durch den GLR-Detektor liefert wiederum die höchsten ARL_0 Schätzwerte. Eine Ausnahme bildet hier das Modell B3. Aufgrund der Wahl von h kann hier mit dem LR-Detektor eine sehr hohe \widehat{ARL}_0 erreicht werden. Aber auch die Ergebnisse des Bayesschen Algorithmus sind unter Einbeziehung einer größeren Referenzmenge ($b = 15$) sehr hoch (siehe Tabelle 4.16).

Die höchste Fehlalarm- und Detektionswahrscheinlichkeit ergab sich wiederum für den Bayes-Algorithmus.

Sowohl der Farrington- als auch der Bayes-Algorithmus liefern sehr hohe Schätzwerte für die Sensitivität. Beide zeigen diesbezüglich einen geringen Vorteil gegenüber dem GLR-Detektor.

Wiederum ist der GLR-Detektor aufgrund der Simulationsstudie hinsichtlich ARL_0 zu empfehlen. Für Modelle mit linearem Trend (Modell B3) kann auf Kosten der Sensitivität mit dem Bayes-Algorithmus ebenfalls ein gutes Ergebnis erzielt werden. Dabei ist jedoch davon auszugehen, dass tatsächliche Änderungen erst einige Wochen nach Eintritt detektiert werden (sehr hohe ARL_1 Schätzwerte).

Der Farrington-Algorithmus liefert für Modelle mit hohen Anzahlen vor allem hinsichtlich Sensitivität sehr gute Ergebnisse. Aufgrund der Simulationsstudie würde deshalb für höhere wöchentliche Fallzahlen der Algorithmus von Farrington et al. (1996) herangezogen werden.

5 Detektion lebensmittelbedingter Krankheitsausbrüche

Anhand gemeldeter Krankheitsfälle von *Campylobacteriose* in Österreich soll in diesem Kapitel die praktische Umsetzbarkeit der zuvor vorgestellten Methoden erläutert werden. Dazu stehen Labordaten der Nationalen Referenzzentrale für Campylobacter (NRZC) für die fünf Jahre 2007-2011 zur Verfügung. Die zu wählenden Parameter sollen an die spezielle Struktur der österreichischen Daten angepasst werden, mit dem Ziel, eine geeignete Methode zur on-line Detektion von Campylobacteriose-Ausbrüchen zu finden. Eine in weiterer Folge mögliche Umsetzung der rechnergestützten Ausbruchsfriherkennung soll die Experten bei der Erkennung von Krankheitsausbrüchen unterstützen bzw. unentdeckte Krankheitsausbrüche aufdecken.

5.1 Gesetzliche Regelungen und Definitionen

Das mit 1.1.2006 in Kraft getretene Bundesgesetz zur Überwachung von Zoonosen und Zoonoseerregern (ZoonG) regelt unter anderem „die Organisation der Überwachung von Zoonosen und Zoonosenerregern“ (§1 ZoonG). Unter Überwachung wird dabei das „System zur Erfassung, Auswertung und Verbreitung von Daten über das Auftreten von Zoonosen und Zoonoseerregern sowie diesbezüglicher Antibiotikaresistenzen“ verstanden (§2 Absatz 2 Satz 3 ZoonG).

Der Begriff Zoonosen umfasst unter anderem Krankheiten, die durch Lebensmittel verursacht werden. Die gesetzliche Definition dazu lautet

(§2 Absatz 2 Satz 4 ZoonG) Zoonosen sind „*Krankheiten und/oder Infektionen, die auf natürlichem Weg direkt oder indirekt zwischen Tieren und Menschen übertragen werden können*“.

Ebenso wird im ZoonG ein lebensmittelbedingter Krankheitsausbruch definiert als

5 Detektion lebensmittelbedingter Krankheitsausbrüche

(§2 Absatz 2 Satz 2 ZoonG) „*das unter gegebenen Umständen festgestellte Auftreten einer mit demselben Lebensmittel oder mit demselben Lebensmittelunternehmen in Zusammenhang stehenden oder wahrscheinlich in Zusammenhang stehenden Krankheit und/oder Infektion in mindestens zwei Fällen beim Menschen oder eine Situation, in der sich die festgestellten Fälle stärker häufen als erwartet*“.

Mit den vorgestellten Methoden kann klarerweise nicht eindeutig erkannt werden, ob Krankheitsfälle mit demselben Lebensmittel oder Lebensmittelunternehmen in Verbindung gebracht werden können. Eine Eingrenzung von möglichen Verbindungen kann aber durch genauere Typisierung des Krankheitserregers erreicht werden. So kann zum Beispiel durch Sero- und Lysotypisierung des Salmonellose-Erregers und die anschließende getrennte Betrachtung der verschiedenen Typen diesbezüglich eine größere Detektionswahrscheinlichkeit erreicht werden. Bei Campylobacteriose-Erregern ist die Spezifizierung der humanen Isolate aufgrund aufwendigerer Methoden nur gröber möglich. Die Analyse gemeldeter Alarme wird durch entsprechende Experten in jedem Fall notwendig sein.

5.2 Zoonoserreger Campylobacter

Campylobacteriose (und ihre Erreger) zählen neben anderen Zoonosen, wie z. B. Salmonellose, Listeriose und Tuberkulose, zu den überwachungspflichtigen Zoonosen (Zoonoserregern) in Österreich (Anl. 1 ZoonG). EU-weit ist Campylobacteriose mit 190 000 Fällen jährlich die am häufigsten auftretende lebensmittelbedingte Krankheit (EFSA 2012). Auch in Österreich hat die Zahl der Campylobacteriose-Erkrankungen jene der Salmonellose bereits überschritten. Experten gehen allerdings davon aus, dass der Anstieg der Infektionen beim Menschen seit dem Jahr 1997 auf „eine höhere Sensibilität der Labors gegenüber Campylobacter als meldepflichtigen Erreger und eine verbesserte Diagnostik“ zurückzuführen ist (BMG 2010).

Der Zoonoserreger ist vor allem in Geflügel und Rohmilch zu finden. Es können aber auch Infektionen durch „kontaminiertes, nicht aufbereitetes Trinkwasser“ ausgelöst werden (vgl. BMG 2010). Die häufigsten Infektionen erfolgen allerdings durch Kreuzkontamination, verursacht durch mangelnde Hygiene beim Bearbeiten von rohem Fleisch.

Mensch-zu-Mensch Übertragungen sind selten und werden, wenn überhaupt, bei Kindern beobachtet (BMG 2010).

Campylobacter-Infektionen „sind gewöhnlich selbstlimitierend“ und haben oft einen asymptomatischen Verlauf. Treten Symptome auf, so sind diese meist „Diarrhöen, Abdominalschmerzen bzw. -krämpfe, Fieber, Müdigkeit“ (vgl. BMG 2010).

Detailliertere allgemeine sowie aktuelle Informationen zum Zoonosenerreger Campylobacter sind u. a. den Web-Sites *www.ages.at*, *www.efsa.europa.eu*, sowie *www.bmg.gv.at* zu entnehmen. Letztere enthält ebenfalls den von der NRZC erstellten Jahresbericht.

5.3 Datensituation in Österreich

Die Datengrundlage für die Analyse liefert seit 1.1.2009 das Epidemiologische Meldesystem für Infektionskrankheiten (EMS). Des Weiteren werden humane Isolate bei Verdacht auf eine Campylobacteriose-Erkrankung von den einzelnen Labors in den jeweiligen Bundesländern an die Nationale Referenzzentrale für Campylobacter (NRZC) übermittelt. Diese Zentrale führt gegebenenfalls eine Spezifizierung der eingesendeten Isolate durch. Die Untersuchungsergebnisse von eingesandten humanen Isolaten werden in Excel-Files erfasst. Die gemeldeten Fälle im EMS, bzw. die Meldungen ans Bundesministerium für Gesundheit (BMG) beinhalten nur sporadisch eine konkretere Spezifizierung des Campylobacter-Erregers. Für die Erkennung eines Krankheitsausbruchs ist dies jedoch eine wesentliche Information. Da im speziellen Fall der Steiermark von nahezu vollständigen Daten hinsichtlich genauerer Spezifizierung ausgegangen werden kann, wird hier die Analyse der gemeldeten Campylobacteriose-Erkrankungen dieses Bundeslandes durch Funktionen des R-package 'surveillance' erläutert werden.

Dem zugrundeliegenden Datensatz können Informationen wie Alter und Geschlecht des Patienten, Laboreingangsdatum sowie die Spezies des Campylobacter-Erregers entnommen werden.

Spezifizierung des Campylobacter-Erregers

Die am häufigsten auftretende Spezies des Campylobacter-Erregers ist *C. jejuni*. Eine Differenzierung der Spezies erfolgt durch die NRZC mithilfe „phänotypischer und molekularbiologischer Methoden“ (vgl. Much et al. 2008). Dabei wird im Wesentlichen eine Differenzierung nach *C. jejuni* und *C. coli* vorgenommen.

Hinsichtlich der Definition von lebensmittelbedingten Krankheitsausbrüchen sollen die gemeldeten Krankheitsfälle verursacht durch *C. jejuni* bzw. *C. coli* getrennt betrachtet werden.

Bestätigte Krankheitsausbrüche

Eine eindeutige Zuordnung der vorliegenden Krankheitsfälle zu Krankheitsausbrüchen kann nicht vorgenommen werden, obwohl in Österreich jährlich ein Bericht von lebensmittelbedingten Krankheitsausbrüchen veröffentlicht wird. Für diesen Bericht und für den jährlichen EU-Zoonosentrendbericht beauftragt das BMG die AGES, die österreichischen Daten zu lebensmittelbedingten Krankheitsausbrüchen zu sammeln. Dazu werden von den jeweiligen Landessanitätsdirektionen und der MA15 in Wien eigens erstellte Tabellen der European Food Safety Authority (EFSA) ausgefüllt. Eine Fallzuordnung wird dabei jedoch nicht vorgenommen. Es werden lediglich Informationen zum Ausbruch selbst gesammelt, wie „Erkrankung im In- bzw. Ausland erworben; Anzahl der erkrankten, hospitalisierten oder verstorbenen Personen; mit dem Ausbruch in Verbindung gebrachtes Lebensmittel; Hinweise zum angegebenen Lebensmittel, ob dieses als Infektionsquelle verdächtigt oder bestätigt wurde“ usw. (Much et al. 2008).

Aus diesem Grund ist eine Evaluierung der anzuwendenden Verfahren durch Ermittlung der in Kapitel 3 erläuterten Performance-Maße nicht möglich. Auch bei eindeutig definierten Krankheitsausbrüchen wäre dies schwierig, da aufgrund der im Allgemeinen geringen Anzahl an Beteiligten an einem Krankheitsausbruch solche Ausbrüche durchaus „übersehen“ werden können.

Die Beurteilung und Auswahl der in Frage kommenden Methoden erfolgt deshalb im Weiteren in Kooperation mit der NRZC.

5.3.1 Datenaufbereitung

Bevor mit der Analyse der gemeldeten Krankheitsfälle begonnen werden konnte, mussten die vorhandenen Excel-Files in R eingelesen und zu einem Datensatz zusammengefügt werden. Dazu wurden die, für die nachfolgende Untersuchung relevanten Spalten extrahiert und benannt.

Für jeden gemeldeten Fall wurde ein Probeneingangsdatum vermerkt. Dieses wird für eine zeitliche Zuordnung herangezogen.

Für eine weitere Anwendung auf aktuell eingehende Daten ist eine wöchentliche Analyse von Interesse. Aus diesem Grund werden die Diagnosedaten (im Format TT.MM.JJJJ) in 52 Kalenderwochen/Jahr eingeteilt. Krankheitsfälle, die in einer Kalenderwoche in ein und demselben Jahr berichtet wurden, werden dabei zu einer wöchentlich berichteten Fallzahl aufsummiert. Es wurde hier bewusst auf die Verwendung der ISO Ein-

teilung in Kalenderwochen verzichtet, um eine bessere Vergleichbarkeit der Wochen in den unterschiedlichen Jahren zu ermöglichen.

Um jedes Jahr in 52 Kalenderwochen einzuteilen, wurde deshalb zuerst festgelegt, dass die zweite Kalenderwoche jedes Jahres mit dem 7.1. beginnt. Danach umfasst jede KW genau sieben Tage. Hier kann es natürlich auch zu 53 Wochen in einem Jahr kommen. Um dies zu standardisieren, werden alle Beobachtungen in der KW 53 des Jahres i den Beobachtungen KW 1 des Jahres $i + 1$ zugeordnet.

Da die vorliegende Datenaufzeichnung im Jahr 2007 beginnt, beginnt auch die KW 1 dieses Jahres mit dem 01.01.2007, wie der Tabelle 5.1 zu entnehmen ist. Kalenderwoche 1 des Jahres 2008 beginnt dagegen mit dem 30.12.2007 und umfasst 8 Tage. Auf diese Weise kann eine Einteilung in 52 Kalenderwochen/Jahr erreicht werden.

KW	Jahr	Beginn-Datum	End-Datum
1	2007	01.01.2007	06.01.2007
2	2007	07.01.2007	13.01.2007
3	2007	14.01.2007	20.01.2007
4	2007	21.01.2007	27.01.2007
5	2007	28.01.2007	03.02.2007
⋮	⋮	⋮	⋮
50	2007	09.12.2007	15.12.2007
51	2007	16.12.2007	22.12.2007
52	2007	23.12.2007	29.12.2007
1	2008	30.12.2007	06.01.2008
2	2008	07.01.2008	13.01.2008

Tabelle 5.1: Einteilung des Jahres 2007 und des Anfangs des Jahres 2008 in Kalenderwochen.

5.3.2 Eine erste deskriptive Analyse

Wie in Abbildung 5.1 zu sehen ist, liegt die Anzahl der gemeldeten Campylobacteriose-Erkrankungen österreichweit in etwa um 5000/Jahr. Im Jahr 2007 gab es mit 6132 gemeldeten Fällen die meisten Krankheitsfälle österreichweit bezogen auf die Jahre 2007 bis 2011. Die Daten für die österreichweiten Fälle stammen aus dem Meldesystem des BMG und sind unter anderem auch dem Jahresbericht 2011¹ der NRZC zu entnehmen.

¹Der Jahresbericht 2011 der Nationalen Referenzzentrale für Campylobacter ist unter www.bmg.gv.at/cms/home/attachments/4/3/2/CH1338/CMS1339605708104/campylobacter_jb_2011_2012_06.11.pdf verfügbar.

5 Detektion lebensmittelbedingter Krankheitsausbrüche

Die Daten für die in der Steiermark gemeldeten Krankheitsfälle von Campylobacteriose entsprechen der verfügbaren Labordatenbank.

In der Steiermark lieferte in der betrachteten Zeitperiode das Jahr 2009 mit 877 Fällen den Spitzenwert. Allerdings sind über die Jahre 2007-2011 keine nennenswerten Schwankungen zu erkennen (735-877 Fälle).

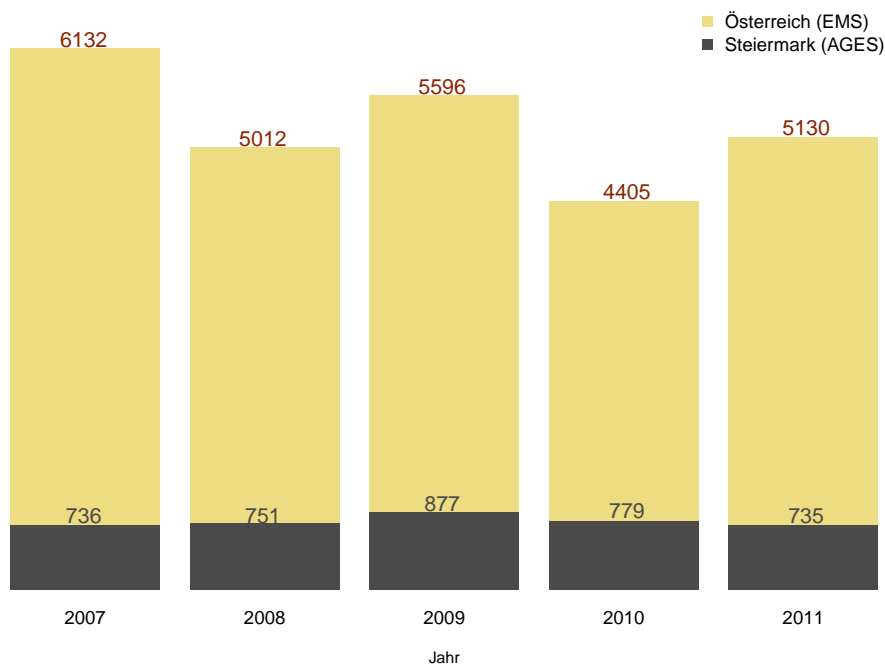


Abbildung 5.1: Anzahl der gemeldeten Campylobacteriose-Erkrankungen in ganz Österreich bzw. im Bundesland Steiermark in den Jahren 2007-2011.

Unterteilung nach Bundesland und Spezies

Um Krankheitsausbrüche effizienter erkennen zu können, werden die Daten getrennt nach Bundesland und nach Spezies untersucht. Mögliche Zusammenhänge verschiedener Bundesländer werden bei dieser univariaten Betrachtung zunächst nicht berücksichtigt.

Wie bereits erwähnt, sind Campylobacteriose-Erkrankungen hauptsächlich auf die Spezies *C. jejuni* und *C. coli* zurückzuführen. Der Campylobacter-Erreger *C. jejuni* ist

aber der mit Abstand am häufigsten auftretende Erreger.

Dies wird in Abbildung 5.2 deutlich. So macht der Anteil der *C. jejuni* Infektionen in der Steiermark zwischen 85% und 92% aus. *C. coli* Infektionen sind mit durchschnittlich 9.6% (entspricht durchschnittlich 75 Fällen) sichtlich seltener vertreten. Vernachlässigbar sind Krankheitsfälle durch andere Campylobacter-Spezies mit durchschnittlich 1.2% pro Jahr.

Kann kein Nährboden gebildet werden, ist es auch möglich, dass aufgrund eines eingesandten humanen Isolats keine Bestimmung des Campylobacter-Erregers vorgenommen werden kann. So lassen sich die durchschnittlich 13 Fälle/Jahr, die nicht spezifizierbar sind, erklären.

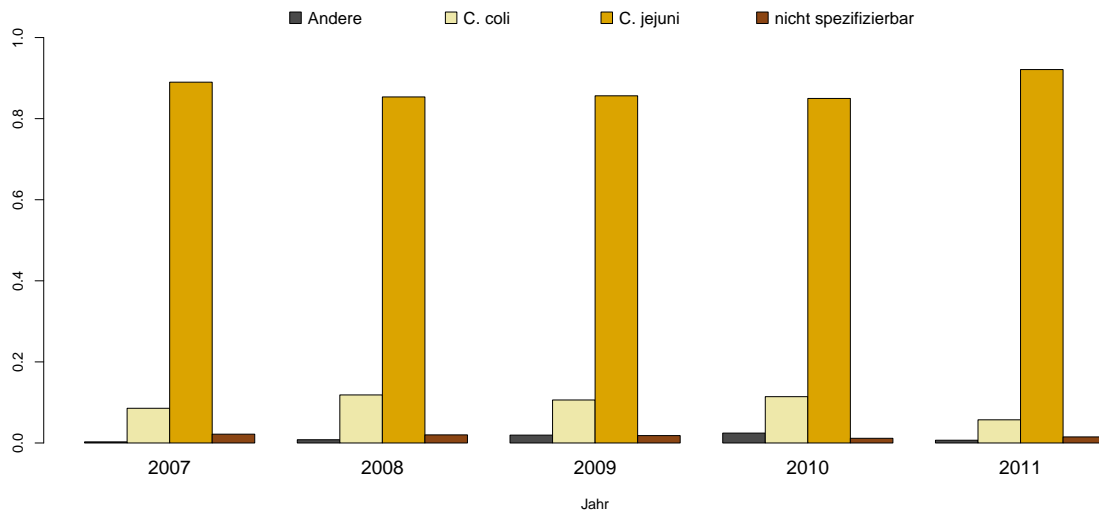


Abbildung 5.2: Relative Anteile der bestimmten Campylobacter-Spezies an der Anzahl der gemeldeten Campylobacteriose-Erkrankungen in der Steiermark.

Betrachtet man die Zeitreihen der wöchentlichen Campylobacter-Infektionen in der Steiermark getrennt nach *C. jejuni* (Abbildung 5.4) und *C. coli* (Abbildung 5.3), wird ersichtlich, dass die saisonale Schwankung bei den *C. coli* Infektionen im Gegensatz zu den *C. jejuni* Infektionen kaum vorhanden ist. Die wöchentlichen Fallzahlen der *C. jejuni* Infektionen liegen zwischen 5 und 10 bzw. in den Sommermonaten Juli, August und September zwischen 10 und 25.

Weiters unterscheiden sich die beiden Zeitreihen, wie erwartet, in den wöchentlichen Anzahlen sehr stark. So gibt es immer wieder Wochen, in denen keine *C. coli* Erkran-

5 Detektion lebensmittelbedingter Krankheitsausbrüche

kung gemeldet wurde.

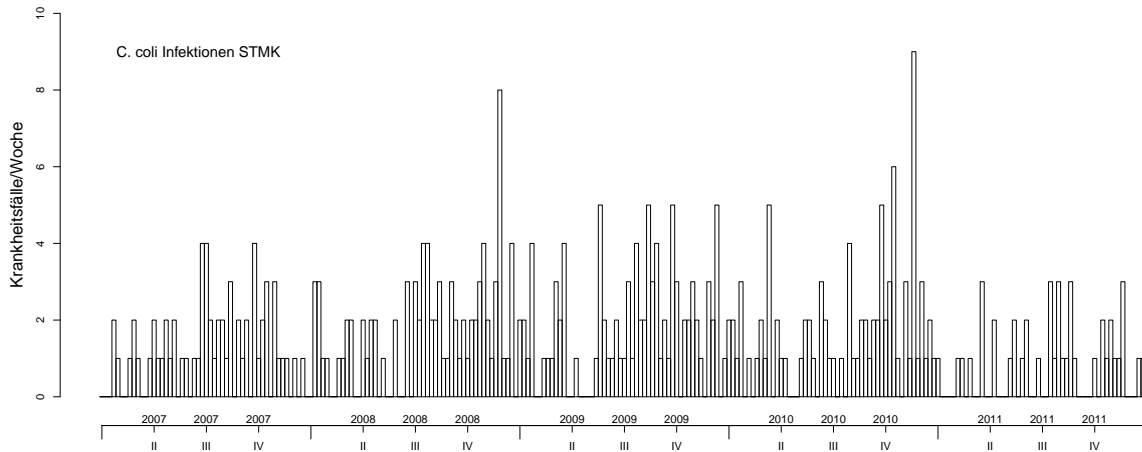


Abbildung 5.3: Zeitreihe der gemeldeten Campylobacter-Erkrankungen steiermarkweit, die auf den Campylobacter-Erreger *C. coli* zurückzuführen sind, für 2007-2011.

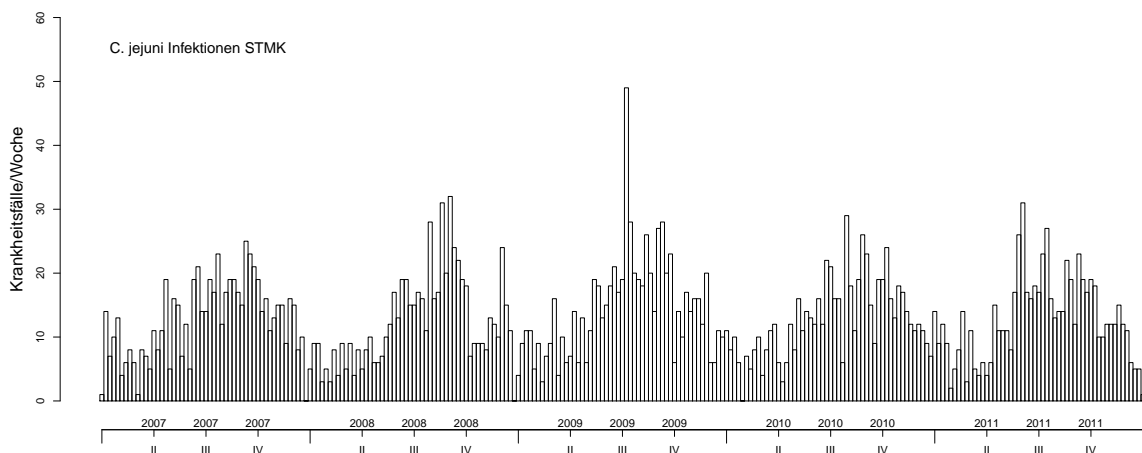


Abbildung 5.4: Zeitreihe der gemeldeten Campylobacter-Erkrankungen steiermarkweit, verursacht durch den Campylobacter-Erreger *C. jejuni* in den Jahren 2007-2011.

Untersuchung auf Poisson-Verteilung

Nach der Datenaufbereitung soll untersucht werden, ob es plausibel ist, dass die Zeitreihen von gemeldeten *C. coli* bzw. *C. jejuni* Krankheitsfällen aus einer Poisson-Verteilung stammen. Dazu wurden in Abbildung 5.5 die Jahresmittelwerte gegen die Jahresvarianzen der *C. coli* bzw. der *C. jejuni* Infektionen gegeneinander aufgetragen. Bei den *C. coli* Infektionen (rechtes Bild) weichen die errechneten Punkte nur gering von der Referenzlinie (Mittelwert = Varianz) ab. Hingegen ist bei Betrachtung der *C. jejuni* Infektionen von einer Überdispersion auszugehen. Abbildung 5.5 zeigt auch die Gesamtmittelwerte bzw. -varianzen der Jahre 2007 bis 2011. Für *C. coli* Erkrankungen ergeben sich hierfür die Werte (1.45, 2.03) sowie für *C. jejuni* Erkrankungen (12.97, 47.04).

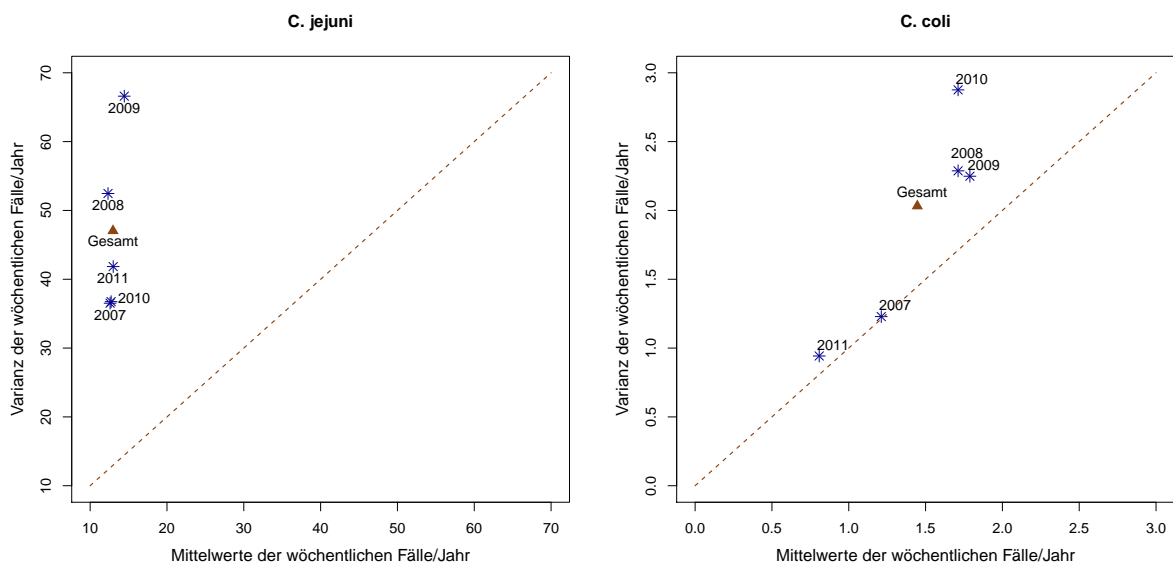


Abbildung 5.5: Jahresmittelwerte und -varianzen der *C. jejuni* bzw. *C. coli* Infektionen in der Steiermark.

5.3.3 Schlussfolgerungen für die praktische Betrachtung

Aus einer ersten Analyse des vorhandenen Datenmaterials ergeben sich bereits einige Punkte, die bei der Anwendung der in Kapitel 2 vorgestellten Methoden zu berücksichtigen sind.

- Sowohl der *in-control* als auch der *out-of-control* Mittelwert sind unbekannt. Des

5 Detektion lebensmittelbedingter Krankheitsausbrüche

Weiteren kann durch das Fehlen von zuordenbaren Ausbruchsinformationen kein „sicherer“ Zeitraum zur Schätzung des *in-control* Parameters bestimmt werden. Es muss in diesem Fall davon ausgegangen werden, dass es in jenem Zeitraum, der zur Schätzung herangezogen wird, Campylobacteriose-Ausbrüche gegeben hat.

- Auch der zu detektierende Parametershift ist unbekannt und kann nicht generell festgelegt werden.
- Durch das vermehrte Auftreten von Campylobacter-Infektionen in den Sommermonaten ist der zu untersuchende Mittelwert zeitabhängig mit saisonalen Schwankungen. Des Weiteren ist ein schwach positiver Trend zu erkennen.
- Bei der Analyse der wöchentlichen Krankheitsfälle von *C. jejuni* Infektionen ist, wenn möglich, auf die Überdispersion einzugehen.

5.4 Anwendung der Algorithmen

Aufgrund der im vorangegangenen Kapitel erläuterten Datensituation eignen sich zur Analyse der Krankheitsfälle Poisson- und Negativ-Binomial-Regression-Charts (Höhle 2006; Höhle und Paul 2008), die Methode von Farrington et al. (1996) sowie der Bayes-basierte Ansatz (Riebler 2004).

Da einerseits der *in-control* Mittelwert und der zu detektierende Shift unbekannt sind, und andererseits ARL_0 und ARL_1 nicht festgelegt werden sollen, werden die CUSUM-Adaptionen nach Rogerson und Yamada (2004) und Rossi et al. (1999) für die Analyse ausgeschlossen.

Im folgenden Abschnitt sollen die Daten einer univariaten Betrachtung unterzogen werden. Das heißt, die Campylobacteriose-Erkrankungen werden getrennt nach *C. coli* und *C. jejuni* Infektionen untersucht.

Beobachtungen der drei Jahre 2007 bis 2009 werden zur Schätzung des *in-control* Mittelwerts herangezogen. Mittels verschiedener Parametereinstellungen sollen dann die jeweiligen Algorithmen auf die Daten der Jahre 2010 und 2011 angewendet werden.

5.4.1 CUSUM - Poisson- und NegBin-Regression Charts

Sowohl der *in-control* Mittelwert als auch der zu detektierende Shift η sind unbekannt. Daher wird der GLR-Detektor herangezogen, der iterativ eine Schätzung für η

ermittelt. Vor Anwendung des Algorithmus wird der *in-control* Mittelwert durch ein generalisiertes lineares Modell geschätzt. Dabei ist eine Verteilungsannahme zu treffen und eine plausible Anzahl I der harmonischen Schwingungen zu wählen, um den saisonalen Effekt widerzuspiegeln.

C. jejuni Infektionen

Aufgrund der obigen deskriptiven Analyse wird für Beobachtungen der *C. jejuni* Infektionen angenommen, dass sie aus einer Negativ-Binomial-Verteilung stammen. Für die Modellauswahl wurden Modelle mit $I = 1, 2, 3, 4$, verbunden mit einer Trendkomponente angepasst. Aufgrund genauerer Betrachtung der berechneten Modelle wurde des Weiteren ein Modell mit $I = 1$, aber ohne Trendkomponente geschätzt. Grundsätzlich lassen sich keine signifikanten Unterschiede in den Schätzungen erkennen (siehe Abbildung 5.6). Die Entscheidung für ein Modell beruht auf dem Wert des *Bayesian Information Criterium* (BIC) und auf allgemeinen Gütekriterien, wie Signifikanz der Parameterschätzer, R^2 etc. Für die Analyse der *C. jejuni* Infektionen wird das Modell mit einer harmonischen Schwingung ($I = 1$) und ohne linearen Trend ($\beta_1 = 0$) zur Mittelwertschätzung herangezogen (entspricht der grünen durchgehenden Linie in Abb. 5.6). Die Parameterschätzer des gewählten Modells sind dem nachfolgendem R-Code zu entnehmen.

```
> t <- 1:156; omega=(2*pi)/52
> mu0 <- jejuni@observed[1:156]
> mod1.ohne <- glm.nb(mu0 ~ 1 + sin(omega*t) + cos(omega*t))
> summary(mod1.ohne)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-3.6571	-0.7131	-0.0725	0.5654	3.4350

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	2.49427	0.03129	79.704	<2e-16
sin(omega * t)	-0.43228	0.04377	-9.875	<2e-16
cos(omega * t)	-0.37428	0.04363	-8.578	<2e-16

(Dispersion parameter for Negative Binomial (15.7503) family taken to be 1)

Null deviance:	348.71	on 155	degrees of freedom
Residual deviance:	175.56	on 153	degrees of freedom
AIC:	934.13		

```
> mod1.ohne$theta
15.7503
```

5 Detektion lebensmittelbedingter Krankheitsausbrüche

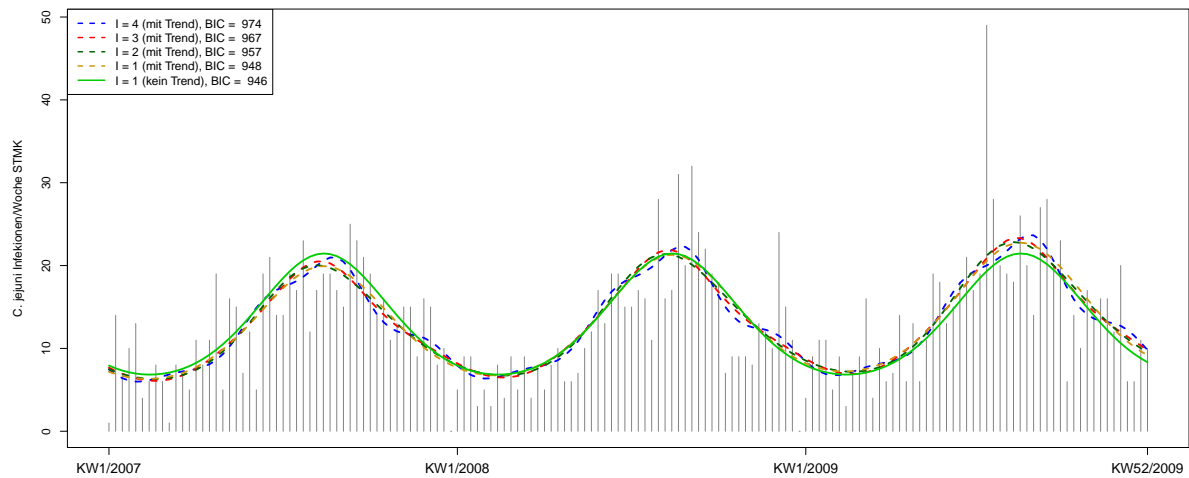


Abbildung 5.6: Modellselektion zur Schätzung des *in-control* Mittelwerts für *C. jejuni*-Infektionen in der Steiermark.

Die Größenordnung des zu wählenden Thresholds h hängt nicht von der Größenordnung der Anzahl der Krankheitsfälle ab, sondern von dem Wert

$$\sum_{i=k}^t \log \left(\frac{f(x_i | \mu_{0k} \exp(\hat{\eta}))}{f(x_i | \mu_{0k})} \right). \quad (5.1)$$

Hier ist dieser Wert sehr gering, wobei erstmal ein Threshold von 0.003 plausibel erscheint. Mit $c.ARL = 0.003$ wird dies bei Aufruf der Funktion `glrnb` festgelegt.

```
> mu0.pred.mod1 <- as.numeric(predict(mod1.ohne, newdata =
  data.frame(t = 157:260), type = "response"))
> glrnb(jejuni, control = list(range = 157:260, mu0 = mu0.pred.mod1,
  alpha = mod1.ohne$theta, Mtilde = 1, M = -1, change = "intercept",
  theta = NULL, dir = "inc", ret = "value", c.ARL = 0.003))
```

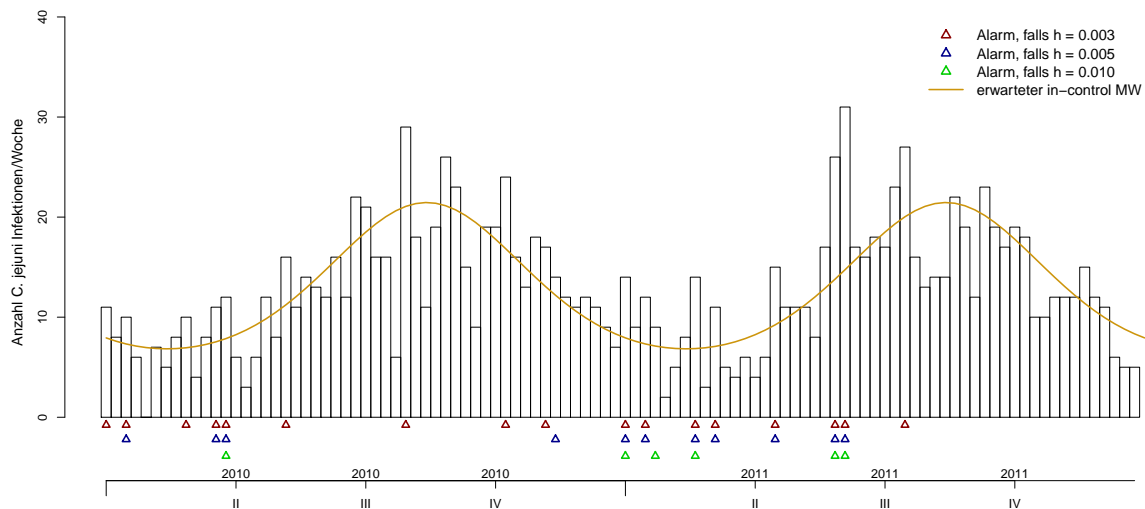



Abbildung 5.7: Die gelbe Linie gibt die Schätzung des ausgewählten Modells für den *in-control* Mittelwert an. Dieses Modell enthält eine harmonische Schwingung und keine Trendkomponente. Die Dreiecke kennzeichnen die Wochen, in denen der gewählte Threshold h überschritten wird.

Die Mittelwert-Schätzung kann ebenfalls im Funktionsaufruf direkt durchgeführt werden, mit $S = 1$, `trend = FALSE`, `refit = FALSE` wird das ausgewählte Modell für die Wochen 1:156 (entspricht den Wochen der Jahre 2007-2009) angepasst. Falls die Option `refit = TRUE` festgelegt wird, wird das GLM nach Ausgabe eines Alarms zur Berechnung des erwarteten *in-control* Mittelwerts neu angepasst. Dabei werden dann nicht nur die Beobachtungen aus den Jahren 2007-2009 herangezogen, sondern alle Beobachtungen bis zum Zeitpunkt des letzten Alarms. Die Liste der angepassten Modelle kann dann mittels `...@control$mu0Model$fitted` abgerufen werden.

```
> glr.jejuni <- glrnb(jejuni, control = list(range = 157:260,
mu0 = list(S = 1, trend = FALSE, refit = TRUE),
alpha = mod1.ohne$theta, Mtilde = 1, M = -1, change = "intercept",
theta = NULL, dir = "inc", ret = "value", c.ARL = 0.003))
```

5 Detektion lebensmittelbedingter Krankheitsausbrüche

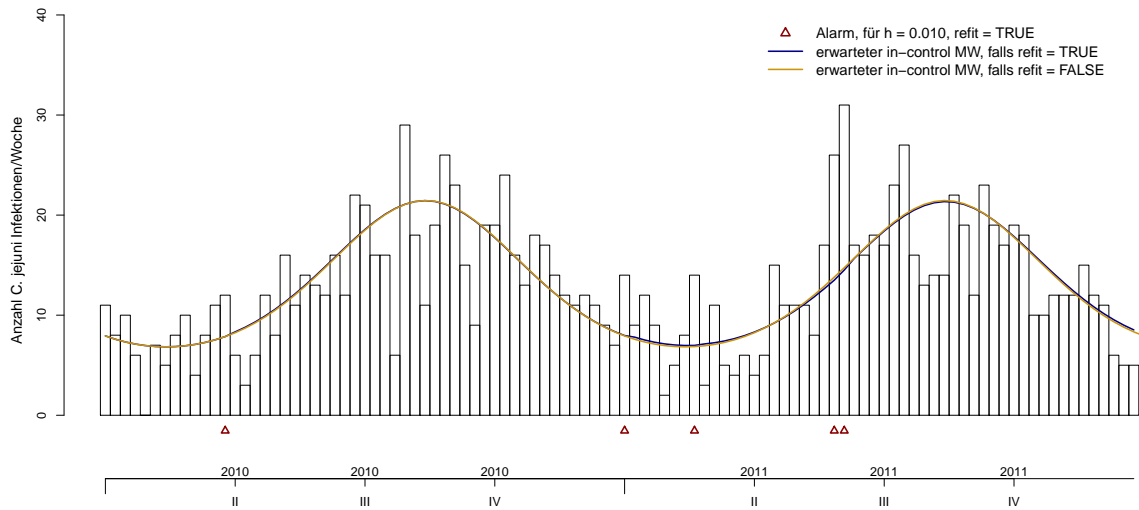


Abbildung 5.8: Vorhersage für *C. jejuni* Infektionen mit bzw. ohne Neuanpassung des Modells nach Ausgabe eines Alarms des CUSUM-Regression-Chart-Detektors.

h	#Alarme
0.001	26
0.002	22
0.003	17
0.004	14
0.005	11
0.006	9
0.007	8
0.008	7
0.009	6
0.010	6
0.011	5
0.012 - 0.016	4
0.017 - 0.020	3
0.021 - 0.038	1
≥ 0.039	0

Tabelle 5.2: Anzahl der angezeigten Alarme bei Analyse der *C. jejuni* Fälle mittels GLR-Detektor und unterschiedlichen h -Werten

Die obige Tabelle 5.2 fasst die Auswirkungen des gewählten Thresholds h zusammen. Einerseits wird für $h = 0.001$ mit 26 Alarmen in den betrachteten 104 Wochen die Schranke zu gering sein, andererseits sind Threshold-Werte größer als 0.039 zu hoch.

Für das zweite und dritte Quartal 2010 werden für $h \geq 0.004$ trotz der auf den ersten Blick hohen Fallzahlen keine Alarme signalisiert. Nahezu für alle gewählten Threshold-Werte (< 0.021) wird die errechnete Schranke in

- KW 1/2011 mit 14 Fällen,
- KW 3/2011 mit 12 Fällen,
- KW 8/2011 mit 14 Fällen und in
- KW 10/2011 mit 11 Fällen

überschritten. Mit einem Höchstwert von 31 Fällen wird in KW 23/2011 jede errechnete Schranke für Threshold-Werte < 0.039 überschritten.

C. coli Infektionen

Für die Analyse der *C. coli* Infektionen wurde ebenfalls eine Modellselektion durchgeführt. Hier kann die Poisson-Verteilung als gerechtfertigt angesehen werden. Dazu wurden Modelle mit $I = 0, 1, 2, 3$ harmonischen Schwingungen miteinander verglichen. Wie bereits bei Betrachtung der Zeitreihe bemerkt, ist hier kaum eine saisonale Schwankung zu erkennen. Sehr wohl lässt sich aber ein leichter positiver Trend feststellen. Dieser ist nicht unbedingt auf eine Zunahme an Krankheitsfällen zurückzuführen, sondern kann auch von der Sensibilisierung der Bevölkerung hinsichtlich Zoonosen herrühren.

Das Modell, das nur eine harmonische Schwingung und eine Trend-Komponente beinhaltet, weist den geringsten BIC-Wert (504) auf (siehe Abbildung 5.9).

```
> t <- 1:156; omega=(2*pi)/52
> mu0.coli <- coli@observed[1:156]
> mod1.coli <- glm(mu0.coli ~ 1 + t + , family="poisson")
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.127580	0.142896	0.893	0.371957
t	0.003414	0.001485	2.298	0.021535
sin(omega * t)	-0.357329	0.096319	-3.710	0.000207
cos(omega * t)	-0.112889	0.091437	-1.235	0.216975

```
Null deviance: 213.10 on 155 degrees of freedom
Residual deviance: 186.49 on 152 degrees of freedom
AIC: 491.84
```

5 Detektion lebensmittelbedingter Krankheitsausbrüche

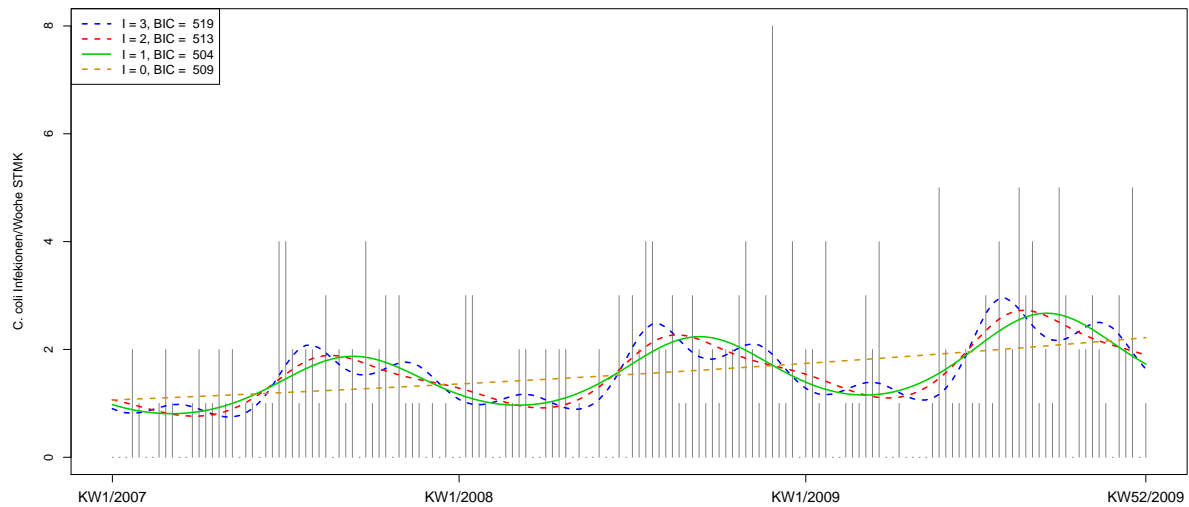


Abbildung 5.9: Modellselektion zur Schätzung des *in-control* Mittelwerts für *C. coli*-Infektionen in der Steiermark. Angegeben sind auch die jeweiligen Werte der Bayesian Information Criteria (BIC).

Mit dem Befehl `dispersiontest` aus dem R-Package 'AER' kann ein Test auf Überdispersion für Poisson-GLMs durchgeführt werden. Dieser bestätigt die Vermutung, dass die Nullhypothese „Dispersionsparameter = 1“ nicht zugunsten der Alternative „Dispersionsparameter > 1“ verworfen werden kann.

```
> dispersiontest(mod1.coli)
      Overdispersion test
data:  mod1.coli
z = 0.6169, p-value = 0.2686
alternative hypothesis: true dispersion is greater than 1
sample estimates:
dispersion
 1.092139
```

Um die Auswirkungen der Wahl von h zu veranschaulichen, sind in Abbildung 5.10 die errechneten Werte (5.1) für gewählte $h = 0.1, 3.1, 5.0$ dargestellt. Bei Überschreiten des festgelegten Grenzwertes h wird ein Alarm für diese Woche ausgegeben (gekennzeichnet durch ein rotes Dreieck) und der Wert aus (5.1) wird auf Null gesetzt. Für $h = 5.0$ wird zum Beispiel die festgelegte Schranke im Beobachtungszeitraum nicht überschritten. Abbildung 5.11 zeigt für $h = 0.1, 3.1, 5.0$ was das hinsichtlich der oberen Schranke für die wöchentliche Anzahl an Krankheitsfällen heißt. Wird etwa in der vierten Woche

5.4 Anwendung der Algorithmen

des Jahres 2010 für $h = 0.1$ ein Alarm angezeigt, so müssten für $h = 3.1$ hingegen ≥ 6 gemeldete Krankheitsfälle vorliegen, damit der Wert aus (5.1) ≥ 3.1 ist.

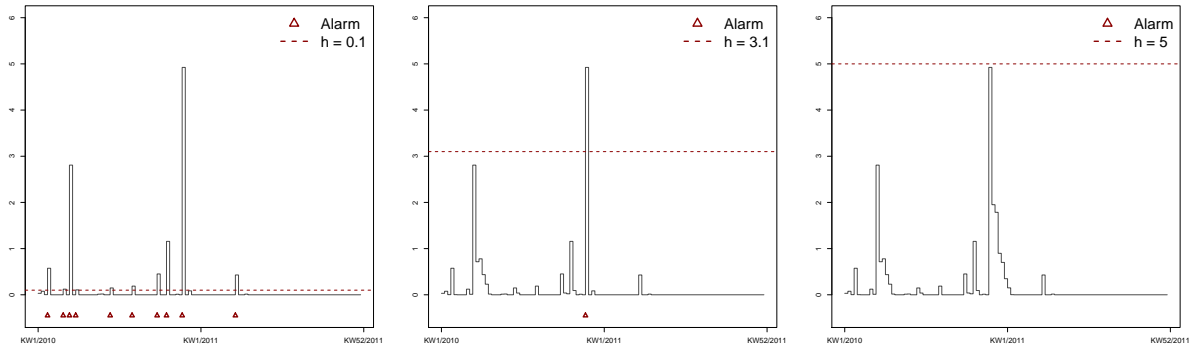


Abbildung 5.10: Für die verschiedenen Threshold-Werte $h = 0.1, 3.1, 5.0$ sind die Werte der Summe in (5.1) zu sehen und die obere Schranke, deren Überschreitung zu einem Alarmsignal (Dreieck) führt.

```
> mu0.coli.mod1 <- as.numeric(predict(mod1.coli, newdata =
  data.frame(t = 157:260), type = "response"))

> cusum.coli <- glrpois(coli, list(range = 157:260,
  mu0 = mu0.coli.mod1, Mtilde = 1, M = -1, change = "intercept",
  theta = NULL, dir = "inc", ret = "cases", c.ARL = 0.1))
```

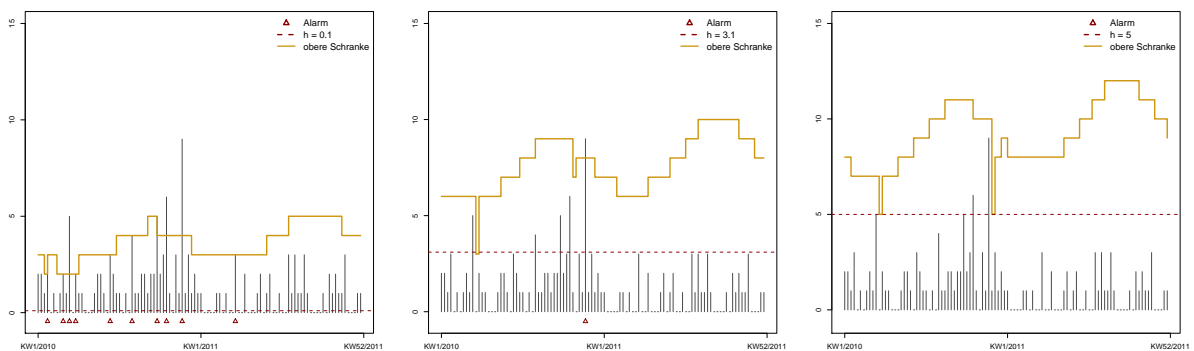


Abbildung 5.11: Die gelbe Linie gibt die Anzahl der Fälle an, die in der jeweiligen Woche notwendig wäre, sodass der Wert der GLR-Statistik (5.1) den gewählten Threshold (strichlierte Linie) überschreitet.

Wiederum werden in Tabelle 5.3 die Anzahl der angezeigten Alarme je nach Threshold aufgelistet. Der Algorithmus wurde für $h \in \{0.1, 0.2, \dots, 4.8, 4.9, 5.0\}$ angewendet. Der

5 Detektion lebensmittelbedingter Krankheitsausbrüche

größte Sprung bzgl. der Alarme ergibt sich bei Erhöhung von h auf 0.2. Hingegen sind die errechneten oberen Grenzen für $h \in \{0.6, 0.7, 0.8, 0.9, 1.0, 1.1\}$ sehr ähnlich. Dementsprechend ist auch die Anzahl der angezeigten Alarme gleichbleibend. Wird $h \geq 5$ gewählt, wird für den gesamten Zeitraum KW 1, 2010 - KW 52, 2011 die errechnete obere Grenze nicht überschritten.

Eine spezielle Woche sticht durch eine besonders hohe Anzahl von gemeldeten Fällen hervor. Mit 9 Fällen liegt die KW 47 im Jahr 2010 weit über der durchschnittlichen Anzahl an gemeldeten *C. coli* Infektionen.

h	#Alarme
0.1	10
0.2 - 0.4	6
0.5	4
0.6 - 1.1	3
1.2 - 2.8	2
2.9 - 4.9	1
≥ 5.0	0

Tabelle 5.3: Anzahl der angezeigten Alarme bei Analyse der *C. coli* Fälle mittels GLR-Detektor und unterschiedlichen h -Werten

Unabhängig von der Wahl des Thresholds liefert der CUSUM-Algorithmus keine zufriedenstellenden Ergebnisse für das Jahr 2011. Da in den Jahren, die zur Datenanpassung herangezogen wurden, ein klarer Trend zu erkennen war, dieser aber im Jahr 2011 nicht mehr erkennbar ist, wird die Zeitreihe deutlich überschätzt. Das zeigt einen deutlichen Nachteil dieser Methode. Die Modellauswahl ist daher laufend zu hinterfragen. Der Algorithmus reagiert bei einmaliger Modellwahl nicht auf Änderungen bzgl. Saisonalität und Trend in „neuen“ Beobachtungen. Soll diese Methode zur *on-line* Detektion herangezogen werden, sollte laufend eine Modellselektion durchgeführt werden. Die Einstellung `refit = TRUE` wäre hier evtl. ebenfalls sinnvoll.

5.4.2 Bayes-basierter Ansatz

C. jejuni Infektionen

Auch der Bayes-basierte Ansatz, der in Riebler (2004) beschrieben wird, wurde zuerst zur Analyse der *C. jejuni* und in weiterer Folge zur Analyse der *C. coli* Infektionen in

der Steiermark herangezogen. Um die Auswirkungen der zu wählenden Parameter zu erkennen, wurden verschiedene Parametereinstellungen gewählt. Tabelle 5.4 fasst die Ergebnisse in Form der angezeigten Alarme für *C. jejuni* Erkrankungen zusammen.

b	w	α	#Alarme
0	4	0.01	6
0	4	0.05	13
0	4	0.10	18
0	6	0.01	5
0	6	0.05	10
0	6	0.10	19
2	4	0.01	3
2	4	0.05	8
2	4	0.10	13
2	6	0.01	2
2	6	0.05	7
2	6	0.10	13

Tabelle 5.4: Anzahl der angezeigten Alarme bei Analyse der *C. jejuni* Fälle mittels Bayes-Algorithmus

Die Abbildung 5.12 zeigt exemplarisch die Ergebnisse der Analyse mittels Bayes-Ansatz für verschiedene Parametereinstellungen. Da mit der Überwachung in der KW 1 des Jahres 2010 begonnen werden soll, stehen zur Bildung der Referenzmenge 156 Beobachtungen zur Verfügung. Um die Auswirkungen der einbezogenen historischen Daten zu verdeutlichen, wurde in Abbildung 5.12 einerseits $\mathbf{b} = 0$ und andererseits $\mathbf{b} = 2$ gewählt. Für die obere Grafik in Abbildung 5.12 wird die Referenzmenge nur durch die jeweiligen 6 Wochen vor der aktuellen Woche gebildet. Für die untere Grafik werden auch Beobachtungen aus zwei Vorjahren berücksichtigt. In beiden Fällen wurden Beobachtungen direkt vor der aktuellen Woche berücksichtigt ($\text{actY} = \text{TRUE}$). Die unterschiedlichen oberen Grenzen (blaue, gelbe und rote strichlierte Linien für $\alpha = 0.01, 0.05, 0.10$) ergeben sich durch die differenzierte Wahl des Signifikanzniveaus α . Die Wochen, in denen die gemeldeten Fallzahlen diese errechneten Schranken überschreiten, werden mit Dreiecken in denselben Farben (blau, gelb, rot) gekennzeichnet. Für $\alpha = 0.10$ wird die Grenze, die durch $P(Y_{n+1} \leq x_{0.10} \mid y_1, \dots, y_n) \geq 0.90$ bestimmt wird, am häufigsten überschritten. Hinsichtlich der Wahl von b ist zu erkennen, dass die obere Grenze sehr sensibel auf aktuell gemeldete Beobachtungen reagiert für $\mathbf{b} = 0$. Der Verlauf der Schranken für $\mathbf{b} = 2$ scheint hingegen bereits etwas geglättet zu sein. Während für $\alpha = 0.10$ und $\alpha = 0.05$ nahezu dieselben Wochen zur Überschreitung der oberen Grenze führen, wird für $\alpha = 0.01$ wesentlich seltener ein Alarm signalisiert.

5 Detektion lebensmittelbedingter Krankheitsausbrüche

Für $b = 2$, $w = 6$, $actY = TRUE$, $alpha = 0.01$ wird nur in KW 22 im Jahr 2011 mit 26 *C. jejuni* Erkrankungen und KW 23 ebenfalls im Jahr 2011 mit 31 Fällen die errechnete obere Grenze überschritten. Für $b = 2$ werden hier durchwegs weniger Alarme signalisiert.

Bei Anwendung des Bayes-basierten Ansatzes mit $b = 0$ werden im Gegensatz zum CUSUM-Algorithmus gerade im zweiten und dritten Quartal des Jahres 2010 relativ viele Alarme signalisiert (je nach Wahl von α). Unter allen betrachteten Parametereinstellungen liefert das Ergebnis für $b = 2$, $w = 6$, $alpha = 0.10$ ein scheinbar realistisches Ergebnis. Auch hier werden in den Kalenderwochen

- 1/2011 mit 14 Fällen,
- 8/2011 mit 14 Fällen,
- 10/2011 mit 11 Fällen,
- 16/2011 mit 15 Fällen,
- 22 und 23/2011 mit 26 und 31 Fällen sowie
- 29/2011 mit 27 Fällen

Ausbrüche vermutet. Zusätzlich führen auch die Spitzenwerte im zweiten und dritten Quartal des Jahres 2010 zu Alarmsignalen.

C. coli Infektionen

Für gemeldete *C. coli* Infektionen der Jahre 2010-2011 liefert Tabelle 5.5 einen Überblick über Ergebnisse des Bayes-Ansatz für verschiedene Parametereinstellungen.

Abbildung 5.13 veranschaulicht wiederum exemplarisch die Ergebnisse für $b = 0$, $w = 4$ (obere Grafik) und $b = 2$, $w = 4$. Aufgrund der geringen Anzahl scheint die Glättung durch Einbeziehung von historischen Daten sogar zu strikt zu sein. So sind vor allem die oberen Grenzen für die letzten Wochen im Jahr 2011 relativ hoch. Auch hier werden für $b = 2$ deutlich weniger Alarme signalisiert. Alle dargestellten Parametereinstellungen führen jedoch für KW 47 im Jahr 2010 mit 9 *C. coli* Fällen und für KW 11 im Jahr 2010 (5 Fälle) zu einem Alarm.

5.4 Anwendung der Algorithmen

b	w	α	#Alarme
0	4	0.01	5
0	4	0.05	13
0	4	0.10	21
0	6	0.01	6
0	6	0.05	12
0	6	0.10	22
2	4	0.01	2
2	4	0.05	6
2	4	0.10	14
2	6	0.01	2
2	6	0.05	5
2	6	0.10	13

Tabelle 5.5: Anzahl der angezeigten Alarme bei Analyse der *C. coli* Fälle mittels Bayes-Algorithmus

5 Detektion lebensmittelbedingter Krankheitsausbrüche

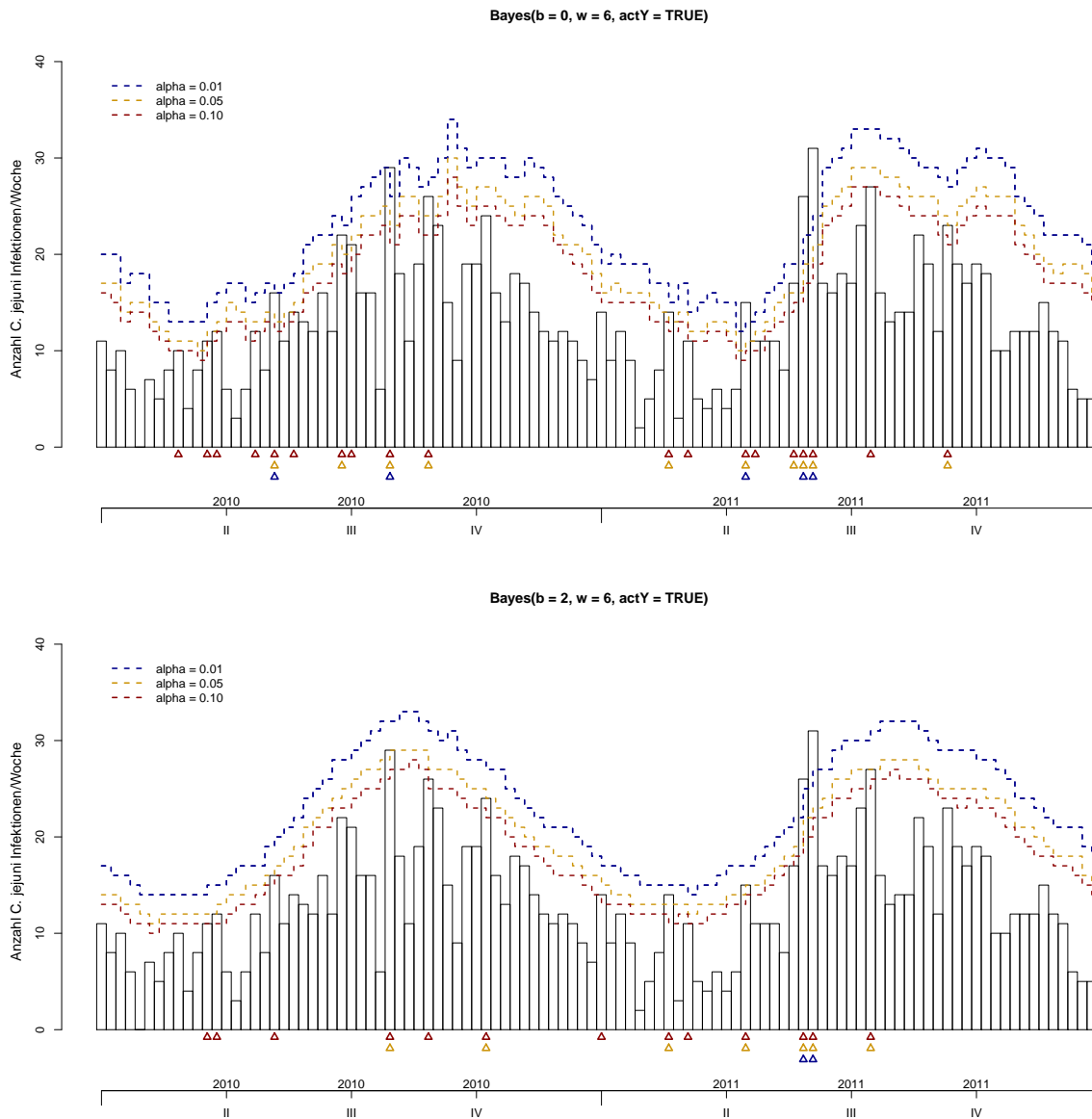


Abbildung 5.12: Ergebnisse der Bayes-Methode für *C. jejuni* Infektionen in der Steiermark für verschiedene Parametereinstellungen.

5.4 Anwendung der Algorithmen

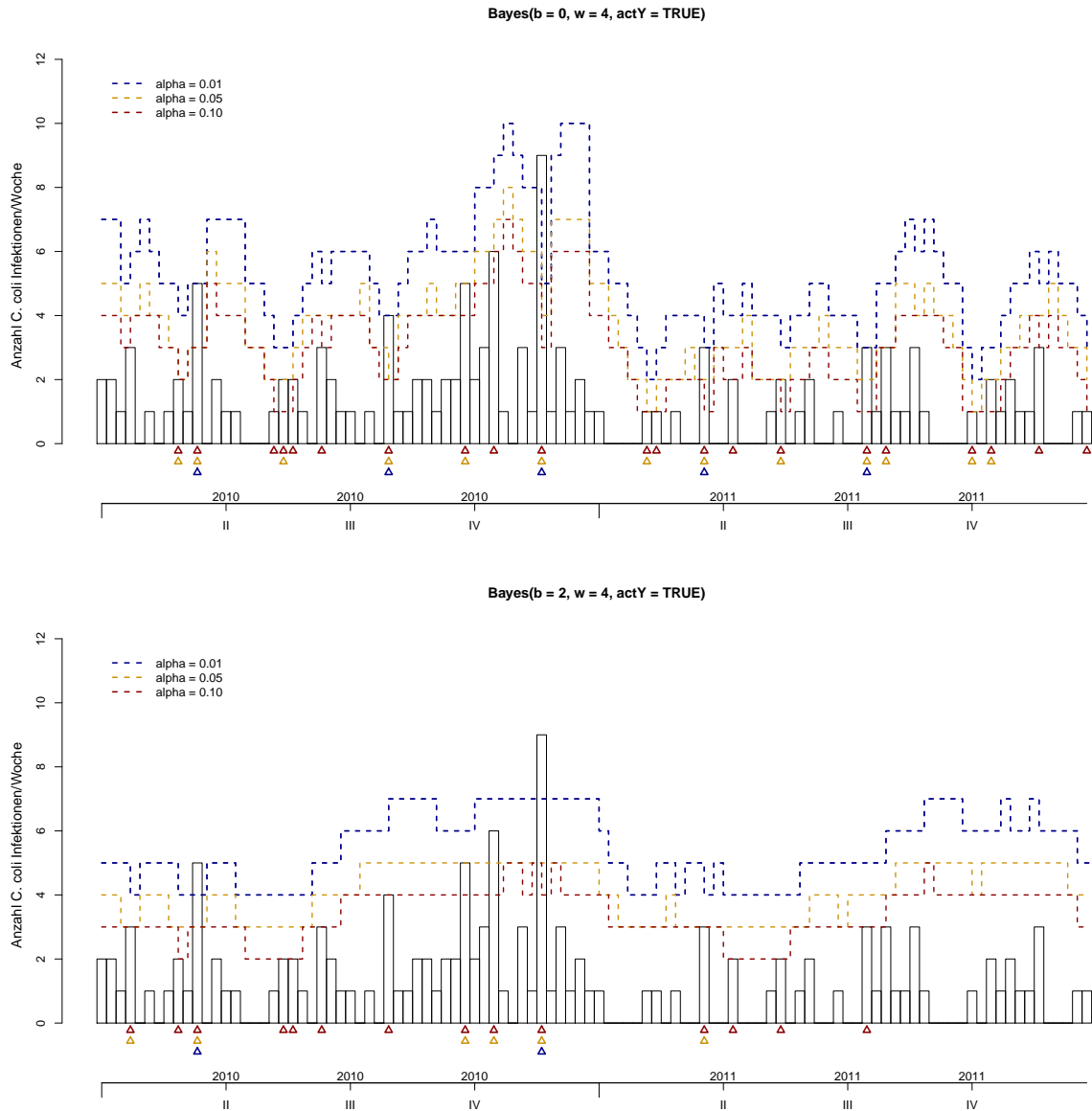


Abbildung 5.13: Ergebnisse der Bayes-Methode für *C. coli* Infektionen in der Steiermark für verschiedene Parametereinstellungen.

5.4.3 Farrington Algorithmus

C. jejuni Infektionen

Der Farrington-Algorithmus ist, wie bereits erwähnt, der flexibelste Algorithmus hinsichtlich der Parameter-Festlegung. Dies kann auch ein Nachteil sein, da durch falsche Wahl der Parameter evtl. nicht die gewünschte Performance erreicht wird. Für die Analyse der steiermarkweit gemeldeten Campylobacteriose-Erkrankungen in den Jahren 2010-2011 wurde aufgrund der Ausführungen in Kapitel 2.3 zunächst ein Transformationssystem festgelegt. Aufgrund der deutlichen Abweichungen zwischen Varianz und Mittelwert bei *C. jejuni* Infektionen wird die varianzstabilisierende Wurzeltransformation als geeignet erachtet. Zum Vergleich sind in Tabelle 5.6 aber auch die Ergebnisse ohne Transformation dargestellt. Betrachtet man Abbildung 5.14 wird allerdings nur ein geringer Unterschied hinsichtlich transformierter (blaue strichlierte Linie) und nicht-transformierter oberer Grenze deutlich. Die Grenze basierend auf dem nicht-transformierten Threshold (gelbe strichlierte Linie) liegt dabei immer etwas niedriger.

Ebenfalls festzulegen ist `limit54`. Dieser Parameter ist entscheidend für die Berechnung des „Überschreitungs-Scores“ X . Da die wöchentliche Anzahl der gemeldeten *C. jejuni* Infektionen in der Steiermark jedoch relativ hoch ist, hat die Änderung der Parameter (5, 4) hier keinen Einfluss auf das Ergebnis.

b	w	α	Transformation	#Alarmer
2	3	0.10	"1/2"	4
2	3	0.10	none	7
2	3	0.05	"1/2"	2
2	3	0.05	none	3
2	4	0.10	"1/2"	3
2	4	0.10	none	4
2	4	0.05	"1/2"	2
2	4	0.05	none	3
2	6	0.10	"1/2"	3
2	6	0.10	none	4
2	6	0.05	"1/2"	1
2	6	0.05	none	3

Tabelle 5.6: Anzahl der angezeigten Alarme bei Analyse der *C. jejuni* Fälle mittels Farrington-Algorithmus

Für die Analyse wurde zwar einerseits $\alpha = 0.05$ und andererseits $\alpha = 0.10$ festgelegt, jedoch zeigen die Ergebnisse, dass $\alpha = 0.10$ hier eine sinnvolle Wahl ist. (Diese resultiert in der Berechnung von 95%-Prädiktionsintervallen.) Deshalb soll nur mehr auf Ergebnisse mit `alpha = 0.10` eingegangen werden.

In den Kalenderwochen

- 8/2011 mit 14 Fällen,
- 22/2011 mit 26 Fällen und
- 23/2011 mit 31 Fällen

wird sowohl für $w = 3$, $w = 4$ als auch für $w = 6$ ein Alarm signalisiert. Für $w = 4$ wird zusätzlich in KW 44/2010 mit 18 Fällen ein *C. jejuni* Ausbruch vermutet. Für $w = 3$ wird die errechnete Grenze ebenfalls in KW 44/2010 sowie in weiteren drei Wochen (KW 13/2010 (12 Fälle), KW 41/2010 (24 Fälle), KW 16/2011 (15 Fälle)) überschritten.

Fließen weniger Wochen (um die Referenzwoche) in die Referenzmenge ein, werden also deutlich mehr Alarme signalisiert.

Auch hier fallen bei allen Parametereinstellungen wieder die Wochen 22 und 23 im Jahr 2011 auf.

C. coli Infektionen

Da die wöchentliche Anzahl der gemeldeten *C. coli* Erkrankungen in der Steiermark sehr gering ist, wird die „2/3“-Transformation in Betracht gezogen. Mit dieser soll die Schiefe ausgeglichen werden, die bekannterweise bei Poisson-Verteilungen mit geringen Anzahlen vorliegt. Da für die Berechnung der Prädiktionsintervalle Normalverteilungs-Quantile herangezogen werden, kann die Transformation für die Zuverlässigkeit des Ergebnisses entscheidend sein.

Aufgrund der geringen Anzahl hat auch die Wahl von `limit54` Auswirkungen auf das Ergebnis. Der Score \mathbb{X} wird nach Farrington et al. (1996) auf 0 gesetzt, falls es weniger als 5 Beobachtungen in den letzten 4 Wochen gab. Denkbar wären für *C. coli* Infektionen auch die Definitionen $\mathbb{X} = 0$, falls ≤ 4 Beobachtungen in den letzten 4 Wochen, oder $\mathbb{X} = 0$, falls ≤ 4 Beobachtungen in den letzten 3 Wochen. Nur für Beobachtungen mit $\mathbb{X} > 1$ wird ein Alarm signalisiert. Der Farrington-Algorithmus wurde mit `limit54 = c(5,4)`, `= c(4,4)`, `= c(4,3)`, `= c(3,4)` und `= c(3,3)` ausgeführt. Die Anzahl der Wochen, in denen $\mathbb{X} > 1$ gilt, unterscheiden sich allerdings gar nicht bzw. kaum. Abbildung 5.15 zeigt zwei Parametereinstellungen, für die sich die Anzahl unterscheidet. Während in der oberen Grafik das von Farrington et al. (1996) vorge-

5 Detektion lebensmittelbedingter Krankheitsausbrüche

schlagene Parameterpaar (5, 4) gewählt wurde, ist in der unteren Grafik das Ergebnis für (4, 3) zu sehen. In KW 15 wird einerseits die berechnete Schranke überschritten, falls $\text{limit54} = c(5,4)$ gilt. Da in den 2 Wochen vor der KW 15 keine Fälle von *C. coli* Infektionen in der Steiermark gemeldet wurden, wird für $\text{limit54} = c(4,3)$ der Score \mathbb{X} auf 0 gesetzt.

b	w	α	Transformation	#Alarme
2	3	0.10	"2/3"	4
2	3	0.10	none	5
2	3	0.05	"2/3"	3
2	3	0.05	none	4
2	4	0.10	"2/3"	4
2	4	0.10	none	5
2	4	0.05	"2/3"	3
2	4	0.05	none	5
2	6	0.10	"2/3"	3
2	6	0.10	none	5
2	6	0.05	"2/3"	3
2	6	0.05	none	3

Tabelle 5.7: Anzahl der angezeigten Alarme bei Analyse der *C. coli* Fälle mittels Farrington-Algorithmus

Des Weiteren ist in Abbildung 5.15 zu erkennen, dass die nicht-transformierte Schranke wiederum unter der Schranke liegt, die aus der „2/3“-Transformation resultiert.

Bei allen Parametereinstellung führen die gemeldeten Krankheitsfälle in

- KW 11/2010 mit 5 Fällen,
- KW 42/2010 mit 6 Fällen und
- KW 47/2010 mit 9 Fällen

zur Überschreitung der berechneten Grenze (sowohl ohne Transformation als auch mit der „2/3“-Transformation). Wird keine Transformation durchgeführt, wird weiters in

- KW 4/2010 mit 3 Fällen ($w = 4$, $\alpha = 0.10$, $w = 6$, $\alpha = 0.10$) und in
- KW 15/2011 mit 2 Fällen ($w = 3$, $\alpha = 0.10$)

ein Alarm signalisiert.

5.4 Anwendung der Algorithmen

Für $b = 2$, $w = 4$, $\alpha = 0.1$, $\text{powertrans} = "2/3"$ liegt auch in KW 39/2010 die Anzahl der tatsächlich gemeldeten Krankheitsfälle weit über der erwarteten Anzahl.

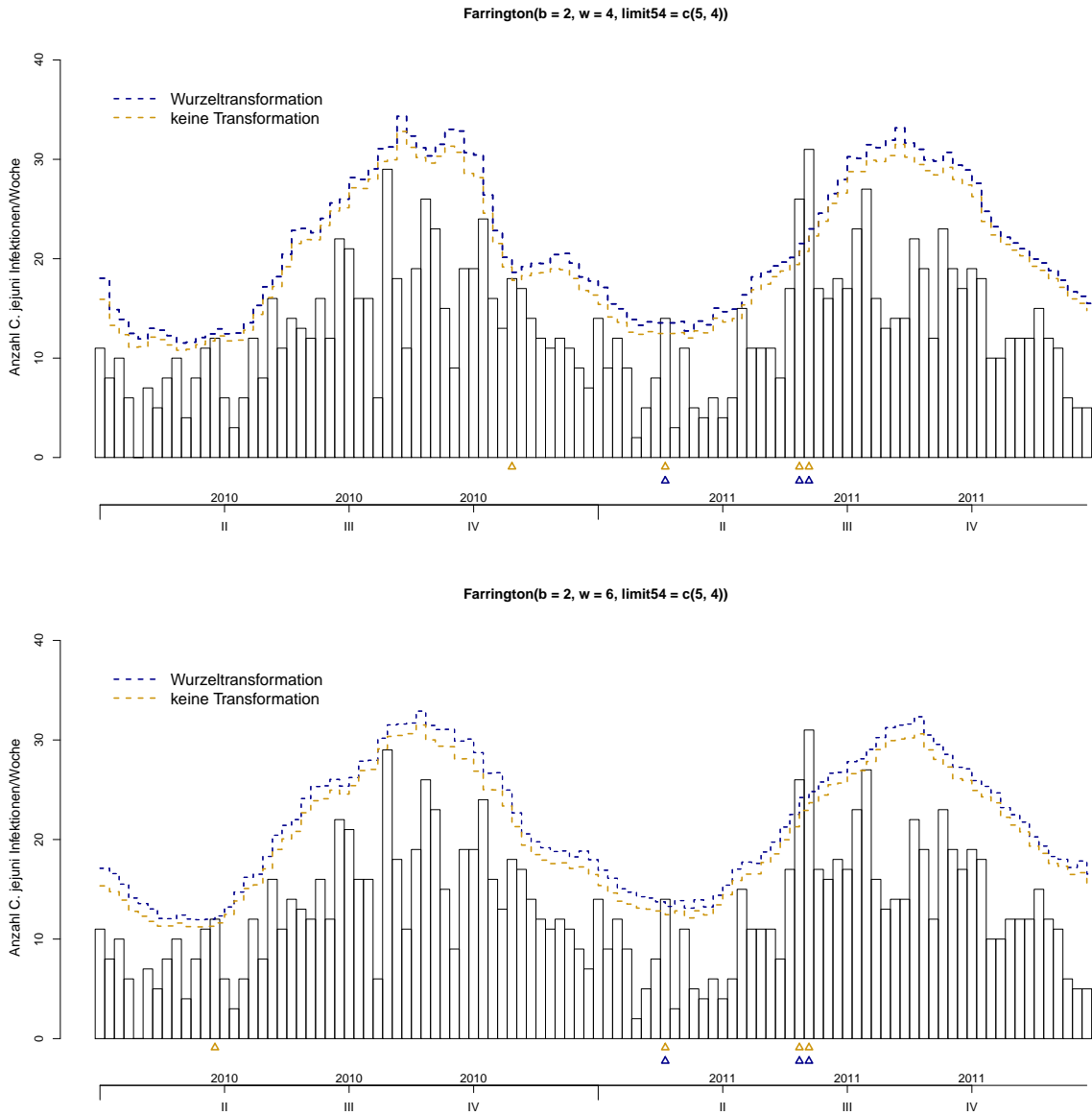


Abbildung 5.14: Ergebnisse Farrington-Algorithmus für *C. jejuni* Infektionen (steiermarkweit). Wochen, in denen die errechnete Grenze (je nach Parametereinstellung) überschritten wurde, sind durch ein Dreieck markiert.

5 Detektion lebensmittelbedingter Krankheitsausbrüche

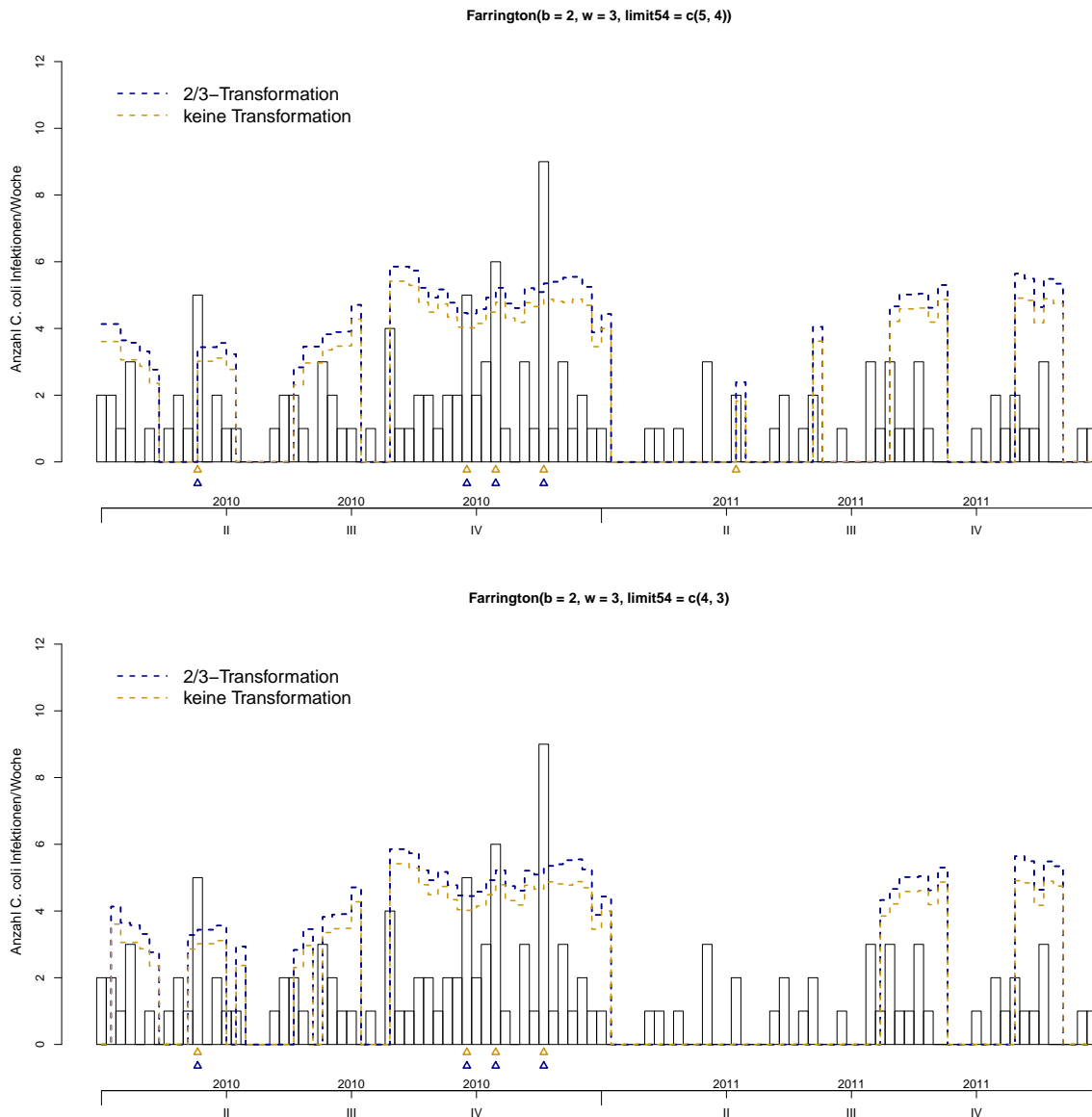


Abbildung 5.15: Ergebnisse Farrington-Algorithmus für *C. coli* Infektionen (steiermarkweit). Wochen, in denen die errechnete Grenze (je nach Parametereinstellung) überschritten wurde, sind durch ein Dreieck markiert.

5.5 Grafische Analyse für die Bezirke der Steiermark

Eine grafische Analyse mittels Bezirkskarten-Plots kann zusätzliche Informationen zur Analyse in Kapitel 5.4 liefern. Vor allem für die Erkennung von Ausbrüchen nach der gesetzlichen Ausbruchsdefinition kann ein örtlicher Zusammenhang der gemeldeten Fälle entscheidend sein.

In den Bezirkskarten-Plots werden die gemeldeten Fälle für ausgewählte Wochen je Bezirk dargestellt. Für die Fallzuordnung zu einem Bezirk wird dabei das im Datensatz vermerkte zuständige Gesundheitsamt herangezogen. Die Größenordnung der Fallzahlen in einem Bezirk wird durch eine Farbskala verdeutlicht. Ein weiß gefärbter Bezirk steht dafür, dass es im betrachteten Zeitraum keinen gemeldeten Krankheitsfall gegeben hat. Je nach maximal gemeldeter Fallzahl der Bezirke steht ein grün gefärbter Bezirk für eine geringe Anzahl an Krankheitsfällen und ein rot gefärbter für eine dementsprechend hohe Anzahl.

Die unterschiedliche Einwohnerzahl verfälscht hier auf den ersten Blick allerdings das Ergebnis. Aus diesem Grund ist es durchaus sinnvoll, die gemeldeten Krankheitsfälle je 100 000 Einwohner (EW) zu betrachten. Für die Einwohnerzahl (für das Jahr 2011) wird die Anzahl aller Personen, die (am 1.1.2011) ihren Hauptwohnsitz im jeweiligen Bezirk hatten, herangezogen.² Die errechnete Anzahl an Fällen je 100 000 EW wurde für die grafische Darstellung (kaufmännisch) gerundet. Des Weiteren ist zu beachten, dass für das Jahr 2011 nicht alle gemeldeten Fälle einem Bezirk zugewiesen werden konnten. Dies trifft bei *C. jejuni* Infektionen auf 11 Beobachtungen und bei *C. coli* Erkrankungen auf 1 Beobachtung zu.

Abbildung 5.16 bzw. Abbildung 5.17 zeigt die Anzahl der gemeldeten *C. jejuni* Erkrankungen in den ersten 12 Kalenderwochen des Jahres 2011 nach Bezirk. Die Anzahlen sind absolut (Abb. 5.16) bzw. je 100 000 EW (Abb. 5.17). Absolut gab es in den Kalenderwochen 1-12 (2011) die meisten *C. jejuni* Infektionen im Bezirk Graz-Stadt. Bezogen auf die Einwohnerzahl liefert der Bezirk Mürzzuschlag mit 15 Infektionen je 100 000 EW den Spitzenwert. Im Bezirk Liezen wurden hingegen keine Fälle von *C. jejuni* Infektionen gemeldet.

Die Abbildungen 5.18 (absolut) und 5.19 (je 100 000 EW) zeigen die Anzahl der gemeldeten *C. jejuni* Erkrankungen in den Kalenderwochen 22-25 im Jahr 2011. Dieser Zeitraum beinhaltet KW 22 und 23 mit Spitzenwerten von 26 bzw. 31 Fällen steiermarkweit. Absolut gesehen, lässt sich eine Infektionsquelle in Graz-Stadt vermuten, die evtl. auch Auswirkungen in Graz-Umgebung zeigt. Relativ gesehen, wären mehrere

²Die entsprechende Tabelle ist unter http://www.verwaltung.steiermark.at/cms/dokumente/10004611_97710/c508d90c/Wbinsgesamt-11.pdf abrufbar (Stand 10.7.2012).

5 Detektion lebensmittelbedingter Krankheitsausbrüche

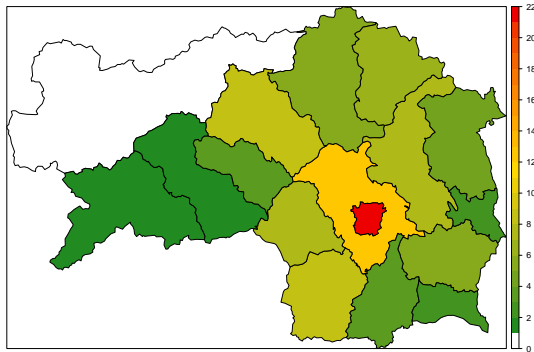


Abbildung 5.16: Anzahl der gemeldeten *C. jejuni* Erkrankungen in den ersten 12 Wochen im Jahr 2011 in der Steiermark nach Bezirk.

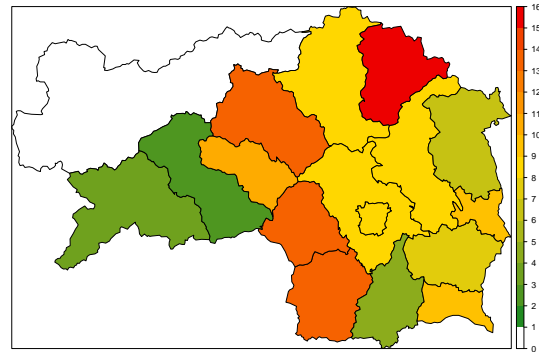


Abbildung 5.17: Anzahl der gemeldeten *C. jejuni* Erkrankungen in den ersten 12 Wochen im Jahr 2011 in der Steiermark nach Bezirk und je 100 000 EW.

Infektionsquellen denkbar. Dazu würde wiederum Graz-Stadt bzw. Graz-Umgebung zählen. Zusätzlich fallen die Bezirke Bruck an der Mur sowie Judenburg (mit Auswirkungen in Knittelfeld und Murau) auf. Auch in diesen Wochen gibt es einen Bezirk, in dem keine Erkrankung gemeldet wurde (Radkersburg).

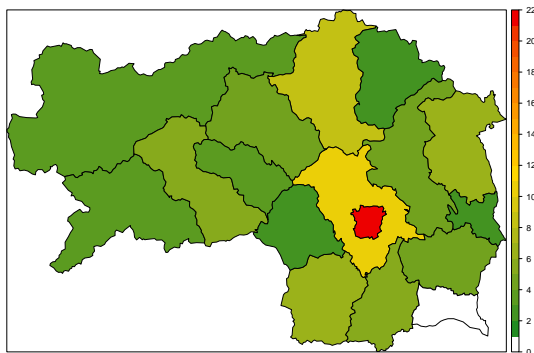


Abbildung 5.18: Anzahl der gemeldeten *C. jejuni* Erkrankungen in KW 22-25 im Jahr 2011 in der Steiermark nach Bezirk.

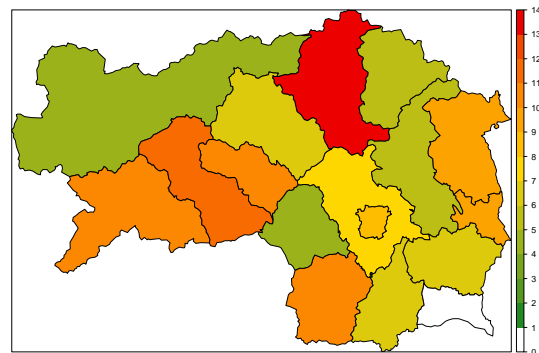


Abbildung 5.19: Anzahl der gemeldeten *C. jejuni* Erkrankungen in KW 22-25 im Jahr 2011 in der Steiermark nach Bezirk und je 100 000 EW.

Da *C. coli* Infektionen wesentlich seltener auftreten als *C. jejuni* Infektionen, deuten hier hohe Fallzahlen in angrenzenden Bezirken noch stärker auf einen örtlichen Zusammenhang hin. Die Kalenderwochen 29-35 decken für das Jahr 2011 alle Sommerwochen, in denen Beobachtungen vorliegen, ab. Aufgrund der Abbildung 5.20 bzw. 5.21 wird ein Ausbruch in den Bezirken Voitsberg/Deutschlandsberg und Bruck a. d. Mur/Liezen vermutet. Die Abbildungen 5.22 und 5.23 zeigen die Anzahl der gemeldeten *C. coli* Erkrankungen für das gesamte Jahr 2011 absolut und je 100 000 Einwohner. Abso-

5.5 Grafische Analyse für die Bezirke der Steiermark

lut gesehen scheinen die Fälle relativ gleichmäßig über die Bezirke verteilt zu sein. Lediglich Graz-Stadt hat eine auffällig hohe Fallzahl. Dies lässt sich einfach mit den hohen Einwohnerzahlen in diesem Bezirk erklären. Aus diesem Grund zeigt Abbildung 5.23 ein klareres Bild. Diese Abbildung lässt Ausbrüche in den Bezirken Voitsberg/Deutschlandsberg sowie Bruck a. d. Mur/Mürzzuschlag (evtl. auch mit Auswirkungen in Liezen und Weiz) vermuten.

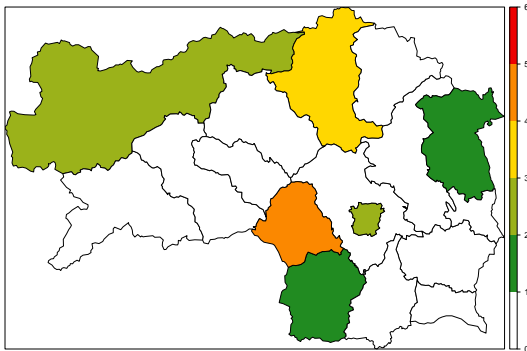


Abbildung 5.20: Anzahl der gemeldeten *C. coli* Erkrankungen in KW 29-35 im Jahr 2011 in der Steiermark nach Bezirk.

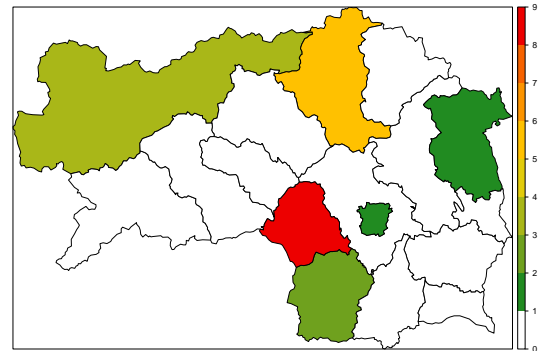


Abbildung 5.21: Anzahl der gemeldeten *C. coli* Erkrankungen in KW 29-35 im Jahr 2011 in der Steiermark nach Bezirk und je 100 000 EW.

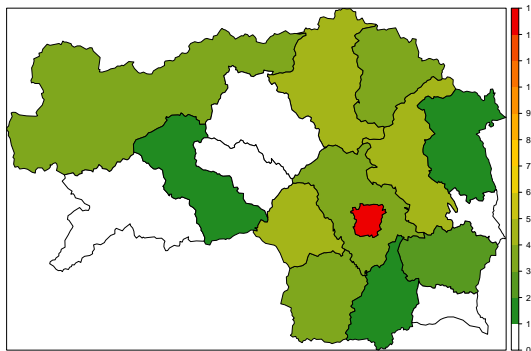


Abbildung 5.22: Anzahl der gemeldeten *C. coli* Erkrankungen im Jahr 2011 in der Steiermark nach Bezirk.

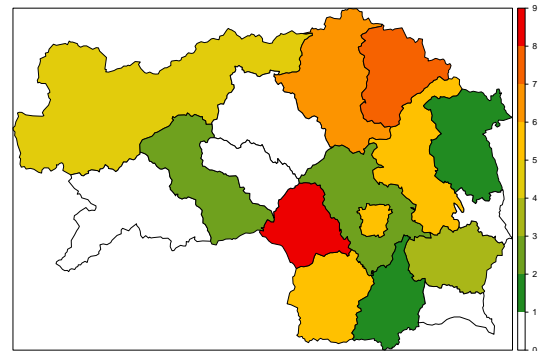


Abbildung 5.23: Anzahl der gemeldeten *C. coli* Erkrankungen im Jahr 2011 in der Steiermark nach Bezirk und je 100 000 EW.

Während in Österreich die Analyse der gemeldeten Campylobacter-Infektionen mittels der vorgestellten Methoden aufgrund der geringen Anzahl auf Bezirksebene nicht zweckmäßig ist, wäre jedoch eine Anwendung auf Regionen denkbar. Alle vorgestellten Methoden sind diesbezüglich durch entsprechende Anpassung der Parameter einfach zu adaptieren. Eine mögliche Zusammenlegung zu Regionen kann aufgrund der Einwohnerzahl erfolgen. Im nachfolgenden Beispiel wird z. B. eine Einteilung in 4 Regionen vorgenommen, wobei der Bezirk Graz-Stadt als eigenständige Region anzusehen ist.

5 Detektion lebensmittelbedingter Krankheitsausbrüche

Beispiel 5.1. Aufgrund der vorhandenen Bezirksinformation wurde die Steiermark in vier Regionen unterteilt.

- *Region 1:* Deutschlandsberg, Feldbach, Leibnitz, Radkersburg und Voitsberg mit insgesamt 280 339 Einwohnern.
- *Region 2:* Bruck a. d. Mur, Judenburg, Knittelfeld, Leoben, Liezen, Mürzzuschlag, Murau mit 349 129 Einwohnern.
- *Region 3:* Fürstenfeld, Graz-Umgebung, Hartberg, Weiz mit 319 606 Einwohnern.
- *Region 4:* Graz-Stadt mit 261 540 Einwohnern.

Abbildung 5.24 zeigt die vier Regionen und die Fallzahlen von *C. jejuni* Infektionen je 100 000 EW für die Kalenderwochen 22-25 im Jahr 2011. Der Abbildung 5.25 ist die Gesamtzahl (je 100 000 EW) der gemeldeten *C. coli* Infektionen in den jeweiligen Regionen für das Jahr 2011 zu entnehmen.

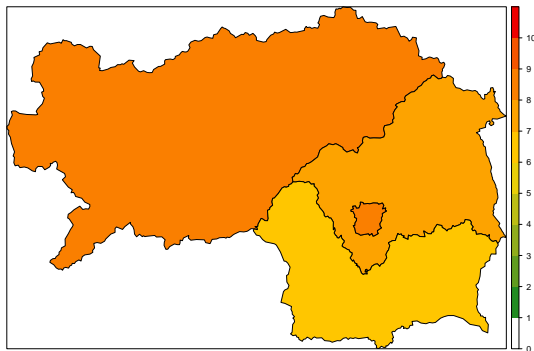


Abbildung 5.24: Anzahl der gemeldeten *C. jejuni* Erkrankungen in KW 22-25 im Jahr 2011 in der Steiermark nach Region und je 100 000 EW.

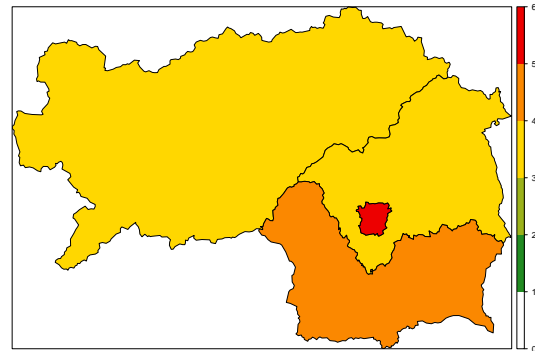


Abbildung 5.25: Anzahl der gemeldeten *C. coli* Erkrankungen im Jahr 2011 in der Steiermark nach Bezirk und je 100 000 EW.

6 Resümee

6.1 Bewertung der Methoden aufgrund der praktischen Anwendung

Alle vorgestellten Methoden bis auf die Adaptierungen nach Rossi et al. (1999) sowie Rogerson und Yamada (2004) eignen sich zur Analyse der Campylobacteriose-Fälle in der Steiermark. Nahezu unabhängig von den verschiedenen Parametereinstellungen werden Wochen mit extrem hohen Fallzahlen gleichermaßen von allen Methoden durch einen Alarm markiert. Die Ergebnisse des GLR-Detektors nach Höhle (2006) bzw. Höhle und Paul (2008) werden stark durch die Modellauswahl zur Modellierung des *in-control* Mittelwerts beeinflusst. Dies führt bei Analyse der Daten für das Jahr 2011 zu einer deutlichen Überschätzung des *in-control* Mittelwerts. Bei Anwendung dieser Methode ist somit die Modellauswahl regelmäßig aufgrund neuer Daten zu hinterfragen. Nach den Ergebnissen der Simulationsstudie könnte der Farrington-Algorithmus für *C. jejuni* Infektionen empfohlen werden. Die Ergebnisse der praktischen Anwendung vermitteln den Eindruck, dass der Bayes-basierte Ansatz mit Parametereinstellungen $b = 2$, $w = 6$, $\alpha = 0.10$ eine realistische Schranke liefert.

Da bei der Analyse von Krankheitsfällen großer Wert auf hohe Sensitivität gelegt wird, dürfte diesbezüglich für *C. coli* Infektionen die Bayes-basierte Methode die besten Ergebnisse erzielen. Hier ist allerdings auch mit einigen Fehlalarmen zu rechnen. Aus statistischer Sicht würde sich auch der Farrington-Algorithmus gut eignen, da mit `limit54` auf die geringen Fallzahlen eingegangen werden und somit Fehlalarmen entgegengewirkt werden kann. Bei entsprechender Wahl eines kleinen Thresholds (um ausreichend hohe Sensitivität zu gewährleisten) und bei ständiger Überprüfung der Modellwahl scheint hier auch der Poisson-GLR-Detektor geeignet.

Eine fundierte Auswahl **einer** Methode ist derzeit aufgrund der praktischen Auswertungen für Campylobacteriose-Fälle in Österreich jedoch nicht möglich. Nach einer Testphase, in der bestenfalls eine Zuordnung von einzelnen Fällen zu einem Krankheitsausbruch erfolgt, könnte aufgrund der Ergebnisse eine Bewertung mit Experten der NRZC durchgeführt werden.

6.2 Bewertung der Realisierbarkeit eines Monitoring-Systems in Österreich

Wie bereits in der Einleitung angedeutet, wird in einigen europäischen Ländern die Ausbruchsfrüherkennung durch Monitoring-Systeme unterstützt.

Als Teil eines schwedischen Projektes über computergestützte Ausbruchsfrüherkennungs-Systeme wurden die verwendeten Systeme bzw. Methoden von fünf europäischen Ländern im Artikel Hulth et al. (2010) erläutert. Weiters wird in Hulth et al. (2010) eine „Checklist for a computer-supported outbreak detection system“ bereit gestellt, um aus Erfahrungen der Umsetzungen in diesen fünf Ländern profitieren zu können. Abschließend sollen nun noch zusammenfassend diese Länder und die benutzten Methoden aufgezählt werden. Für detaillierte Angaben (z. B. zur Datensituation) siehe Hulth et al. (2010).

Das *Statens Serum Institut* (Dänemark) überwacht seit 2001 die Anzahl an positiven Fällen pathogener gastrointestinaler Bakterien. Zur wöchentlichen Kontrolle wird eine SAS Implementation des Algorithmus von Farrington et al. (1996) herangezogen.

Dieser Algorithmus findet bereits seit 1993 beim *Centre for Infections of the Health Protection Agency* (HPA) Anwendung. Damit werden die von klinischen Labors in England, Wales und Nordirland gemeldeten Krankheitsfälle einer automatischen Ausbruchserkennung mittels Algorithmus von Farrington et al. (1996) unterzogen.

Das *National Institute for Public Health and the Environment* (RIVM) in den Niederlanden verwendet für alle meldepflichtigen Krankheiten, mit Ausnahme von Salmonellose- und Campylobacteriose-Erkrankungen, die Methode von Stroup et al. (1993). Für Salmonellose-Erkrankungen wird ebenfalls die Methode von Farrington et al. (1996) bemüht.

Das Programmpaket *CASE Computed Assisted Search for Epidemics* wurde vom schwedischen *Institute for Infectious Disease Control* entwickelt. Damit können alle erdenklichen Algorithmen zur Ausbruchsfrüherkennung implementiert und angewendet werden.

Während das Robert-Koch-Institut in Deutschland die Methode von Stroup et al. (1993) mit leichten Adaptionen zur wöchentlichen Analyse von gemeldeten Krankheitsfällen heranzieht, wird vom Niedersächsischen Landesgesundheitsamt (NLGA) das R-package 'surveillance' herangezogen. Neben der Anwendung der Algorithmen nach Stroup et al. (1993) sowie Farrington et al. (1996) wird mittels SaTScan Software¹ auch versucht,

¹SaTScan ist eine freie Software. Nähere Informationen unter www.satscan.org

6.2 Bewertung der Realisierbarkeit eines Monitoring-Systems in Österreich

räumliche Cluster zu identifizieren.

In jedem der vorhin aufgezählten Länder erfolgt eine Diskussion bzw. Beurteilung der berichteten Alarme durch Epidemiologen oder andere Fachexperten (vgl. Hulth et al. 2010).

Für die Bewertung der Realisierbarkeit eines Monitoring-Systems in Österreich ist vor allem die Datensituation entscheidend. Durch die laufende Erfassung im Epidemiologischen Meldesystem (EMS) seit 2009 wäre dieses als zentrale Datenbank denkbar. Hinsichtlich Datums-, Orts- und vor allem Speziesangabe ist die Datenbank zurzeit jedoch noch unvollständig. Vor allem die Spezifizierung der Spezies kann, wie in Kapitel 5 erläutert, dazu beitragen Krankheitsausbrüche nach Definition des ZoonG zu erkennen. Eine Vervollständigung bzgl. *Campylobacter*-Spezies wird somit für ein gut funktionierendes Monitoring-System unerlässlich sein.

Denkbar wäre dann die laufende Analyse der gemeldeten Krankheitsfälle durch wöchentlichen Zugriff auf Daten des EMS. Ein automatisch erstellter wöchentlicher Bericht, der eine Übersicht der gemeldeten Krankheitsfälle und Alarme sowie Bezirkskarten-Plots beinhaltet, soll dann dabei helfen, Krankheitsausbrüche schnellstmöglich zu erkennen.

Zu Beginn wäre es durchaus sinnvoll, mehrere Methoden zur Analyse anzuwenden. Nach einer Testphase könnte mit entsprechenden Experten der NRZC eine Methode für die weitere *on-line* Überwachung entwickelt werden.

Zum jetzigen Zeitpunkt wäre eine laufende Analyse für gemeldete *Campylobacter*-Erkrankungen für die Steiermark realisierbar. Vor allem aufgrund der Tatsache, dass die vorhandenen historischen Daten keine definierten Ausbrüche enthalten, ist eine Testphase mit den drei vorgestellten Methoden zu empfehlen.

Literaturverzeichnis

BASSEVILLE, M. ; NIKIFOROV, I. V.: *Detection of Abrupt Changes: Theory and Application*. Prentice-Hall, Inc., Englewood Cliffs, N.J. (1993)

BENNEYAN, J. C. ; LLOYD, R. C. ; PLSEK, P. E.: Statistical process control as a tool for research and healthcare improvement. *Quality Safety Health Care* 12:458-464. (2003)

BMG, Bundesministerium für Gesundheit: Lebensmittelbedingter Krankheitsausbruch: Handbuch. (2010)

BROOK, D. ; EVANS, D. A.: An approach to the probability distribution of CUSUM run length. *Biometrika* 59(3):539-549. (1972)

CASELLA, G. ; BERGER, R. L.: *Statistical Inference*. 2nd Edition. Duxbury Press, Pacific Grove, California. (2002)

DAVISON, A. C. ; SNELL, E. J.: *Statistical Theory and Modelling*. Chapter: Residuals and Diagnostics, Chapman and Hall, London. (1991)

DNEM, Deutsches Netzwerk evidenzbasierter Medizin e. V.: Glossar zur Evidenzbasierten Medizin. (2011). – <http://www.ebm-netzwerk.de/was-ist-ebm/grundbegriffe/glossar/>, Stand 27.06.2012

EFSA, European Food Safety Authority: *EFSA explain zoonotic diseases: Campylobacter*. EFSA, Parma. (2012). – <http://www.efsa.europa.eu/en/corporate/pub/factsheetcampylobacter.htm> Stand 16.06.2012

FARRINGTON, C. P. ; ANDREWS, N. J. ; BEALE, A. D. ; CATCHPOLE, M. A.: A statistical algorithm for the early detection of outbreaks of infectious disease. *Journal of the Royal Statistical Society. Series A* 159(3):547-563. (1996)

FRIEDL, H.: Skriptum zur Vorlesung Generalisierte Lineare Modelle. (2000)

FRISÉN, M.: Statistical surveillance. Optimality and methods. *International Statistical Review* 71(2):403-434. (2003)

Literaturverzeichnis

- FRISÉN, M.: Optimal sequential surveillance for finance, public health, and other areas. *Statistical Research Unit, Department of Economics, School of Business, Economics and Law, Göteborg University*. (2007). – Research Report.
- FRISÉN, M.: Properties and use of the Shewart method and followers. *Sequential Analysis* 26:171-193. (2007)
- GAN, F. F.: An optimal design of EWMA control charts based on median run length. *Journal of Statistical Computation and Simulation* 45:169-184. (1993)
- HAWKINS, D. M.: Evaluation of average run lengths of cumulative sum charts for an arbitrary data distribution. *Communications in Statistics - Simulation and Computation* 21(4):1001-1020. (1992)
- HAWKINS, D. M. ; OLWELL, D. H.: *Cumulative Sum Charts and Charting for Quality Improvement*. Springer-Verlag, New York. (1998)
- HELD, L.: *Methoden der statistischen Inferenz: Likelihood und Bayes*. Spektrum Akademischer Verlag, Heidelberg. (2008)
- HÖHLE, M.: Poisson regression charts for the monitoring of surveillance time series. *Sonderforschungsbereich 386, Paper 500* (2006). – <http://epub.ub.uni-muenchen.de>
- HÖHLE, M.: surveillance: An R package for the monitoring of infectious diseases. *Computational Statistics* 22(4):571-582. (2007)
- HÖHLE, M.: *Statistical Modelling and Regression Structures - Festschrift in Honour of Ludwig Fahrmeir, pages 377-397*. Chapter: Online Change-Point Detection in Categorical Time Series, Springer-Verlag, Berlin u.a. (2010)
- HÖHLE, M. ; CORREA, T. ; HOFMANN, M. ; LANG, C. ; PAUL, M. ; RIEBLER, A. ; STEINER, S. ; VIRTANEN, M. ; WIMMER, V.: Package 'surveillance'. (2012). – <http://cran.r-project.org/web/packages/surveillance/index.html>
- HÖHLE, M. ; MAZICK, A.: *Biosurveillance: Methods and Case Studies*. Chapter: Aberration Detection in R Illustrated by Danish Mortality Monitoring, Chapman & Hall/CRC, Boca Raton. (2010)
- HÖHLE, M. ; PAUL, M.: Count data regression charts for the monitoring of surveillance time series. *Computational Statistics and Data Analysis* 52:4357-4368. (2008)
- HÖHLE, M. ; PAUL, M. ; HELD, L.: Statistical approaches to the surveillance of infectious diseases for veterinary public health. *Technical Report Number 014* (2007)
- HÖHLE, M. ; RIEBLER, A.: The R-Package 'surveillance'. *Discussion Paper 422, SFB 386* (2005)

- HULTH, A. ; ANDREWS, N. ; ETHELBERG, S. ; DREESMAN, J. ; FAENSEN, D. ; PELT, W van ; SCHNITZLER, J.: Practical usage of computer-supported outbreak detection in five european countries. *Euro Surveillance* 15(36). (2010). – <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=19658>
- LAI, T. ; SHAN, J.: Efficient recursive algorithms for detection of abrupt changes in signals and control systems. *IEEE Transactions on Automatic Control* 44(5):952-966. (1999)
- LAI, T. L.: Sequential changepoint detection in quality control and dynamical systems. *Journal of the Royal Statistical Society. Series B* 57(4):613-658. (1995). – <http://www.jstor.org/stable/2345934>
- LAWLESS, J. F.: Negative binomial and mixed Poisson regression. *Canadian Journal of Statistics* 15(3):209-225. (1987). – <http://dx.doi.org/10.2307/3314912>. – ISSN 1708-945X
- LORDEN, G.: Procedures for reacting to a change in distribution. *The Annals of Mathematical Statistics* 42(6):1897-1908. (1971)
- LUCAS, J. M. ; CROSIER, R. B.: Fast initial response for CUSUM quality-control schemes: Give your CUSUM a head start. *Technometrics* 24(3):199-205. (1982). – <http://www.jstor.org/stable/1268679>
- MONTGOMERY, D.C.: *Statistical Quality Control: A Modern Introduction, 6th Edition*. John Wiley & Sons, Inc., Hoboken, N.J. (2009)
- MOUSTAKIDES, G. V.: Optimal stopping times for detecting changes in distributions. *The Annals of Statistics* 14(4):1379-1387. (1986)
- MUCH, P. ; PICHLER, J. ; FRETZ, R. ; ALLERBERGER, F.: Lebensmittelbedingte Krankheitsausbrüche, Österreich 2008. *Bericht der Österreichischen Agentur für Gesundheit und Ernährungssicherheit (AGES)*. (2008)
- PAGE, E. S.: Continuous inspection schemes. *Biometrika* 41(1/2):100-115. (1954)
- PAUL, M. ; HÖHLE, M. ; HELD, L.: Surveillance of Infectious Disease Data using Cumulative Sum Methods. (2007). – Präsentation am Swiss Statistics Meeting 2007. - <http://www.statoo.ch/sst07/presentations/Paul.pdf>
- R DEVELOPMENT CORE TEAM: R: A Language and Environment for Statistical Computing. (2012). – R Foundation for Statistical Computing. - Vienna, Austria. - <http://www.R-project.org/>
- RIEBLER, A.: Empirischer Vergleich von statistischen Methoden zur Ausbruchserkennung bei Surveillance Daten. *Bachelor-Arbeit an der Universität München*. (2004)

Literaturverzeichnis

ROGERSON, P.A. ; YAMADA, I.: Approaches to syndromic surveillance when data consists of small regional counts. *Morbidity and Mortality Weekly Report* 53 pages 79-85. (2004)

ROSSI, G. ; DEL SARTO, S. ; MARCHI, M.: Approximate Poisson CUSUM charts for the monitoring of time series with time-varying mean. *45th Scientific Meeting of the Italian Statistical Society, Contributed Paper* (2010). – <http://homes.stat.unipd.it/mgri/SIS2010/>

ROSSI, G. ; LAMPUGNANI, L. ; MARCHI, M.: An approximate CUSUM procedure for surveillance of health events. *Statistics in Medicine* 18:2111-2122. (1999)

SAKAR, P. ; MEEKER, W. Q.: A Bayesian on-line change detection algorithm with process monitoring applications. *Quality Engineering* 10(3):539-549. (1998)

STROUP, D. F. ; WHARTON, M. ; KAFADAR, K. ; DEAN, A. G.: Evaluation of a method for detecting aberrations in public health surveillance data. *American Journal of Epidemiology* 137(3):373-380. (1993). – <http://aje.oxfordjournals.org/content/137/3/373.abstract>

UNKEL, S. ; FARRINGTON, C. P. ; GARTHWAITE, P. H. ; ROBERTSON, C. ; ANDREWS, N.: Statistical methods for the prospective detection of infectious disease outbreaks: A review. *Journal of the Royal Statistical Society. Series A* 175:49-82. (2012)

VAN DOBBEN DE BRUYN, C. S.: *Cumulative Sum Test: Theory and Practice*. Griffin, London (1968)

WATKINS, R. E. ; EAGLESON, S. ; VEENENDAAL, B. ; WRIGHT, G. ; PLANT, A. J.: Applying CUSUM-based methods for the detection of outbreaks of ross river virus disease in Western Australia. *BMC Medical Informatics and Decision Making* 8(37). (2008)

WILLSKY, A. ; JONES, H.: A generalized likelihood ratio approach to the detection and estimation of jumps in linear systems. *IEEE Transactions on Automatic Control* 21(1):108-112. (1976)

WIMMER, V. ; HÖHLE, M.: The function 'algo.glrnb' in the R-Package 'surveillance'. *Vignette* (2008)

WINKEL, P. ; ZHANG, N. F.: *Statistical Development of Quality in Medicine*. John Wiley & Sons Ltd, N.J. (2007)

Anhang

Für alle praktischen Auswertung sowie für die Simulationsstudie wurde das Statistik-Software Paket GNU R, Version 2.14.2 (64-Bit) verwendet. Der R-Code liegt beim Autor auf und kann unter Umständen auf Anfrage eingesehen werden. Alle Auswertungen wurden mit einem HP Laptop mit Intel(R) Core(TM) i7 CPU Q720 @1.60GHz mit Windows 7 (64-Bit) Betriebssystem und 4 GB Arbeitsspeicher durchgeführt.

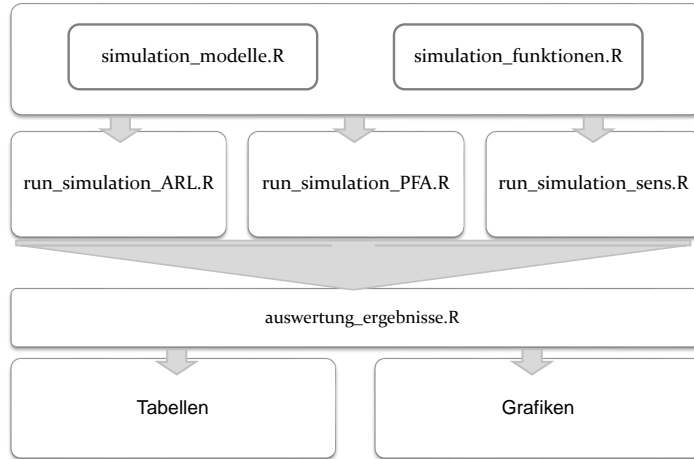
Beschreibung des Datensatzes

Der Datensatz für die praktische Anwendung wurde von der NRZC zur Verfügung gestellt und entstammt deren Labordatenbank. Die nachfolgende Tabelle fasst die Variablen nach der Datenaufbereitung, samt einer kurzen Beschreibung zusammen.

Name	Struktur in R	Beschreibung
ID	character	ID, damit eindeutige Zuordnung der Fälle möglich ist
Labordatum	POSIXlt	Tag, an dem das humane Isolat in der NRZC eingegangen ist
Jahr	numeric	Jahr des Labordatums
Monat	numeric	Monat des Labordatums
Bundesland	character	Bundesland, von dem die Probe eingesandt wurde
Bezirk	character	Bezirk, in dem das zuständige Gesundheitsamt liegt
PLZ	character	PLZ der Wohnadresse des Patienten
Geschlecht	character	Geschlecht des Patienten
Alter	numeric	$[(\text{Labordatum} - \text{Geburtsdatum})/365.25]$
Species	character	Gibt die Species des Campylobacter-Erregers an (andere, unbekannt, jejuni oder coli)

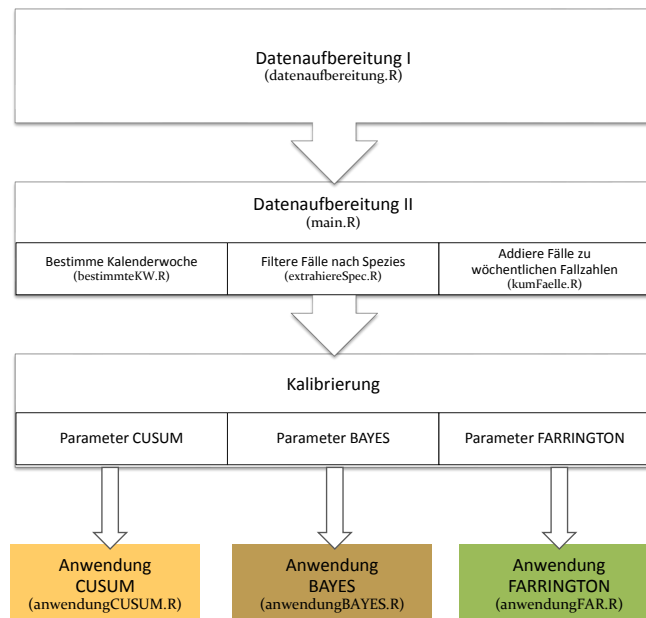
Variablen-Beschreibung des R-Datensatzes

Prozessdiagramm für die Simulationstudie R



Prozessdiagramm der Simulationstudie R.

Prozessdiagramm für die praktischen Auswertungen in R



Prozessdiagramm der praktischen Anwendung in R.