



---

seit 1558

# Die bedingte Power in der Überlebenszeitanalyse

M A S T E R A R B E I T

zur Erlangung des akademischen Grades

Master of Science (M. Sc.)

im Studiengang Mathematik

FRIEDRICH-SCHILLER-UNIVERSITÄT JENA

Fakultät für Mathematik und Informatik

eingereicht von Andreas Kühnapfel

geboren am 8.3.1989 in Oschatz

Betreuer:

Prof. Dr. Markus Scholz

(Universität Leipzig)

Prof. Dr. Michael Neumann

(Friedrich-Schiller-Universität Jena)

Jena, 27.9.2013

# Zusammenfassung

Im Rahmen von Zwischenauswertungen bei randomisierten klinischen Studien mit der Überlebenszeit als Zielgröße stellt die bedingte Power ein mögliches Mittel dar, um zu beurteilen, ob eine Studie fortgeführt oder vorzeitig abgebrochen werden sollte. Auch wenn noch weitere Faktoren in solch eine Entscheidung mit eingehen, ist die bedingte Power ein wichtiger Indikator bezüglich dieser Fragestellung.

Im ersten Kapitel werden die Ausgangssituation und deren Problematik erläutert und grundlegende Begriffe wie Überlebensfunktion, Hazardfunktion und Hazard-Ratio eingeführt.

Der zweite Abschnitt beschreibt zunächst das Exponentialmodell mit konstanten Hazards. Aufbauend auf diesem parametrischen Modell zur Überlebenszeitanalyse werden Maximum-Likelihood-Schätzer bei zensierten Beobachtungen sowie die Power-Funktion im unbedingten und bedingten Fall bezüglich der Fragestellung „Gibt es einen Unterschied zwischen den Hazards oder nicht?“ hergeleitet.

Das nächste Kapitel befasst sich mit Überlebensfunktionen mit Plateau und dementsprechend nicht-konstanten Hazards. Das Hauptaugenmerk wird dabei auf sogenannte Non-Mixture-Modelle gelegt. Unter der Annahme, dass die Ereigniszeiten einer Exponential-, Weibull- beziehungsweise Gammaverteilung mit Überlebensfraktion folgen, werden auch hier Schätzer nach der Maximum-Likelihood-Methode unter Zensierung berechnet und im Fall des Non-Mixture-Exponentialmodells die bedingte Power-Funktion hergeleitet.

In Kapitel 4 erfolgt nun explizit die Anwendung der hergeleiteten Theorie auf einen zufällig generierten Datensatz. Dabei werden die bisher behandelten Verfahren in R implementiert.

Im fünften Abschnitt dieser Arbeit werden die vier hier behandelten Modelle auf Daten aus zwei randomisierten klinischen Studien bezüglich der Behandlung hochmaligner Non-Hodgkin-Lymphome aus den Zeiträumen 2000 bis 2003 und 2003 bis 2009 angewendet und bedingte Power-Analysen durchgeführt.

Schließlich wird im sechsten und letzten Kapitel das R-Paket CP vorgestellt.

# Inhaltsverzeichnis

<b>Zusammenfassung</b>	<b>2</b>
<b>Bezeichnungen</b>	<b>5</b>
<b>Vorwort</b>	<b>6</b>
<b>1 Grundlagen</b>	<b>7</b>
1.1 Motivation . . . . .	7
1.2 Definitionen . . . . .	8
<b>2 Überlebenszeitmodell mit konstanten Hazards: Exponentialmodell</b>	<b>12</b>
2.1 Herleitung der Power-Funktion . . . . .	13
2.1.1 Maximum-Likelihood-Schätzung . . . . .	14
2.1.2 Teststatistik . . . . .	17
2.1.3 Power-Funktion . . . . .	22
2.2 Herleitung der bedingten Power-Funktion . . . . .	22
2.2.1 Bedingte Teststatistik . . . . .	24
2.2.2 Bedingte Power-Funktion . . . . .	25
<b>3 Überlebenszeitmodelle mit nicht-konstanten Hazards: Überlebensfunktionen mit Plateau</b>	<b>27</b>
3.1 Mixture-Modell . . . . .	27
3.1.1 Allgemein . . . . .	28
3.1.2 Exponentialverteilung . . . . .	29
3.2 Non-Mixture-Modell . . . . .	31
3.2.1 Allgemein . . . . .	32
3.2.2 Exponentialverteilung . . . . .	33
3.2.3 Alternative: Weibull-Verteilung . . . . .	41
3.2.4 Alternative: Gammaverteilung . . . . .	45
3.2.5 Unterschiede zwischen den Verteilungen und deren Einfluss auf die bedingte Power . . . . .	48
<b>4 Praktische Anwendung</b>	<b>50</b>
4.1 Berechnung der zukünftigen Personenzeiten . . . . .	50
4.1.1 Exponentialmodell . . . . .	50
4.1.2 Non-Mixture-Modell . . . . .	54
4.2 Sensitivitätsanalyse der bedingte Power-Funktionen . . . . .	58
<b>5 Futility-Analyse am Beispiel zweier klinischer Studien</b>	<b>64</b>
5.1 Inhalt der klinischen Studien . . . . .	64
5.2 Futility-Analyse . . . . .	65
<b>6 R-Paket CP</b>	<b>74</b>

6.1	Paket CP . . . . .	74
6.2	Funktion ConPwrExp . . . . .	76
6.3	Funktion ConPwrNonMixExp . . . . .	78
6.4	Funktion ConPwrNonMixWB . . . . .	80
6.5	Funktion ConPwrNonMixGamma . . . . .	82
6.6	Funktion CompSurvMod . . . . .	84
6.7	Funktion ConPwrExpAndersen . . . . .	86
6.8	Funktion GenerateDataFrame . . . . .	88
	<b>Literaturverzeichnis</b>	<b>89</b>
	<b>Diskussion</b>	<b>91</b>
	<b>Selbstständigkeitserklärung</b>	<b>93</b>

## Bezeichnungen

$\Omega$	Basismenge
$\mathcal{A}$	Sigma-Algebra über $\Omega$
$P$	Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{A})$
$T$	zufällige Zeit bis zum Eintreten des Ereignisses
$\mathcal{B}([0, \infty))$	Borel-Sigma-Algebra über $[0, \infty)$
$F$	kumulative Verteilungsfunktion von $T$
$f$	Wahrscheinlichkeitsdichtefunktion von $T$
$S$	Überlebensfunktion von $T$
$h$	Hazardfunktion von $T$
$r$	Hazard-Ratio der Hazardfunktionen $h_2$ und $h_1$ ( $h_2/h_1$ )
$\theta$	zu schätzender Parameter (in der Regel Hazard-Ratio)
$\alpha$	Signifikanzniveau; Wahrscheinlichkeit eines Fehlers erster Art
$\beta$	Wahrscheinlichkeit eines Fehlers zweiter Art
$\mathbb{N}$	natürliche Zahlen; $\{1, 2, \dots\}$
$\hat{\theta}$	Schätzer für $\theta$
$C$	zufällige Zeit bis zur Zensierung
$Y$	$\min\{T, C\}$
$\Delta$	$\mathbb{1}(T \leq C)$
$l$	logarithmierte Likelihood-Funktion
$W$	Teststatistik
$\gamma(\cdot)$	Power-Funktion
$\Phi$	kumulative Verteilungsfunktion der Standardnormalverteilung
$q_\alpha$	$\alpha$ -Quantil der Standardnormalverteilung
$W^*$	bedingte Teststatistik
$\gamma_c$	bedingte Power-Funktion
$c$	Niveau des Plateaus
$\gamma(\cdot, \cdot)$	unvollständige Gammafunktion der oberen Grenze
$\Gamma$	Gammafunktion
$\Gamma^{(0)}$	regularisierte unvollständige Gammafunktion der oberen Grenze

## Vorwort

Prof. Dr. Markus Scholz möchte ich für den Themenvorschlag und die Möglichkeit danken, die vorliegende Masterarbeit am Institut für Medizinische Informatik, Statistik und Epidemiologie (IMISE) der Universität Leipzig anzufertigen. Dabei halfen mir unsere Gespräche sehr, in denen wir auch über die pure Mathematik hinausschauten und die Anwendung in der Medizin und sich ergebende Probleme betrachteten.

Danken möchte ich weiter Prof. Dr. Michael Neumann, der mir schon während meines bisherigen Studiums an der Friedrich-Schiller-Universität in Jena für Gespräche stets offen gegenüberstand. Die Konsultationen waren mir vor allem hinsichtlich der mathematischen Korrektheit eine große Hilfe.

Schließlich bedanke ich mich bei der Deutschen Studiengruppe Hochmaligne Non-Hodgkin-Lymphome (DSHNHL) unter der Leitung von Prof. Dr. Michael Pfreundschuh für die Bereitstellung der Daten aus der High-CHOEP-Studie aus dem Zeitraum 2000 bis 2003 und der R-Mega-CHOEP-Studie aus dem Zeitraum 2003 bis 2009.

# 1 Grundlagen

Das erste Kapitel gibt zunächst eine Einführung in die Thematik.

Dabei soll im ersten Abschnitt auf den Ursprung der hier behandelten Problematik eingegangen werden, ohne den Fokus zu sehr auf die Mathematik zu legen.

Dies wird sich im zweiten Teil dieses Kapitels ändern. Hier werden die benötigten wahrscheinlichkeitstheoretischen Größen eingeführt und wichtige Begriffe erläutert.

## 1.1 Motivation

Das Ziel klinischer Studien ist es im einfachsten Fall, zu überprüfen, ob eine neue Therapie einer bisher verwendeten Standardtherapie überlegen ist. Hier kann mit Therapie auch ein Medikament gemeint sein.

Dabei wird zum Beispiel die Überlebenszeit als Zielgröße betrachtet. Statt dem Tod im Allgemeinen kann aber auch ein anderes Ereignis hinsichtlich der Überlebenszeit gemeint sein. Das sind zum Beispiel der Tod durch eine bestimmte Krankheit, das Nachlassen von Krankheitssymptomen - auch als „Remission“ bezeichnet - oder das Wiederauftreten einer Krankheit, was man als „Rezidiv“ bezeichnet. Somit ist nicht zwangsläufig vom Tod im gewöhnlichen Sinne die Rede, wenn man das kritische Ereignis als „Tod“ bezeichnet (Weiß (2010), 7.3.1 Einige wichtige Begriffe, Seiten 146/147).

Wie genau das Studiendesign angelegt wird, hängt dabei von der jeweiligen Fragestellung ab (Weiß (2010), 12.3 Studiendesigns, Seiten 243-247).

Hier sei angenommen, dass es zwei Gruppen gibt, zu denen die Patienten zufällig zugeordnet werden, was man auch als „randomisierte“ klinische Studie bezeichnet. Damit soll gewährleistet werden, dass sowohl bekannte als auch nicht bekannte Einflussfaktoren eliminiert werden. Man erreicht damit eine gewisse Strukturgleichheit (Weiß (2010), 15.1.3 Randomisation, Seite 307).

Nun legt man in Zusammenarbeit mit medizinischem Fachpersonal fest, was für ein Effekt als klinisch relevant angesehen wird. Würde dies nicht auf diesem Wege geschehen, könnte durch die Studie ein signifikantes Ergebnis erzielt werden, wobei aber der zugrundeliegende erwartete Effekt medizinisch nicht bedeutsam wäre. Damit könnte die Studie an Aussagekraft beziehungsweise Nutzen verlieren. Was hierbei konkret mit „Effekt“ gemeint ist, wird sich in dem folgenden Kapitel ergeben.

Anschließend gibt man das Signifikanzniveau und die Power vor. Die Power gibt dabei die Wahrscheinlichkeit dafür an, ein signifikantes Studienergebnis zu erhalten, wenn der unbekannte wahre Effekt gleich dem erwarteten Effekt ist.

Mittels dieser dreier Größen kann dann im Vorfeld die benötigte Fallzahl für die Studie ermittelt werden (Weiß (2010), 9.1.4 Der Stichprobenumfang, Seite 183). Einerseits ist

man aufgrund ethischer Aspekte an einer Bestimmung der minimalen Fallzahl interessiert. Auch würden bei einem zu großen Stichprobenumfang Ressourcen verschwendet werden. Allerdings darf die Fallzahl auch nicht zu klein gewählt werden, da sonst die Möglichkeit bestünde, dass es von vornherein unmöglich ist im Rahmen der Studie ein signifikantes Ergebnis zu erzielen.

Da klinische Studien oft auf längere Zeiträume ausgelegt und zudem aufwendig und kostenintensiv sind, finden häufig Zwischenauswertungen statt. Hier wird der tatsächliche Effekt aus den bisher erhobenen Daten berechnet. Sollte dabei festgestellt werden, dass eine Therapie der anderen überlegen ist, wird die Studie aus ethischen Gründen abgebrochen. Falls der tatsächliche Effekt auf Grundlage der bisherigen Daten jedoch geringer ist als der erwartete Effekt, kann man mit verschiedenen Hilfsmitteln untersuchen, ob die Studie vorfristig abgebrochen werden sollte. Dieses Vorgehen bezeichnet man auch als „Futility-Analyse“ (deutsch: Vergeblichkeit). Eines dieser Hilfsmittel ist die bedingte Power. Sie beschreibt - wie die Power - die Wahrscheinlichkeit, ein signifikantes Ergebnis der Studie zu erhalten, wenn der tatsächliche Effekt gleich dem erwarteten Effekt ist, allerdings bedingt auf die bisher erhobenen Daten. Wenn diese Wahrscheinlichkeit sehr klein ist, kann es aufgrund der zu Beginn dieses Absatzes genannten Umstände sinnvoll sein, die Studie abzubrechen. Dabei sei aber auch angemerkt, dass die bedingte Power eben eine statistische Kenngröße ist und nicht allein darauf beruhend Entscheidungen getroffen werden. Vielmehr müssen auch weitere Aspekte aus Medizin, Ethik und Wirtschaftlichkeit betrachtet werden, welche aber nicht Bestandteil dieser Arbeit sein werden.

Ziel soll es hier sein, ausgehend von der Klasse von Überlebensfunktionen mit konstanten Hazards, eine Methode zur Berechnung der bedingten Power für Klassen von Überlebensfunktionen mit nicht-konstanten Hazards herzuleiten.

Als Erstes auf diesem Weg werden nun grundlegende Begriffe ein- und Definitionen aufgeführt.

## 1.2 Definitionen

Im Folgenden sei  $(\Omega, \mathcal{A}, P)$  der zugrundeliegende Wahrscheinlichkeitsraum.

$T: (\Omega, \mathcal{A}) \rightarrow ([0, \infty), \mathcal{B}([0, \infty)))$  sei eine stetige, nicht-negative Zufallsvariable, wobei man mit  $\mathcal{B}([0, \infty))$  die Borel-Sigma-Algebra über den positiven reellen Zahlen einschließlich der Null bezeichnet. Hier kann  $\Omega = [0, \infty)$  und  $\mathcal{A} = \mathcal{B}([0, \infty))$  gewählt werden.  $T$  soll dabei den zufälligen Zeitpunkt des Eintretens eines Ereignisses wie Tod, Remission oder Rezidiv repräsentieren, was zu Beginn dieses Kapitels bereits erwähnt wurde.

Weiter bezeichnen  $F$  und  $f$  die zu  $T$  gehörende kumulative Verteilungsfunktion beziehungsweise Wahrscheinlichkeitsdichtefunktion. Dabei gilt:

$$F(t) = P(T \leq t)$$

und

$$f(t) = F'(t),$$

jeweils für alle  $t \geq 0$ .

Eine zentrale Funktion im Rahmen dieser Arbeit stellt nun folgende dar:

**Definition 1** (Überlebensfunktion).  $S: [0, \infty) \rightarrow [0, 1]$  bezeichnet die Überlebensfunktion bezüglich  $T$ , welche gemäß

$$S(t) = P(T > t)$$

definiert ist.

Die Überlebensfunktion zu einem bestimmten Zeitpunkt gibt also einfach die Wahrscheinlichkeit dafür an, bis zu diesem bestimmten Zeitpunkt nicht zu sterben. Hierbei gilt:

$$S(t) = 1 - F(t)$$

für alle  $t \geq 0$ . Die Bezeichnung „ $S$ “ stammt dabei von dem englischen Wort „survival“ (deutsch: Überleben).

**Definition 2** (Hazardfunktion). Die Hazardfunktion  $h: [0, \infty) \rightarrow [0, \infty)$  bezüglich  $T$  wird definiert durch

$$h(t) = \lim_{s \downarrow 0} \frac{P(t < T \leq t + s \mid T > t)}{s}.$$

Die Existenz des Grenzwertes folgt hier wegen der Differenzierbarkeit von  $F$ . Dies wird bei der äquivalenten Darstellung der Hazardfunktion (1) in diesem Abschnitt (Seite 10) deutlich.

Dabei kann die Hazardfunktion in folgendem Sinne interpretiert werden:  $s h(t)$  gibt annähernd die Wahrscheinlichkeit an, im Zeitraum  $(t, t + s]$  zu sterben, gegeben das Überleben

bis zur Zeit  $t$ , mit  $t \geq 0$  und  $s > 0$  hinreichend klein.

Als alternative Bezeichnungen für die Hazardfunktion werden auch die Begriffe „Hazardrate“ beziehungsweise „Hazard“ verwendet.

Nun kann man  $h$  auch folgendermaßen darstellen:

$$h(t) = \frac{f(t)}{S(t)}$$

für alle  $t \geq 0$  mit  $S(t) > 0$ .

Dies ergibt sich durch eine kurze Rechnung unter Verwendung der Definition der bedingten Wahrscheinlichkeit, der Definition von  $S$  und wegen  $F' = f$ :

$$\begin{aligned} \frac{P(t < T \leq t + s \mid T > t)}{s} &= \frac{P(t < T \leq t + s)}{P(T > t) s} \\ &= \frac{1}{S(t)} \frac{F(t + s) - F(t)}{(t + s) - t} \\ &\rightarrow \frac{1}{S(t)} f(t) \end{aligned} \tag{1}$$

für  $s \downarrow 0$  und für alle  $t \geq 0$ .

Damit gilt aber nun:

$$h(t) = \frac{f(t)}{S(t)} = \frac{F'(t)}{1 - F(t)} = -(\log S)'(t)$$

und man erhält mit  $S(0) = 1$ :

$$S(t) = \exp \left( - \int_0^t h(s) ds \right) \tag{2}$$

jeweils für alle  $t \geq 0$ .

Weiter setzt man:

**Definition 3** (Kumulative Hazardfunktion). *Zu einer Hazardfunktion  $h: [0, \infty) \rightarrow [0, \infty)$  definiert man die kumulative Hazardfunktion  $H: [0, \infty) \rightarrow [0, \infty)$  gemäß*

$$H(t) = \int_0^t h(s) ds.$$

Schließlich wird noch folgende Größe benötigt:

**Definition 4** (Hazard-Ratio). *Seien  $h_1, h_2: [0, \infty) \rightarrow [0, \infty)$  zwei Hazardfunktionen.*

*Dann definiert man das Hazard-Ratio  $r$  von  $h_1$  und  $h_2$  durch*

$$r(t) = \frac{h_2(t)}{h_1(t)}$$

*für alle  $t \geq 0$  mit  $h_1(t) > 0$ .*

Damit sind nun die Grundsteine für die weiteren Betrachtungen gelegt.

## 2 Überlebenszeitmodell mit konstanten Hazards: Exponentialmodell

In diesem Abschnitt der Arbeit wird ein einfaches Verfahren zur Berechnung der bedingten Power vorgestellt. Die Grundlage dafür bildet eine Arbeit von Per Kragh Andersen aus dem Jahr 1987.

Die Ausgangssituation ist dabei folgende: Verglichen werden zwei verschiedene Therapien - eine Standardtherapie (Index 1) und eine Alternativtherapie (Index 2). Die Zielgröße ist dabei die Überlebenszeit.

Man nimmt an, dass diese bei beiden Therapien exponentialverteilt ist, mit den Raten  $\lambda_1 > 0$  und  $\lambda_2 > 0$ . Also hat man:

$$S_i(t) = \exp(-\lambda_i t)$$

und

$$f_i(t) = \lambda_i \exp(-\lambda_i t),$$

jeweils für alle  $t \geq 0$  und  $i \in \{1,2\}$ . Die Hazardfunktionen sind in diesem Fall konstant, denn mit der alternativen Darstellung der Hazardfunktionen gilt:

$$h_i(t) = \frac{f_i(t)}{S_i(t)} = \frac{\lambda_i \exp(-\lambda_i t)}{\exp(-\lambda_i t)} = \lambda_i$$

für alle  $t \geq 0$  und  $i \in \{1,2\}$ . Hierbei wird die sogenannte „Nichtalterungseigenschaft“ der Exponentialverteilung deutlich (Weiß (2010), 7.3.2 Die Exponentialverteilung, Seite 149).

Als zu untersuchenden Parameter wählt man das Hazard-Ratio:

$$r(t) = \frac{h_2(t)}{h_1(t)} = \frac{\lambda_2}{\lambda_1}$$

für alle  $t \geq 0$ , setzt

$$\theta = \frac{\lambda_2}{\lambda_1} > 0$$

und betrachtet die Hypothesen

$$H_0: \theta = 1 \quad \text{versus} \quad H_1: \theta \neq 1,$$

das heißt, man betrachtet die Fragestellung: Gibt es einen Unterschied zwischen den Hazards oder nicht?

Sei  $\alpha \in (0,1)$  das im Vorhinein festgelegte Signifikanzniveau, das heißt die Wahrscheinlichkeit, dass man sich für die Alternative  $H_1$  entscheidet, obwohl  $H_0$  gilt. Der übliche Wert ist dabei  $\alpha = 0,05$ .

Weiter wählt man für den Fehler zweiter Art  $\beta \in (0,1)$  und erhält für die Power  $1 - \beta \in (0,1)$ . Die Power gibt in Abhängigkeit vom zu interessierenden Parameter die Wahrscheinlichkeit dafür an, dass man einen vorhandenen Unterschied erkennt. Eine gebräuchliche Wahl ist hier  $1 - \beta = 0,8$ .

Die dritte Größe, die man benötigt, ist die Effektgröße. Diese sollte - wie in Abschnitt 1.1 erwähnt - mindestens so groß wie der kleinste klinisch relevante Effekt sein. Dabei ist hier mit Effekt das Verhältnis der Hazards gemeint, also das Hazard-Ratio  $\theta = \lambda_2/\lambda_1$ .

Damit lässt sich nun a priori der optimale Stichprobenumfang  $n \in \mathbb{N} = \{1,2,\dots\}$  berechnen. Auf dessen Bestimmung soll allerdings nicht weiter eingegangen werden.

Auch muss bei den folgenden statistischen Untersuchungen immer berücksichtigt werden, dass man es sehr wahrscheinlich mit zensierten Daten zu tun hat. Das heißt, bei Patienten tritt das zu interessierende Ereignis während der Studie nicht ein. Es handelt sich dann um sogenannte „rechtszensierte“ Daten, was der in der Praxis am häufigsten auftretende Fall der Zensierung ist.

Außerdem kann es vorkommen, dass Patienten die Studie vorzeitig abbrechen, was man auch als „Drop-Out“ (deutsch: Ausfall) bezeichnet.

Würde man diese Personen bei der statistischen Auswertung einfach weglassen oder als solche mit Ereignis zählen, könnte dies zu einem Bias und damit zu verfälschten Ergebnissen der Studie führen. Wie man stattdessen mit solchen Daten umgeht, wird im Weiteren erläutert werden.

## 2.1 Herleitung der Power-Funktion

Man befindet sich am Anfang der klinische Studie.

Mit  $D_1$  und  $D_2$  werden im Folgenden die Anzahl der Patienten mit dem Ereignis Tod bei Behandlung mit der Standardtherapie beziehungsweise der Alternativtherapie bezeichnet. Dabei seien  $D_1$  und  $D_2$  voneinander unabhängig. Der Buchstabe „ $D$ “ stammt hierbei von dem englischen Wort „dead“ (deutsch: Tote).

$O_1$  und  $O_2$  seien die Personenzeiten bei der Standardtherapie beziehungsweise bei der Alternativtherapie, also jeweils die Anzahl der Zeiteinheiten (zum Beispiel Monate) pro Patient, aufsummiert über alle Patienten der jeweiligen Gruppe. Auch hier seien  $O_1$  und  $O_2$  voneinander unabhängig. Das „ $O$ “ ist dabei in Anlehnung an den englischen Aus-

druck „observation time“ (deutsch: Beobachtungszeit) gewählt.

Über die Maximum-Likelihood-Methode erhält man als Schätzer  $\widehat{\lambda}_i$  für die Hazards:

$$\widehat{\lambda}_i = \frac{D_i}{O_i}$$

für alle  $i \in \{1,2\}$  und definiert als Schätzer  $\widehat{\theta}$  für das Hazard-Ratio:

$$\widehat{\theta} = \frac{\widehat{\lambda}_2}{\widehat{\lambda}_1}.$$

Dabei sei angenommen, dass  $D_1(\omega) > 0$ ,  $D_2(\omega) > 0$ ,  $O_1(\omega) > 0$  und  $O_2(\omega) > 0$  jeweils für P-fast alle  $\omega \in \Omega$  gelte. Dies setzt dann implizit  $n \geq 2$  voraus.

Dass hier die Maximum-Likelihood-Methode zur Schätzung verwendet werden kann, ist dadurch möglich, da die Verteilung der Überlebenszeit bis auf den zu schätzenden Parameter spezifiziert ist. Aufgrund der Vorteile (Greene (2003), 17 Maximum Likelihood Estimation, Seiten 468, 470-472, 521) gegenüber anderen Schätzverfahren, worauf hier nicht näher eingegangen werden soll, wird diese Methode nun zum Schätzen der Parameter verwendet.

### 2.1.1 Maximum-Likelihood-Schätzung

An dieser Stelle soll etwas genauer auf die Herleitung der Schätzer für die Hazards über die Maximum-Likelihood-Methode (Klein und Moeschberger (2005), 3.5 Likelihood Construction for Censored and Truncated Data, Seiten 74-76) eingegangen werden, da auf die Vorgehensweise im nächsten Kapitel zurückgegriffen wird.

Es seien  $n_1 \in \mathbb{N}$  Personen in Gruppe 1 und  $n_2 \in \mathbb{N}$  Personen in Gruppe 2, mit  $n_1 + n_2 = n$ . Dabei wird jeder Person  $(1, j_1)$  aus Gruppe 1 die Überlebensfunktion  $S_{1, j_1} = S_1$  mit  $j_1 \in \{1, \dots, n_1\}$  und jeder Person  $(2, j_2)$  aus Gruppe 2 die Überlebensfunktion  $S_{2, j_2} = S_2$  mit  $j_2 \in \{1, \dots, n_2\}$  unterstellt, wobei die entsprechenden Überlebenszeiten voneinander unabhängig seien.

Was nun beachtet werden muss, sind Rechtszensierungen. Das heißt es wird nicht die Überlebenszeit  $T_{i, j_i}$  einer Person  $(i, j_i)$  der Gruppe  $i$ , sondern die Überlebenszeit  $T_{i, j_i}$  oder die Zensierungszeit  $C_{i, j_i}$  - je nachdem, was zuerst eintritt - beobachtet, mit  $j_i \in \{1, \dots, n_i\}$  und  $i \in \{1, 2\}$ .

Hier gelte, dass  $C_{i, j_i}: (\Omega, \mathcal{A}) \rightarrow ([0, \infty), \mathcal{B}([0, \infty)))$  stetig, nicht-negativ, unabhängig von  $C_{i, k_i}$  für alle  $k_i \in \{1, \dots, n_i\}$  mit  $k_i \neq j_i$  und unabhängig von  $T_{i, k_i}$  für alle  $k_i \in \{1, \dots, n_i\}$

sei, für alle  $j_i \in \{1, \dots, n_i\}$  und  $i \in \{1, 2\}$ .

Weiter sei die Verteilung von  $C_{i,j_i}$  nicht von den Parametern der Verteilung von  $T_{i,j_i}$  abhängig, für alle  $j_i \in \{1, \dots, n_i\}$  und  $i \in \{1, 2\}$ .

Dann definiert man  $Y_{i,j_i}: (\Omega, \mathcal{A}) \rightarrow ([0, \infty), \mathcal{B}([0, \infty)))$  durch

$$Y_{i,j_i} = \min\{T_{i,j_i}, C_{i,j_i}\}$$

und setzt  $\Delta_{i,j_i}: (\Omega, \mathcal{A}) \rightarrow (\{0, 1\}, \{\emptyset, \{0\}, \{1\}, \{0, 1\}\})$  als

$$\Delta_{i,j_i} = \mathbb{1}(T_{i,j_i} \leq C_{i,j_i}),$$

jeweils für alle  $j_i \in \{1, \dots, n_i\}$  und  $i \in \{1, 2\}$ .

Damit gilt für die kumulative Verteilungsfunktion  $F_{Y_{i,j_i}}$  von  $Y_{i,j_i}$  mit  $F_{C_{i,j_i}}$  als kumulative Verteilungsfunktion von  $C_{i,j_i}$  und  $F_i = 1 - S_i$ :

$$\begin{aligned} F_{Y_{i,j_i}}(t_{i,j_i}, \delta_{i,j_i}) &= P(Y_{i,j_i} \leq t_{i,j_i}, \Delta_{i,j_i} = \delta_{i,j_i}) \\ &= \begin{cases} \int_0^{t_{i,j_i}} [1 - F_{C_{i,j_i}}(s)] dF_i(s), & \text{falls } \delta_{i,j_i} = 1 \\ \int_0^{t_{i,j_i}} [1 - F_i(s)] dF_{C_{i,j_i}}(s), & \text{falls } \delta_{i,j_i} = 0 \end{cases} \end{aligned}$$

für alle  $t_{i,j_i} \geq 0$ ,  $j_i \in \{1, \dots, n_i\}$  und  $i \in \{1, 2\}$ .

Nun erhält man für die Wahrscheinlichkeitsdichtefunktion  $f_{Y_{i,j_i}}$  von  $Y_{i,j_i}$  mit  $f_{C_{i,j_i}}$  als Wahrscheinlichkeitsdichtefunktion von  $C_{i,j_i}$ :

$$\begin{aligned} f_{Y_{i,j_i}}(t_{i,j_i}, \delta_{i,j_i}) &= \begin{cases} [1 - F_{C_{i,j_i}}(t_{i,j_i})] f_i(t_{i,j_i}), & \text{falls } \delta_{i,j_i} = 1 \\ [1 - F_i(t_{i,j_i})] f_{C_{i,j_i}}(t_{i,j_i}), & \text{falls } \delta_{i,j_i} = 0 \end{cases} \\ &= [1 - F_{C_{i,j_i}}(t_{i,j_i})]^{\delta_{i,j_i}} f_{C_{i,j_i}}^{1-\delta_{i,j_i}}(t_{i,j_i}) f_i^{\delta_{i,j_i}}(t_{i,j_i}) [1 - F_i(t_{i,j_i})]^{1-\delta_{i,j_i}} \end{aligned}$$

für alle  $t_{i,j_i} \geq 0$ ,  $j_i \in \{1, \dots, n_i\}$  und  $i \in \{1, 2\}$ , mit der Konvention  $0^0 = 1$ .

Seien der Beobachtungszeitpunkt  $Y_{i,j_i} = t_{i,j_i} \geq 0$  und der Zensierungsindikator  $\Delta_{i,j_i} = \delta_{i,j_i} \in \{0, 1\}$  von Person  $(i, j_i)$  aus Gruppe  $i$  für alle  $j_i \in \{1, \dots, n_i\}$  und  $i \in \{1, 2\}$  gegeben.

Es ergibt sich für die Likelihood-Funktion  $L_i: (0, \infty) \rightarrow (0, \infty)$  (positiv mit Wahrscheinlichkeit Eins) der Gruppe  $i$ :

$$\begin{aligned} L_i(\lambda_i) &= \prod_{j_i=1}^{n_i} f_{Y_{i,j_i}}(t_{i,j_i}) \\ &= a_i \prod_{j_i=1}^{n_i} f_i^{\delta_{i,j_i}}(t_{i,j_i}) S_i^{1-\delta_{i,j_i}}(t_{i,j_i}) \end{aligned}$$

mit

$$a_i = \prod_{j_i=1}^{n_i} [1 - F_{C_{i,j_i}}(t_{i,j_i})]^{\delta_{i,j_i}} f_{C_{i,j_i}}^{1-\delta_{i,j_i}}(t_{i,j_i}) > 0$$

mit Wahrscheinlichkeit Eins, jeweils für alle  $i \in \{1,2\}$ . Diese Möglichkeit der Abspaltung ergibt sich aufgrund der Annahme, dass die Verteilung von  $C_{i,j_i}$  nicht von den Parametern der Verteilung von  $T_{i,j_i}$  abhängt, für alle  $j_i \in \{1, \dots, n_i\}$  und  $i \in \{1,2\}$ .

Mit  $f_i = h_i S_i$  erhält man dann

$$L_i(\lambda_i) = a_i \prod_{j_i=1}^{n_i} h_i^{\delta_{i,j_i}}(t_{i,j_i}) S_i(t_{i,j_i}) \quad (3)$$

für alle  $i \in \{1,2\}$ .

Betrachtet man nun die logarithmierte Likelihood-Funktion  $l_i: (0,\infty) \rightarrow \mathbb{R}$  der Gruppe  $i$ , die konkrete Hazardfunktion  $h_i \equiv \lambda_i$  und die allgemeine Darstellung der Überlebensfunktion (2) aus Abschnitt 1.2 (Seite 10), so erhält man:

$$l_i(\lambda_i) = \log L_i(\lambda_i) = \log a_i + \sum_{j_i=1}^{n_i} (\delta_{i,j_i} \log \lambda_i - \lambda_i t_{i,j_i})$$

für alle  $i \in \{1,2\}$ .

Nach dem Differenzieren der logarithmierten Likelihood-Funktion bezüglich  $\lambda_i$  erhält man dann die Score-Funktion  $l'_i: (0,\infty) \rightarrow \mathbb{R}$  mit

$$l'_i(\lambda_i) = \frac{1}{\lambda_i} \sum_{j_i=1}^{n_i} \delta_{i,j_i} - \sum_{j_i=1}^{n_i} t_{i,j_i}$$

für alle  $i \in \{1,2\}$ .

Durch Nullsetzen und Einsetzen der Zufallsvariablen  $Y_{i,j_i}$  für  $t_{i,j_i}$  und  $\Delta_{i,j_i}$  für  $\delta_{i,j_i}$  für alle  $j_i \in \{1, \dots, n_i\}$  gelangt man schließlich mit

$$D_i = \sum_{j_i=1}^{n_i} \Delta_{i,j_i}$$

und

$$O_i = \sum_{j_i=1}^{n_i} Y_{i,j_i}$$

für alle  $i \in \{1,2\}$  zu den Maximum-Likelihood-Schätzern für die Hazards  $\lambda_1$  und  $\lambda_2$ .

### 2.1.2 Teststatistik

Als Teststatistik  $W$  wählt man nun

$$\begin{aligned} W &= \log \hat{\theta} \\ &= \log \hat{\lambda}_2 - \log \hat{\lambda}_1 \\ &= \log \left( \frac{D_2}{O_2} \right) - \log \left( \frac{D_1}{O_1} \right). \end{aligned}$$

Es wird im Folgenden gezeigt, dass  $W$  für fixiertes  $O_1$  und  $O_2$  approximativ normalverteilt ist mit Erwartungswert  $\log \theta$  und Varianz  $(\lambda_1 O_1)^{-1} + (\lambda_2 O_2)^{-1}$ .

Die Verwendung des Logarithmus hilft dahingehend, dass statt dem Schätzer für das Hazard-Ratio nun die Differenz der Logarithmen der Schätzer für die Hazards betrachtet wird. Da die Schätzer für die Hazards über die Maximum-Likelihood-Methode gewonnen wurden, können bereits bestehende Resultate bezüglich deren Asymptotik verwendet werden.

Dazu seien die Personenzeiten  $O_1$  und  $O_2$  fixiert. Die einzigen Zufallsvariablen, die jetzt noch betrachtet werden müssen, sind die Anzahlen der Todesfälle, also der nichtzensierten Beobachtungen, in jeder der beiden Gruppen, das heißt  $D_1$  und  $D_2$ .

Nun lässt man gedanklich den Stichprobenumfang  $n = n_1 + n_2$  gegen Unendlich gehen, wobei dies für die Fallzahl jeder der beiden Gruppen gelte, also  $n_i = n_i(n) \rightarrow \infty$  für  $n \rightarrow \infty$ , mit  $D_i = D_i(n_i)$  und  $O_i = O_i(n_i)$ , jeweils für alle  $i \in \{1, 2\}$ . Außerdem sei angenommen, dass zwischen den Fallzahlen in den beiden Gruppen die Beziehung

$$\lim_{n \rightarrow \infty} \frac{n_1}{n_2} = \text{const} \in (0, \infty)$$

besteht. Dies ist eine mögliche Annahme, welche für eine gewisse Gleichmäßigkeit zwischen den Gruppen sorgen soll. Dass dies praktisch auch erfüllt ist, muss nicht der Fall sein, kann jedoch durch sogenannte „Rekrutierungen“ gewährleistet werden. Das bedeutet, dass während der Studie eine gewisse Anzahl neuer Patienten aufgenommen und entsprechend in die beiden Gruppen verteilt wird, um die Ausfälle durch Tod oder Drop-Outs zu kompensieren.

An dieser Stelle sei darauf hingewiesen, dass bei den weiteren theoretischen Überlegungen Rekrutierungen nicht betrachtet werden. Dies soll erst später beim Übergang von der Theorie zur praktischen Anwendung erfolgen.

Aufgrund der Darstellung  $D_i = \sum_{j_i=1}^{n_i} \Delta_{i,j_i}$  und der Unabhängigkeit von  $\Delta_{i,j_i}$  für alle  $j_i \in \{1, \dots, n_i\}$  folgt, dass  $D_i$  binomialverteilt ist zu den Parametern  $n_i$  und  $p_i = P_{\lambda_i}(T_{i,j_i} \leq C_{i,j_i})$  für alle  $j_i \in \{1, \dots, n_i\}$  und  $i \in \{1, 2\}$ .

Damit gilt:

$$E_{\lambda_i} D_i = n_i p_i$$

und für hinreichend großes  $n_i$  und fixiertes  $O_i$ :

$$E_{\lambda_i} D_i \approx \lambda_i O_i,$$

was aus der asymptotischen Erwartungstreue der Maximum-Likelihood-Schätzer (Greene (2003), 17.4.5.a Consistency, Seiten 477/478) folgt:

$$E_{\lambda_i} \hat{\lambda}_i = E_{\lambda_i} \frac{D_i}{O_i} \rightarrow \lambda_i$$

für  $n \rightarrow \infty$ , für alle  $i \in \{1, 2\}$ . Auf die notwendigen Regularitätsbedingungen für die asymptotische Erwartungstreue (Greene (2003), Definition 17.3 (Regularity Conditions), 17.4.1 Regularity Conditions, Seite 474) soll jetzt und auch im Weiteren nicht weiter eingegangen werden.

Hier bezieht sich  $P_{\lambda_i}$  auf die Verteilung bei angenommenem Parameter  $\lambda_i$ , wobei dann  $E_{\lambda_i}$  den Erwartungswert bezüglich  $P_{\lambda_i}$  bezeichnet, für alle  $i \in \{1, 2\}$ .

Die asymptotische Normalität folgt jetzt aus der sogenannten „Delta-Methode“.

Dabei bezeichnet  $\mathcal{B}(\mathbb{R})$  die Borel-Sigma-Algebra über den reellen Zahlen,  $\xrightarrow{d}$  die Konvergenz in Verteilung und  $\xrightarrow{P}$  die Konvergenz in Wahrscheinlichkeit.

**Lemma 1** (Delta-Methode; Serfling (1980), Theorem A, 3.3 Functions of Asymptotically Normal Vectors, Seite 122). *Sei  $(X_n)_{n \in \mathbb{N}}$  eine Folge von Zufallsvariablen mit  $X_n: (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$  für alle  $n \in \mathbb{N}$ , sodass gilt:*

$$\sqrt{n} (X_n - \mu) \xrightarrow{d} Z \sim N(0, \sigma^2)$$

mit  $\mu \in \mathbb{R}$  und  $\sigma \geq 0$ . Weiter sei  $g: \mathbb{R} \rightarrow \mathbb{R}$  eine stetig differenzierbare Funktion.

Dann gilt:

$$\sqrt{n} [g(X_n) - g(\mu)] \xrightarrow{d} g'(\mu) Z \sim N(0, [g'(\mu)]^2 \sigma^2).$$

*Beweis.* In der angegebenen Quelle findet sich eine allgemeinere Version der Aussage von Lemma 1. Hier wird aber nur ein Spezialfall benötigt. Entsprechend wird auch nur dieser hier genannt und bewiesen.

Es gilt nach dem Mittelwertsatz der Differentialrechnung (Bartle (2001), 12.4 First Mean Value Theorem, 12. Applications to Calculus, Part 1. Integration on Compact Intervals, Seite 193):

$$g(x) = g(\mu) + g'(\xi)(x - \mu)$$

für alle  $x \in \mathbb{R}$  mit  $\xi = \xi(x, \mu)$  zwischen  $x$  und  $\mu$ .

Wegen der Voraussetzung des Lemmas folgt, dass  $X_n - \mu$  in Verteilung gegen die Konstante Null konvergiert, was wiederum die Konvergenz in Wahrscheinlichkeit von  $X_n - \mu$  gegen Null impliziert. Das sieht man folgendermaßen:

$$\begin{aligned} \mathbb{P}(|X_n - \mu| > \varepsilon) &\leq \mathbb{P}(X_n > \mu + \varepsilon) + \mathbb{P}(X_n < \mu - \varepsilon) \\ &\rightarrow \mathbb{P}(\mu > \mu + \varepsilon) + \mathbb{P}(\mu < \mu - \varepsilon) \\ &= 0 \end{aligned}$$

für  $n \rightarrow \infty$ , für alle  $\varepsilon > 0$ , da  $X_n \xrightarrow{d} \mu$  gilt.

Aus  $|\xi - \mu| \leq |X_n - \mu| \xrightarrow{P} 0$  erhält man mit dem Continuous Mapping Theorem (Billingsley (1999), Theorem 2.7, The Mapping Theorem, Section 2: Properties of Weak Convergence, Seite 21):

$$g'(\xi) \xrightarrow{P} g'(\mu).$$

Damit folgt unter Anwendung des Lemmas von Slutsky (Cramér (1946), 20.6 A convergence theorem, Seiten 254/255) auf die rechte Seite von

$$\sqrt{n} [g(X_n) - g(\mu)] = g'(\xi) \sqrt{n} (X_n - \mu)$$

die Behauptung. □

Zerlegt man also  $\log \hat{\theta}$  in die Differenz von  $\log \hat{\lambda}_2$  und  $\log \hat{\lambda}_1$ , folgt mit  $g = \log$  und der asymptotischen Normalität (Greene (2003), 17.4.5.b Asymptotic Normality, Seiten 478/479) und Unabhängigkeit der Maximum-Likelihood-Schätzer  $\hat{\lambda}_1$  und  $\hat{\lambda}_2$  die asymptotische Normalität der Teststatistik. Auf eine Überprüfung der Regularitätsbedingungen für die asymptotische Normalität der Maximum-Likelihood-Schätzer (Greene (2003), Definition 17.3 (Regularity Conditions), 17.4.1 Regularity Conditions, Seite 474) sei an dieser Stelle und auch im Folgenden verzichtet. Weiter ist zu beachten, dass die Zentrierung mittels  $\log \theta$  unter  $H_0$  entfällt, da hier  $\theta = 1$  gilt.

Zur Berechnung der Varianz von  $W$  schreibt man nun:

$$\begin{aligned} \log \hat{\theta} &= \log \hat{\lambda}_2 - \log \hat{\lambda}_1 \\ &= (\log D_2 - \log O_2) - (\log D_1 - \log O_1) \end{aligned}$$

und

$$\begin{aligned}
\log D_i &= \log \left[ E_{\lambda_i} D_i \left( 1 + \frac{D_i - E_{\lambda_i} D_i}{E_{\lambda_i} D_i} \right) \right] \\
&= \log E_{\lambda_i} D_i + \log \left( 1 + \frac{D_i - E_{\lambda_i} D_i}{E_{\lambda_i} D_i} \right) \\
&\approx \log E_{\lambda_i} D_i + \frac{D_i - E_{\lambda_i} D_i}{E_{\lambda_i} D_i}
\end{aligned} \tag{4}$$

für  $i \in \{1,2\}$ , was aus der Approximation  $\log(1 + \varepsilon) \approx \varepsilon$  mit  $\varepsilon > 0$  hinreichend klein folgt (Freedman (2008), A: Appendix: The delta method in more detail, Seiten 118/119). Dies gilt hier für hinreichend großes  $n_i$ , denn:

$$\begin{aligned}
\frac{D_i - E_{\lambda_i} D_i}{E_{\lambda_i} D_i} &= \frac{\sum_{j_i=1}^{n_i} \Delta_{i,j_i} - \sum_{j_i=1}^{n_i} E_{\lambda_i} \Delta_{i,j_i}}{\sum_{j_i=1}^{n_i} E_{\lambda_i} \Delta_{i,j_i}} \\
&= \frac{1}{E_{\lambda_i} \Delta_{i,1}} \frac{1}{n_i} \sum_{j_i=1}^{n_i} (\Delta_{i,j_i} - E_{\lambda_i} \Delta_{i,j_i})
\end{aligned}$$

und

$$\frac{1}{n_i} \sum_{j_i=1}^{n_i} (\Delta_{i,j_i} - E_{\lambda_i} \Delta_{i,j_i}) \xrightarrow{\text{f. s.}} 0$$

nach dem starken Gesetz der großen Zahlen (Etemadi (1981), Theorem 1, Seiten 119/120), wenn die Zensierungszeiten  $C_{i,j_i}$  für alle  $j_i \in \{1, \dots, n_i\}$  identisch verteilt sind, für alle  $i \in \{1,2\}$ .

Hier bezeichnet  $\xrightarrow{\text{f. s.}}$  die fast sicher-Konvergenz.

Weiter gilt:

$$\text{Var}_{\lambda_i} D_i = n_i p_i (1 - p_i) \approx n_i p_i = E_{\lambda_i} D_i \approx \lambda_i O_i,$$

falls  $p_i$  hinreichend klein,  $n_i$  hinreichend groß und  $O_i$  fixiert sind, für alle  $i \in \{1,2\}$ . Auch hier bezieht sich  $\text{Var}_{\lambda_i}$  auf die Varianz bezüglich  $P_{\lambda_i}$  für alle  $i \in \{1,2\}$ .

**Bemerkung 1.** *Im Fall des hinreichend kleinen  $p_i$  lässt sich die Binomial-Verteilung zu den Parametern  $n_i$  und  $p_i$  für kleines  $n_i p_i^2$  auch durch die Poisson-Verteilung mit dem Parameter  $n_i p_i$  approximieren. Der Approximationsfehler kann dann durch den Totalvariationsabstand  $d$  quantifiziert werden, welcher folgendermaßen definiert ist:*

$$d(P, Q) = \sup_{A \in \mathcal{A}} |P(A) - Q(A)|$$

für zwei Verteilungen  $P$  und  $Q$  auf  $(\Omega, \mathcal{A})$ . Damit erhält man:

$$d(\text{Poisson}(n_i p_i), \text{Binomial}(n_i, p_i)) \leq 2n_i p_i^2$$

(Le Cam (1960), Theorem 2 (1), 5. Approximation of the binomial by a Poisson distribution, Seiten 1194/1195). Gemäß dieser Abschätzung sollte als Bedingung  $n_i p_i^2 \leq 1/2$  gelten.

Damit hat man nun unter Zuhilfenahme von (4) aus diesem Abschnitt (Seite 20),  $\text{Var}_{\lambda_i} D_i \approx \text{E}_{\lambda_i} D_i$  und  $\text{E}_{\lambda_i} D_i \approx \lambda_i O_i$  für hinreichend kleines  $p_i$ , hinreichend großes  $n_i$  und fixiertes  $O_i$ :

$$\text{Var}_{\lambda_i} \log D_i \approx \frac{1}{\text{E}_{\lambda_i} D_i} \approx \frac{1}{\lambda_i O_i}$$

für alle  $i \in \{1,2\}$ .

Dass man bezüglich der Varianz von  $\log D_i$  die Bedingung des hinreichend kleinen  $p_i$ ,  $i \in \{1,2\}$ , sogar komplett wegfällen lassen kann, sieht man folgendermaßen ein. Mit (4) aus diesem Abschnitt (Seite 20) gilt zunächst:

$$\text{Var}_{\lambda_i} \log D_i \approx \frac{\text{Var}_{\lambda_i} D_i}{(\text{E}_{\lambda_i} D_i)^2}$$

für alle  $i \in \{1,2\}$ .

Damit hat man aber:

$$\begin{aligned} \left| \frac{1}{\text{E}_{\lambda_i} D_i} - \frac{\text{Var}_{\lambda_i} D_i}{(\text{E}_{\lambda_i} D_i)^2} \right| &= \left| \frac{\text{E}_{\lambda_i} D_i - \text{Var}_{\lambda_i} D_i}{(\text{E}_{\lambda_i} D_i)^2} \right| \\ &= \left| \frac{n_i p_i - n_i p_i (1 - p_i)}{(n_i p_i)^2} \right| \\ &= \frac{p_i}{n_i p_i} \\ &= \frac{1}{n_i} \end{aligned} \tag{5}$$

für alle  $i \in \{1,2\}$ .

Mit der Unabhängigkeit von  $D_1$  und  $D_2$  folgt schließlich:

$$\text{Var}_{\lambda_1, \lambda_2} \log \hat{\theta} \approx \frac{1}{\text{E}_{\lambda_1} D_1} + \frac{1}{\text{E}_{\lambda_2} D_2} \approx \frac{1}{\lambda_1 O_1} + \frac{1}{\lambda_2 O_2}$$

für hinreichend große  $n_1, n_2$  und fixierte  $O_1, O_2$ .

Zusammenfassend erhält man:

**Satz 1** (Konvergenz in Verteilung der Teststatistik; Andersen (1987), Results, Seiten 68/69). *Unter den Annahmen aus Abschnitt 2.1 gilt für hinreichend großes  $n$  und fixierte  $O_1, O_2$ :*

$$\frac{W - \log \theta}{\sqrt{\frac{1}{\lambda_1 O_1} + \frac{1}{\lambda_2 O_2}}} \xrightarrow{d} N(0,1).$$

**Bemerkung 2.** *Bei Satz 1 wurde bei der Konvergenz die verkürzte Schreibweise genutzt, das heißt ohne Zufallsvariable auf der rechten Seite, da diese im Folgenden nicht weiter benötigt wird. Diese Konvention soll auch im Weiteren beibehalten werden.*

### 2.1.3 Power-Funktion

Mit  $\Phi$  als kumulative Verteilungsfunktion der Standardnormalverteilung und  $q_\alpha = \Phi^{-1}(\alpha)$  als  $\alpha$ -Quantil der Standardnormalverteilung mit  $\alpha \in (0,1)$  kann man nun die Power-Funktion  $\gamma: (0,\infty) \rightarrow [0,1]$  folgendermaßen definieren (Andersen (1987), Results, Seite 69):

$$\begin{aligned} \gamma(\theta) &= P_{\lambda_1, \lambda_2} \left( |W| \geq q_{1-\frac{\alpha}{2}} \sqrt{\text{Var}_{\lambda_1, \lambda_1} W} \right) \\ &\approx \Phi \left( \frac{q_{\frac{\alpha}{2}} \sqrt{\frac{1}{\lambda_1 O_1} + \frac{1}{\lambda_1 O_2}} - \log \theta}{\sqrt{\frac{1}{\lambda_1 O_1} + \frac{1}{\lambda_2 O_2}}} \right) + \left[ 1 - \Phi \left( \frac{q_{1-\frac{\alpha}{2}} \sqrt{\frac{1}{\lambda_1 O_1} + \frac{1}{\lambda_1 O_2}} - \log \theta}{\sqrt{\frac{1}{\lambda_1 O_1} + \frac{1}{\lambda_2 O_2}}} \right) \right] \end{aligned}$$

unter den Annahmen aus Abschnitt 2.1 für hinreichend großes  $n$  und fixierte  $O_1, O_2$ . Dabei bezieht sich  $\text{Var}_{\lambda_1, \lambda_1}$  auf die Varianz unter  $H_0$  mit  $\theta = 1$ , das heißt unter  $\lambda_1 = \lambda_2$ .

## 2.2 Herleitung der bedingten Power-Funktion

Man stellt sich nun vor, man befindet sich an einem beliebigen Zeitpunkt  $t^* > 0$  innerhalb der Studie.

Es wird eine Zwischenauswertung durchgeführt, in der der Effekt aus den bisher erhobenen Daten berechnet wird. Dann gibt es zwei Möglichkeiten: Entweder es liegt ein signifikanter Unterschied zwischen den Hazards vor oder nicht. Sollte Letzteres der Fall sein und entsprechend der beobachtete Effekt wesentlich geringer als der erwartete Effekt ausfallen, so kann man untersuchen, wie wahrscheinlich es ist, dass man zum geplanten

Ende der Studie noch ein signifikantes Ergebnis erhält. Und genau dies führt nun zur bedingten Power.

Dazu seien die Daten  $(D_1, O_1, D_2, O_2) = (d_1, o_1, d_2, o_2)$  gegeben.

$D_1^*$  und  $D_2^*$  bezeichnen die Anzahlen der zukünftigen Todesfälle in den beiden Therapiegruppen, also vom Zeitpunkt der Zwischenauswertung bis zum Studienende, wobei die beiden Größen unabhängig voneinander seien.

$O_1^*$  und  $O_2^*$  seien die Personenzeiten der jeweiligen Therapiegruppe von der Zwischenauswertung bis zum Ende der Studie, welche ebenfalls voneinander unabhängig seien.

Auch hier ist  $D_i^* = D_i^*(n_i^*)$  und  $O_i^* = O_i^*(n_i^*)$ , wobei  $n_i^* \in \mathbb{N}$  die Anzahl der Patienten in Gruppe  $i$  zum Zeitpunkt der Zwischenauswertung bezeichnet und bei gegebenen Daten  $(D_1, O_1, D_2, O_2) = (d_1, o_1, d_2, o_2)$  nicht-zufällig ist, für alle  $i \in \{1, 2\}$ .

Dabei gelte für die gesamte Fallzahl:  $n^* = n_1^* + n_2^*$  und  $n_i^* = n_i^*(n^*) \rightarrow \infty$  für  $n^* \rightarrow \infty$ , für alle  $i \in \{1, 2\}$ . Außerdem sei die Bedingung  $\lim_{n^* \rightarrow \infty} n_1^*/n_2^* = \text{const} \in (0, \infty)$  wieder erfüllt.

Nun fixiert man wieder  $O_1^*$  und  $O_2^*$ .

Da  $D_i^*$  analog zu Abschnitt 2.1.2 binomialverteilt ist zu den Parametern  $n_i^*$  und  $p_i^*$ , und wegen der asymptotischen Erwartungstreue der Maximum-Likelihood-Schätzer (Greene (2003), 17.4.5.a Consistency, Seiten 477/478), gilt:

$$E_{\lambda_i} D_i^* = n_i^* p_i^* \approx \lambda_i O_i^*$$

und

$$\text{Var}_{\lambda_i} D_i^* = n_i^* p_i^* (1 - p_i^*) \approx n_i^* p_i^* = E_{\lambda_i} D_i^*$$

für hinreichend kleines  $p_i^*$ , hinreichend großes  $n_i^*$  und fixiertes  $O_i^*$ , für alle  $i \in \{1, 2\}$ .

Die Wahrscheinlichkeit  $p_i^*$  ist hier auf ähnliche Weise wie in Abschnitt 2.1.2 zu verstehen:  $p_i^* = P_{\lambda_i}(T_{i,j_i} \leq C_{i,j_i} \mid T_{i,j_i} > t^*)$  für alle  $j_i \in \{1, \dots, n_i^*\}$  und  $i \in \{1, 2\}$ .

Es sei daran erinnert, dass die Bedingung des hinreichend kleinen  $p_i^*$  für alle  $i \in \{1, 2\}$  analog zu (5) aus Abschnitt 2.1.2 (Seite 21) weggelassen werden kann.

### 2.2.1 Bedingte Teststatistik

Jetzt wählt man als Teststatistik  $W^*$ :

$$W^* = \log \left( \frac{D_2 + D_2^*}{O_2 + O_2^*} \right) - \log \left( \frac{D_1 + D_1^*}{O_1 + O_1^*} \right),$$

welche für hinreichend großes  $n_i^*$  und fixiertes  $O_i^*$  für alle  $i \in \{1,2\}$  unter  $(D_1, O_1, D_2, O_2) = (d_1, o_1, d_2, o_2)$  auch wieder approximativ normalverteilt ist mit Erwartungswert

$$\log \left( \frac{d_2 + \lambda_2 O_2^*}{o_2 + O_2^*} \right) - \log \left( \frac{d_1 + \lambda_1 O_1^*}{o_1 + O_1^*} \right)$$

und Varianz

$$\frac{1}{d_1 + \lambda_1 O_1^*} + \frac{1}{d_2 + \lambda_2 O_2^*}.$$

Dies sieht man folgendermaßen ein:

Die asymptotische Normalität unter  $(D_1, O_1, D_2, O_2) = (d_1, o_1, d_2, o_2)$  folgt wegen der asymptotischen Normalität (Greene (2003), 17.4.5.b Asymptotic Normality, Seiten 478/479) und Unabhängigkeit der Maximum-Likelihood-Schätzer und damit von  $D_1^*$  und  $D_2^*$ , und wieder mittels Delta-Methode (Lemma 1, Abschnitt 2.1.2, Seite 18) mit  $g = \log$ .

Mit (4) aus Abschnitt 2.1.2 (Seite 20) gilt für hinreichend großes  $n_i^*$  und fixiertes  $O_i^*$ :

$$\begin{aligned} E_{\lambda_i} \log \left( \frac{D_i + D_i^*}{O_i + O_i^*} \right) &\approx \log E_{\lambda_i} \left( \frac{D_i + D_i^*}{O_i + O_i^*} \right) \\ &\approx \log \left( \frac{d_i + \lambda_i O_i^*}{o_i + O_i^*} \right) \end{aligned}$$

für alle  $i \in \{1,2\}$  unter  $(D_1, O_1, D_2, O_2) = (d_1, o_1, d_2, o_2)$ .

Ebenfalls analog zu (4) in Abschnitt 2.1.2 (Seite 20) erhält man für hinreichend großes

$n_i^*$  und fixiertes  $O_i^*$  für die Varianz:

$$\begin{aligned}
\text{Var}_{\lambda_i} \log \left( \frac{D_i + D_i^*}{O_i + O_i^*} \right) &\approx \text{Var}_{\lambda_i} \left[ \frac{\left( \frac{D_i + D_i^*}{O_i + O_i^*} \right) - \mathbb{E}_{\lambda_i} \left( \frac{D_i + D_i^*}{O_i + O_i^*} \right)}{\mathbb{E}_{\lambda_i} \left( \frac{D_i + D_i^*}{O_i + O_i^*} \right)} \right] \\
&= \text{Var}_{\lambda_i} \left[ \frac{(D_i + D_i^*) - \mathbb{E}_{\lambda_i}(D_i + D_i^*)}{\mathbb{E}_{\lambda_i}(D_i + D_i^*)} \right] \\
&= \frac{\text{Var}_{\lambda_i}(D_i + D_i^*)}{[\mathbb{E}_{\lambda_i}(D_i + D_i^*)]^2} \\
&\approx \frac{\mathbb{E}_{\lambda_i}(D_i + D_i^*)}{[\mathbb{E}_{\lambda_i}(D_i + D_i^*)]^2} \\
&= \frac{1}{\mathbb{E}_{\lambda_i}(D_i + D_i^*)} \\
&= \frac{1}{d_i + \lambda_i O_i^*}
\end{aligned}$$

für alle  $i \in \{1,2\}$  unter  $(D_1, O_1, D_2, O_2) = (d_1, o_1, d_2, o_2)$ .

Somit gilt:

**Satz 2** (Konvergenz in Verteilung der bedingten Teststatistik; Andersen (1987), Results, Seite 69). *Unter den Annahmen aus Abschnitt 2.1 und Abschnitt 2.2 gilt für hinreichend großes  $n^*$  und fixierte  $O_1^*, O_2^*$  unter  $(D_1, O_1, D_2, O_2) = (d_1, o_1, d_2, o_2)$ :*

$$\frac{W^* - \left[ \log \left( \frac{D_2 + \lambda_2 O_2^*}{O_2 + O_2^*} \right) - \log \left( \frac{D_1 + \lambda_1 O_1^*}{O_1 + O_1^*} \right) \right]}{\sqrt{\frac{1}{D_1 + \lambda_1 O_1^*} + \frac{1}{D_2 + \lambda_2 O_2^*}}} \xrightarrow{d} N(0,1).$$

### 2.2.2 Bedingte Power-Funktion

Die bedingte Power-Funktion  $\gamma_c: (0, \infty) \rightarrow [0,1]$  definiert man nun gemäß (Andersen (1987), Results, Seite 70):

$$\begin{aligned}
\gamma_c(\theta) &= P_{\lambda_1, \lambda_2} \left( |W^*| \geq q_{1-\frac{\alpha}{2}} \sqrt{\text{Var}_{\lambda_1, \lambda_1} W^*} + \mathbb{E}_{\lambda_1, \lambda_1} W^* \mid (D_1, O_1, D_2, O_2) = (d_1, o_1, d_2, o_2) \right) \\
&\approx \Phi \left( \frac{q_{\frac{\alpha}{2}} \sigma_{\lambda_1, \lambda_1} + \mu_{\lambda_1, \lambda_1} - \mu_{\lambda_1, \lambda_2}}{\sigma_{\lambda_1, \lambda_2}} \right) + \left[ 1 - \Phi \left( \frac{q_{1-\frac{\alpha}{2}} \sigma_{\lambda_1, \lambda_1} + \mu_{\lambda_1, \lambda_1} - \mu_{\lambda_1, \lambda_2}}{\sigma_{\lambda_1, \lambda_2}} \right) \right]
\end{aligned}$$

unter den Annahmen aus Abschnitt 2.1 und Abschnitt 2.2 für hinreichend großes  $n^*$  und fixierte  $O_1^*, O_2^*$ , mit

$$\begin{aligned}\sigma_{\lambda_1, \lambda_1} &= \sqrt{\frac{1}{d_1 + \lambda_1 O_1^*} + \frac{1}{d_2 + \lambda_1 O_2^*}}, \\ \mu_{\lambda_1, \lambda_1} &= \log\left(\frac{d_2 + \lambda_1 O_2^*}{o_2 + O_2^*}\right) - \log\left(\frac{d_1 + \lambda_1 O_1^*}{o_1 + O_1^*}\right), \\ \mu_{\lambda_1, \lambda_2} &= \log\left(\frac{d_2 + \lambda_2 O_2^*}{o_2 + O_2^*}\right) - \log\left(\frac{d_1 + \lambda_1 O_1^*}{o_1 + O_1^*}\right)\end{aligned}$$

und

$$\sigma_{\lambda_1, \lambda_2} = \sqrt{\frac{1}{d_1 + \lambda_1 O_1^*} + \frac{1}{d_2 + \lambda_2 O_2^*}}.$$

Hierbei sei nochmals angemerkt, dass sich  $\text{Var}_{\lambda_1, \lambda_1}$  auf die Varianz und  $E_{\lambda_1, \lambda_1}$  auf den Erwartungswert unter  $H_0$  mit  $\theta = 1$  und entsprechend  $\lambda_1 = \lambda_2$  beziehen.

$\mu_{\lambda_1, \lambda_1}$  und  $\sigma_{\lambda_1, \lambda_1}$  entsprechen also dem approximierten Erwartungswert beziehungsweise der approximierten Standardabweichung der bedingten Teststatistik unter der Nullhypothese, und  $\mu_{\lambda_1, \lambda_2}$  und  $\sigma_{\lambda_1, \lambda_2}$  dem approximierten Erwartungswert beziehungsweise der approximierten Standardabweichung der bedingten Teststatistik unter der Alternativhypothese.

Das „c“ soll verdeutlichen, dass es sich um die bedingte (englisch: conditional) Powerfunktion handelt.

Somit ist eine erste Formel, hier bezüglich des Exponentialmodells, hergeleitet.

Nun ist es aber so, dass sich das Exponentialmodell nicht immer gut genug an die Daten anpassen lässt. Solch ein Fall kann zum Beispiel auftreten, wenn die Überlebenswahrscheinlichkeiten bezüglich einer Krankheit und deren unterschiedlichen Therapien betrachtet werden und die Überlebenskurven in ein sogenanntes „Plateau“ übergehen.

Wie man dann in solch einer Situation vorgehen könnte, soll das nun folgende Kapitel klären.

### 3 Überlebenszeitmodelle mit nicht-konstanten Hazards: Überlebensfunktionen mit Plateau

Bisher ging es um Verteilungen im üblichen Sinne. Das heißt für die gewählten Überlebensfunktionen galt:

$$S(0) = 1$$

und

$$\lim_{t \rightarrow \infty} S(t) = 0.$$

Untersucht man aber nun die Überlebenswahrscheinlichkeit eines Patienten bezüglich einer bestimmten Krankheit, so ist es möglich, dass die Überlebenskurve mit der Zeit in ein sogenanntes „Plateau“ übergeht. Beispielsweise kann sich solch ein Bild bei der Überlebenszeitanalyse von Patienten einstellen, die an einer gut kurierbaren Form von Krebs erkrankt sind. Kommt es hier bei einem Patienten zur Remission, tritt das Ereignis bei diesem Patienten nicht ein. Dies ist zumindest solange der Fall, wie der Patient keinem Rezidiv ausgesetzt ist.

Im nächsten Schritt wird nun ein erstes mögliches Modell für Überlebensfunktionen mit Plateau allgemein beschrieben. Danach werden für einen Spezialfall die Eigenschaften dieses Modells genauer untersucht, ohne jedoch die asymptotischen Eigenschaften näher zu beleuchten. Anschließend wird zu einem zweiten Modell übergegangen, welches dann ausführlich - auch hinsichtlich der Asymptotik - analysiert wird.

Eine Arbeit von Lambert aus dem Jahr 2007 gibt zu beiden Modellen eine gute Einführung und kann entsprechend hinzugezogen werden.

#### 3.1 Mixture-Modell

Das im Folgende dargestellte Modell geht zurück auf Arbeiten von Boag aus dem Jahr 1949 und von Berkson und Gage aus dem Jahr 1952. Der Grundgedanke ist hier, dass sich die Verteilung für das Überleben aus zwei Komponenten zusammensetzt. Eine Komponente stellt dabei den Anteil derer dar, die eine Remission erfahren haben, während das Überleben des restlichen Anteils über eine Überlebensfunktion im üblichen Sinne ausgedrückt wird.

Weitere Literatur findet sich beispielsweise in Sposto und Sather aus dem Jahr 1985 und in Sposto aus dem Jahr 2002.

### 3.1.1 Allgemein

Man definiert  $S: [0, \infty) \rightarrow [0, 1]$  gemäß

$$S(t) = c + (1 - c)S^{(0)}(t)$$

mit  $c \in (0, 1)$  und  $S^{(0)}$  einer Überlebensfunktion im üblichen Sinne.

Der Anteil  $1 - c$  der Patienten, die weiterhin spezifische Krankheitssymptome aufweisen, wird über eine Überlebensfunktion im üblichen Sinne charakterisiert.

Die Bezeichnung „ $c$ “ rührt dabei von dem englischen Wort „cured“ (deutsch: geheilt) her.

Damit ist nun ersichtlich:

$$S(0) = 1$$

und

$$\lim_{t \rightarrow \infty} S(t) = c,$$

was dem Anteil derjenigen entspricht, bei denen es zu einer Remission kommt.

Da es sich hier wegen der Konvergenz gegen eine positive Konstante nicht um eine Überlebensfunktion im üblichen Sinne handelt, muss man im Weiteren zusätzlich beachten, in wie weit benötigte Resultate aus der Wahrscheinlichkeitstheorie ihre Gültigkeit behalten.

Damit man nun zu der zu  $S$  entsprechenden Hazardfunktion  $h$  gelangt, berechnet man mit  $f = -S' = -(1 - c)S^{(0)'} = (1 - c)f^{(0)}$  und  $h^{(0)} = f^{(0)}/S^{(0)}$ :

$$\begin{aligned} h(t) &= \frac{f(t)}{S(t)} \\ &= \frac{(1 - c)f^{(0)}(t)}{c + (1 - c)S^{(0)}(t)} \\ &= h^{(0)}(t) \frac{(1 - c)S^{(0)}(t)}{c + (1 - c)S^{(0)}(t)} \end{aligned}$$

für alle  $t \geq 0$ . Dabei bezeichnen  $f$  und  $f^{(0)}$  die zu  $S$  beziehungsweise  $S^{(0)}$  entsprechenden Wahrscheinlichkeitsdichtefunktionen und  $h^{(0)}$  die zu  $S^{(0)}$  gehörende Hazardfunktion.

Somit hat man für die kumulative Hazardfunktion  $H$  von  $S$  (Definition 3, Abschnitt 1.2,

Seite 10) unter Verwendung von  $S'/S = (\log S)'$  und  $c + (1 - c)S^{(0)}(0) = 1$ :

$$\begin{aligned}
 H(t) &= \int_0^t h(s) \, ds \\
 &= \int_0^t \frac{(1 - c)f^{(0)}(s)}{c + (1 - c)S^{(0)}(s)} \, ds \\
 &= - \int_0^t \frac{(1 - c)S^{(0)'}(s)}{c + (1 - c)S^{(0)}(s)} \, ds \\
 &= - \int_0^t \left\{ \log [c + (1 - c)S^{(0)}(s)] \right\}' \, ds \\
 &= - \log [c + (1 - c)S^{(0)}(t)] + \log [c + (1 - c)S^{(0)}(0)] \\
 &= - \log [c + (1 - c)S^{(0)}(t)]
 \end{aligned}$$

für alle  $t \geq 0$ .

Auch das kann man unter Einbeziehung der zu  $S^{(0)}$  entsprechenden kumulativen Hazardfunktion  $H^{(0)}$  ausdrücken. Man erhält mit der Definition der kumulativen Hazardfunktion (Definition 3, Abschnitt 1.2, Seite 10) und der allgemeinen Darstellung der Überlebensfunktion (2) aus Abschnitt 1.2 (Seite 10):

$$\begin{aligned}
 H(t) &= - \log S(t) \\
 &= - \log S^{(0)}(t) + \log S^{(0)}(t) - \log S(t) \\
 &= H^{(0)}(t) + \log \left( \frac{S^{(0)}(t)}{S(t)} \right) \\
 &= H^{(0)}(t) + \log \left( \frac{S^{(0)}(t)}{c + (1 - c)S^{(0)}(t)} \right)
 \end{aligned}$$

für alle  $t \geq 0$ .

### 3.1.2 Exponentialverteilung

Wie zu Beginn dieses Kapitels angedeutet, soll es nun um eine bestimmte Klasse von Verteilungsfunktionen für  $S_i^{(0)}$  der Personen aus Gruppe  $i$ , mit  $i \in \{1, 2\}$ , gehen. Konkret betrachtet man dabei die Klasse der Exponentialverteilungen. Damit haben die

Überlebensfunktion  $S_i$  und die Hazardfunktion  $h_i$  nun folgende Gestalt:

$$S_i(t) = c_i + (1 - c_i) \exp(-\lambda_i t),$$

$$h_i(t) = \lambda_i \frac{(1 - c_i) \exp(-\lambda_i t)}{c_i + (1 - c_i) \exp(-\lambda_i t)}$$

und man erhält für das Hazard-Ratio  $r$  von  $h_1$  und  $h_2$ :

$$\begin{aligned} r(t) &= \frac{h_2(t)}{h_1(t)} \\ &= \frac{\lambda_2 (1 - c_2) \exp(-\lambda_2 t)}{\lambda_1 (1 - c_1) \exp(-\lambda_1 t)} \frac{c_1 + (1 - c_1) \exp(-\lambda_1 t)}{c_2 + (1 - c_2) \exp(-\lambda_2 t)} \\ &= \frac{\lambda_2}{\lambda_1} \frac{1 - c_2}{1 - c_1} \frac{c_1 \exp(\lambda_1 t) + (1 - c_1)}{c_2 \exp(\lambda_2 t) + (1 - c_2)} \end{aligned}$$

für alle  $t \geq 0$  mit der Rate  $\lambda_i > 0$  und  $c_i \in (0,1)$  für alle  $i \in \{1,2\}$ .

Das Problem, welches sich hier ergibt, ist, dass nicht nur die Hazardrate  $\lambda_i$  der Nicht-Geheilten von Gruppe  $i$  sondern auch der Anteil  $c_i$  der Geheilten der Gruppe  $i$  geschätzt werden muss, mit  $i \in \{1,2\}$ .

Dies soll wieder über die Maximum-Likelihood-Methode erfolgen. Dazu wird die Darstellung der Likelihood-Funktion (3) aus Abschnitt 2.1.1 (Seite 16) verwendet, wobei für die Zensierungszeitpunkte dieselben Annahmen getroffen werden wie in Abschnitt 2.1.1. Weiterhin wird bei der Likelihood-Funktion der für die Maximierung überflüssige Vorfaktor weggelassen. Dann erhält man bei gegebenen  $(t_{i,j_i}, \delta_{i,j_i}) \in [0, \infty) \times \{0,1\}$  für alle  $j_i \in \{1, \dots, n_i\}$  für die logarithmierte Likelihood-Funktion  $l_i$  der Gruppe  $i$ :

$$\begin{aligned} l_i(\lambda_i, c_i) &= \sum_{j_i=1}^{n_i} [\delta_{i,j_i} \log h_i(t_{i,j_i}) + \log S_i(t_{i,j_i})] \\ &= \sum_{j_i=1}^{n_i} \{ \delta_{i,j_i} \{ \log \lambda_i + \log(1 - c_i) - \lambda_i t_{i,j_i} - \log [c_i + (1 - c_i) \exp(-\lambda_i t_{i,j_i})] \} \\ &\quad + \log [c_i + (1 - c_i) \exp(-\lambda_i t_{i,j_i})] \} \\ &= \sum_{j_i=1}^{n_i} \{ \delta_{i,j_i} (\log \lambda_i + \log(1 - c_i) - \lambda_i t_{i,j_i}) \\ &\quad + (1 - \delta_{i,j_i}) \log [c_i + (1 - c_i) \exp(-\lambda_i t_{i,j_i})] \} \end{aligned}$$

für alle  $i \in \{1,2\}$ .

Für die partiellen Ableitungen der logarithmierten Likelihood-Funktion  $l_i$  gilt dann:

$$\frac{\partial}{\partial \lambda_i} l_i(\lambda_i, c_i) = \frac{1}{\lambda_i} \sum_{j_i=1}^{n_i} \delta_{i,j_i} - \sum_{j_i=1}^{n_i} \delta_{i,j_i} t_{i,j_i} - \sum_{j_i=1}^{n_i} (1 - \delta_{i,j_i}) \frac{(1 - c_i) t_{i,j_i} \exp(-\lambda_i t_{i,j_i})}{c_i + (1 - c_i) \exp(-\lambda_i t_{i,j_i})}$$

und

$$\frac{\partial}{\partial c_i} l_i(\lambda_i, c_i) = -\frac{1}{1-c_i} \sum_{j_i=1}^{n_i} \delta_{i,j_i} + \sum_{j_i=1}^{n_i} (1-\delta_{i,j_i}) \frac{1-\exp(-\lambda_i t_{i,j_i})}{c_i + (1-c_i) \exp(-\lambda_i t_{i,j_i})}$$

für alle  $i \in \{1,2\}$ .

Nach dem Nullsetzen, Umstellen und Einsetzen der Zufallsvariablen  $Y_{i,j_i}$  für  $t_{i,j_i}$  und  $\Delta_{i,j_i}$  für  $\delta_{i,j_i}$  für alle  $j_i \in \{1, \dots, n_i\}$  wie in Abschnitt 2.1.1 erhält man die implizit gegebenen Schätzer

$$\hat{\lambda}_i = \frac{\sum_{j_i=1}^{n_i} \Delta_{i,j_i}}{\sum_{j_i=1}^{n_i} \Delta_{i,j_i} Y_{i,j_i} + \sum_{j_i=1}^{n_i} (1-\Delta_{i,j_i}) \frac{(1-\hat{c}_i) Y_{i,j_i}}{\hat{c}_i \exp(\hat{\lambda}_i Y_{i,j_i}) + (1-\hat{c}_i)}}$$

und

$$\hat{c}_i = 1 - \frac{\sum_{j_i=1}^{n_i} \Delta_{i,j_i}}{\sum_{j_i=1}^{n_i} (1-\Delta_{i,j_i}) \frac{1-\exp(-\hat{\lambda}_i Y_{i,j_i})}{\hat{c}_i + (1-\hat{c}_i) \exp(-\hat{\lambda}_i Y_{i,j_i})}}$$

für alle  $i \in \{1,2\}$ .

Da die Schätzer nicht explizit gegeben sind, müssen diese numerisch bestimmt werden.

Auf das Mixture-Modell soll nun nicht weiter eingegangen werden. Das bedeutet insbesondere, dass keine weiteren Aussagen zur Berechnung der Power-Funktion beziehungsweise bedingten Power-Funktion folgen werden. Vielmehr soll dies im Rahmen eines weiteren Modells erfolgen, um welches es sich im nächsten Abschnitt handeln wird. Dennoch soll das Mixture-Modell als eine Möglichkeit der Modellierung von Überleben mit einer Überlebensfraktion der Vollständigkeit halber und als eine Alternative zu dem nun folgenden, häufiger genutzten, Modell genannt bleiben.

## 3.2 Non-Mixture-Modell

Eine andere Möglichkeit der Modellierung von Überleben, welche in der Krebsforschung ihren Ursprung hat, soll in diesem Abschnitt erläutert werden. Dazu können Arbeiten von Tsodikov, Löffler und Yakovlev aus dem Jahr 1998 und von Tsodikov, Ibrahim und Yakovlev aus dem Jahr 2003 hinzugezogen werden.

### 3.2.1 Allgemein

Für die Überlebensfunktion mit Plateau gilt:

$$S(0) = 1$$

und

$$\lim_{t \rightarrow \infty} S(t) = c$$

für  $c \in (0,1)$ .

Mit der Definition der kumulativen Hazardfunktion (Definition 3, Abschnitt 1.2, Seite 10) und der allgemeinen Darstellung der Überlebensfunktion (2) aus Abschnitt 1.2 (Seite 10) bedeutet dies für die kumulative Hazardfunktion:

$$\lim_{t \rightarrow \infty} H(t) = -\log c,$$

das heißt, die kumulative Hazardfunktion ist beschränkt.

Damit kann man  $H$  folgendermaßen definieren:

$$H(t) = -[1 - S^{(0)}(t)] \log c$$

für alle  $t \geq 0$  mit  $S^{(0)}$  als Überlebensfunktion im üblichen Sinne.

Nun erhält man als Modell:

$$S(t) = c^{1-S^{(0)}(t)}$$

für alle  $t \geq 0$ .

Dies lässt sich äquivalent in Form eines Mixture-Modells darstellen:

$$S(t) = d + (1-d) \left[ \frac{c^{1-S^{(0)}(t)} - d}{1-d} \right]$$

für alle  $t \geq 0$  und  $d \in (0,1)$ . Zu beachten ist hier allerdings, dass  $(c^{1-S^{(0)}} - d)/(1-d)$  eventuell keine Überlebensfunktion im üblichen Sinne ist, denn es gilt zwar:

$$\frac{c^{1-S^{(0)}(0)} - d}{1-d} = \frac{1-d}{1-d} = 1$$

zum Zeitpunkt  $t = 0$ , aber

$$\lim_{t \rightarrow \infty} \frac{c^{1-S^{(0)}(t)} - d}{1-d} = \frac{c-d}{1-d},$$

was nur für  $c = d$  verschwindet. Darauf müsste bei den folgenden Überlegungen gegebenenfalls Rücksicht genommen werden.

Hier soll es aber nun mit der ursprünglichen Form des Non-Mixture-Modells weitergehen.

Zur Berechnung der Hazardfunktion  $h$  gemäß  $S$  benötigt man zunächst die zu  $S$  entsprechende Wahrscheinlichkeitsdichtefunktion  $f$ , die folgende Gestalt hat:

$$\begin{aligned}
 f(t) &= -S'(t) \\
 &= -\left[c^{1-S^{(0)}(t)}\right]' \\
 &= -\left[\exp\left([1 - S^{(0)}(t)] \log c\right)\right]' \\
 &= -\exp\left([1 - S^{(0)}(t)] \log c\right) f^{(0)}(t) \log c \\
 &= -c^{1-S^{(0)}(t)} f^{(0)}(t) \log c
 \end{aligned}$$

für alle  $t \geq 0$  mit  $f^{(0)}$  als Wahrscheinlichkeitsdichtefunktion bezüglich  $S^{(0)}$ .

Damit gilt für  $h$ :

$$\begin{aligned}
 h(t) &= \frac{f(t)}{S(t)} \\
 &= \frac{-c^{1-S^{(0)}(t)} f^{(0)}(t) \log c}{c^{1-S^{(0)}(t)}} \\
 &= -f^{(0)}(t) \log c \\
 &= -h^{(0)}(t) S^{(0)}(t) \log c
 \end{aligned}$$

für alle  $t \geq 0$  mit  $h^{(0)} = f^{(0)}/S^{(0)}$  als zu  $S^{(0)}$  entsprechender Hazardfunktion.

Schließlich erhält man für die kumulative Hazardfunktion  $H$  (Definition 3, Abschnitt 1.2, Seite 10) mit der allgemeinen Darstellung der Überlebensfunktion (2) aus Abschnitt 1.2 (Seite 10) unter Einbeziehung der zu  $S^{(0)}$  entsprechenden kumulativen Hazardfunktion  $H^{(0)}$ :

$$H(t) = -\log S(t) = -\left[1 - S^{(0)}(t)\right] \log c = -\left[1 - \exp\left(-H^{(0)}(t)\right)\right] \log c$$

für alle  $t \geq 0$ .

### 3.2.2 Exponentialverteilung

Auch hier soll  $S_i^{(0)}$  speziell aus der Klasse der Exponentialverteilungen gewählt werden, wobei Gruppe  $i$  wieder aus  $n_i$  Personen besteht, mit  $i \in \{1,2\}$ .

Das heißt, es gelten für die Überlebensfunktion  $S_i$  und die Hazardfunktion  $h_i$ :

$$\begin{aligned} S_i(t) &= c_i^{1-\exp(-\lambda_i t)}, \\ h_i(t) &= -\lambda_i \exp(-\lambda_i t) \log c_i \end{aligned}$$

und entsprechend für das Hazard-Ratio  $r$  von  $h_1$  und  $h_2$ :

$$r(t) = \frac{h_2(t)}{h_1(t)} = \frac{\lambda_2}{\lambda_1} \frac{\log c_2}{\log c_1} \exp(-(\lambda_2 - \lambda_1)t) \quad (6)$$

für alle  $t \geq 0$  mit den Parametern  $\lambda_i > 0$  und  $c_i \in (0,1)$  für alle  $i \in \{1,2\}$ .

### Maximum-Likelihood-Schätzung

Wie in dem entsprechenden vorangegangenen Abschnitt 3.1.2 soll die Schätzung der Parameter  $\lambda_i$  und  $c_i$  über die Maximum-Likelihood-Methode erfolgen, für alle  $i \in \{1,2\}$ . Dazu soll wie in Abschnitt 2.1.1 vorgegangen werden, wobei der für die Maximierung nicht benötigte Vorfaktor wieder weggelassen wird.

Somit ergibt sich für die logarithmierte Likelihood-Funktion  $l_i$  von Gruppe  $i$  bei gegebenen  $(t_{i,j_i}, \delta_{i,j_i}) \in [0, \infty) \times \{0,1\}$  für alle  $j_i \in \{1, \dots, n_i\}$ :

$$\begin{aligned} l_i(\lambda_i, c_i) &= \sum_{j_i=1}^{n_i} [\delta_{i,j_i} \log h_i(t_{i,j_i}) + \log S_i(t_{i,j_i})] \\ &= \sum_{j_i=1}^{n_i} \{ \delta_{i,j_i} [\log \lambda_i - \lambda_i t_{i,j_i} + \log(-\log c_i)] + [1 - \exp(-\lambda_i t_{i,j_i})] \log c_i \} \end{aligned}$$

für alle  $i \in \{1,2\}$ .

Hier gilt für die partiellen Ableitungen der logarithmierten Likelihood-Funktion  $l_i$ :

$$\frac{\partial}{\partial \lambda_i} l_i(\lambda_i, c_i) = \frac{1}{\lambda_i} \sum_{j_i=1}^{n_i} \delta_{i,j_i} - \sum_{j_i=1}^{n_i} \delta_{i,j_i} t_{i,j_i} + \log(c_i) \sum_{j_i=1}^{n_i} t_{i,j_i} \exp(-\lambda_i t_{i,j_i})$$

und

$$\frac{\partial}{\partial c_i} l_i(\lambda_i, c_i) = \frac{1}{c_i \log c_i} \sum_{j_i=1}^{n_i} \delta_{i,j_i} + \frac{1}{c_i} \sum_{j_i=1}^{n_i} [1 - \exp(-\lambda_i t_{i,j_i})]$$

für alle  $i \in \{1,2\}$ .

Durch Nullsetzen mit  $c_i \in (0,1)$  und Einsetzen der Zufallsvariablen  $Y_{i,j_i}$  für  $t_{i,j_i}$  und  $\Delta_{i,j_i}$  für  $\delta_{i,j_i}$  für alle  $j_i \in \{1, \dots, n_i\}$  wie in Abschnitt 2.1.1 erhält man die wieder nur implizit gegebenen Maximum-Likelihood-Schätzer

$$\hat{\lambda}_i = \frac{\sum_{j_i=1}^{n_i} \Delta_{i,j_i}}{\sum_{j_i=1}^{n_i} \Delta_{i,j_i} Y_{i,j_i} - \log(\hat{c}_i) \sum_{j_i=1}^{n_i} Y_{i,j_i} \exp(-\hat{\lambda}_i Y_{i,j_i})}$$

und

$$\hat{c}_i = \exp\left(-\frac{\sum_{j_i=1}^{n_i} \Delta_{i,j_i}}{\sum_{j_i=1}^{n_i} [1 - \exp(-\hat{\lambda}_i Y_{i,j_i})]}\right) \quad (7)$$

für alle  $i \in \{1,2\}$ .

Auch hier müssten die Maximum-Likelihood-Schätzer numerisch bestimmt werden.

## Annahmen

Für die weiteren Betrachtungen wird nun folgende Annahme getroffen:

$$\lambda_1 = \lambda_2.$$

Das heißt, dass sich die Anteile der Geheilten der beiden Gruppen unterscheiden können, jedoch nicht die Geschwindigkeiten, mit der die Gruppen auf ihre jeweilige Anzahl von Geheilten schrumpfen.

Wie plausibel diese Annahme aus medizinischer Sicht ist, muss im konkreten Fall natürlich hinreichend geklärt und statistisch getestet werden.

Der Vorteil dieser Annahme ist, dass man es nun mit zueinander proportionalen Hazardraten zu tun hat, da das mit (6) aus diesem Abschnitt (Seite 34) gegebene Hazard-Ratio  $r$  unabhängig von der Zeit ist:

$$r(t) = \frac{\log c_2}{\log c_1}$$

für alle  $t \geq 0$ .

Man setzt als Maß für den Therapieeffekt

$$\theta = \frac{\log c_2}{\log c_1} > 0,$$

wählt als Schätzer für das Hazard-Ratio

$$\hat{\theta} = \frac{\log \hat{c}_2}{\log \hat{c}_1}$$

und betrachtet wie in Kapitel 2 die Hypothesen

$$H_0: \theta = 1 \quad \text{versus} \quad H_1: \theta \neq 1.$$

Dabei sei angenommen, dass mit  $D_i = \sum_{j_i=1}^{n_i} \Delta_{i,j_i}$  und  $O'_i = \sum_{j_i=1}^{n_i} [1 - \exp(-\hat{\lambda}_i Y_{i,j_i})]$  für alle  $i \in \{1,2\}$  folgendes gelte:  $D_1(\omega) > 0$ ,  $D_2(\omega) > 0$ ,  $O'_1(\omega) > 0$  und  $O'_2(\omega) > 0$ , jeweils für P-fast alle  $\omega \in \Omega$ .

### Teststatistik

Man wählt hier als Teststatistik  $W$  analog zu Abschnitt 2.1.2 den logarithmierten Schätzer für das Hazard-Ratio:

$$\begin{aligned} W &= \log \hat{\theta} \\ &= \log \left( \frac{-\log \hat{c}_2}{-\log \hat{c}_1} \right) \\ &= \log(-\log \hat{c}_2) - \log(-\log \hat{c}_1) \\ &= \log \left( \frac{D_2}{O'_2} \right) - \log \left( \frac{D_1}{O'_1} \right), \end{aligned}$$

wobei in der zweiten und entsprechend dritten Zeile der Gleichung die beiden Minuszeichen als Vorzeichen eingefügt werden müssen, da definitionsgemäß und nach Voraussetzung  $\hat{c}_i(\omega) \in (0,1)$  und damit  $\log(\hat{c}_i(\omega)) \in (-\infty,0)$  für P-fast alle  $\omega \in \Omega$  und für alle  $i \in \{1,2\}$  gilt.

Nun seien die zufälligen Zeiten des Eintretens des Ereignisses beziehungsweise der Zensurierung von Person  $(i,j_i)$  aus Gruppe  $i$ ,  $Y_{i,j_i}$ , und die Maximum-Likelihood-Schätzer der Gruppe  $i$ ,  $\hat{\lambda}_i$ , fixiert, für alle  $i \in \{1,2\}$ .

Dann lässt man gedanklich den Gesamtstichprobenumfang  $n = n_1 + n_2$  gegen Unendlich gehen, mit  $n_i = n_i(n) \rightarrow \infty$  für  $n \rightarrow \infty$  und  $\lim_{n \rightarrow \infty} n_1/n_2 = \text{const} \in (0,\infty)$ , wobei  $D_i = D_i(n_i)$  und  $O'_i = O'_i(n_i)$  gilt, für alle  $i \in \{1,2\}$ .

Dass die Teststatistik  $W$  nun approximativ normalverteilt ist, folgt wegen der asymptotischen Normalität und Unabhängigkeit der Maximum-Likelihood-Schätzer  $\hat{c}_1$  und  $\hat{c}_2$  (Greene (2003), 17.4.5.b Asymptotic Normality, Seiten 478/479) und Anwendung der

Delta-Methode (Lemma 1, Abschnitt 2.1.2, Seite 18) mit  $g = \log(-\log)$ .

Dabei ist festzuhalten, dass die Zentrierung mittels  $\log \theta$  entfallen kann, da  $\theta = 1$  unter  $H_0$  gilt.

Weiter gilt für unabhängige  $\Delta_{i,j_i}$  für alle  $j_i \in \{1, \dots, n_i\}$ , dass  $D_i = \sum_{j_i=1}^{n_i} \Delta_{i,j_i}$  binomialverteilt ist zu den Parametern  $n_i$  und  $p_i = P_{\lambda_i, c_i}(T_{i,j_i} \leq C_{i,j_i})$  für alle  $j_i \in \{1, \dots, n_i\}$  und  $i \in \{1, 2\}$ .

Dann hat man

$$\text{Var}_{\lambda_i, c_i} D_i = n_i p_i (1 - p_i) \approx n_i p_i = E_{\lambda_i, c_i} D_i$$

für hinreichend kleines  $p_i$  und

$$E_{\lambda_i, c_i} D_i \approx -\log(c_i) O'_i$$

für hinreichend großes  $n_i$  unter Zuhilfenahme von Gleichung (7) aus diesem Abschnitt (Seite 35) und der asymptotischen Erwartungstreue der Maximum-Likelihood-Schätzer (Greene (2003), 17.4.5.a Consistency, Seiten 477/478).

Damit gilt mit (4) aus Abschnitt 2.1.2 (Seite 20),  $E_{\lambda_i, c_i} D_i \approx -\log(c_i) O'_i$  und (5) aus Abschnitt 2.1.2 (Seite 21)

$$\text{Var}_{\lambda_i, c_i} \log D_i \approx \frac{1}{E_{\lambda_i, c_i} D_i} \approx \frac{1}{-\log(c_i) O'_i}$$

für hinreichend großes  $n_i$  und fixiertes  $O'_i$ , für alle  $i \in \{1, 2\}$ .

Somit folgt mit der Unabhängigkeit von  $D_1$  und  $D_2$  für die Varianz der Teststatistik:

$$\text{Var}_{\lambda_1, c_1, \lambda_2, c_2} \log \hat{\theta} \approx \frac{1}{-\log(c_1) O'_1} + \frac{1}{-\log(c_2) O'_2}$$

für hinreichend große  $n_1, n_2$  und fixierte  $O'_1, O'_2$ .

Zusammenfassend hat man:

**Satz 3** (Konvergenz in Verteilung der Teststatistik). *Unter den Annahmen aus Abschnitt 3.2.2 gilt für hinreichend großes  $n$  und fixierte  $O'_1, O'_2$ :*

$$\frac{W - \log \theta}{\sqrt{\frac{1}{-\log(c_1) O'_1} + \frac{1}{-\log(c_2) O'_2}}} \xrightarrow{d} N(0, 1).$$

Hier sei nochmals hervorgehoben, dass die bisherigen asymptotischen Überlegungen auf der Annahme  $\lambda_1 = \lambda_2$  beruhen.

### Power-Funktion

Wenn man wie in Abschnitt 2.1.3 mit  $\Phi$  die kumulative Verteilungsfunktion der Standardnormalverteilung und mit  $q_\alpha$  das  $\alpha$ -Quantil der Standardnormalverteilung mit  $\alpha \in (0,1)$  bezeichnet, kann man die Power-Funktion  $\gamma: (0,\infty) \rightarrow [0,1]$  folgendermaßen definieren:

$$\begin{aligned} \gamma(\theta) &= P_{\lambda_1, c_1, \lambda_2, c_2} \left( |W| \geq q_{1-\frac{\alpha}{2}} \sqrt{\text{Var}_{\lambda_1, c_1, \lambda_1, c_1} W} \right) \\ &\approx \Phi \left( \frac{q_{\frac{\alpha}{2}} \sqrt{\frac{1}{-\log(c_1)O'_1} + \frac{1}{-\log(c_1)O'_{2, \widehat{\lambda}_1}} - \log \theta}}{\sqrt{\frac{1}{-\log(c_1)O'_1} + \frac{1}{-\log(c_2)O'_2}}} \right) \\ &\quad + \left[ 1 - \Phi \left( \frac{q_{1-\frac{\alpha}{2}} \sqrt{\frac{1}{-\log(c_1)O'_1} + \frac{1}{-\log(c_1)O'_{2, \widehat{\lambda}_1}} - \log \theta}}{\sqrt{\frac{1}{-\log(c_1)O'_1} + \frac{1}{-\log(c_2)O'_2}}} \right) \right] \end{aligned}$$

mit  $O'_{2, \widehat{\lambda}_1} = \sum_{j_2=1}^{n_2} \left[ 1 - \exp \left( -\widehat{\lambda}_1 Y_{2, j_2} \right) \right]$  unter den Annahmen aus Abschnitt 3.2.2 für hinreichend großes  $n$  und fixierte  $O'_1, O'_2, O'_{2, \widehat{\lambda}_1}$ .

$\text{Var}_{\lambda_1, c_1, \lambda_1, c_1}$  drückt hier die Varianz unter der Annahme  $\lambda_1 = \lambda_2$  und unter  $H_0$  mit  $\log c_1 = \log c_2$ , also  $c_1 = c_2$ , aus.

### Bedingte Teststatistik

Man befindet sich wieder am Zeitpunkt  $t^* > 0$  innerhalb der Studie und möchte eine Zwischenauswertung durchführen. Dazu seien die Daten  $(\Delta_{i, j_i}, Y_{i, j_i}) = (\delta_{i, j_i}, t_{i, j_i})$  für alle  $j_i \in \{1, \dots, n_i\}$  und  $i \in \{1, 2\}$  gegeben.

Mit  $\Delta_{i, j_i}^*$  sei der Zensierungsindikator und mit  $Y_{i, j_i}^*$  der Zeitpunkt des Todes oder der Zensierung von Person  $(i, j_i)$  aus Gruppe  $i$  bezeichnet, wobei die Paare  $(\Delta_{i, j_i}^*, Y_{i, j_i}^*)$  voneinander unabhängig seien, für alle  $j_i \in \{1, \dots, n_i^*\}$  und  $i \in \{1, 2\}$ .

Hier bezieht sich  $n_i^*$  wieder auf die Anzahl der Patienten der Gruppe  $i$  zum Zeitpunkt  $t^*$  der Zwischenauswertung, mit der Gesamtanzahl  $n^* = n_1^* + n_2^*$  und  $n_i^* = n_i^*(n^*) \rightarrow \infty$

für  $n^* \rightarrow \infty$ , für alle  $i \in \{1,2\}$ , wobei  $\lim_{n^* \rightarrow \infty} n_1^*/n_2^* = \text{const} \in (0,\infty)$  gelte.

Dann wählt man in Anlehnung an die Teststatistik  $W$  als bedingte Teststatistik  $W^*$  mit  $D_i^* = \sum_{j_i=1}^{n_i^*} \Delta_{i,j_i}^*$  und  $O_i^* = \sum_{j_i=1}^{n_i^*} \left[ 1 - \exp\left(-\widehat{\lambda}_i Y_{i,j_i}^*\right) \right]$  für alle  $i \in \{1,2\}$ :

$$W^* = \log\left(\frac{D_2 + D_2^*}{O_2' + O_2'^*}\right) - \log\left(\frac{D_1 + D_1^*}{O_1' + O_1'^*}\right).$$

Nun seien die Zeitpunkte des Todes beziehungsweise der Zensierung,  $Y_{i,j_i}^*$ , und die Maximum-Likelihood-Schätzer  $\widehat{\lambda}_i$  fixiert, für alle  $j_i \in \{1, \dots, n_i^*\}$  und  $i \in \{1,2\}$ .

Dann gilt für  $D_i^* = \sum_{j_i=1}^{n_i^*} \Delta_{i,j_i}^*$  analog zu Abschnitt 2.2, dass  $D_i^*$  binomialverteilt ist zu den Parametern  $n_i^*$  und  $p_i^* = P_{\lambda_i, c_i}(T_{i,j_i} \leq C_{i,j_i} \mid T_{i,j_i} > t^*)$  mit:

$$\text{Var}_{\lambda_i, c_i} D_i^* = n_i^* p_i^* (1 - p_i^*) \approx n_i^* p_i^* = E_{\lambda_i, c_i} D_i^*$$

für hinreichend kleines  $p_i^*$  und hinreichend großes  $n_i^*$  für alle  $i \in \{1,2\}$ .

Über die Delta-Methode (Lemma 1, Abschnitt 2.1.2, Seite 18) mit  $g = \log$  gelangt man dann mit der asymptotischen Normalität (Greene (2003), 17.4.5.b Asymptotic Normality, Seiten 478/479) und Unabhängigkeit der Maximum-Likelihood-Schätzer und damit von  $D_1^*$  und  $D_2^*$  zur asymptotischen Normalität der Teststatistik  $W^*$ .

Unter Verwendung von  $d_i = \sum_{j_i=1}^{n_i} \delta_{i,j_i}$ ,  $o_i' = \sum_{j_i=1}^{n_i} [1 - \exp(-\lambda_i t_{i,j_i})]$  und

$$E_{\lambda_i, c_i} D_i^* \approx -\log(c_i) O_i'^*$$

gilt mit (4) aus Abschnitt 2.1.2 (Seite 20) und (5) aus Abschnitt 2.1.2 (Seite 21):

$$\begin{aligned} E_{\lambda_i, c_i} \log\left(\frac{D_i + D_i^*}{O_i' + O_i'^*}\right) &\approx \log E_{\lambda_i, c_i} \left(\frac{D_i + D_i^*}{O_i' + O_i'^*}\right) \\ &= \log\left(\frac{d_i + E_{\lambda_i, c_i} D_i^*}{o_i' + O_i'^*}\right) \\ &\approx \log\left(\frac{d_i - \log(\widehat{c}_i) O_i'^*}{o_i' + O_i'^*}\right) \end{aligned}$$

und

$$\begin{aligned}
\text{Var}_{\lambda_i, c_i} \log \left( \frac{D_i + D_i^*}{O'_i + O_i'^*} \right) &\approx \text{Var}_{\lambda_i, c_i} \left[ \frac{\left( \frac{D_i + D_i^*}{O'_i + O_i'^*} \right) - \mathbb{E}_{\lambda_i, c_i} \left( \frac{D_i + D_i^*}{O'_i + O_i'^*} \right)}{\mathbb{E}_{\lambda_i, c_i} \left( \frac{D_i + D_i^*}{O'_i + O_i'^*} \right)} \right] \\
&= \text{Var}_{\lambda_i, c_i} \left[ \frac{(D_i + D_i^*) - \mathbb{E}_{\lambda_i, c_i}(D_i + D_i^*)}{\mathbb{E}_{\lambda_i, c_i}(D_i + D_i^*)} \right] \\
&= \frac{\text{Var}_{\lambda_i, c_i}(D_i + D_i^*)}{[\mathbb{E}_{\lambda_i, c_i}(D_i + D_i^*)]^2} \\
&\approx \frac{\mathbb{E}_{\lambda_i, c_i}(D_i + D_i^*)}{[\mathbb{E}_{\lambda_i, c_i}(D_i + D_i^*)]^2} \\
&= \frac{1}{\mathbb{E}_{\lambda_i, c_i}(D_i + D_i^*)} \\
&= \frac{1}{d_i + \mathbb{E}_{\lambda_i, c_i} D_i^*} \\
&\approx \frac{1}{d_i - \log(c_i) O_i'^*}
\end{aligned}$$

für hinreichend großes  $n_i^*$  und fixiertes  $O_i'^*$  unter  $(\Delta_{i, j_i}, Y_{i, j_i}) = (\delta_{i, j_i}, t_{i, j_i})$  für alle  $j_i \in \{1, \dots, n_i\}$  und  $i \in \{1, 2\}$ .

Als Resultat sei festgehalten:

**Satz 4** (Konvergenz in Verteilung der bedingten Teststatistik). *Unter den Annahmen aus Abschnitt 3.2.2 gilt für hinreichend großes  $n^*$  und fixierte  $O_1^*, O_2^*$  unter  $(\Delta_{i, j_i}, Y_{i, j_i}) = (\delta_{i, j_i}, t_{i, j_i})$  für alle  $j_i \in \{1, \dots, n_i\}$  und  $i \in \{1, 2\}$ :*

$$\frac{W^* - \left[ \log \left( \frac{D_2 - \log(c_2) O_2^*}{O_2' + O_2'^*} \right) - \log \left( \frac{D_1 - \log(c_1) O_1^*}{O_1' + O_1'^*} \right) \right]}{\sqrt{\frac{1}{D_1 - \log(c_1) O_1^*} + \frac{1}{D_2 - \log(c_2) O_2^*}}} \xrightarrow{d} N(0, 1).$$

Schließlich sei auch hier betont, dass die Annahme  $\lambda_1 = \lambda_2$  der Ausgangspunkt der asymptotischen Betrachtungen war.

## Bedingte Power-Funktion

Man definiert nun die bedingte Power-Funktion  $\gamma_c: (0, \infty) \rightarrow [0, 1]$  gemäß:

$$\begin{aligned} \gamma_c(\theta) &= P_{\lambda_1, c_1, \lambda_2, c_2} \left( |W^*| \geq q_{1-\frac{\alpha}{2}} \sqrt{\text{Var}_{\lambda_1, c_1, \lambda_1, c_1} W^*} + E_{\lambda_1, c_1, \lambda_1, c_1} W^* \mid \right. \\ &\quad \left. ((\Delta_{i,j_i}, Y_{i,j_i}) = (\delta_{i,j_i}, t_{i,j_i}))_{\substack{j_i \in \{1, \dots, n_i\}, \\ i \in \{1, 2\}}} \right) \\ &\approx \Phi \left( \frac{q_{\frac{\alpha}{2}} \sigma_{\lambda_1, c_1, \lambda_1, c_1} + \mu_{\lambda_1, c_1, \lambda_1, c_1} - \mu_{\lambda_1, c_1, \lambda_2, c_2}}{\sigma_{\lambda_1, c_1, \lambda_2, c_2}} \right) \\ &\quad + \left[ 1 - \Phi \left( \frac{q_{1-\frac{\alpha}{2}} \sigma_{\lambda_1, c_1, \lambda_1, c_1} + \mu_{\lambda_1, c_1, \lambda_1, c_1} - \mu_{\lambda_1, c_1, \lambda_2, c_2}}{\sigma_{\lambda_1, c_1, \lambda_2, c_2}} \right) \right] \end{aligned}$$

unter den Annahmen aus Abschnitt 3.2.2 für hinreichend großes  $n^*$  und fixierte  $O_1^*, O_2^*, O_{2, \widehat{\lambda}_1}^*$ , mit

$$\begin{aligned} \sigma_{\lambda_1, c_1, \lambda_1, c_1} &= \sqrt{\frac{1}{d_1 - \log(c_1)O_1^*} + \frac{1}{d_2 - \log(c_1)O_{2, \widehat{\lambda}_1}^*}}, \\ \mu_{\lambda_1, c_1, \lambda_1, c_1} &= \log \left( \frac{d_2 - \log(c_1)O_{2, \widehat{\lambda}_1}^*}{o_2' + O_2^*} \right) - \log \left( \frac{d_1 - \log(c_1)O_1^*}{o_1' + O_1^*} \right), \\ \mu_{\lambda_1, c_1, \lambda_2, c_2} &= \log \left( \frac{d_2 - \log(c_2)O_2^*}{o_2' + O_2^*} \right) - \log \left( \frac{d_1 - \log(c_1)O_1^*}{o_1' + O_1^*} \right) \end{aligned}$$

und

$$\sigma_{\lambda_1, c_1, \lambda_2, c_2} = \sqrt{\frac{1}{d_1 - \log(c_1)O_1^*} + \frac{1}{d_2 - \log(c_2)O_2^*}}.$$

Hier beziehen sich  $\text{Var}_{\lambda_1, c_1, \lambda_1, c_1}$  auf die Varianz und  $E_{\lambda_1, c_1, \lambda_1, c_1}$  auf den Erwartungswert unter der Annahme  $\lambda_1 = \lambda_2$  und  $H_0$  mit  $\log c_1 = \log c_2$ , was gleichbedeutend ist mit  $c_1 = c_2$ .

Dabei bezeichnen  $\mu_{\lambda_1, c_1, \lambda_1, c_1}$  den approximierten Erwartungswert und  $\sigma_{\lambda_1, c_1, \lambda_1, c_1}$  die approximierte Standardabweichung der bedingten Teststatistik unter der Nullhypothese, während sich  $\mu_{\lambda_1, c_1, \lambda_2, c_2}$  und  $\sigma_{\lambda_1, c_1, \lambda_2, c_2}$  auf den approximierten Erwartungswert und die approximierte Standardabweichung der bedingten Teststatistik unter der Alternativhypothese beziehen.

### 3.2.3 Alternative: Weibull-Verteilung

Auch wenn die Annahme der exponentialverteilten Überlebenszeiten häufig gerechtfertigt ist, sucht man nach geeigneteren Verteilungen, die eine im Vergleich zur Exponenti-

alverteilung größere Flexibilität aufweisen und sich damit besser an die Daten anpassen lassen.

Eine mögliche Option stellt hierbei die Weibull-Verteilung dar. Deren Überlebensfunktion, Wahrscheinlichkeitsdichtefunktion und Hazardfunktion sind folgendermaßen charakterisiert:

$$\begin{aligned} S^{(0)}(t) &= \exp(-\lambda t^k), \\ f^{(0)}(t) &= \lambda k t^{k-1} \exp(-\lambda t^k) \end{aligned}$$

und

$$h^{(0)}(t) = \lambda k t^{k-1}$$

für alle  $t \geq 0$  mit dem Skalenparameter  $\lambda > 0$  und dem Formparameter  $k > 0$ , mit der Konvention  $0^0 = 1$  für  $t = 0$  und  $k = 1$ .

Hierbei wird ersichtlich, dass man für  $k = 1$  die Exponentialverteilung mit der Rate  $\lambda$  erhält, wodurch die Weibull-Verteilung als Verallgemeinerung der Exponentialverteilung zu verstehen ist und mit dem Formparameter  $k$  für die erwähnte höhere Flexibilität sorgt.

Im Rahmen des Non-Mixture-Modells erhält man schließlich für die Überlebensfunktion und die Hazardfunktion der Gruppe  $i$ :

$$\begin{aligned} S_i(t) &= c_i^{1 - \exp(-\lambda_i t^{k_i})}, \\ h_i(t) &= -\lambda_i k_i t^{k_i-1} \exp(-\lambda_i t^{k_i}) \log c_i \end{aligned}$$

und für das Hazard-Ratio:

$$r(t) = \frac{h_2(t)}{h_1(t)} = \frac{\lambda_2}{\lambda_1} \frac{k_2}{k_1} t^{k_2-k_1} \exp(-(\lambda_2 t^{k_2} - \lambda_1 t^{k_1})) \frac{\log c_2}{\log c_1}$$

für alle  $t \geq 0$  mit dem Skalenparameter  $\lambda_i > 0$ , dem Formparameter  $k_i > 0$  und  $c_i \in (0,1)$  für alle  $i \in \{1,2\}$ .

## Maximum-Likelihood-Schätzung

Wie bisher schätzt man die Parameter wieder über die Maximum-Likelihood-Methode und erhält bei gegebenen  $(t_{i,j_i}, \delta_{i,j_i}) \in [0, \infty) \times \{0,1\}$  für alle  $j_i \in \{1, \dots, n_i\}$  für die logarithmierte Likelihood-Funktion der Gruppe  $i$  gemäß (3) aus Abschnitt 2.1.1 (Seite

16) nach Weglassen des für die Maximierung irrelevanten Vorfaktors:

$$\begin{aligned}
l_i(\lambda_i, k_i, c_i) &= \sum_{j_i=1}^{n_i} [\delta_{i,j_i} \log h_i(t_{i,j_i}) + \log S_i(t_{i,j_i})] \\
&= \sum_{j_i=1}^{n_i} \{ \delta_{i,j_i} [\log \lambda_i + \log k_i + (k_i - 1) \log t_{i,j_i} - \lambda_i t_{i,j_i}^{k_i} + \log(-\log c_i)] \\
&\quad + [1 - \exp(-\lambda_i t_{i,j_i}^{k_i})] \log c_i \}
\end{aligned}$$

für alle  $i \in \{1, 2\}$ .

Um nun zu den Maximum-Likelihood-Schätzern zu gelangen, betrachtet man die partiellen Ableitungen:

$$\begin{aligned}
\frac{\partial}{\partial \lambda_i} l_i(\lambda_i, k_i, c_i) &= \frac{1}{\lambda_i} \sum_{j_i=1}^{n_i} \delta_{i,j_i} - \sum_{j_i=1}^{n_i} \delta_{i,j_i} t_{i,j_i}^{k_i} + \log(c_i) \sum_{j_i=1}^{n_i} t_{i,j_i}^{k_i} \exp(-\lambda_i t_{i,j_i}^{k_i}), \\
\frac{\partial}{\partial k_i} l_i(\lambda_i, k_i, c_i) &= \frac{1}{k_i} \sum_{j_i=1}^{n_i} \delta_{i,j_i} + \sum_{j_i=1}^{n_i} \delta_{i,j_i} \log t_{i,j_i} - \lambda_i \sum_{j_i=1}^{n_i} \delta_{i,j_i} t_{i,j_i}^{k_i} \log t_{i,j_i} \\
&\quad + \lambda_i \log(c_i) \sum_{j_i=1}^{n_i} \exp(-\lambda_i t_{i,j_i}^{k_i}) t_{i,j_i}^{k_i} \log t_{i,j_i}
\end{aligned}$$

und

$$\frac{\partial}{\partial c_i} l_i(\lambda_i, k_i, c_i) = \frac{1}{c_i \log c_i} \sum_{j_i=1}^{n_i} \delta_{i,j_i} + \frac{1}{c_i} \sum_{j_i=1}^{n_i} [1 - \exp(-\lambda_i t_{i,j_i}^{k_i})]$$

für alle  $i \in \{1, 2\}$ .

Schließlich setzt man die partiellen Ableitungen Null, setzt die Zufallsvariablen  $Y_{i,j_i}$  für  $t_{i,j_i}$  und  $\Delta_{i,j_i}$  für  $\delta_{i,j_i}$  für alle  $j_i \in \{1, \dots, n_i\}$  ein und erhält:

$$\begin{aligned}
\hat{\lambda}_i &= \frac{\sum_{j_i=1}^{n_i} \Delta_{i,j_i}}{\sum_{j_i=1}^{n_i} \Delta_{i,j_i} Y_{i,j_i}^{\hat{k}_i} - \log(\hat{c}_i) \sum_{j_i=1}^{n_i} Y_{i,j_i}^{\hat{k}_i} \exp(-\hat{\lambda}_i Y_{i,j_i}^{\hat{k}_i})}, \\
\hat{k}_i &= \left( \sum_{j_i=1}^{n_i} \Delta_{i,j_i} \right) / \left( \hat{\lambda}_i \sum_{j_i=1}^{n_i} \Delta_{i,j_i} Y_{i,j_i}^{\hat{k}_i} \log Y_{i,j_i} - \sum_{j_i=1}^{n_i} \Delta_{i,j_i} \log Y_{i,j_i} \right. \\
&\quad \left. - \hat{\lambda}_i \log(\hat{c}_i) \sum_{j_i=1}^{n_i} \exp(-\hat{\lambda}_i Y_{i,j_i}^{\hat{k}_i}) Y_{i,j_i}^{\hat{k}_i} \log Y_{i,j_i} \right)
\end{aligned}$$

und

$$\hat{c}_i = \exp \left( - \frac{\sum_{j_i=1}^{n_i} \Delta_{i,j_i}}{\sum_{j_i=1}^{n_i} [1 - \exp(-\hat{\lambda}_i Y_{i,j_i}^{\hat{k}_i})]} \right)$$

für alle  $i \in \{1,2\}$ .

### Annahmen

Um ein zeitunabhängiges Hazard-Ratio zu erhalten, setzen wir voraus, dass folgendes gilt:

$$\lambda_1 = \lambda_2$$

und

$$k_1 = k_2.$$

Auch hier müssten diese Annahmen sowohl medizinisch als auch statistisch geprüft werden.

Dann aber erhält man für das Hazard-Ratio:

$$r(t) = \frac{\log c_2}{\log c_1}$$

für alle  $t \geq 0$ , mit der Konvention  $0^0 = 1$  für  $t = 0$ .

Mit

$$\theta = \frac{\log c_2}{\log c_1} > 0$$

und dem entsprechend definierten Schätzer

$$\hat{\theta} = \frac{\log \hat{c}_2}{\log \hat{c}_1}$$

kann nun wie bisher weitergearbeitet werden.

Die Asymptotik soll hier allerdings nicht untersucht werden. Es lässt sich jedoch bei genauerem Hinsehen erkennen, dass man mit der asymptotischen Normalität der Maximum-Likelihood-Schätzer (Greene (2003), 17.4.5.b Asymptotic Normality, Seiten 478/479) und der Delta-Methode (Lemma 1, Abschnitt 2.1.2, Seite 18) wieder zu einem Ziel kommen wird.

### 3.2.4 Alternative: Gammaverteilung

Eine weitere Möglichkeit der Modellierung von Hazardraten, welche sowohl wachsen als auch fallen, erreicht man durch die Verwendung der Gammaverteilung. Die Überlebensfunktion, Wahrscheinlichkeitsdichtefunktion und Hazardfunktion dieser ebenfalls häufig verwendeten Verteilung haben folgende Gestalt:

$$S^{(0)}(t) = 1 - \Gamma^{(0)}(a, bt) = \frac{\Gamma(a) - \gamma(a, bt)}{\Gamma(a)},$$

$$f^{(0)}(t) = \frac{b^a}{\Gamma(a)} t^{a-1} \exp(-bt)$$

und

$$h^{(0)}(t) = \frac{\frac{b^a}{\Gamma(a)} t^{a-1} \exp(-bt)}{1 - \Gamma^{(0)}(a, bt)} = \frac{b^a}{\Gamma(a) - \gamma(a, bt)} t^{a-1} \exp(-bt)$$

für alle  $t \geq 0$  mit dem Formparameter  $a > 0$  und der Rate  $b > 0$ , mit der Konvention  $0^0 = 1$  für  $t = 0$  und  $a = 1$ .

Hier bezeichnen  $\gamma: (0, \infty) \times [0, \infty) \rightarrow [0, \infty)$  die unvollständige Gammafunktion der oberen Grenze,  $\Gamma: (0, \infty) \rightarrow (0, \infty)$  die Gammafunktion und  $\Gamma^{(0)}: (0, \infty) \times [0, \infty) \rightarrow [0, \infty)$  die regularisierte unvollständige Gammafunktion der oberen Grenze gemäß:

$$\gamma(x, y) = \int_0^y s^{x-1} \exp(-s) ds,$$

$$\Gamma(x) = \int_0^\infty s^{x-1} \exp(-s) ds$$

und

$$\Gamma^{(0)}(x, y) = \frac{\gamma(x, y)}{\Gamma(x)}.$$

Für  $a = 1$  erhält man auch hier wieder die Exponentialverteilung, dann mit der Rate  $b$ . Somit stellt auch die Gammaverteilung eine verallgemeinerte Form der Exponentialverteilung dar, wobei die größere Flexibilität der Gammaverteilung über den Formparameter  $a$  gesteuert wird.

Überträgt man dies nun in das Non-Mixture-Modell, so ergeben sich für die Überlebensfunktion und die Hazardfunktion der Gruppe  $i$ :

$$S_i(t) = c_i^{\Gamma^{(0)}(a_i, b_i t)},$$

$$h_i(t) = -\frac{b_i^{a_i}}{\Gamma(a_i)} t^{a_i-1} \exp(-b_i t) \log c_i$$

und für das Hazard-Ratio:

$$r(t) = \frac{h_2(t)}{h_1(t)} = \frac{b_2^{a_2}}{b_1^{a_1}} \frac{\Gamma(a_1)}{\Gamma(a_2)} t^{a_2-a_1} \exp(-(b_2 - b_1)t) \frac{\log c_2}{\log c_1}$$

für alle  $t \geq 0$  mit dem Formparameter  $a_i > 0$ , der Rate  $b_i > 0$  und  $c_i \in (0,1)$  für alle  $i \in \{1,2\}$ .

### Maximum-Likelihood-Schätzung

Nach (3) aus Abschnitt 2.1.1 (Seite 16) ergibt sich bei gegebenen  $(t_{i,j_i}, \delta_{i,j_i}) \in [0, \infty) \times \{0,1\}$  für alle  $j_i \in \{1, \dots, n_i\}$  für die logarithmierte Likelihood-Funktion für die Gruppe  $i$  unter Vernachlässigung des für die Maximierung nicht benötigten Vorfaktors:

$$\begin{aligned} l_i(a_i, b_i, c_i) &= \sum_{j_i=1}^{n_i} [\delta_{i,j_i} \log h_i(t_{i,j_i}) + \log S_i(t_{i,j_i})] \\ &= \sum_{j_i=1}^{n_i} \{ \delta_{i,j_i} [a_i \log b_i - \log \Gamma(a_i) + (a_i - 1) \log t_{i,j_i} - b_i t_{i,j_i} + \log(-\log c_i)] \\ &\quad + \Gamma^{(0)}(a_i, b_i t_{i,j_i}) \log c_i \} \end{aligned}$$

für alle  $i \in \{1,2\}$ .

Dann folgt für die partiellen Ableitungen:

$$\begin{aligned} \frac{\partial}{\partial a_i} l_i(a_i, b_i, c_i) &= \log(b_i) \sum_{j_i=1}^{n_i} \delta_{i,j_i} - \frac{\Gamma'(a_i)}{\Gamma(a_i)} \sum_{j_i=1}^{n_i} \delta_{i,j_i} + \sum_{j_i=1}^{n_i} \delta_{i,j_i} \log t_{i,j_i} \\ &\quad + \log(c_i) \sum_{j_i=1}^{n_i} \frac{\partial}{\partial a_i} \Gamma^{(0)}(a_i, b_i t_{i,j_i}), \\ \frac{\partial}{\partial b_i} l_i(a_i, b_i, c_i) &= \frac{a_i}{b_i} \sum_{j_i=1}^{n_i} \delta_{i,j_i} - \sum_{j_i=1}^{n_i} \delta_{i,j_i} t_{i,j_i} + \log(c_i) \sum_{j_i=1}^{n_i} \frac{\partial}{\partial b_i} \Gamma^{(0)}(a_i, b_i t_{i,j_i}) \end{aligned}$$

und

$$\frac{\partial}{\partial c_i} l_i(a_i, b_i, c_i) = \frac{1}{c_i \log c_i} \sum_{j_i=1}^{n_i} \delta_{i,j_i} + \frac{1}{c_i} \sum_{j_i=1}^{n_i} \Gamma^{(0)}(a_i, b_i t_{i,j_i})$$

für alle  $i \in \{1,2\}$ , mit  $(\log \Gamma)' = \Gamma'/\Gamma: (0, \infty) \rightarrow \mathbb{R}$  als Digammafunktion.

Schließlich erhält man durch Nullsetzen der partiellen Ableitungen und Einsetzen der Zufallsvariablen  $Y_{i,j_i}$  für  $t_{i,j_i}$  und  $\Delta_{i,j_i}$  für  $\delta_{i,j_i}$  für alle  $j_i \in \{1, \dots, n_i\}$ :

$$\frac{\Gamma(\hat{a}_i)}{\Gamma'(\hat{a}_i)} = \frac{\sum_{j_i=1}^{n_i} \Delta_{i,j_i}}{\log(\hat{b}_i) \sum_{j_i=1}^{n_i} \Delta_{i,j_i} + \sum_{j_i=1}^{n_i} \Delta_{i,j_i} \log Y_{i,j_i} + \log(\hat{c}_i) \sum_{j_i=1}^{n_i} \frac{\partial}{\partial a_i} \Gamma^{(0)}(\hat{a}_i, \hat{b}_i Y_{i,j_i})},$$

$$\hat{b}_i = \frac{\hat{a}_i \sum_{j_i=1}^{n_i} \Delta_{i,j_i}}{\sum_{j_i=1}^{n_i} \Delta_{i,j_i} Y_{i,j_i} - \log(\hat{c}_i) \sum_{j_i=1}^{n_i} \frac{\partial}{\partial b_i} \Gamma^{(0)}(\hat{a}_i, \hat{b}_i Y_{i,j_i})}$$

und

$$\hat{c}_i = \exp\left(-\frac{\sum_{j_i=1}^{n_i} \Delta_{i,j_i}}{\sum_{j_i=1}^{n_i} \Gamma^{(0)}(\hat{a}_i, \hat{b}_i Y_{i,j_i})}\right)$$

für alle  $i \in \{1, 2\}$ .

## Annahmen

Unter den medizinisch und statistisch zu prüfenden Annahmen

$$a_1 = a_2$$

und

$$b_1 = b_2$$

ist das Hazard-Ratio wieder zeitunabhängig:

$$r(t) = \frac{\log c_2}{\log c_1}$$

für alle  $t \geq 0$ , mit der Konvention  $0^0 = 1$  für  $t = 0$ .

Letztlich setzt man für das weitere Vorgehen wieder

$$\theta = \frac{\log c_2}{\log c_1} > 0$$

und wählt als Schätzer

$$\hat{\theta} = \frac{\log \hat{c}_2}{\log \hat{c}_1}.$$

Auch hier sollen keine weiteren asymptotischen Betrachtungen erfolgen. Dazu sei wieder auf die asymptotische Normalität der Maximum-Likelihood-Schätzer (Greene (2003),

17.4.5.b Asymptotic Normality, Seiten 478/479) und die Delta-Methode (Lemma 1, Abschnitt 2.1.2, Seite 18) verwiesen.

Somit ist nun das theoretische Fundament gelegt. Der nächste Schritt besteht jetzt darin, die behandelten Modelle auf gegebene Daten anzuwenden und Berechnungen der bedingten Power durchzuführen.

### 3.2.5 Unterschiede zwischen den Verteilungen und deren Einfluss auf die bedingte Power

Dass es sich bei der Exponential-, Weibull- und Gammaverteilung um verschiedene Verteilungen handelt, ist klar. Doch an welcher Stelle macht sich deren Unterschied nun konkret bemerkbar? Dazu können die entsprechenden Hazardfunktionen betrachtet werden. Diese seien dafür an dieser Stelle ohne Index der Gruppe nochmals aufgeführt:

$$\begin{array}{ll}
 \text{Exponential:} & h(t) = \lambda \\
 \text{Non-Mixture-Exponential:} & h(t) = -\lambda \exp(-\lambda t) \log c \\
 \text{Non-Mixture-Weibull:} & h(t) = -\lambda k t^{k-1} \exp(-\lambda t^k) \log c \\
 \text{Non-Mixture-Gamma:} & h(t) = -\frac{b^a}{\Gamma(a)} t^{a-1} \exp(-bt) \log c
 \end{array}$$

für alle  $t \geq 0$ ,  $\lambda > 0$ ,  $k > 0$ ,  $a > 0$ ,  $b > 0$  und  $c \in (0,1)$ .

**Bemerkung 3.** Die Bezeichnungen „Exponential“, „Non-Mixture-Exponential“, „Non-Mixture-Weibull“ und „Non-Mixture-Gamma“ ergeben sich entsprechend der behandelten Modelle aus den Kapiteln 2 und 3.

Hierbei wird ersichtlich, dass die Hazardfunktion im Non-Mixture-Exponentialmodell im Vergleich zum Exponentialmodell nun nicht mehr konstant sondern zeitabhängig ist. Noch flexibler sind dann die Hazardfunktionen bei dem Non-Mixture-Weibull- und Non-Mixture-Gammamodell, was sich aufgrund des jeweils zusätzlichen Verteilungsparameters ergibt. Dabei unterscheiden sich das Non-Mixture-Weibull- und das Non-Mixture-Gammamodell vor allem dadurch, dass bei letzterem der Ausdruck in der Exponentialfunktion lediglich linear ist.

Um zu sehen, wie sich die Wahl der Verteilung auf die bedingte Power auswirkt, betrachtet man das in die Teststatistik eingehende geschätzte Hazard-Ratio unter den Annahmen aus Abschnitt 3.2.2, 3.2.3 beziehungsweise 3.2.4:

$$\hat{\theta} = \frac{\log \hat{c}_2}{\log \hat{c}_1}.$$

Nun schaut man sich an, wie die Schätzer im Rahmen der Maximum-Likelihood-Methode bestimmt wurden:

$$\begin{aligned}
 \text{Non-Mixture-Exponential: } \hat{c} &= \exp\left(-\frac{\sum_{j=1}^n \Delta_j}{\sum_{j=1}^n [1 - \exp(-\hat{\lambda} Y_j)]}\right) \\
 \text{Non-Mixture-Weibull: } \hat{c} &= \exp\left(-\frac{\sum_{j=1}^n \Delta_j}{\sum_{j=1}^n [1 - \exp(-\hat{\lambda} Y_j^k)]}\right) \\
 \text{Non-Mixture-Gamma: } \hat{c} &= \exp\left(-\frac{\sum_{j=1}^n \Delta_j}{\sum_{j=1}^n \Gamma^{(0)}(\hat{a}, \hat{b} Y_j)}\right),
 \end{aligned}$$

wobei der Index der Gruppe wieder weggelassen wurde.

Nimmt man schließlich die Power- und bedingte Power-Funktion des Non-Mixture-Exponentialmodells aus Abschnitt 3.2.2 hinzu, wird deutlich, dass die Wahl der Verteilung und die daraus resultierende unterschiedliche Gestalt des Schätzers für die Überlebensfraktion die bedingte Power beeinflusst. Dabei sei jedoch angemerkt, dass die Power- beziehungsweise bedingte Power-Funktionen für das Non-Mixture-Weibull- und Non-Mixture-Gammamodell in dieser Arbeit nicht angegeben sind und noch berechnet werden müssten.

## 4 Praktische Anwendung

Dieser Teil der Arbeit soll die Verbindung zwischen der Theorie und der Anwendung herstellen. Die hergeleiteten Formeln wurden dabei mit dem Statistikprogramm R (<http://www.r-project.org>) in Form des Paketes CP implementiert.

Zunächst wird im ersten Abschnitt dieses Kapitels noch ein Nachtrag zu den Kapiteln 2 und 3 getätigt. Darin geht es um die Berechnung der zukünftigen Personenzeiten innerhalb der jeweiligen Modelle. Während die zukünftigen Personenzeiten bei dem Exponentialmodell mit konstanten Hazards aus dem zweiten Kapitel, welche mit  $O_1^*$  für Gruppe 1 und  $O_2^*$  für Gruppe 2 bezeichnet wurden, dort explizit zur Berechnung der bedingten Power benötigt werden, braucht man diese Größen bei den Non-Mixture-Modellen im dritten Kapitel nicht direkt. Trotzdem sollen die zukünftigen Personenzeiten in den Überlebenszeitmodellen mit Plateau berechnet werden und als zusätzliche Information dienen. Hierfür werden entsprechende Formeln hergeleitet.

Anschließend wird anhand eines einfachen Beispieldatensatzes getestet, inwiefern sich Änderungen bei den Randparametern der Berechnungen, also zum Beispiel Änderungen bei der Restlaufzeit der Studie oder bei der Rekrutierungsrate, auf die bedingte Power auswirken.

### 4.1 Berechnung der zukünftigen Personenzeiten

Die konkrete Berechnung der zukünftigen Personenzeiten vom Zeitpunkt der Zwischenauswertung bis zum Ende der Studie anhand der Daten aus der Zwischenauswertung spielte bei den theoretischen Überlegungen noch keine Rolle. Allerdings benötigt man diese Größen im praktischen Teil nun explizit zur Berechnung der bedingten Power beziehungsweise als Zusatzinformation, wie zu Beginn dieses Kapitels erwähnt.

#### 4.1.1 Exponentialmodell

Den folgenden Überlegungen liegt eine Arbeit von Andersen aus dem Jahr 1987 zugrunde.

Man befindet sich zeitlich gesehen bei einer Zwischenauswertung und betrachtet genau eine der beiden Gruppen, sodass man bei den folgenden Rechnungen den Gruppenindex weglassen kann. Die Überlegungen gelten dabei natürlich für jede der beiden Gruppen,

gegebenenfalls mit verschiedenen Parametern.

Zur Berechnung der zukünftigen Personenzzeit verwendet man zwei stetig differenzierbare Funktionen  $f_{\text{old}}, f_{\text{new}} : [0, \infty) \rightarrow [0, \infty)$ .

Dabei bezeichnet  $f_{\text{old}}$  die Anzahl der „alten“ Patienten, also derjenigen, welche zu Beginn der Studie dabei waren und dies auch immer noch sind. Mit  $n_{\text{al}} \in \mathbb{N} \cup \{0\}$  als Anzahl der Patienten, welche von Beginn an dabei sind und bei denen noch kein Ereignis eingetreten ist, gilt dann:

$$f_{\text{old}}(0) = n_{\text{al}}.$$

Dabei steht „al“ für das englische Wort „alive“ (deutsch: am Leben). Hier sei angemerkt, dass mit dem Zeitpunkt  $t = 0$  auf den Zeitpunkt der Zwischenauswertung Bezug genommen wird.

Mit  $f_{\text{new}}$  gibt man die Anzahl der „neuen“ Patienten an, also die Anzahl der seit der Zwischenauswertung rekrutierten und noch lebenden Patienten vom Zeitpunkt der Zwischenauswertung an. Hier gilt:

$$f_{\text{new}}(0) = 0,$$

da zum Zeitpunkt der Zwischenauswertung noch keine neuen Patienten rekrutiert wurden.

Weiter sei die Konstante  $q > 0$  die Summe der Drop-Out-Rate und der Todesrate.

Die Idee ist nun, die zukünftigen Personenzeiten mit den berechneten Raten aus den Daten der Zwischenauswertung zu bestimmen.

Dabei sei angemerkt, dass die Anzahl der Patienten eine natürliche Zahl einschließlich der Null ist und im Gegensatz dazu die Funktionen  $f_{\text{old}}$  und  $f_{\text{new}}$  als stetig differenzierbar, also insbesondere stetig, vorausgesetzt sind. Dies ist jedoch aus rechentechnischen Gründen nötig. Die berechneten Werte müssten gegebenenfalls gerundet werden.

## Alte Patienten

Dann erhält man für die alten Patienten folgende Gleichung:

$$f_{\text{old}}(t + s) = f_{\text{old}}(t) - q f_{\text{old}}(t)s$$

für alle  $t \geq 0$  und  $s \geq 0$ .

Damit folgt:

$$\frac{f_{\text{old}}(t + s) - f_{\text{old}}(t)}{(t + s) - t} = -q f_{\text{old}}(t)$$

für alle  $t \geq 0$  und  $s > 0$  und mit der stetigen Differenzierbarkeit von  $f_{\text{old}}$ :

$$f'_{\text{old}}(t) = -q f_{\text{old}}(t)$$

für alle  $t \geq 0$  und  $s \downarrow 0$ .

Die Lösung dieser homogenen Differentialgleichung mit obigem Anfangswert ergibt sich dann zu:

$$f_{\text{old}}(t) = n_{\text{al}} \exp(-q t)$$

für alle  $t \geq 0$ .

Schließlich erhält man als zukünftige Personenzeit  $o_{\text{old}}^*: [0, \infty) \rightarrow [0, \infty)$  der alten Patienten:

$$\begin{aligned} o_{\text{old}}^*(t) &= \int_0^t f_{\text{old}}(s) \, ds \\ &= n_{\text{al}} \int_0^t \exp(-q s) \, ds \\ &= \frac{n_{\text{al}}}{q} [1 - \exp(-q t)]. \end{aligned}$$

## Neue Patienten

Für die neuen Patienten hingegen erhält man mit der konstanten Rekrutierungsrate  $r \geq 0$  die Gleichung:

$$f_{\text{new}}(t + s) = f_{\text{new}}(t) - q f_{\text{new}}(t) s + r s$$

für alle  $t \geq 0$  und  $s \geq 0$  oder analog:

$$\frac{f_{\text{new}}(t + s) - f_{\text{new}}(t)}{(t + s) - t} = -q f_{\text{new}}(t) + r$$

für alle  $t \geq 0$  und  $s > 0$ .

Für den Limes  $s \downarrow 0$  folgt dann mit der stetigen Differenzierbarkeit von  $f_{\text{new}}$ :

$$f'_{\text{new}}(t) = -q f_{\text{new}}(t) + r$$

für alle  $t \geq 0$ .

Diese inhomogene Differentialgleichung löst man nun in zwei Schritten:

Zunächst erhält man für die homogene Differentialgleichung in äquivalenter Weise zu der Differentialgleichung der alten Patienten:

$$f_{\text{new}}^{\text{hom}}(t) = d \exp(-q t)$$

für alle  $t \geq 0$  und eine Konstante  $d \geq 0$ .

Im zweiten Schritt lässt man die Konstante  $d$  variieren, das heißt man betrachtet  $d: [0, \infty) \rightarrow [0, \infty)$  als stetig differenzierbare Funktion und berechnet mit

$$f_{\text{new}}^{\text{inh}}(t) = d(t) \exp(-q t)$$

für alle  $t \geq 0$ :

$$\begin{aligned} (f_{\text{new}}^{\text{inh}})'(t) &= d'(t) \exp(-q t) - q d(t) \exp(-q t) \\ &= -q f_{\text{new}}^{\text{inh}}(t) + d'(t) \exp(-q t) \end{aligned}$$

für alle  $t \geq 0$ .

Nach der Gleichung für die neuen Patienten muss gelten:

$$r = d'(t) \exp(-q t)$$

für alle  $t \geq 0$ .

Dies ist für

$$d(t) = \frac{r}{q} \exp(q t) + e$$

für alle  $t \geq 0$  und eine Konstante  $e \geq 0$  erfüllt.

Die Lösung ergibt sich letztlich mit obigem Anfangswert, und entsprechend  $e = -r/q$ , als:

$$\begin{aligned} f_{\text{new}}(t) &= f_{\text{new}}^{\text{inh}}(t) \\ &= d(t) \exp(-q t) \\ &= \frac{r}{q} + e \exp(-q t) \\ &= \frac{r}{q} - \frac{r}{q} \exp(-q t) \\ &= \frac{r}{q} [1 - \exp(-q t)] \end{aligned}$$

für alle  $t \geq 0$ .

Als Personenzeit  $o_{\text{new}}^* : [0, \infty) \rightarrow [0, \infty)$  der neuen Patienten erhält man:

$$\begin{aligned}
 o_{\text{new}}^*(t) &= \int_0^t f_{\text{new}}(s) \, ds \\
 &= \frac{r}{q} \int_0^t [1 - \exp(-q s)] \, ds \\
 &= \frac{r}{q} t - \frac{r}{q} \int_0^t \exp(-q s) \, ds \\
 &= \frac{r}{q} t - \frac{r}{q^2} [1 - \exp(-q t)].
 \end{aligned}$$

### Alle Patienten

Somit lässt sich die zukünftige Personenzeit  $o^* : [0, \infty) \rightarrow [0, \infty)$  aller Patienten darstellen als:

$$\begin{aligned}
 o^*(t) &= o_{\text{old}}^*(t) + o_{\text{new}}^*(t) \\
 &= \frac{n_{\text{al}}}{q} [1 - \exp(-q t)] \\
 &\quad + \frac{r}{q} t - \frac{r}{q^2} [1 - \exp(-q t)].
 \end{aligned}$$

#### 4.1.2 Non-Mixture-Modell

Legt man nun eines der Non-Mixture-Modelle zugrunde, so ergibt sich mit den bisherigen Bezeichnungen:

$$f_{\text{old}}(t + s) = f_{\text{old}}(t) - q(t)f_{\text{old}}(t)s, \quad f_{\text{old}}(0) = n_{\text{al}}$$

und

$$f_{\text{new}}(t + s) = f_{\text{new}}(t) - q(t)f_{\text{new}}(t)s + r s, \quad f_{\text{new}}(0) = 0$$

für alle  $t \geq 0$  und  $s \geq 0$ .

Der Unterschied zu dem Exponentialmodell mit konstanten Hazards besteht darin, dass hier die Summe aus Drop-Out-Rate und Todesrate aufgrund der Zeitabhängigkeit der Hazardrate im Non-Mixture-Modell ebenfalls abhängig von der Zeit ist. Konkret betrachtet man eine stetige Funktion  $q: [0, \infty) \rightarrow [0, \infty)$ . Damit werden sich die entsprechenden Differentialgleichungen zwar analog zu dem vorangegangenen Abschnitt lösen lassen, jedoch auch eine komplexere Struktur aufweisen.

### Alte Patienten

Man betrachtet zunächst wieder die zukünftige Personenzeit der alten Patienten. Die entsprechende Differentialgleichung lautet dann:

$$f'_{\text{old}}(t) = -q(t)f_{\text{old}}(t)$$

für alle  $t \geq 0$ .

Die entsprechende Lösung ergibt sich mit dem obigen Anfangswert als:

$$f_{\text{old}}(t) = n_{\text{al}} \exp \left( - \int_0^t q(s) \, ds \right)$$

für alle  $t \geq 0$ .

Damit erhält man für die zukünftige Personenzeit der alten Patienten:

$$\begin{aligned} o_{\text{old}}^*(t) &= \int_0^t f_{\text{old}}(s) \, ds \\ &= n_{\text{al}} \int_0^t \exp \left( - \int_0^s q(u) \, du \right) \, ds \end{aligned}$$

für alle  $t \geq 0$ .

### Neue Patienten

Für die neuen Patienten ergibt sich wieder eine inhomogene Differentialgleichung, dieses Mal in der Form:

$$f'_{\text{new}}(t) = -q(t)f_{\text{new}}(t) + r$$

für alle  $t \geq 0$ .

Dann erhält man für die homogene Differentialgleichung:

$$f_{\text{new}}^{\text{hom}}(t) = d \exp \left( - \int_0^t q(s) \, ds \right)$$

für alle  $t \geq 0$  und eine Konstante  $d \geq 0$ .

Als Ansatz für die inhomogene Differentialgleichung wählt man mit  $d: [0, \infty) \rightarrow [0, \infty)$

$$f_{\text{new}}^{\text{inh}}(t) = d(t) \exp \left( - \int_0^t q(s) \, ds \right)$$

und berechnet

$$\begin{aligned} (f_{\text{new}}^{\text{inh}})'(t) &= d'(t) \exp \left( - \int_0^t q(s) \, ds \right) - q(t) d(t) \exp \left( - \int_0^t q(s) \, ds \right) \\ &= -q(t) f_{\text{new}}^{\text{inh}}(t) + d'(t) \exp \left( - \int_0^t q(s) \, ds \right) \end{aligned}$$

für alle  $t \geq 0$ .

Nun muss folgendes gelten:

$$r = d'(t) \exp \left( - \int_0^t q(s) \, ds \right)$$

für alle  $t \geq 0$ .

Dies erfüllt:

$$d(t) = r \int_0^t \exp \left( \int_0^s q(u) \, du \right) \, ds + e$$

für alle  $t \geq 0$  und eine Konstante  $e \geq 0$ .

Dann ergibt sich als Lösung mit obigem Anfangswert und gemäß diesem mit  $e = 0$ :

$$\begin{aligned}
 f_{\text{new}}(t) &= f_{\text{new}}^{\text{inh}}(t) \\
 &= d(t) \exp\left(-\int_0^t q(s) \, ds\right) \\
 &= r \exp\left(-\int_0^t q(s) \, ds\right) \left[ \int_0^t \exp\left(\int_0^s q(u) \, du\right) \, ds \right] + e \exp\left(-\int_0^t q(s) \, ds\right) \\
 &= r \exp\left(-\int_0^t q(s) \, ds\right) \left[ \int_0^t \exp\left(\int_0^s q(u) \, du\right) \, ds \right]
 \end{aligned}$$

für alle  $t \geq 0$ .

Somit hat man hinsichtlich der Personenzeit der neuen Patienten:

$$\begin{aligned}
 o_{\text{new}}^*(t) &= \int_0^t f_{\text{new}}(s) \, ds \\
 &= r \int_0^t \exp\left(-\int_0^s q(u) \, du\right) \left[ \int_0^s \exp\left(\int_0^u q(v) \, dv\right) \, du \right] \, ds
 \end{aligned}$$

für alle  $t \geq 0$ .

### Alle Patienten

Damit ergibt sich schließlich insgesamt:

$$\begin{aligned}
 o^*(t) &= o_{\text{old}}^*(t) + o_{\text{new}}^*(t) \\
 &= n_{\text{al}} \int_0^t \exp\left(-\int_0^s q(u) \, du\right) \, ds \\
 &\quad + r \int_0^t \exp\left(-\int_0^s q(u) \, du\right) \left[ \int_0^s \exp\left(\int_0^u q(v) \, dv\right) \, du \right] \, ds
 \end{aligned}$$

für alle  $t \geq 0$ .

## 4.2 Sensitivitätsanalyse der bedingte Power-Funktionen

Hier soll nun anhand des Beispieldatensatzes `test` aus dem R-Paket `CP` untersucht werden, welche Änderungen sich bei der bedingten Power bei Variation der Restlaufzeit der Studie, der Anzahl der rekrutierten Patienten, des erwarteten Effekts und des Signifikanzniveaus ergeben.

Aus dem Datensatz `test` erhält man folgende Daten zur Zwischenauswertung, welche in Tabelle 1 zusammengefasst sind:

Tabelle 1: Daten aus der Zwischenauswertung

	A	B	Gesamt
Patienten	99	102	201
Todesfälle	51	60	111
Zensierungen	48	42	90
Personenzeit (Monate)	159	173	332

Für die folgenden Betrachtungen wird von einer Rekrutierungsrate von 2,5 Patienten pro Gruppe und Monat ausgegangen mit einer Restlaufzeit der Studie von 12 Monaten. Hierbei wird angenommen, dass es keine Drop-Outs gibt. Dann erhält man am Studienende eine Gesamtzahl von 261 Patienten.

Als klinisch relevant sei ein Hazard-Ratio (Hazard der Gruppe B im Verhältnis zum Hazard der Gruppe A) von 0,75 angesehen.

Das Signifikanzniveau liege bei dem üblichen Wert von 0,05.

Die entsprechenden Kaplan-Meier-Kurven, die geschätzten Überlebenskurven gemäß der hier behandelten vier Modelle und die bedingte Power-Kurven sind in Abbildung 1 dargestellt:

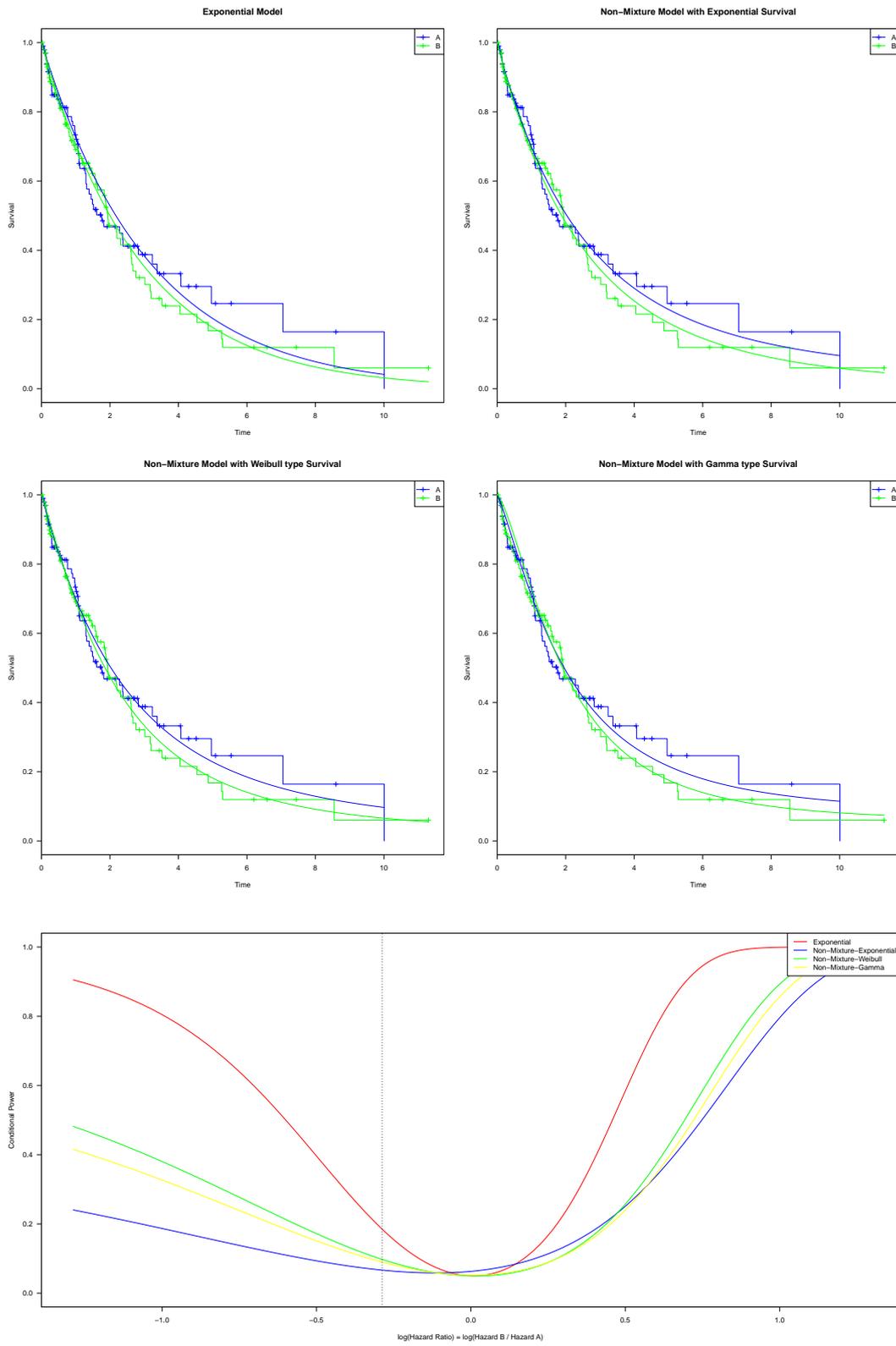


Abbildung 1: Überlebenskurven und bedingte Power-Kurven

Die berechneten Werte können Tabelle 2 entnommen werden:

Tabelle 2: Weitere Daten aus der Zwischenauswertung

		A	B
Exponential $\sum \text{AIC} = 469,6533$ $\hat{\theta} = 1,0869$	$l(\text{MLE})$	-109,2725	-123,5542
	AIC	220,5450	249,1083
	$\hat{\lambda}$	0,3190	0,3467
	$O^*$	217	185
Non-Mixture-Exponential $\sum \text{AIC} = 471,1232$ $\hat{\theta} = 1,5639$	$l(\text{MLE})$	-108,5107	-123,0509
	AIC	221,0213	250,1019
	$\hat{\lambda}$	0,1041	0,0691
	$\hat{c}$	0,0265	0,0034
Non-Mixture-Weibull $\sum \text{AIC} = 474,6836$ $\hat{\theta} = 1,0787$	$l(\text{MLE})$	-108,5070	-122,8348
	AIC	223,0140	251,6696
	$\hat{\lambda}$	0,1093	0,1042
	$\hat{k}$	1,0124	1,0992
Non-Mixture-Gamma $\sum \text{AIC} = 479,3775$ $\hat{\theta} = 1,1449$	$l(\text{MLE})$	-109,1316	-124,5571
	AIC	224,2633	255,1142
	$\hat{a}$	1,2357	1,4347
	$\hat{b}$	0,2481	0,2896
	$\hat{c}$	0,0847	0,0592
	$O^*$	262	219

Hier soll die Bezeichnung  $l(\text{MLE})$  für den Wert der logarithmierten Likelihood-Funktion mit den berechneten Maximum-Likelihood-Schätzern als Argument stehen.

Desweiteren wird als Entscheidungshilfe bezüglich der Modellwahl Akaikes Informationskriterium (Akaike (1974)), kurz AIC, verwendet. Dabei ist das Modell mit dem niedrigsten Gesamt-AIC-Wert zu bevorzugen.

Zu beachten ist, dass in Tabelle 2 alle Werte - bis auf die zukünftigen Personenzeiten - auf vier Nachkommastellen gerundet sind.

Wiederholend sei erwähnt, dass sich  $O^*$  jeweils auf die zukünftige Personenzzeit bezieht.

Die Werte der bedingten Power sind in Tabelle 3 aufgeführt:

Tabelle 3: Bedingte Power

	Bedingte Power
Exponential	0,1851
Non-Mixture-Exponential	0,0665
Non-Mixture-Weibull	0,0976
Non-Mixture-Gamma	0,0899

Nach Akaikes Informationskriterium wäre das Exponentialmodell zu bevorzugen. Dies ist plausibel, da bei Betrachtung der Kaplan-Meier-Kurven in Abbildung 1 Überlebensfraktionen kaum zu erkennen sind. Dadurch können die Überlebenskurven der Non-Mixture-Modelle nach dem AIC-Wert zu urteilen nur unwesentlich besser an die Daten angepasst werden, bei gleichzeitig erhöhter Modellkomplexität im Vergleich zum Exponentialmodell. Allerdings sind die Unterschiede zwischen den AIC-Werten in Tabelle 2, insbesondere beim Vergleich des Exponential-, Non-Mixture-Exponential- und Non-Mixture-Weibull-Modells eher gering.

Weiter ist zu erwähnen, dass die Voraussetzungen bei den asymptotischen Untersuchungen der Non-Mixture-Modelle annähernd erfüllt sind, wenn man die geschätzten Parameter in Tabelle 2 betrachtet. Zu den Voraussetzungen zählten bei dem Non-Mixture-Exponentialmodell die Annahme  $\lambda_1 = \lambda_2$ , bei dem Non-Mixture-Weibull-Modell die Annahmen  $\lambda_1 = \lambda_2$  und  $k_1 = k_2$  und beim Non-Mixture-Gammamodell die Annahmen  $a_1 = a_2$  und  $b_1 = b_2$ .

Bei der berechneten bedingten Power in Tabelle 3 wird ersichtlich, dass es eine etwas größere Abweichung zwischen dem Wert bei dem Exponentialmodell und den Werten bei den Non-Mixture-Modellen gibt. Somit geht der bisher beobachtete negative Effekt bei dem Exponentialmodell weniger stark in die bedingte Power ein als bei den Non-Mixture-Modellen.

Im Folgenden werden nun nacheinander verschiedene Parameter variiert, wobei die restlichen Parameter entsprechend der Ausgangswerte fixiert bleiben. Dabei wird jeweils die bedingte Power unter obigen Randparametern zum Vergleich mit angegeben.

Es ergibt sich für die bedingte Power bei Variation der Restlaufzeit der Studie:

Tabelle 4: Bedingte Power in Abhängigkeit von der Restlaufzeit der Studie

	6 Monate	<b>12 Monate</b>	18 Monate
Exponential	0,1275	<b>0,1851</b>	0,2376
Non-Mixture-Exponential	0,0635	<b>0,0665</b>	0,0700
Non-Mixture-Weibull	0,0879	<b>0,0976</b>	0,1080
Non-Mixture-Gamma	0,0818	<b>0,0899</b>	0,0987

Bei Änderung der Rekrutierungsraten ergeben sich für die bedingte Power:

Tabelle 5: Bedingte Power in Abhängigkeit von den Rekrutierungsraten

	0 Patienten pro Gruppe und Monat	<b>2,5 Patienten pro Gruppe und Monat</b>	5 Patienten pro Gruppe und Monat
Exponential	0,1170	<b>0,1851</b>	0,2632
Non-Mixture-Exponential	0,0604	<b>0,0665</b>	0,0746
Non-Mixture-Weibull	0,0777	<b>0,0976</b>	0,1211
Non-Mixture-Gamma	0,0733	<b>0,0899</b>	0,1098

Betrachtet man die bedingte Power, die sich bei Änderung des im vorhinein festgelegten klinisch relevanten Effekts einstellt, so erhält man:

Tabelle 6: Bedingte Power in Abhängigkeit vom Hazard-Ratio

	Hazard-Ratio = 0,5	<b>Hazard-Ratio = 0,75</b>	Hazard-Ratio = 0,9
Exponential	0,5903	<b>0,1851</b>	0,0711
Non-Mixture-Exponential	0,1271	<b>0,0665</b>	0,0584
Non-Mixture-Weibull	0,2527	<b>0,0976</b>	0,0580
Non-Mixture-Gamma	0,2184	<b>0,0899</b>	0,0578

Schließlich berechnet sich die bedingte Power bei Variation des Signifikanzniveaus zu:

Tabelle 7: Bedingte Power in Abhängigkeit vom Signifikanzniveau

	Signifikanz- niveau = 0,01	<b>Signifikanz- niveau = 0,05</b>	Signifikanz- niveau = 0,1
Exponential	0,0672	<b>0,1851</b>	0,2795
Non-Mixture-Exponential	0,0159	<b>0,0665</b>	0,1237
Non-Mixture-Weibull	0,0274	<b>0,0976</b>	0,1671
Non-Mixture-Gamma	0,0246	<b>0,0899</b>	0,1563

Hier wird bei Betrachtung der Tabellen 4 bis 7 vor allem deutlich, dass die bedingte Power bei dem Exponentialmodell wesentlich sensibler auf die Änderungen der Randparameter reagiert als dies bei den Non-Mixture-Modellen der Fall ist. Dies kann damit begründet werden, dass die bedingte Power-Funktionen bei den Non-Mixture-Modellen entsprechend diesen eine komplexere aber zugleich stabilere Struktur aufweisen.

## 5 Futility-Analyse am Beispiel zweier klinischer Studien

In diesem Kapitel sollen nun die theoretischen Resultate auf reale Daten angewendet werden.

Dazu werden zwei verschiedene randomisierte, klinische Studien herangezogen: die High-CHOEP-Studie, welche vom 1. Juli 2000 bis 4. Dezember 2003 lief, (Pfreundschuh et al. (2007)) und die R-Mega-CHOEP-Studie, die im Zeitraum vom 3. März 2003 bis 7. April 2009 erfolgte (Schmitz et al. (2012)).

Was jeweils Gegenstand der genannten Studien ist, soll dabei im Folgenden etwas näher erläutert werden.

Anschließend finden jeweils bedingte Power-Analysen statt, in der die verschiedenen hier behandelten Modelle zur Berechnung der bedingten Power miteinander verglichen werden.

### 5.1 Inhalt der klinischen Studien

Non-Hodgkin-Lymphome bezeichnen alle malignen Lymphome, welche keine Hodgkin-Lymphome sind. Dabei werden mit malignen Lymphomen alle bösartigen Erkrankungen des lymphatischen Systems bezeichnet. Damit sind dann konkret Erkrankungen der lymphatischen Organe oder der lymphatischen Zellen gemeint. Zu den lymphatischen Organen zählen zum Beispiel die Lymphknoten und die Milz, während mit lymphatischen Zellen die sogenannten B- und T-Zellen, welche zu den weißen Blutkörperchen gehören, gemeint sind. Hodgkin-Lymphome sind hierbei durch das Vorkommen von sogenannten Hodgkin-Reed-Sternberg-Zellen charakterisiert, wodurch sie von Non-Hodgkin-Lymphomen, bei denen solche Zellen nicht vorkommen, unterschieden werden.

Man unterscheidet bei Non-Hodgkin-Lymphomen zwischen den Begriffen niedrigmaligne und hochmaligne Lymphome. Niedrigmaligne Lymphome sind langsam wachsende maligne Lymphome. Die Chancen auf Remission sind in diesem Fall gering, zumal kaum kurative Ansätze zur Behandlung existieren. Im Gegensatz dazu bezeichnet man mit hochmalignen Lymphomen schnell wachsende maligne Lymphome. Diese sind gut therapierbar, mit einer guten Aussicht auf Remission, selbst in fortgeschrittenen Stadien dieser Krebserkrankung.

Hochmaligne Non-Hodgkin-Lymphome werden in der Regel nach einem CHOP-artigen Schema behandelt, wobei die Bezeichnung von den verwendeten Arzneistoffen Cyclophosphamid, Hydroxydaunorubicin, Oncovin und Prednison herrührt. Bei den beiden

hier betrachteten Studien wurde zusätzlich die Substanz Etoposid verabreicht. In diesem Fall verwendet man dann häufig die Bezeichnung CHOEP.

In den Studien wurde nun untersucht, ob sich mit einer moderaten (High-CHOEP) beziehungsweise starken Erhöhung der Arzneimitteldosierung (R-Mega-CHOEP) im Vergleich zur Standarddosierung der Medikamente bessere Therapieergebnisse bei hochmalignen Non-Hodgkin-Lymphomen erzielen lassen. Das „R“ bei der R-Mega-CHOEP-Studie verdeutlicht, dass hier zusätzlich der Arzneistoff Rituximab verwendet wurde. Mit diesem Antikörper hat sich in verschiedenen Studien bezüglich der Behandlung von hochmalignen Non-Hodgkin-Lymphomen mittels CHOP- beziehungsweise CHOP-artiger Schemata eine bessere Wirksamkeit gezeigt.

Die hier verwendeten Informationen stammen von den Internetseiten des Kompetenznetz Maligne Lymphome e. V. (<https://www.lymphome.de>). Für ausführlichere und weitere Fakten bezüglich des Krankheitsbildes von hochmalignen Non-Hodgkin-Lymphomen und deren Therapiemöglichkeiten sei ebenfalls auf das Online-Angebot des Kompetenznetz Maligne Lymphome e. V. (<https://www.lymphome.de>) verwiesen.

## 5.2 Futility-Analyse

Die in den Kapiteln 2 und 3 behandelten Modelle zur Berechnung der bedingten Power sollen nun auf die Daten angewendet werden, die bei beiden Studien jeweils im Rahmen von Zwischenauswertungen erhoben wurden.

Die bedingte Power gibt dabei die Wahrscheinlichkeit an, gegeben die Daten aus der Zwischenauswertung, ein signifikantes Ergebnis am Ende der Studie zu erhalten, unter der Annahme, dass der unbekannte, wahre Effekt gleich dem angenommenen Effekt ist. Die bedingte Power kann daher als ein möglicher Indikator hinsichtlich der Fragestellung verwendet werden, ob die Studie vor dem geplanten Studienende abgebrochen oder fortgeführt werden sollte.

Im Folgenden werden zunächst für die High-CHOEP-Studie die Daten der Zwischenauswertung zusammengefasst, die Kaplan-Meier-Kurven mit den geschätzten Überlebenskurven bezüglich der Modelle verglichen und die bedingte Power-Funktionen gemäß des Exponential-, Non-Mixture-Exponential-, Non-Mixture-Weibull- und Non-Mixture-Gammamodells berechnet. Anschließend erfolgen analoge Betrachtungen für die R-Mega-CHOEP-Studie. Für die Berechnungen wird jeweils die Funktion `CompSurvMod` aus dem R-Paket `CP` verwendet.

## High-CHOEP

Im Zuge der Zwischenauswertung ergaben sich folgende Daten, welche in Tabelle 8 dargestellt sind:

Tabelle 8: Daten aus der Zwischenauswertung

	CHOEP	highCHOEP	Gesamt
Patienten	118	115	233
Todesfälle	33	36	69
Zensierungen	85	79	164
Personenzeit (Monate)	2191	2115	4306

Im Weiteren geht man hier für die Berechnung der bedingten Power von einer durchschnittlichen Rekrutierung von 11 Patienten pro Monat (5,5 pro Gruppe) über einen Zeitraum von 15 Monaten aus. Dabei wird angenommen, dass es zu keinen Drop-Outs kommt. Dies ergibt dann eine Gesamtzahl von 398 Patienten am Ende der folgenden 15 Monate.

Als klinisch relevanter Effekt wird ein Hazard-Ratio (Hazard der Gruppe highCHOEP im Verhältnis zum Hazard der Gruppe CHOEP) von 0,653 angenommen.

Das Signifikanzniveau liegt bei 0,05.

Man erhält gemäß der Modelle in Abbildung 2 folgende geschätzte Überlebenskurven, jeweils im Vergleich mit den Kaplan-Meier-Kurven, und bedingte Power-Kurven:

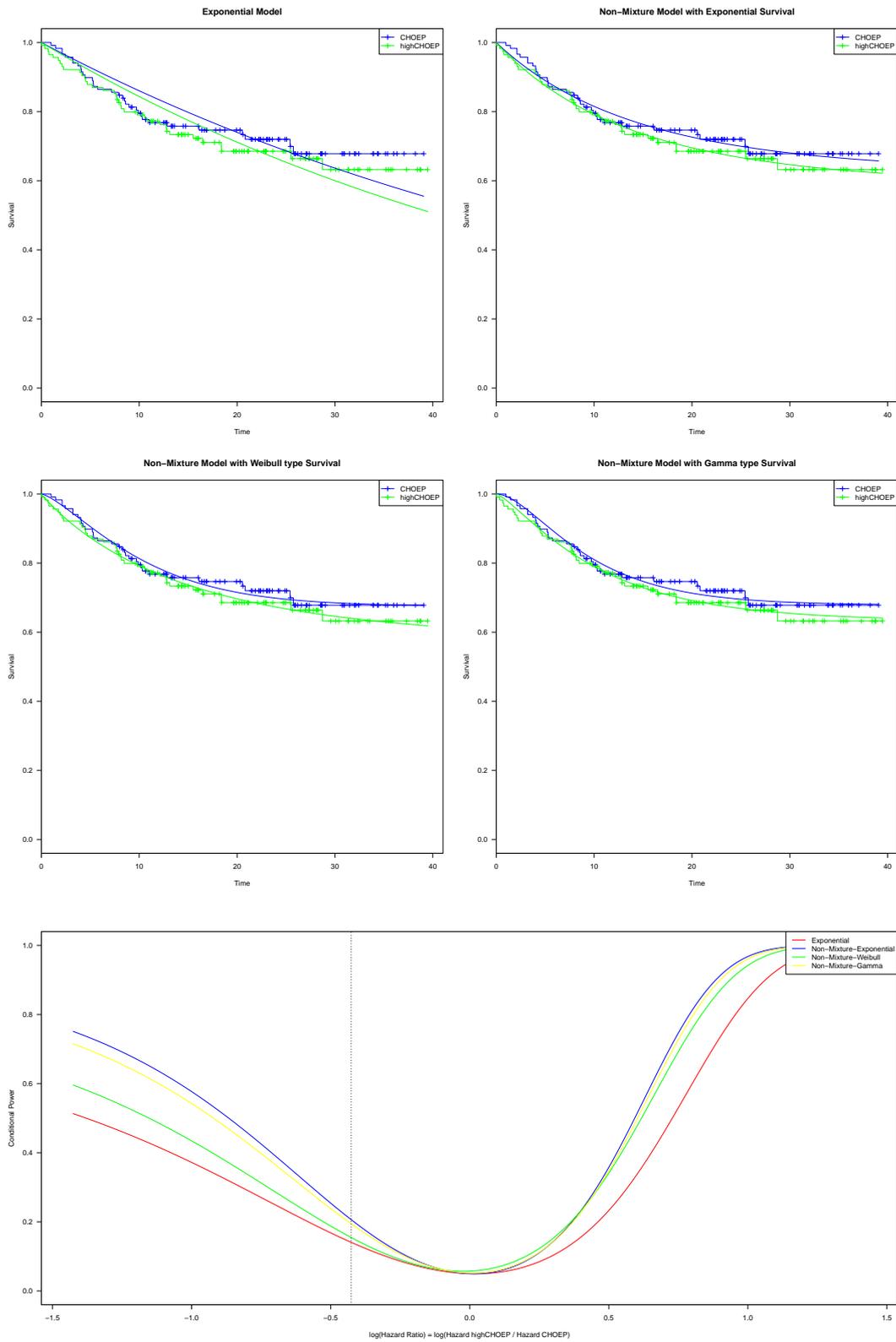


Abbildung 2: Überlebenskurven und bedingte Power-Kurven

Entsprechende Daten enthält Tabelle 9:

Tabelle 9: Weitere Daten aus der Zwischenauswertung

		CHOEP	highCHOEP
Exponential $\sum \text{AIC} = 712,2286$ $\hat{\theta} = 1,1300$	$l(\text{MLE})$	-171,4636	-182,6507
	AIC	344,9271	367,3015
	$\hat{\lambda}$	0,0151	0,0170
	$O^*$	1715	1615
Non-Mixture-Exponential $\sum \text{AIC} = 702,9311$ $\hat{\theta} = 1,1239$	$l(\text{MLE})$	-168,3723	-179,0933
	AIC	340,7445	362,1866
	$\hat{\lambda}$	0,0571	0,0582
	$\hat{c}$	0,6250	0,5896
	$O^*$	1663	1558
Non-Mixture-Weibull $\sum \text{AIC} = 704,0468$ $\hat{\theta} = 1,4055$	$l(\text{MLE})$	-166,9485	-179,0749
	AIC	339,8970	364,1498
	$\hat{\lambda}$	0,0324	0,0599
	$\hat{k}$	1,3638	0,9613
	$\hat{c}$	0,6756	0,5763
	$O^*$	1662	1558
Non-Mixture-Gamma $\sum \text{AIC} = 705,8888$ $\hat{\theta} = 1,1804$	$l(\text{MLE})$	-166,5747	-180,3697
	AIC	339,1494	366,7394
	$\hat{a}$	1,9086	1,3468
	$\hat{b}$	0,1725	0,1068
	$\hat{c}$	0,6785	0,6326
	$O^*$	1661	1558

Der Ausdruck  $l(\text{MLE})$  bezeichnet wieder den Wert der logarithmierten Likelihood-Funktion mit den berechneten Maximum-Likelihood-Schätzern als Argument.

Hinsichtlich der Modellwahl wird Akaikes Informationskriterium (Akaike (1974)) herangezogen. Dabei sollte das Modell mit dem kleinsten Gesamt-AIC-Wert gewählt werden.

Auch hier sei darauf hingewiesen, dass in Tabelle 9 alle Werte - außer die zukünftigen Personenzeiten - auf vier Nachkommastellen gerundet sind.

$O^*$  bezieht sich wieder auf die jeweilige zukünftige Personenzzeit.

Hier ergeben sich für die bedingte Power für ein Hazard-Ratio von 0,653 beziehungsweise das entsprechend logarithmierte Hazard-Ratio  $\log(0,653) \approx -0,426$  in Tabelle 10:

Tabelle 10: Bedingte Power

	Bedingte Power
Exponential	0,1406
Non-Mixture-Exponential	0,2062
Non-Mixture-Weibull	0,1546
Non-Mixture-Gamma	0,1943

Bei Verwendung der AIC-Werte aus Tabelle 9 als Indikator hinsichtlich der Modellwahl entscheidet man sich für das Non-Mixture-Exponentialmodell. Während sich die Überlebenskurven des Exponentialmodells aufgrund der fehlenden Überlebensfraktion in diesem Modell nur unzureichend an die Daten anpassen lassen, bringt die höhere Flexibilität innerhalb des Non-Mixture-Weibull- beziehungsweise Non-Mixture-Gammamodells keine wesentliche Verbesserung hinsichtlich der Anpassungsgüte.

Die für die Asymptotik der Non-Mixture-Modelle benötigten Annahmen sind bei Betrachtung der geschätzten Werte in Tabelle 9 nahezu erfüllt. Dabei wurde bei dem Non-Mixture-Exponentialmodell  $\lambda_1 = \lambda_2$ , bei dem Non-Mixture-Weibull-Modell  $\lambda_1 = \lambda_2$  und  $k_1 = k_2$  und bei dem Non-Mixture-Gammamodell  $a_1 = a_2$  und  $b_1 = b_2$  vorausgesetzt.

Bezüglich der bedingten Power in Tabelle 10 ist festzuhalten, dass die Non-Mixture-Modelle jeweils eine geringfügig höhere bedingte Power aufweisen als dies bei dem Exponentialmodell der Fall ist. Trotz der deutlich besseren Anpassungsgüte der Non-Mixture-Modelle gegenüber dem Exponentialmodell an die Daten lässt sich hier kein wesentlicher Unterschied bei der bedingten Power feststellen.

## R-Mega-CHOEP

Zur Zwischenauswertung konnten die in Tabelle 11 zusammengefassten Daten erhoben werden:

Tabelle 11: Daten aus der Zwischenauswertung

	R-CHOEP	R-MegaCHOEP	Gesamt
Patienten	91	94	185
Todesfälle	26	39	65
Zensierungen	65	55	120
Personenzeit (Monate)	2264	1916	4180

Bei der R-Mega-CHOEP-Studie legt man eine mittlere Rekrutierungsrate von 5 Patienten pro Monat (2,5 pro Gruppe) zugrunde. Der Folgezeitraum umfasst dabei 29 Monate. Auch hier wird angenommen, dass es keine Drop-Outs gibt. Dies führt dann zu einer Patientenzahl von 330.

Als klinisch bedeutsam wird ein Hazard-Ratio (Hazard der Gruppe R-MegaCHOEP im Verhältnis zum Hazard der Gruppe R-CHOEP) von 0,440 angesehen.

Auch hier liege das Signifikanzniveau bei 0,05.

Zum Vergleich der berechneten Überlebenskurven mit den Kaplan-Meier-Kurven betrachtet man zusammen mit den bedingte Power-Kurven Abbildung 3:

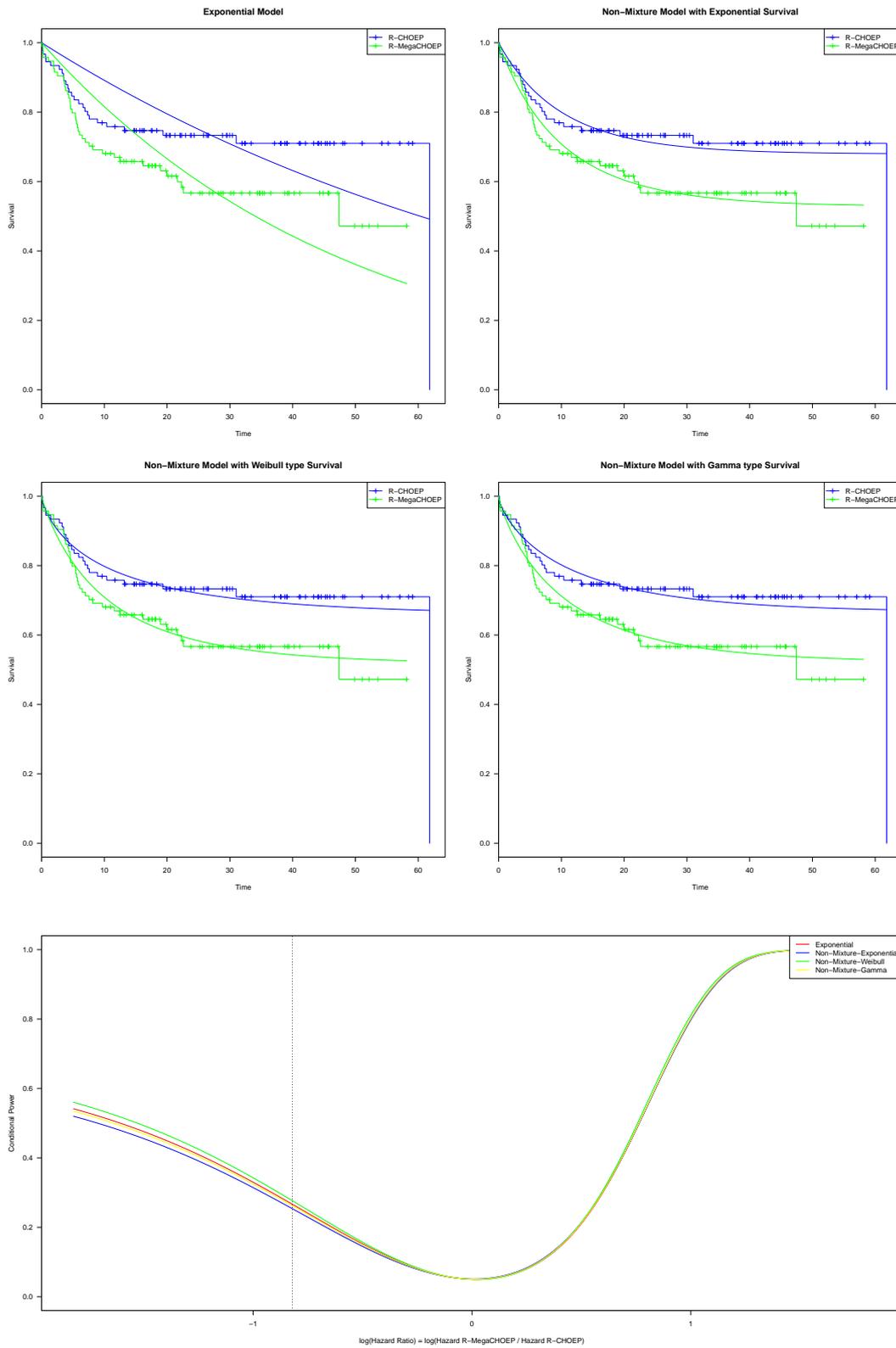


Abbildung 3: Überlebenskurven und bedingte Power-Kurven

Die berechneten Werte enthält Tabelle 12:

Tabelle 12: Weitere Daten aus der Zwischenauswertung

		R-CHOEP	R-MegaCHOEP
Exponential $\sum \text{AIC} = 670,0457$ $\hat{\theta} = 1,7726$	$l(\text{MLE})$	-142,1394	-190,8834
	AIC	286,2788	383,7668
	$\hat{\lambda}$	0,0115	0,0204
	$O^*$	2546	2076
Non-Mixture-Exponential $\sum \text{AIC} = 636,1145$ $\hat{\theta} = 1,6500$	$l(\text{MLE})$	-133,2190	-180,8383
	AIC	270,4380	365,6766
	$\hat{\lambda}$	0,0861	0,0779
	$\hat{c}$	0,6793	0,5284
	$O^*$	2430	1989
Non-Mixture-Weibull $\sum \text{AIC} = 636,5083$ $\hat{\theta} = 1,5672$	$l(\text{MLE})$	-131,5952	-180,6589
	AIC	269,1904	367,3179
	$\hat{\lambda}$	0,1397	0,0900
	$\hat{k}$	0,7406	0,9131
	$\hat{c}$	0,6565	0,5172
	$O^*$	2443	1996
Non-Mixture-Gamma $\sum \text{AIC} = 637,4943$ $\hat{\theta} = 1,6022$	$l(\text{MLE})$	-131,9968	-180,7504
	AIC	269,9936	367,5007
	$\hat{a}$	0,7096	0,8435
	$\hat{b}$	0,0484	0,0580
	$\hat{c}$	0,6659	0,5213
	$O^*$	2444	2003

Es sei nochmals darauf hingewiesen, dass mit  $l(\text{MLE})$  der Wert der logarithmierten Likelihood-Funktion bezüglich der berechneten Maximum-Likelihood-Schätzer gemeint ist.

Als Hilfsmittel zur Entscheidung, welches Modell den besten Kompromiss zwischen Anpassungsgüte des Modells an die Daten und der Komplexizität des Modells eingeht, wird wieder Akaikes Informationskriterium (Akaike (1974)) verwendet.

Auch in Tabelle 12 gilt, dass sämtliche Werte - bis auf die zukünftigen Personenzeiten - auf vier Nachkommastellen gerundet sind.

$O^*$  bezieht sich wieder auf die jeweilige zukünftige Personenzzeit.

Man erhält für ein Hazard-Ratio von 0,440 beziehungsweise  $\log(0,440) \approx -0,821$  in

Tabelle 13 folgende bedingte Power:

Tabelle 13: Bedingte Power

	Bedingte Power
Exponential	0,2666
Non-Mixture-Exponential	0,2537
Non-Mixture-Weibull	0,2767
Non-Mixture-Gamma	0,2622

Betrachtet man zunächst wieder die AIC-Werte in Tabelle 12, so lässt sich erkennen, dass diese bei allen Non-Mixture-Modellen nahezu identisch sind und die Non-Mixture-Modelle klar dem Exponentialmodell zu bevorzugen sind.

Zudem sind auch hier die Voraussetzungen der asymptotischen Untersuchungen der Non-Mixture-Modelle annähernd erfüllt, wenn man diesbezüglich die entsprechenden geschätzten Werte in Tabelle 12 betrachtet.

Hinsichtlich der bedingten Power ergibt sich in Tabelle 13 ein noch extremeres Bild als in Tabelle 10 in dem Sinne, dass sich die einzelnen Werte kaum noch unterscheiden, obwohl das Exponentialmodell eine wesentlich schlechtere Anpassungsgüte als die Non-Mixture-Modelle aufweist.

## 6 R-Paket CP

In diesem abschließenden Kapitel soll nun das R-Paket `CP` vorgestellt werden. Dabei sind die folgenden Beschreibungen des Paketes und der Funktionen an die entsprechenden Hilfeseiten der Implementierung in R angelehnt.

### 6.1 Paket CP

#### Beschreibung

Dieses Paket kann eine Hilfe bei der Berechnung der bedingten Power bei randomisierten klinischen Studien zum Vergleich zweier Therapien mit Überlebenszeit als Zielgröße sein.

Die bedingte Power ist definiert als Wahrscheinlichkeit, ein signifikantes Ergebnis am Ende der Studie zu erhalten, wenn der wahre Effekt gleich dem erwarteten Effekt ist, gegeben die Daten aus der Zwischenauswertung.

Es stehen Funktionen für das Exponential-, Non-Mixture-Exponential-, Non-Mixture-Weibull- und Non-Mixture-Gammamodell zur Verfügung.

Weiterhin gibt es eine Funktion, um die vier Modelle miteinander zu vergleichen.

Zusätzlich wird eine Funktion für das Exponentialmodell zur Verfügung gestellt, bei der die Formeln aus der Arbeit von Andersen aus dem Jahr 1987 verwendet werden.

Schließlich ist es für den Benutzer möglich, über eine Funktion zufällige Datensätze zu generieren.

Zum Testen wird ein Beispieldatensatz zur Verfügung gestellt.

#### Funktionen

`ConPwrExp`  
`ConPwrNonMixExp`  
`ConPwrNonMixWB`  
`ConPwrNonMixGamma`  
`CompSurvMod`  
`ConPwrExpAndersen`  
`GenerateDataFrame`

**Datensatz**

test

## 6.2 Funktion ConPwrExp

### Beschreibung

Diese Funktion berechnet die bedingte Power auf Grundlage des Exponentialmodells.

### Verwendung

```
ConPwrExp(df, cont.time, new.pat = c(0, 0),  
          theta.0 = 1, alpha = 0.05,  
          disp.data = FALSE, plot.km = FALSE)
```

### Argumente

<code>df</code>	Datensatz, welcher aus mindestens drei Spalten besteht: die Gruppe (zwei verschiedene Bezeichnungen) als erste, der Status (1 = Ereignis, 0 = Zensierung) als zweite und der Zeitpunkt des Ereignisses als dritte Spalte
<code>cont.time</code>	Restlaufzeit der Studie
<code>new.pat</code>	2-dimensionaler Vektor, welcher aus den Rekrutierungsraten von Gruppe 1 und Gruppe 2 besteht; Default bei <code>c(0, 0)</code>
<code>theta.0</code>	Erwarteter Effekt (Hazard-Ratio = Hazard Gruppe 2 / Hazard Gruppe 1); Default bei 1
<code>alpha</code>	Signifikanzniveau für Berechnungen der bedingten Power; Default bei 0.05
<code>disp.data</code>	Logischer Wert, ob alle Daten ausgegeben werden sollen; Default bei <code>FALSE</code>
<code>plot.km</code>	Logischer Wert, ob Kaplan-Meier-Kurven und die geschätzten Überlebenskurven gemäß dem Exponentialmodell gezeichnet werden sollen; Default bei <code>FALSE</code>

### Rückgabe

Es wird kein spezieller Wert zurückgegeben.

Es erfolgt die Ausgabe der bedingten Power und optional weiterer berechneter Größen (log-Likelihoods, AICs, berechnete Schätzer, weitere Personenzeiten), der Kaplan-Meier-Kurven und der geschätzten Überlebenskurven gemäß dem Exponentialmodell.

### Beispiel

```
# mit 'GenerateDataFrame' erzeugter Datensatz 'test'  
  
# bedingte Power-Berechnung gemäß dem Exponentialmodell  
ConPwrExp(df = test, cont.time = 12, new.pat = c(2.5, 2.5),  
          theta.0 = 0.75, alpha = 0.05,  
          disp.data = TRUE, plot.km = TRUE)
```

## 6.3 Funktion ConPwrNonMixExp

### Beschreibung

Diese Funktion berechnet die bedingte Power auf Grundlage des Non-Mixture-Exponentialmodells.

### Verwendung

```
ConPwrNonMixExp(df, cont.time, new.pat = c(0, 0),  
                theta.0 = 1, alpha = 0.05,  
                disp.data = FALSE, plot.km = FALSE)
```

### Argumente

**df** Datensatz, welcher aus mindestens drei Spalten besteht: die Gruppe (zwei verschiedene Bezeichnungen) als erste, der Status (1 = Ereignis, 0 = Zensierung) als zweite und der Zeitpunkt des Ereignisses als dritte Spalte

**cont.time** Restlaufzeit der Studie

**new.pat** 2-dimensionaler Vektor, welcher aus den Rekrutierungsraten von Gruppe 1 und Gruppe 2 besteht;  
Default bei `c(0,0)`

**theta.0** Erwarteter Effekt (Hazard-Ratio = Hazard Gruppe 2 / Hazard Gruppe 1);  
Default bei 1

**alpha** Signifikanzniveau für Berechnungen der bedingten Power;  
Default bei 0.05

**disp.data** Logischer Wert, ob alle Daten ausgegeben werden sollen;  
Default bei `FALSE`

**plot.km** Logischer Wert, ob Kaplan-Meier-Kurven und die geschätzten Überlebenskurven gemäß dem Non-Mixture-Exponentialmodell gezeichnet werden sollen;  
Default bei `FALSE`

### Rückgabe

Es wird kein spezieller Wert zurückgegeben.

Es erfolgt die Ausgabe der bedingten Power und optional weiterer berechneter Größen (log-Likelihoods, AICs, berechnete Schätzer, weitere Personenzeiten), der Kaplan-Meier-Kurven und der geschätzten Überlebenskurven gemäß dem Non-Mixture-Exponentialmodell.

### Beispiel

```
# mit 'GenerateDataFrame' erzeugter Datensatz 'test'  
  
# bedingte Power-Berechnung gemäß dem Non-Mixture-Exponentialmodell  
ConPwrNonMixExp(df = test, cont.time = 12, new.pat = c(2.5, 2.5),  
                theta.0 = 0.75, alpha = 0.05,  
                disp.data = TRUE, plot.km = TRUE)
```

## 6.4 Funktion ConPwrNonMixWB

### Beschreibung

Diese Funktion berechnet die bedingte Power auf Grundlage des Non-Mixture-Weibull-Modells.

### Verwendung

```
ConPwrNonMixWB(df, cont.time, new.pat = c(0, 0),  
                theta.0 = 1, alpha = 0.05,  
                disp.data = FALSE, plot.km = FALSE)
```

### Argumente

**df** Datensatz, welcher aus mindestens drei Spalten besteht: die Gruppe (zwei verschiedene Bezeichnungen) als erste, der Status (1 = Ereignis, 0 = Zensierung) als zweite und der Zeitpunkt des Ereignisses als dritte Spalte

**cont.time** Restlaufzeit der Studie

**new.pat** 2-dimensionaler Vektor, welcher aus den Rekrutierungsraten von Gruppe 1 und Gruppe 2 besteht;  
Default bei `c(0,0)`

**theta.0** Erwarteter Effekt (Hazard-Ratio = Hazard Gruppe 2 / Hazard Gruppe 1);  
Default bei 1

**alpha** Signifikanzniveau für Berechnungen der bedingten Power;  
Default bei 0.05

**disp.data** Logischer Wert, ob alle Daten ausgegeben werden sollen;  
Default bei `FALSE`

**plot.km** Logischer Wert, ob Kaplan-Meier-Kurven und die geschätzten Überlebenskurven gemäß dem Non-Mixture-Weibull-Modell gezeichnet werden sollen;  
Default bei `FALSE`

### Rückgabe

Es wird kein spezieller Wert zurückgegeben.

Es erfolgt die Ausgabe der bedingten Power und optional weiterer berechneter Größen (log-Likelihoods, AICs, berechnete Schätzer, weitere Personenzeiten), der Kaplan-Meier-Kurven und der geschätzten Überlebenskurven gemäß dem Non-Mixture-Weibull-Modell.

### Beispiel

```
# mit 'GenerateDataFrame' erzeugter Datensatz 'test'  
  
# bedingte Power-Berechnung gemäß dem Non-Mixture-Weibull-Modell  
ConPwrNonMixWB(df = test, cont.time = 12, new.pat = c(2.5, 2.5),  
               theta.0 = 0.75, alpha = 0.05,  
               disp.data = TRUE, plot.km = TRUE)
```

## 6.5 Funktion ConPwrNonMixGamma

### Beschreibung

Diese Funktion berechnet die bedingte Power auf Grundlage des Non-Mixture-Gammamodells.

### Verwendung

```
ConPwrNonMixGamma(df, cont.time, new.pat = c(0, 0),  
                  theta.0 = 1, alpha = 0.05,  
                  disp.data = FALSE, plot.km = FALSE)
```

### Argumente

**df** Datensatz, welcher aus mindestens drei Spalten besteht: die Gruppe (zwei verschiedene Bezeichnungen) als erste, der Status (1 = Ereignis, 0 = Zensierung) als zweite und der Zeitpunkt des Ereignisses als dritte Spalte

**cont.time** Restlaufzeit der Studie

**new.pat** 2-dimensionalen Vektor, welcher aus den Rekrutierungsraten von Gruppe 1 und Gruppe 2 besteht;  
Default bei  $c(0,0)$

**theta.0** Erwarteter Effekt (Hazard-Ratio = Hazard Gruppe 2 / Hazard Gruppe 1);  
Default bei 1

**alpha** Signifikanzniveau für Berechnungen der bedingten Power;  
Default bei 0.05

**disp.data** Logischer Wert, ob alle Daten ausgegeben werden sollen;  
Default bei FALSE

**plot.km** Logischer Wert, ob Kaplan-Meier-Kurven und die geschätzten Überlebenskurven gemäß dem Non-Mixture-Gammamodell gezeichnet werden sollen;  
Default bei FALSE

### Rückgabe

Es wird kein spezieller Wert zurückgegeben.

Es erfolgt die Ausgabe der bedingten Power und optional weiterer berechneter Größen (log-Likelihoods, AICs, berechnete Schätzer, weitere Personenzeiten), der Kaplan-Meier-Kurven und der geschätzten Überlebenskurven gemäß dem Non-Mixture-Gammamodell.

### Beispiel

```
# mit 'GenerateDataFrame' erzeugter Datensatz 'test'  
  
# bedingte Power-Berechnung gemäß dem Non-Mixture-Gammamodell  
ConPwrNonMixGamma(df = test, cont.time = 12, new.pat = c(2.5, 2.5),  
                  theta.0 = 0.75, alpha = 0.05,  
                  disp.data = TRUE, plot.km = TRUE)
```

## 6.6 Funktion CompSurvMod

### Beschreibung

Diese Funktion berechnet die bedingte Power auf Grundlage des Exponential-, Non-Mixture-Exponential-, Non-Mixture-Weibull- und Non-Mixture-Gammamodells.

### Verwendung

```
CompSurvMod(df, cont.time, new.pat = c(0, 0),  
            theta.0 = 1, alpha = 0.05,  
            disp.data = FALSE, plot.km = FALSE)
```

### Argumente

**df** Datensatz, welcher aus mindestens drei Spalten besteht: die Gruppe (zwei verschiedene Bezeichnungen) als erste, der Status (1 = Ereignis, 0 = Zensierung) als zweite und der Zeitpunkt des Ereignisses als dritte Spalte

**cont.time** Restlaufzeit der Studie

**new.pat** 2-dimensionaler Vektor, welcher aus den Rekrutierungsraten von Gruppe 1 und Gruppe 2 besteht;  
Default bei  $c(0,0)$

**theta.0** Erwarteter Effekt (Hazard-Ratio = Hazard Gruppe 2 / Hazard Gruppe 1);  
Default bei 1

**alpha** Signifikanzniveau für Berechnungen der bedingten Power;  
Default bei 0.05

**disp.data** Logischer Wert, ob alle Daten ausgegeben werden sollen;  
Default bei FALSE

**plot.km** Logischer Wert, ob Kaplan-Meier-Kurven und die geschätzten Überlebenskurven gemäß der vier erwähnten Modelle gezeichnet werden sollen;  
Default bei FALSE

### Rückgabe

Es wird kein spezieller Wert zurückgegeben.

Es erfolgt die Ausgabe der bedingten Power und optional weiterer berechneter Größen (log-Likelihoods, AICs, berechnete Schätzer, weitere Personenzeiten), der Kaplan-Meier-Kurven und der geschätzten Überlebenskurven gemäß der vier erwähnten Modelle.

### Beispiel

```
# mit 'GenerateDataFrame' erzeugter Datensatz 'test'  
  
# bedingte Power-Berechnungen gemäß der vier erwähnten Modelle  
CompSurvMod(df = test, cont.time = 12, new.pat = c(2.5, 2.5),  
            theta.0 = 0.75, alpha = 0.05,  
            disp.data = TRUE, plot.km = TRUE)
```

## 6.7 Funktion ConPwrExpAndersen

### Beschreibung

Diese Funktion berechnet die bedingte Power auf Grundlage des Exponentialmodells unter Verwendung der Formeln aus der Arbeit von Andersen.

### Verwendung

```
ConPwrExpAndersen(df, cont.time, new.pat = c(0, 0),  
                  theta.0 = 1, alpha = 0.05,  
                  disp.data = FALSE, plot.km = FALSE)
```

### Argumente

**df** Datensatz, welcher aus mindestens drei Spalten besteht: die Gruppe (zwei verschiedene Bezeichnungen) als erste, der Status (1 = Ereignis, 0 = Zensierung) als zweite und der Zeitpunkt des Ereignisses als dritte Spalte

**cont.time** Restlaufzeit der Studie

**new.pat** 2-dimensionaler Vektor, welcher aus den Rekrutierungsraten von Gruppe 1 und Gruppe 2 besteht;  
Default bei  $c(0,0)$

**theta.0** Erwarteter Effekt (Hazard-Ratio = Hazard Gruppe 2 / Hazard Gruppe 1);  
Default bei 1

**alpha** Signifikanzniveau für Berechnungen der bedingten Power;  
Default bei 0.05

**disp.data** Logischer Wert, ob alle Daten ausgegeben werden sollen;  
Default bei FALSE

**plot.km** Logischer Wert, ob Kaplan-Meier-Kurven und die geschätzten Überlebenskurven gemäß dem Exponentialmodell gezeichnet werden sollen;  
Default bei FALSE

### Rückgabe

Es wird kein spezieller Wert zurückgegeben.

Es erfolgt die Ausgabe der bedingten Power und optional weiterer berechneter Größen (log-Likelihoods, AICs, berechnete Schätzer, weitere Personenzeiten), der Kaplan-Meier-Kurven und der geschätzten Überlebenskurven gemäß dem Exponentialmodell unter Verwendung der Formeln aus der Arbeit von Andersen.

### Beispiel

```
# mit 'GenerateDataFrame' erzeugter Datensatz 'test'  
  
# bedingte Power-Berechnung gemäß dem Exponentialmodell  
# unter Verwendung der Formeln aus der Arbeit von Andersen  
ConPwrExpAndersen(df = test, cont.time = 12, new.pat = c(2.5, 2.5),  
                  theta.0 = 0.75, alpha = 0.05,  
                  disp.data = TRUE, plot.km = TRUE)
```

## 6.8 Funktion `GenerateDataFrame`

### Beschreibung

Diese Funktion erzeugt einen Datensatz für bedingte Power-Berechnungen.

Die Generierung erfolgt dabei auf folgende Art und Weise:

Die Anzahl der Patienten ist die Realisierung einer Poisson-verteilten Zufallsvariable mit Parameter 200.

Die Wahrscheinlichkeit einer Zensierung ist die Realisierung einer gleichverteilten Zufallsvariable auf dem Intervall  $[2/5, 3/5]$ , mit je einer Zufallsvariable pro Gruppe.

Die Patienten werden je mit Wahrscheinlichkeit  $1/2$  Gruppe A beziehungsweise Gruppe B zugewiesen.

Der Status ( $1 = \text{Ereignis}$ ,  $0 = \text{Zensierung}$ ) ist die Realisierung einer Bernoulli-verteilten Zufallsvariable mit Parameter  $1 - \text{Wahrscheinlichkeit einer Zensierung}$ .

Die Ereigniszeit ist die Realisierung einer exponentialverteilten Zufallsvariable mit Parameter  $1 - \text{Wahrscheinlichkeit einer Zensierung}$ .

### Rückgabe

Diese Funktion gibt einen Datensatz zurück, welcher aus drei Spalten besteht. Dabei steht die Gruppe ('A' oder 'B') in der ersten ('group'), der Status (0 oder 1) in der zweiten ('stat') und die Zeit des Ereignisses in der dritten Spalte ('time').

### Beispiel

```
# Erzeugen eines Datensatzes  
df <- GenerateDataFrame()
```

## Literaturverzeichnis

- [1] Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19** (6), 716-723.
- [2] Andersen, P. K. (1987). Conditional power calculations as an aid in the decision whether to continue a clinical trial. *Controlled Clinical Trials* **8**, 67-74.
- [3] Bartle, R. G. (2001). *A Modern Theory of Integration*. American Mathematical Society, Providence.
- [4] Berkson, J. und Gage, R. P. (1952). Survival curve for cancer patients following treatment. *Journal of the American Statistical Association* **47** (259), 501-515.
- [5] Billingsley, P. (1999). *Convergence of Probability Measures*. John Wiley & Sons, New York.
- [6] Boag, J. W. (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of the Royal Statistical Society. Series B (Methodological)* **11** (1), 15-53.
- [7] Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton University Press, Princeton.
- [8] Etemadi, N. (1981). An elementary proof of the strong law of large numbers. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **55**, 119-122.
- [9] Freedman, D. A. (2008). Survival analysis: a primer. *The American Statistician* **62** (2), 110-119.
- [10] Greene, W. H. (2003). *Econometric Analysis*. Prentice Hall, Upper Saddle River.
- [11] Klein, J. P. und Moeschberger, M. L. (2005). *Survival Analysis. Techniques for Censored and Truncated Data*. Springer Science+Business Media, New York.
- [12] Kompetenznetz Maligne Lymphome e. V. <https://www.lymphome.de> (25.9.2013).
- [13] Lambert, P. C. (2007). Modeling of the cure fraction in survival studies. *The Stata Journal* **7** (3), 351-375.
- [14] Le Cam, L. (1960). An approximation theorem for the Poisson binomial distribution. *Pacific Journal of Mathematics* **10** (4), 1181-1197.
- [15] Pfreundschuh, M., Zwick, C., Zeynalova, S., Dührsen, U., Pflüger, K.-H., Vrieling, T., Mesters, R., Mergenthaler, H.-G., Einsele, H., Bentz, M., Lengfelder, E., Trümper, L., Rube, C., Schmitz, N. und Löffler, M. (2007). Dose-escalated CHOEP for the treatment of young patients with aggressive non-Hodgkin's lymphoma: II. Results of the randomized high-CHOEP trial of the German High-Grade Non-Hodgkin's Lymphoma Study Group (DSHNHL). *Annals of Oncology* **19**, 545-552.

- [16] Schmitz, N., Nickelsen, M., Ziepert, M., Hänel, M., Borchmann, P., Schmidt, C., Viardot, A., Bentz, M., Peter, N., Ehninger, G., Dölken, G., Rube, C., Trümper, L., Rosenwald, A., Pfreundschuh, M., Löffler, M. und Glass, B. (2012). Conventional chemotherapy (CHOEP-14) with rituximab or high-dose chemotherapy (MegaCHOEP) with rituximab for young, high-risk patients with aggressive B-cell lymphoma: an open-label, randomised, phase 3 trial (DSHNHL 2002-1). *The Lancet Oncology* **13** (12), 1250-1259.
- [17] Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons, New York.
- [18] Sposto, R. (2002). Cure model analysis in cancer: an application to data from the Children's Cancer Group. *Statistics in Medicine* **21**, 293-312.
- [19] Sposto, R. und Sather, H. N. (1985). Determining the duration of comparative clinical trials while allowing for cure. *Journal of Chronic Diseases* **38** (8), 683-690.
- [20] The R Project for Statistical Computing. <http://www.r-project.org> (25.9.2013).
- [21] Tsodikov, A. D., Ibrahim, J. G. und Yakovlev, A. Y. (2003). Estimating cure rates from survival data: an alternative to two-component mixture models. *Journal of the American Statistical Association* **98** (464), 1063-1078.
- [22] Tsodikov, A., Löffler, M. und Yakovlev, A. (1998). A cure model with time-changing risk factor: an application to the analysis of secondary leukaemia. *Statistics in Medicine* **17**, 27-40.
- [23] Weiß, C. (2010). *Basiswissen Medizinische Statistik*. Springer Medizin Verlag, Heidelberg.

## Diskussion

Ziel der Arbeit war es, Formeln zur Berechnung der bedingten Power im Rahmen von Überlebenszeitmodellen mit Plateau herzuleiten und ein entsprechendes Simulationstool in R zu entwickeln.

Als Grundlage dazu diente eine Arbeit von Per Kragh Andersen aus dem Jahr 1987 über das Exponentialmodell mit konstanten Hazards. Hierzu wurden über die Maximum-Likelihood-Methode mit zensierten Beobachtungen entsprechende Schätzer für die Hazards hergeleitet und die daraus konstruierte Teststatistik in Form des Hazard-Ratios, bestehend aus den geschätzten Hazards, im unbedingten und bedingten Fall hinsichtlich der Verteilungskonvergenz untersucht. Dabei wurden die Power-Funktion und bedingte Power-Funktion gemäß des Exponentialmodells hergeleitet. Es ergab sich ein Unterschied zwischen der Formel zur Berechnung der bedingten Power aus der Arbeit von Andersen und der hier hergeleiteten bedingte Power-Funktion. Dieser besteht darin, dass Andersen auf eine Zentrierung der betrachteten Teststatistik verzichtet. Was im unbedingten Fall unter der Nullhypothese mit den hier hergeleiteten Formeln übereinstimmt, führt im bedingten Fall bei gegebenen Daten zu einer von Andersen nicht näher begründeten fehlenden Zentrierung. Aus diesem Grund wurde innerhalb dieser Arbeit und innerhalb des erzeugten R-Paketes mit den in dieser Arbeit hergeleiteten zentrierten Formeln gearbeitet. Auch wurden teils andere Abschätzungen vorgenommen als dies in der Arbeit von Andersen der Fall war. Vergleichend wird in dem R-Paket eine Funktion zur Berechnung der bedingten Power nach dem Schema von Andersen bereitgestellt.

Der folgende Schritt bestand darin, Überlebenszeitmodelle mit Überlebensfraktion zu betrachten. Dazu wurde zunächst auf die Mixture-Modelle eingegangen, die das Überleben mit Überlebensfraktion anschaulich modellieren. Für asymptotische Untersuchungen hinsichtlich der Verteilungskonvergenz der in Analogie an das Exponentialmodell angelehnten Teststatistik und zur Herleitung der bedingten Power-Funktion wurden Non-Mixture-Modelle betrachtet. Dabei wurde als Spezialfall das Non-Mixture-Modell mit exponentiellem Überleben im Detail untersucht. Als mögliche Alternativen wurden die Non-Mixture-Modelle mit Weibull-verteilttem und Gammaverteilttem Überleben genannt und deren Maximum-Likelihood-Schätzer hergeleitet. Da im Fall der Überlebenszeitmodelle mit Plateau in Form der Non-Mixture-Modelle die Hazards nicht länger proportional zueinander sind, wurden entsprechende Annahmen getroffen, welche für ein zeitunabhängiges Hazard-Ratio sorgten. Welche weiteren Möglichkeiten es im Umgang mit zeitabhängigen Hazard-Ratios gibt, wurde nicht erläutert. Interessant ist auch die Frage, welche weiteren Teststatistiken für die Problematik geeignet wären und was für Eigenschaften diese aufweisen.

Anschließend wurden die theoretischen Resultate in R implementiert und zu dem Paket CP zusammengefasst. Dieses wurde anhand eines zufällig generierten und zweier realer Datensätze getestet.

Da die Maximum-Likelihood-Schätzer im Non-Mixture-Modell mit exponentiellem, Weibull-verteilttem- beziehungsweise Gammaverteilttem Überleben nur implizit gegeben sind und numerisch bestimmt werden müssen, ergibt sich die Frage, wie man entsprechende Startwerte berechnet. Dabei wurde in der Implementierung von CP so vorgegangen, dass als Startwerte für die Überlebensfraktionen das Verhältnis der Anzahl der Überlebenden zu der Gesamtfallzahl der jeweiligen Gruppe verwendet wird. Die Startwerte für die jeweils restlichen Parameter schätzt man wie in den simplen Überlebenszeitmodellen ohne Plateau. Im Fall des Non-Mixture-Exponentialmodells sind dies die Maximum-Likelihood-Schätzer des Exponentialmodells. Bei dem Non-Mixture-Weibull-Modell werden die Kleinste-Quadrate-Schätzer aus dem Weibull-Modell verwendet. Startwerte für das Non-Mixture-Gammamodell erhält man schließlich über die Momentenschätzer aus dem Gammamodell. Die Frage, inwiefern diese Werte noch umskaliert werden sollten und ob dies der optimale Weg der Bestimmung der Startwerte für die numerischen Berechnungen ist, bleibt offen. Eine mögliche Alternative besteht jedoch darin, die Startpunkte zufällig zu wählen und die Berechnungen dann mehrmals mit verschiedenen zufälligen Startwerten durchzuführen. Denn in der Regel werden lediglich lokale aber eben nicht unbedingt globale Maxima gefunden. Dass aufgrund der Verwendung von numerischen Verfahren mit Fehlern in der Approximation zu rechnen ist, sei hier auch erwähnt. Eine genaue Quantifizierung dieser Fehler sollte jedoch nicht Gegenstand dieser Arbeit sein.

Weiterhin wird zur Berechnung der bedingten Power die asymptotische Verteilung der Teststatistik zugrunde gelegt. Das dies für hinreichend große Fallzahlen gerechtfertigt ist, wurde gezeigt. In praktischen Anwendungen ist hier jedoch ein Approximationsfehler vorhanden, den es noch genau zu spezifizieren gilt.

Ausgangspunkt der Überlegungen ist stets die Annahme eines parametrischen Modells der Überlebenszeiten. Welche Möglichkeiten es im Rahmen von nicht-parametrischen oder semi-parametrischen Modellen gibt, könnte für weitere Untersuchungen relevant sein.

Schließlich könnte auch die Thematik der Nichtunterlegenheitsstudien für bedingte Power-Berechnungen interessant sein. Hier hat man es aber dann in der Regel nicht mehr wie in dieser Arbeit mit einfachen sondern mit zusammengesetzten Nullhypothesen zu tun. Inwieweit die Vorgehensweisen hinsichtlich der asymptotischen Untersuchungen aus dieser Arbeit adaptiert werden können oder ob neue Wege beschritten werden müssen, ist noch unklar.

# Selbstständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt habe.

Seitens des Verfassers bestehen keine Einwände, die vorliegende Masterarbeit für die öffentliche Benutzung im Universitätsarchiv zur Verfügung zu stellen.

Jena, 27.9.2013,