

Die Analyse von Überlebenszeiten

Heinz Tüchler

Einleitung

Welche Vorgänge lassen sich untersuchen?

Überlebenskurven und damit zusammenhängende Berechnungen dienen der Darstellung und dem Vergleich von Überlebenszeiten.

Dieselben Verfahren lassen sich zweckmäßig auch auf Remissionsdauer und viele andere Daten anwenden, wenn letztere Vorgänge beschreiben, die als eine geordnete Folge von Abschnitten betrachtet werden können.

Wesentlich ist dabei, daß das betreffende Merkmal durch einen Vorgang zustande kommt, der nicht umkehrbar ist in dem Sinn, wie es für Überlebenszeiten zutrifft. Damit ist gemeint, daß jemand, der z.B. 5 Monate überlebt hat, insgesamt nur mehr länger nicht aber kürzer überleben kann. Dasselbe gilt für Remissionsdauer und alle anderen kumulativen Vorgänge.

Beispiele wären die Ehedauer, die Zahl von Löchern in einem Schlauchboot bis es untergeht, die Zahl der Zahnarztbesuche bis zum dritten Gebiß, die Größe eines Tumors bis zur Diagnose.

Unvollständige Beobachtungen - sogenannte zensierte Daten

Erfaßt man z.B. die Remissionsdauer von Patienten, dann ist der Rückfall das Zielereignis. Tritt er ein, so ist die Remissionsdauer bestimmt und es liegt für diesen Patienten eine vollständige Beobachtung vor. Es liegt in der Natur der beobachteten Vorgänge, daß sie oft nicht bei allen interessierenden Fällen bis zu ihrem Ende, dem Erreichen des Zielereignisses, beobachtet werden.

Das hat drei mögliche Ursachen:

1. Die Studie und damit die Beobachtung wird beendet, bevor alle im Rahmen der Studie untersuchten Fälle ihr Zielereignis erreicht haben.
2. Fälle gehen verloren. Z.B. Patienten wandern aus, und ihr Schicksal kann nicht weiter verfolgt werden.
3. Jemand kann das Zielereignis nicht mehr erreichen. Wenn z.B. ein Patient während seiner Remission an einem Unfall stirbt, dann ist die beobachtete rückfallfreie Zeit nur eine untere Grenze für seine mögliche rückfallfreie Zeit, und nicht zwangsläufig jene Remissionsdauer, die aufgrund der Therapie zu erwarten gewesen wäre.

Es ist einzusehen, daß unvollständige Beobachtungen weder wie vollständige, noch wie nicht vorhandene behandelt werden sollen. Sie enthalten zweifelsfrei die Information über die untere Schranke, die der Betreffende erreicht hat.

Die Methoden der Überlebenskurvenberechnung nehmen darauf Rücksicht.

Ein ganz einfaches Beispiel (Beispiel 1)

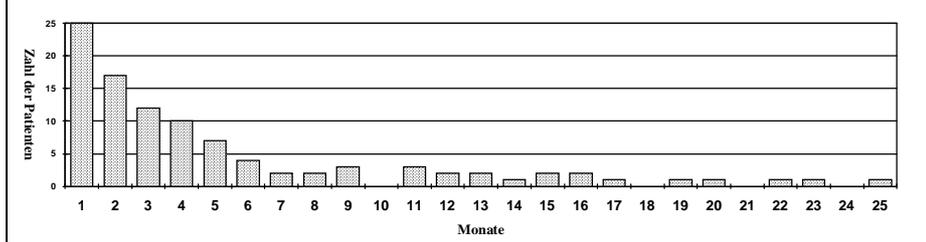
Angenommen es wären 100 Patienten mit einer bestimmten Erkrankung vom Tag der Diagnosestellung bis zu ihrem Tod beobachtet worden. Dann wäre die Überlebenszeit jedes Patienten bekannt. Man würde zuerst eine Tabelle erstellen wie Tabelle 1/1.

Tabelle 1/1: Überlebenszeiten in Monaten für Beispiel 1

Monate	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
Zahl der Patienten	25	17	12	10	7	4	2	2	3	0	3	2	2	1	2	2	1	0	1	1	0	1	1	0	1	100
Prozentanteil	25	17	12	10	7	4	2	2	3	0	3	2	2	1	2	2	1	0	1	1	0	1	1	0	1	100%
Kumulativer %-anteil	25	42	54	64	71	75	77	79	82	82	85	87	89	90	92	94	95	95	96	97	97	98	99	99	100	
Hazard in Prozent	25	23	21	22	19	14	8	9	14	0	17	13	15	9	20	25	17	0	20	25	0	33	50	0	100	

Dort notiert man zu jedem Monat die Anzahl der Patienten mit entsprechender Überlebenszeit. Die nächste Zeile gibt die entsprechende Prozentzahl an. In unserem Beispiel unterscheiden sich diese Zahlen nicht, da wir zur Vereinfachung von 100 Patienten ausgegangen sind.

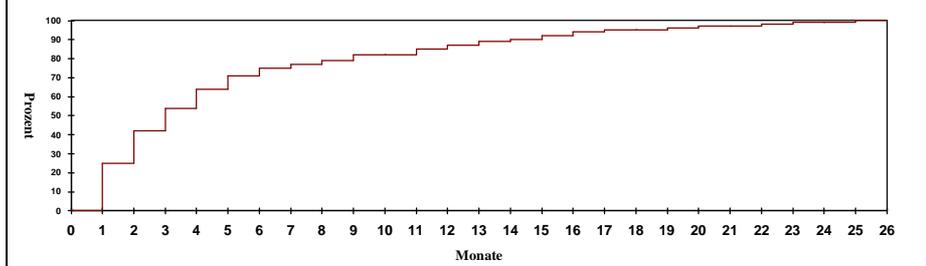
Bild 1/1: Häufigkeitsverteilung der Überlebenszeiten (zu Beispiel 1)
Schätzung für die Dichtefunktion $f(t)$



entnehmen, daß in den ersten Monaten viele Patienten sterben und später immer weniger.

Bild 1/1 zeigt ein Balkendiagramm der Überlebenszeiten. Jeder Monat wird durch einen Balken dargestellt. Seine Höhe entspricht der Zahl von Patienten, die gerade so viele Monate gelebt haben. Bild 1/1 zeigt die sogenannte Häufigkeitsverteilung der Überlebenszeiten. Man kann dem Bild

Bild 1/2: Kumulative Verteilung der Überlebenszeiten (zu Beispiel 1)
Schätzung für die Verteilungsfunktion $F(t)$



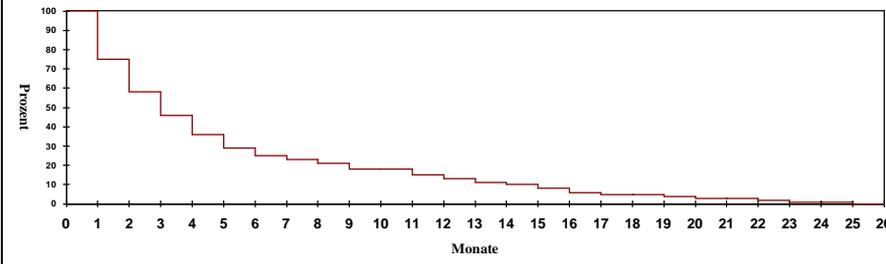
aufgetragen. Die Höhe der Linien entspricht also der Zeile "kumulativer Prozentanteil" der Tabelle 1/1. Das ergibt eine Schätzung für die Verteilungsfunktion oder Summenfunktion der Überlebenszeit. Man sieht z.B., daß in den ersten drei Monaten 54% der Patienten (25%+17%+12%) gestorben sind, oder am Ende des 14. Monats bereits 90%.

Zwei Anmerkungen sind zu machen:

1. Die Linie in Bild 2 ist so gezeichnet, daß die Gestorbenen immer erst am Ende des Monats berücksichtigt werden.
2. Wenn wir vom 1., 3. oder irgendeinem Monat sprechen, handelt es sich um den 1., 3. oder irgendwievielten "Überlebensmonat" eines Patienten und nicht um einen bestimmten Kalendermonat wie z.B. Mai 1986. Wir tun so, als hätten wir alle Patienten zugleich in die Studie aufgenommen. Das ist im allgemeinen zulässig, solange der wirkliche Zeitpunkt keinen Einfluß auf das Geschehen hat. (Unzulässig wäre es bei jahreszeitlichen Schwankungen oder Katastrophenfolgen.)

Die Überlebenskurve

Bild 1/3: Überlebenskurve (zu Beispiel 1)
Schätzung für die Überlebensfunktion $S(t)$
 $S(t) = 1 - F(t)$



Von Bild 1/2 zu Bild 1/3 ist es ein kleiner Schritt. Während Bild 1/2 den Prozentsatz der bis zu einem bestimmten Monat Gestorbenen zeigt, sieht man in Bild 1/3 den Prozentsatz der Lebenden am Beginn jeden Monats. Bild 1/3 ist die Darstellung einer Überlebenskurve.

Zur Wiederholung

Wählen wir als Beispiel den Monat 5. Aus Bild 1 entnehmen wir, daß 7 Personen im 5. Überlebensmonat gestorben sind. Bild 2 zeigt, daß zu Beginn des 5. Monats bereits 64% der Patienten verstorben sind, Bild 1/3 gibt dementsprechend an, daß noch 36% der Patienten leben.

Die durchschnittliche Überlebenszeit

Oft besteht der Wunsch, eine vorliegende Überlebenskurve durch einen Kennwert, also durch eine einzige Zahl zu charakterisieren. Nun enthält aber die ganze Gestalt der Überlebenskurve Information und die läßt sich nicht so ohne weiteres "auf einen Nenner bringen".

Will man die Kurve trotzdem mit einem einzigen Kennwert beschreiben, muß man einen finden, der "die wichtigere Information" enthält und weniger wichtige Details vernachlässigt.

Welcher Kennwert das ist, hängt entscheidend von der Fragestellung und zusätzlich gemachten Annahmen ab.

(Entspricht eine Überlebenskurve einer mathematisch definierten Funktion, dann läßt sie sich unter Umständen durch einen Parameter dieser Funktion charakterisieren. Die Exponentialfunktion, die in bestimmten Zusammenhängen als Überlebenskurve auftritt, wird beispielsweise bereits durch einen Kennwert vollständig bestimmt.

Sehr oft macht man aber in der Therapieforchung keine Annahmen über die Form der Kurven. Dann gibt es auch keine funktionsspezifischen Kennwerte.)

Man könnte sagen: Wenn es keine speziellen Kennwerte gibt, dann bestimmen wir einfach den Durchschnitt der Überlebenszeit.

Als Durchschnittswerte bieten sich das arithmetische Mittel und der Median an. Beide stellen einen Durchschnitt dar, aber das Ergebnis ist im allgemeinen nicht dasselbe.

Wir wollen erst beide Kennwerte kurz getrennt besprechen und dann miteinander vergleichen.

Das arithmetische Mittel

Das ist der Kennwert, den man am ehesten mit der Vorstellung eines Durchschnittes verbindet. Man errechnet ihn (falls keine zensierten Daten vorliegen), indem man die Überlebenszeiten aller Personen addiert und diese Summe durch die Zahl der Personen dividiert.

Der Median

Der Median ist jene Überlebenszeit, die 50% der Patienten nicht erreichen und die die anderen 50% der Patienten überschreiten.

Wenn wir Bild 1/3 betrachten, sehen wir, daß am Beginn des 2. Monats 75% der Personen leben, am Beginn des 3. Monats 58% und am Beginn des 4. Monats 46%. Daraus ergibt sich, daß irgendwann im

3. Monat die 50% Marke erreicht wurde, d.h. die Hälfte der Patienten bereits gestorben war. Wann genau, läßt sich nicht feststellen und so betrachten wir 3 Monate als den Median der Überlebenszeit.

Vergleich von arithmetischem Mittel und Median

Angenommen, wir hätten eine Überlebenskurve für Patienten mit einer bestimmten Diagnose (z.B. eine Krankheit, die ab der Diagnose bis zum Tod einen Spitalsaufenthalt erfordert).

Interessiert sich etwa die Spitalsverwaltung dafür, wie lange ein Patient dieser Diagnose durchschnittlich ein Bett belegt, werden wir mit dem arithmetischen Mittel antworten.

Im Einzelfall wird das zwar eine ungenaue Prognose sein. Ein Patient lebt länger, der andere kürzer als der Durchschnitt. So weichen viele Patienten von der Prognose ab. Über längere Zeit werden sich die Abweichungen jedoch ausgleichen und die Zeit, die einige Patienten länger als der Durchschnitt leben, wird der entsprechen, die andere Patienten früher als der Durchschnitt sterben.

Das arithmetische Mittel ist so bestimmt, daß (wenigstens in der Stichprobe, für die es berechnet wird) die Summe aller positiven und negativen Abweichungen Null ergibt.

Wenn jedoch am Beginn einer Krankheit viele sterben und andererseits einige sehr lang leben, dann gleichen sich zwar die Abweichungen immer noch auf null aus, aber es sterben **viele** vor Erreichen des Durchschnittes und **wenige** sterben lange nach dessen Überschreitung.

Für die Spitalsverwaltung ist das gleich. Es macht ja keinen Unterschied, ob ein Bett durch einen Patienten ein Jahr oder durch zwölf Patienten je einen Monat belegt ist.

Der Mediziner könnte aber den Median als Durchschnittsangabe vorziehen, denn wenn er diesen als Prognose verwendet, dann gleichen sich zwar die Abweichungen **nicht** auf Null aus, dafür hat er aber die Überlebenszeit genau bei **50%** der Patienten **unterschätzt** und bei den anderen **50%** **überschätzt**.

Dazu kommt noch, daß man bei beiden Werten Abweichungen in Kauf nehmen muß. Wie oben gesagt, gleichen sich diese Abweichungen beim arithmetischen Mittel auf null aus, aber nur wenn man ihre Richtung berücksichtigt. (So wie "einmal eine Stunde zu früh kommen und zweimal eine halbe Stunde zuspätkommen ist drei mal pünktlich dort sein.")

Will man aber, daß die Summe aller Abweichungen ohne Rücksicht auf deren Richtung möglichst gering ist, dann erfüllt der Median diese Forderung, das arithmetische Mittel aber nicht. (Letzteres minimiert die Summe der quadrierten Abweichungen, aber das betrifft uns hier nicht.)

Anmerkung: Außerdem läßt sich der Median bereits dann bestimmen, wenn die Hälfte der Patienten gestorben ist (bzw. wenn die Überlebenskurve auf das 50%-Niveau abgesunken ist). Das arithmetische Mittel kann man erst bei Vorliegen aller Überlebenszeiten errechnen, also wenn alle gestorben sind (bzw. wenn der Patient mit der längsten Überlebenszeit gestorben ist).

Die x-Jahres-Überlebensrate

Ein anderer Kennwert, der zur Beschreibung von Überlebenskurven verwendet wird, ist die Überlebensrate nach einer bestimmten Zeit, z.B. die 5-Jahres-Überlebensrate. Diese entspricht einfach der Höhe der Kurve zu diesem Zeitpunkt.

Die Hazardfunktion

Gibt es gefährliche und weniger gefährliche Monate in unserem Beispiel? Gefährlich z.B. in dem Sinn, daß man die Patienten in bestimmten Monaten intensiver betreuen müßte, als in anderen. Betrachtet man Bild 1/1, so sieht man, daß im ersten Monat die meisten sterben und dann, abgesehen von kleineren Schwankungen, immer weniger.

Zum Zeitpunkt der Diagnose stellt es sich so dar, daß die Wahrscheinlichkeit im ersten Monat zu sterben 25% ist, im zweiten Monat 17% ist und so weiter.

Das stimmt, ist aber nur eine der möglichen Antworten auf die Frage nach der Gefährlichkeit einzelner Monate. Im 10. Monat konnten z.B. nicht so viele sterben wie im 1. Monat, weil gar nicht mehr so viele gelebt haben.

Um zu einer Abschätzung des Risikos zu kommen, setzt man deshalb die Todesfälle in einem Zeitabschnitt in Beziehung zu jenen, die in diesem Intervall sterben können, also zu jenen, die am Beginn des Intervalls

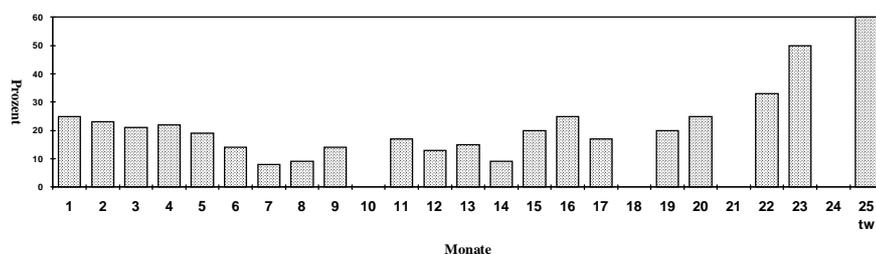
noch gelebt haben und somit "unter Risiko standen". Genau das macht die in Bild 1/4 dargestellte Hazardfunktion.

Bild 1/4: Hazardkurve (zu Beispiel 1)

Schätzung für die Hazardfunktion $\lambda(t)$

$$\lambda(t) = f(t)/(1 - F(t)); \text{ d.h.: } \lambda(t) = f(t)/S(t)$$

$\lambda(t)$ ist der Anteil jener, die in einem Zeitintervall sterben, bezogen auf alle, die bis zu diesem Zeitintervall überlebt haben, d.h. in diesem Intervall unter Risiko stehen.



Sie zeigt den Anteil der Todesfälle pro Monat bezogen auf die Zahl der unter Risiko stehenden Patienten, also auf die Zahl jener, die am Monatsbeginn noch leben.

Sie gibt die lokale oder bedingte Sterberate an.

Wir sehen, daß in unserem Beispiel die ersten Monate gar nicht wesentlich gefährlicher sind als die späteren.

Die Hazardfunktion bringt genau dieselbe Information zum Ausdruck wie die

Überlebenskurve, allerdings in einer anderen Form.

Zweites Beispiel (Quelle: Embury et al., Western J. Medicine 1977 zit. nach Neiß et al., Analyse von Überlebenskurven)

Wie eingangs erwähnt, läßt sich die Methodik der Überlebenskurvenberechnung auch auf andere Daten anwenden. Das nun folgende Beispiel beschäftigt sich mit der Remissionsdauer von Patienten. Das Zielereignis ist hier der Rückfall.

Wir kennen die Remissionsdauer von 11 Patienten. Sie war 9, 13, 13+, 18, 23, 28+, 31, 34, 45+, 48, 161+. Vier dieser Angaben sind mit "+" gekennzeichnet. Diese Patienten wurden 13, 28, 45 bzw. 161 Monate beobachtet ohne daß sie einen Rückfall gehabt hätten. Von ihnen liegen nur diese zensierten Daten vor.

Tabelle 2/1: Remissionsdauer in Monaten zu Beispiel 2

Monate	Zahl der Patienten mit Rückfall	Zahl zensierter Patienten	Zahl bis dahin rückfallfreier Patienten	monatlicher Anteil rückfallfreier Patienten	Gesamter Anteil rückfallfreier Patienten
9	1	0	11	10/11=0.909	0.909
13	1	1	10	9/10=0.900	0.818
18	1	0	8	7/8=0.875	0.716
23	1	0	7	6/7=0.857	0.613
28	0	1	6	6/6=1.000	0.613
31	1	0	5	4/5=0.800	0.491
34	1	0	4	3/4=0.750	0.368
45	0	1	3	3/3=1.000	0.368
48	1	0	2	1/2=0.500	0.184
161	0	1	1	1/1=1.000	0.184

Es stellt sich die Frage, was man mit zensierten Beobachtungen tun soll. Grob gesehen gibt es vier Möglichkeiten mit unvollständigen Daten umzugehen:

Lösung 1: Sie werden weggelassen.

Lösung 2: Sie werden dem Verfahren von Kaplan und Meier entsprechend berücksichtigt.

Lösung 3: Sie werden behandelt wie die anderen, also so, als hätten sie am Ende der Beobachtungszeit einen Rückfall gehabt.

Lösung 4: Man behandelt sie, als würden sie nie einen Rückfall haben (d.h. die Remissionsdauer wäre unendlich)

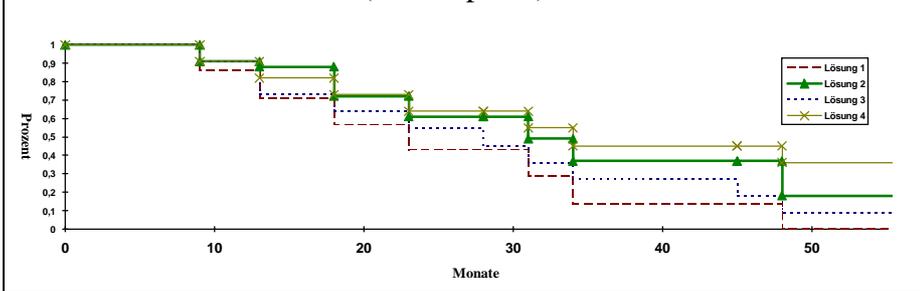
Bild 2/1: Überlebenskurve (zu Beispiel 2)

Bild 2/1 zeigt, wie die "Überlebenskurven" aussehen, je nach dem, welche der vier Lösung man wählt. Sowohl die erste, als auch die dritte Möglichkeit ist nicht zufriedenstellend. Im ersten Fall vernachlässigt man den Informationsgehalt der Tatsache, daß 4 Patienten, solange sie

beobachtet wurden, keinen Rückfall hatten, im anderen Fall verfälscht man die Daten, indem man fehlende Beobachtung mit Rückfall gleichsetzt.

Die vierte Lösung ist wohl auch nicht überzeugend, da man dabei ja unterstellt, daß jene Patienten, die bis zum Ende der Beobachtung keinen Rückfall hatten, auch in Zukunft nie einen solchen haben werden. Eine einleuchtende Lösung zur Berücksichtigung zensierter Daten bietet das Verfahren von Kaplan und Meier(1958).

Das Verfahren von Kaplan und Meier

Erinnern wir uns an das erste Beispiel. Wir hatten alle Patienten bis zu deren Tod beobachtet und damit vollständige Daten. Daraus ließ sich die Überlebenskurve einfach konstruieren.

Man brauchte nur für jeden Monat den Prozentsatz der Überlebenden berechnen, indem man die Zahl der noch Lebenden auf die Gesamtzahl der Patienten in der Studie bezog. Am Beginn des ersten Monats waren es 100%, am Beginn des zweiten Monats 75% und so fort. Diese Prozentsätze nennt man **kumulative** Überlebensrate.

Die Hazardfunktion hingegen gab an, welcher Anteil von Patienten, die am Anfang eines Monats lebten, am Ende desselben Monats gestorben sein würde - die **lokale** Sterberate.

Der Verlust an Patienten in einem bestimmten Monat läßt sich offensichtlich mit der "Gefährlichkeit" dieses Monats erklären.

Um das Folgende einfacher erklären zu können, überlegen wir, daß jeder Monat eine "Gefährlichkeit" und eine "Überlebensfreundlichkeit" besitzt. Letztere sei das Gegenteil der "Gefährlichkeit". Der "Gefährlichkeit" entspricht die lokale Sterberate, der "Überlebensfreundlichkeit" die **lokale** Überlebensrate.

Schauen wir z.B. den 7.Monat aus Beispiel 1 an. Die Hazardfunktion hat den Wert 8%, weil von den 27 Patienten, die den 7.Monat erreicht haben ($25 \cdot 0.08 = 2$), 2 Patienten in diesem Monat gestorben sind. Andererseits haben 92% ($25 \cdot 0.92 = 23$), das sind 23 Patienten, den Monat überlebt.

Die 92% wollen wir die "Überlebensfreundlichkeit" oder lokale Überlebensrate des 7.Monats nennen. Man sieht, daß "Gefährlichkeit" und "Überlebensfreundlichkeit" eines Monats zusammen 100% ergeben.

Multipliziert man also die Ausgangszahl an Patienten mit der **lokalen** Überlebensrate

("Überlebensfreundlichkeit") des ersten Monats, erhält man die Überlebendenzahl am Ende des ersten Monats. Multipliziert man diese Zahl mit der lokalen Überlebensrate des zweiten Monats, erhält man die Überlebendenzahl (oder **kumulative** Überlebensrate) am Ende dieses Monats etc.

Geht man statt von der Patientenzahl von der Zahl 100 aus, dann erhält man die Zahl der Überlebenden in Prozent bis zum Ende des jeweiligen Monats.

Das sieht nach unnötiger Umständlichkeit aus, ist es aber nicht. Im ersten Beispiel ging es zwar einfacher, aber was wir jetzt überlegt haben ist dann interessant, wenn auch zensierte Daten vorliegen.

Tabelle 2/1 zeigt in der fünften Spalte unter der Bezeichnung "monatlicher Anteil rückfallfreier Patienten" die lokale Überlebensrate (das, was wir vorher "Überlebensfreundlichkeit" genannt haben) in Teilen von 1.0, berechnet nach Kaplan und Meier.

Man berechnet den Überlebendenanteil (die kumulative Überlebensrate in Teilen von 1.0) einfach als Zahl derer, die nicht im betreffenden Monat gestorben sind, dividiert durch die Zahl aller, die am Beginn des Monats gelebt haben. Personen, die nicht bis zum Erreichen des Zielereignisses (Tod oder Rückfall, etc.) beobachtet wurden (zensierte Daten), werden dabei für die ganze Zeit, in der sie beobachtet wurden, zu den Lebenden gezählt und danach ganz weggelassen, ohne irgendwann als gestorben verrechnet zu werden.

Derart wird die Information, die in der unvollständigen Beobachtung steckt, zur Gänze genutzt.

Anmerkung: Wenn in einen Monat sowohl Todesfälle als auch zensierte Daten fallen, dürfte dieses Verfahren die Überlebenskurve zu optimistisch schätzen. Kaplan und Meier empfehlen, in so einem Fall bei der Berechnung

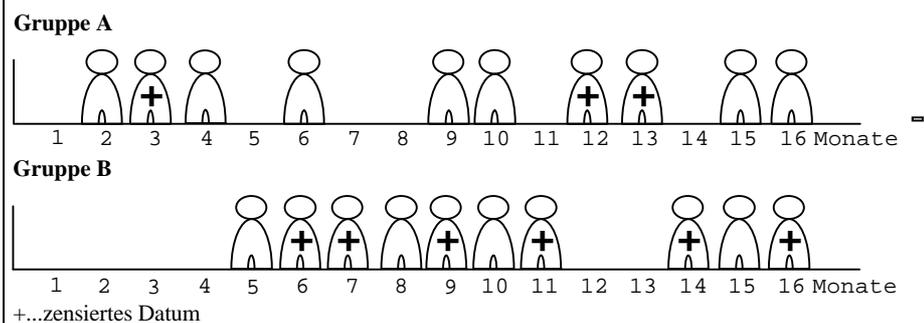
der lokalen Überlebensrate die Personen, deren zensierte Überlebenszeit im betreffenden Monat endet, bei den Überlebenden nur mit halbem Gewicht einzubeziehen. Es hat sich aber zu Unrecht eingebürgert, zensierte Daten auch in jenem Intervall, in dem die Zensierung erfolgt, mit vollem Gewicht zu verrechnen.

Drittes Beispiel - zum Anschauen und Mitrechnen

Wir gehen von folgenden erfundenen Überlebenszeiten für zwei Gruppen aus:

Gruppe A 2, 3+, 4, 6, 9, 10, 12+, 13+, 15, 16
 Gruppe B 5, 6+, 7+, 8, 9+, 10, 11+, 14+, 15, 16+

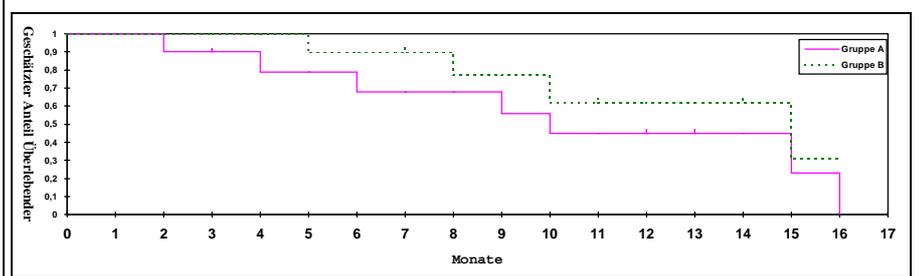
Bild 3/1: Häufigkeitsverteilungen der Überlebenszeiten (zu Beispiel 3)



Fürs erste konzentrieren wir uns auf Gruppe A. Sie umfaßt 10 Personen und die "+" hinter den Zahlen 3, 12 und 13 bedeuten zensierte Daten (und nicht, wie man im christlichen Abendland vermuten könnte, "gestorben"). Somit sind in Gruppe A 7 Personen gestorben und von 3 liegen zensierte Beobachtungen

vor. Es spielt zwar für die Berechnung keine Rolle, aber nehmen wir an, die Überlebenszeiten seien in Monaten angegeben.

Bild 3/2: Überlebenskurven zweier Gruppen (zu Beispiel 3)



Zuerst wollen wir die **Vorgangsweise des Kaplan-Meier-Verfahrens** bildlich darstellen.

Wir werden dazu die Bilder 3/3 und 3/4 miteinander vergleichen. Bild 3/3 zeigt am linken Rand, sozusagen am Beginn des ersten Monats, unsere 10 Personen als kleine (geschlechtlich nicht

spezifizierte, d.h. potentiell auch weibliche!) Männchen, eines über dem anderen aufgestellt.

Jene Person mit der längsten Überlebenszeit steht ganz unten, die mit der kürzesten ganz oben.

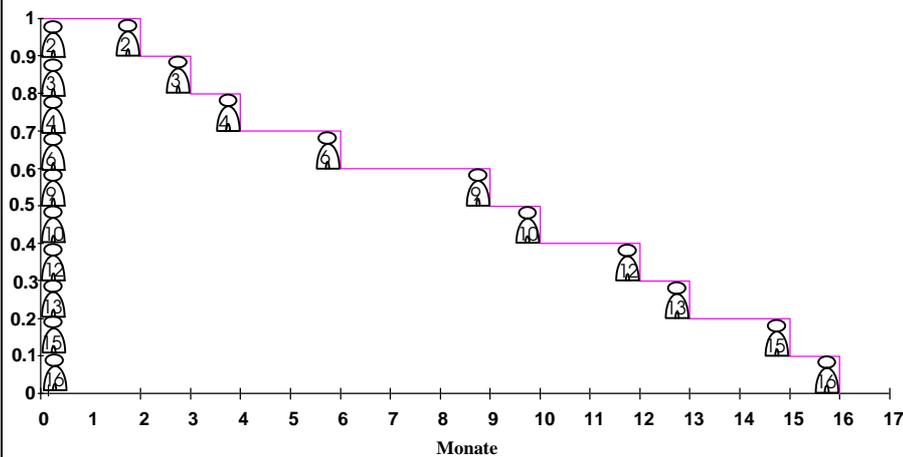
Da es sich um ein erfundenes Beispiel handelt, gestatten wir uns den Vergleich mit einem Brettspiel.

Wir lassen nun jedes Männchen auf seiner Höhe so viele "Monatseinheiten" nach rechts ziehen, wie es seiner Überlebenszeit entspricht.

Am Ende seiner Zeit zeichnen wir es nochmals ein.

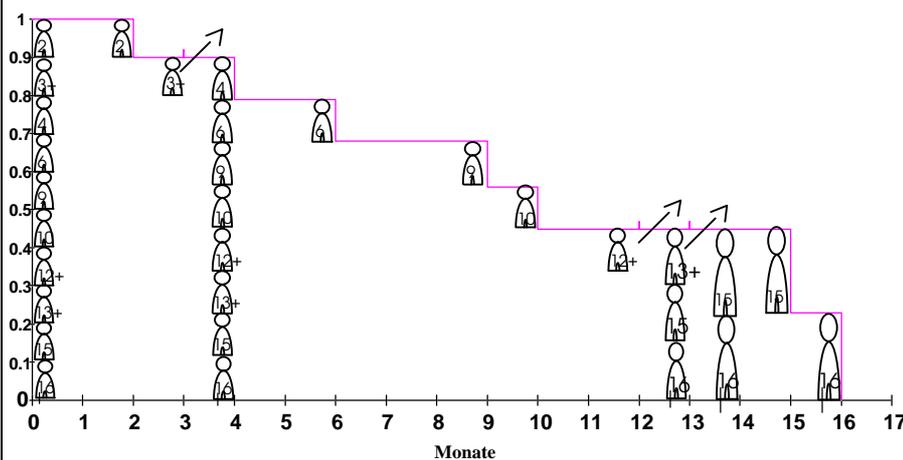
In Bild 3/3 lassen wir, wider besseren Wissens, auch die mit "+" gekennzeichneten Personen sterben. Das ergibt eine einfache Überlebenskurve und entspricht der Lösung 3 im Sinne des vorigen Beispiels.

Bild 3/3: Überlebenskurve für Gruppe A (zu Beispiel 3)
unter Vernachlässigung der Unvollständigkeit zensierter Daten



Das klingt nach Kindergarten, ist so unernst aber nicht! Es bedeutet soviel wie die Überlebenschancen des Verschwundenen gleichmäßig auf die Verbleibenden aufteilen. Mit jedem, der in Zukunft stirbt, stirbt auch "ein Teil" des nicht mehr beobachtbaren Männchens.

Bild 3/4: Überlebenskurve für Gruppe A nach der Methode von Kaplan und Meier



zeichnen der Männchen - genau dem Ergebnis einer Überlebenskurvenberechnung nach dem Verfahren von Kaplan und Meier.

Zum Berechnen der Überlebenskurve, wie sie in Bild 3/4 dargestellt ist, legen wir folgende Tabelle an (siehe Tabelle 3/1).

Tabelle 3/1: Überlebenskurvenberechnung für Gruppe A in Beispiel 3

Zeitpunkt	t	2	3	4	6	9	10	12	13	15	16
Zahl der Pers. mit erreichtem Zielereignis	d	1	0	1	1	1	1	0	0	1	1
Zahl der Personen mit zensiertem Datum	z	0	1	0	0	0	0	1	1	0	0
Zahl der Personen mit Überlebenszeit $\geq t$	n	10	9	8	7	6	5	4	3	2	1
lokale Überlebensrate	$1-\lambda$	9/10	9/9	7/8	6/7	5/6	4/5	4/4	3/3	1/2	0/1
		.90	1.00	.87	.85	.83	.80	1.00	1.00	.50	.00
kumulative Überlebensrate	S	.90	.90	.79	.67	.56	.45	.45	.45	.23	.00
Streuungsrechnungen											
SE nach Kaplan und Meier		.09		.13	.16	.16	.16			.18	.00
SE nach TNM-Klassifikation		.10		.14	.18	.20	.23			.29	.00
SE nach Peto et al.		.09	.09	.13	.15	.15	.15	.17	.19	.14	.00

In die erste Zeile schreiben wir die Zeitpunkte, zu denen jemand gestorben ist oder zensiert wurde, in aufsteigender Reihe.

Das entspricht den nach Größe geordneten Überlebenszeiten. Dabei werden mehrfach auftretende Zeiten (das kommt in diesem Beispiel ohnehin nicht vor) aber nur einmal notiert.

Die zweite Zeile gibt an, wieviele Personen zum entsprechenden Zeitpunkt das Zielereignis erreicht haben, also gestorben sind. Im Beispiel sind das zu jedem Zeitpunkt eine Person oder niemand - dann nämlich, wenn zu diesem Zeitpunkt jemand ausgeschieden ist ohne zu sterben.

Die Zahl dieser zensierten Beobachtungen gibt die dritte Zeile an.

Die vierte Zeile zeigt, wieviele Personen unmittelbar vor jedem Zeitpunkt unter Beobachtung standen und somit auch dem Risiko, zu sterben, ausgesetzt waren.

Am Anfang waren es alle 10 Personen und mit jedem Toten oder Zensierten wurden es um einen weniger. Diese drei Zeilen erhält man, wie zu sehen, durch abzählen. Die lokale Überlebensrate in der fünften Zeile muß man ausrechnen. Sie nennt sich "lokal", weil sie den Überlebendenanteil derer angibt, die unmittelbar vor dem entsprechenden Moment "unter Risiko" standen.

Beispielsweise standen vor dem 2.Monat alle 10 Personen unter Sterberisiko. Eine ist im 2.Monat gestorben, daher ist das lokale Sterberisiko im 2.Monat $9/10=0.9$. Vor dem dritten Monat stehen 9 Personen unter Risiko. Eine scheidet zwar aus, aber niemand stirbt, und das ergibt eine lokale Überlebensrate von $9/9=1$.

Zeitpunkte, zu denen zwar jemand ausscheidet aber niemand stirbt, haben daher immer eine lokale Überlebensrate von 1. Sie haben, wie wir sehen werden, keinen Einfluß auf die Höhe der Überlebenskurve. So soll es auch sein.

Die nächste Zeile - die kumulative Überlebensrate - gibt die Höhe der Überlebenskurve an.

Die Überlebenskurve ist ja die graphische Darstellung der kumulativen Überlebensrate. Wie die lokale Überlebensrate drückt man auch die kumulative Überlebensrate üblicherweise in Teilen von 1.0 aus. Wer Prozentangaben vorzieht, multipliziert das Ergebnis mit 100.

Man errechnet die kumulative Überlebensrate für einen Zeitpunkt, indem man den Wert, den sie unmittelbar vor dem Zeitpunkt hat, mit der lokalen Überlebensrate des betreffenden Zeitpunktes multipliziert.

Am Beginn hat die kumulative Überlebensrate den Wert 1.0, denn am Anfang leben noch alle. Also ist der Überlebendenanteil 1.0. Zum Zeitpunkt des ersten Todesfalles sinkt sie auf die lokale Überlebensrate dieses Zeitpunktes ab, in unserem Beispiel auf 0.9. Im vierten Monat sinkt die kumulative Überlebensrate auf $0.90 \cdot 0.87 = 0.79$ und so weiter.

Zu den Zeitpunkten mit Todesfällen sinkt sie auf die errechnete Höhe ab. Diese Höhe erhält man, indem man die aktuelle Höhe mit der lokalen Überlebensrate multipliziert. Da die lokale Überlebensrate an Zeitpunkten von Todesfällen immer kleiner 1.0 ist, sinkt die Überlebenskurve dort. Andererseits bleibt sie an Stellen, an denen nur zensierte Daten anfallen, gleich hoch, da dort die lokale Überlebensrate 1.0 beträgt und daher eine Multiplikation keine Veränderung bewirkt.

Ist die längste beobachtete Überlebenszeit eine zensierte Beobachtung, so ist die Höhe der Überlebenskurve ab diesem Zeitpunkt unbestimmt in den Grenzen ihrer letzten Höhe und 0.

Ist die längste Überlebenszeit keine zensierte Beobachtung dann fällt die Kurve auf 0.0, da die lokale Überlebensrate dieses Zeitpunktes 0.0 beträgt und damit auch die mit 0.0 multiplizierte kumulative Überlebensrate.

Bestimmung des Medianes der Überlebenszeit

In der Überlebenskurvenberechnung definiert man den Median - etwas abweichend vom sonstigen Gebrauch - als den Zeitpunkt bzw. die Mitte des Zeitintervalles wo die Überlebenskurve die 50%-Höhe durchschreitet, bzw. erreicht.

Bild 3/2 zeigt die Überlebenskurve für Beispiel 3/Gruppe A in der allgemein üblichen Form (durchgezogene Linie). Wie zu sehen, durchtritt die Kurve die 50%-Höhe am Ende des 10.Monates.

Bestimmung des Mittelwertes der Überlebenszeit

Der Mittelwert der Überlebenszeit entspricht der Fläche unter der Überlebenskurve. Diese ist nur dann eindeutig bestimmt, wenn die längste Beobachtung eine unzensierte ist, und die Kurve somit auf Null absinkt.

Andernfalls behilft man sich mit der Begrenzung der Fläche zum Zeitpunkt der längsten Beobachtung. Das führt zu einer systematischen Unterschätzung des wahren Mittelwertes.

Für Gruppe A im Beispiel 4 beträgt der Mittelwert 10.43.

Konfidenzintervall einer Überlebenskurve

Wenn wir für eine bestimmte Patientengruppe eine Überlebenskurve berechnen, tun wir das z.B., um den natürlichen Verlauf einer Krankheit oder jenen unter Einfluß einer speziellen Therapie darzustellen. Wir interessieren uns für etwas, wovon die beobachteten Daten ein Teil - eine Stichprobe sind. Wir erwarten bei anderen Patienten ähnliche, aber doch nicht exakt dieselben Ergebnisse, auch wenn sie der gleichen Diagnosegruppe angehören oder der gleichen Therapie unterzogen werden. Wir erwarten also zufällige Schwankungen der Ergebnisse.

Je mehr Patienten wir beobachten, desto kleiner werden diese Schwankungen sein.

Wir stellen uns vor, daß es eine wahre Überlebenskurve gibt, die wir erhalten, wenn wir alle in Betracht kommenden Patienten (die sogenannte Grundgesamtheit) beobachten. (Das ist normalerweise nicht möglich, da viele Grundgesamtheiten als unendlich groß gedacht sind.)

Alle Überlebenskurven, die wir aus den einzelnen Stichproben erhalten, schwanken um diese (gedachte) Überlebenskurve.

Unter gewissen Annahmen kann man den Abstand zwischen "wahrer" und aus der Stichprobe stammender Überlebenskurve berechnen, der nur mehr von einem bestimmten kleinen Teil von "Stichprobenkurven" überschritten wird.

Zeichnet man nun über und unter der errechneten Kurve diesen Abstand ein - wie es in Bild 3/5 geschehen ist (gestrichelte Linien), dann kann man sagen, daß der so begrenzte Bereich mit einer Wahrscheinlichkeit von 95% die "wahre" Überlebenskurve miteinschließt.

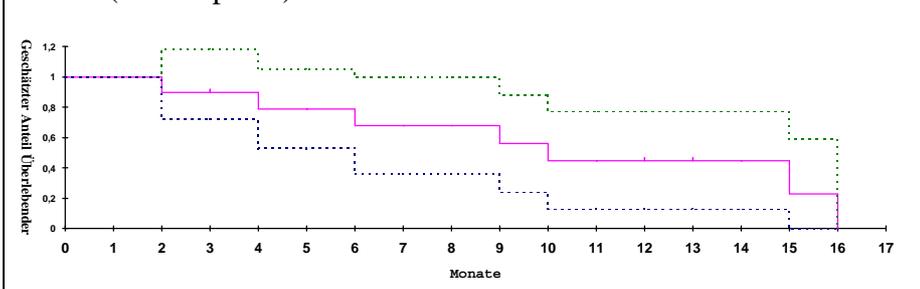
Anmerkung: Leider ist diese Erklärung in mehrfacher Hinsicht nicht ganz (genau genommen gar nicht) korrekt. Es überschreitet aber den gegebenen Rahmen, die Diskussion rund um die richtige Interpretation eines Konfidenzintervalles hier darzustellen.

Dieses Konfidenzintervall ist nach einer Näherungsformel berechnet. Das sieht man auch daran, daß das Konfidenzintervall Bereiche außerhalb von 0% und 100% einschließt, was natürlich nicht richtig sein kann, denn eine Überlebenskurve kann nur zwischen 0% und 100% liegen.

Trotzdem gibt das Konfidenzintervall einen guten Eindruck von den möglichen Zufallsschwankungen, die das Ergebnis mitbeeinflusst haben.

Der zweite Teil der Tabelle 3/1 zeigt die Ergebnisse von Streuungsberechnungen nach drei verschiedenen Formeln, die in der Literatur zu finden sind.

Bild 3/5: Konfidenzbereich der Überlebenskurve für Gruppe A
(zu Beispiel 3)



In Bild 3/5 wurde das Konfidenzintervall aufgrund der Streuungsformel, die Kaplan und Meier angeben, eingezeichnet. Will man zu einer vorhandenen Überlebenskurve das Konfidenzintervall ohne Computerprogramm berechnen, empfiehlt es sich die Formel nach Peto et al. (1977) zu verwenden. Sie lautet:

$$SE(S) = S \cdot \sqrt{\frac{1-S}{n}}$$

dabei bedeutet:

- SE(S) Standardfehler der Überlebenskurve
- S Höhe der Überlebenskurve zum betreffenden Zeitpunkt
- n Zahl der Personen unter Risiko zum betreffenden Zeitpunkt

in Worten: der Standardfehler der Überlebenskurve ergibt sich aus der Höhe der Überlebenskurve multipliziert mit der Wurzel aus 1-S durch die Zahl der Personen unter Risiko.

Diesen Ausdruck berechnet man für jeden Zeitpunkt, an dem sich die Zahl der Personen unter Risiko ändert.

Die Formel gibt den Standardfehler an. Diesen Standardfehler muß man, wenn man ein 95%-Konfidenzintervall erhalten will, mit 2 (oder scheinbar oder anscheinend genauer mit 1.96) multiplizieren.

Die maximale Höhe der Überlebenskurve zu einem bestimmten Zeitpunkt beträgt dann $S+SE(S)*1.96$ und die Minimalhöhe $S-SE(S)*1.96$.

Konfidenzintervall des Medianes

Ein Konfidenzintervall läßt sich auch für den Median bestimmen. Dazu gibt es verschiedene Ansätze. Sie führen im allgemeinen auch zu unterschiedlichen Ergebnissen, und die Berechnungen sind mühsam. Wir wollen hier darauf verzichten. Man kann sich aber überlegen, daß der Median in Beispiel 3 zwischen 6 Monaten und 16 Monaten schwanken würde, wenn die Überlebenskurve im gestrichelten Bereich, also im Rahmen ihres Konfidenzintervalles, schwankt.

Vergleich mehrerer Gruppen

Die Patienten in Beispiel 3 Gruppe A hatten die Überlebenszeiten 2, 3+, 4, 6, 9, 10, 12+, 13+, 15, 16. Nun wollen wir sie mit den Überlebenszeiten von Gruppe B vergleichen. Die 10 Patienten der Gruppe B hatten folgende Überlebenszeiten: 5, 6+, 7+, 8, 9+, 10, 11+, 14+, 15, 16+.

Wir zeichnen die Überlebenskurven wie in Bild 3/2 zu sehen.

Hier sieht man, daß die Gruppe B zu jedem Zeitpunkt über der Gruppe A liegt.

Das Sterberatenverhältnis

Um die unterschiedlichen Überlebensaussichten der Patienten der beiden Gruppen abschätzen zu können, überlegen wir folgendes: wenn wir beide Gruppen gemeinsam betrachten, stellen wir im 2., 4., 5., ... Monat einen Todesfall fest. Angenommen, die Patienten in beiden Gruppen hätten gleich große Überlebenschancen, dann sollten sich auch die Todesfälle gleichmäßig auf beide Gruppen aufteilen. Würden in einem Monat 2 Patienten sterben, so sollte einer davon aus der Gruppe A und einer aus der Gruppe B sein.

Selbstverständlich werden sich die Todesfälle nie genau aufteilen und, wenn gerade nur einer stirbt, gar nicht aufteilen lassen, aber über die ganze Studie gesehen, wäre doch ein Gleichgewicht zu erwarten. Dabei ist noch zu bedenken, daß sich die Todesfälle nur dann 1:1 aufteilen sollen, wenn in beiden Gruppen gleich viele Personen unter Risiko stehen.

Bestünde hingegen eine Gruppe zu einem bestimmten Zeitpunkt aus 60 Personen und die andere aus 20, dann wäre eine Aufteilung von 3:1 entsprechend.

Nennen wir die Zahl von erwarteten Todesfällen in einer Gruppe "erwartete Häufigkeit" - im Gegensatz zur tatsächlichen Zahl der Todesfälle in der Gruppe der "beobachteten Häufigkeit".

Wir berechnen diese erwartete Häufigkeit (in Tabelle 3/2 wird sie E_a für Gruppe A und E_b für Gruppe B genannt), indem wir für jeden Zeitpunkt eines Todesfalles, egal in welcher Gruppe er eintritt, den "gerechten" Anteil jeder Gruppe bestimmen (in der Tabelle in den mit $e(a)$ und $e(b)$ bezeichneten Zeilen zu finden), und dann alle diese erwarteten Anteile für jede Gruppe zusammenzählen. Es stört beim Errechnen der einzelnen Anteile nicht, daß oft "halbe Tote" herauskommen.

Nun bilden wir für jede Gruppe den Bruch D_a/E_a bzw. D_b/E_b . Wir vergleichen die wirkliche Zahl der Todesfälle mit der erwarteten.

Bei Gruppe A gibt das 1.333, bei Gruppe B 0.695.

Bei Gruppe A sind somit um ein Drittel mehr Patienten gestorben als erwartet, bei Gruppe B um ein Drittel weniger.

Das Sterberatenverhältnis zwischen beiden Gruppen ist $1.33/0.69=1.92$. Die Chance auf einen Todesfall steht zwischen den Gruppen auf 1.92:1, also rund 2:1 d.h. wenn in beiden Gruppen gleich viele Personen unter Risiko stehen und zu einem Zeitpunkt eine stirbt, so steht die Chance 2:1, daß sie aus Gruppe A ist. (Wie man sieht, wird der Begriff Chance hier wertfrei gebraucht.)

Anmerkung: Dieses Sterberatenverhältnis ist eine gute Näherung für den Proportionalitätsfaktor im Proportional-Hazard-Modell von Cox. Das Modell wird später erklärt, dann wird diese Anmerkung verständlich werden.

Zufallskritische Beurteilung von Unterschieden

Wir haben nun festgestellt, daß die Aussichten in beiden Gruppen nicht gleich sind. Es stellt sich aber die Frage, ob dieser Unterschied nicht auf zufällige Schwankungen zurückzuführen ist.

Zwar ist die Überlebenskurve von Gruppe B immer höher als die von Gruppe A (siehe Bild 3/2), bedenkt man aber, wie sehr diese Ergebnisse schwanken können (siehe das Konfidenzintervall in Bild 3/5), dann scheint der Unterschied nicht gesichert.

Um das zu prüfen, rechnen wir einen statistischen Test, den Logrank-Test (auch Mantel-Haenszel-Test, generalized Savage-Test oder Mantel-Cox-Test genannt).

- Anmerkungen:
1. Hier wird die einfachste approximative Berechnung des Logrank-Tests nach Peto et al. (1977) dargestellt. Die Testgröße X -Quadrat ist näherungsweise Chi-Quadrat-verteilt mit Gruppenzahl - 1 Freiheitsgraden.
 2. Es gibt weitere Tests, die unter denselben Bedingungen anwendbar sind, z.B. der generalized Wilcoxon (Gehan-Wilcoxon) oder Breslow-Test und der Taron-Ware Test. Sie unterscheiden sich in der Gewichtung, die sie vornehmen. Der Wilcoxon-Test ist auf Unterschiede zwischen den Kurven nahe dem Zeitpunkt Null besonders empfindlich, während der Logranktest Unterschiede gegen Ende der Kurven stärker als die anderen Tests berücksichtigt.
 3. Für den Fall von mehr als zwei Gruppen, die in eine interpretierbare Reihung gebracht werden können (z.B. drei unterschiedlich intensive Therapieprotokolle), gibt es zu allen genannten Tests auch Versionen, die auf das Vorliegen eines Trends prüfen.

Gesamtgruppenvergleich

Wir brauchen dabei nur die Ergebnisse von vorhin zu verwenden. Mit den Werten E_a , D_a , E_b , D_b bilden wir die Testgröße X -Quadrat, wie sie in der Mitte von Tabelle 3/2 beschrieben ist.

Wenn X -Quadrat größer als 3.84 ist, dann kann man sagen, daß ein so großer oder noch größerer Unterschied zwischen zwei Überlebenskurven in nur 5% aller Untersuchungen zu erwarten ist, **wenn sich die Grundgesamtheiten, aus denen die beiden Gruppen stammen, in Wirklichkeit nicht unterscheiden.**

Tabelle 3/2: Vergleich zweier Gruppen - Sterberatenverhältnis und Logranktest für die Gruppen A und B des Beispiels 3

t	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Gruppe A															
d	1	0	1	0	1	0	0	1	1	0	0	0	0	1	1
z	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0
n(a)	10	9	8	7	7	6	6	6	5	4	4	3	2	2	1
e(a)	.50	.00	.44	.41	.44	.00	.46	.50	1.00	.00	.00	.00	.00	1.00	.50
Gruppe B															
d	0	0	0	1	0	0	1	0	1	0	0	0	0	1	0
z	0	0	0	0	1	1	0	1	0	0	0	0	1	0	1
n(b)	10	10	10	10	9	8	7	6	5	4	3	3	3	2	1
e(b)	.50	.00	.56	.59	.56	.00	.54	.50	1.00	.00	.00	.00	.00	1.00	.50
gesamt															
d(ges)	1	0	1	1	1	0	1	1	2	0	0	0	0	2	1
n(ges)	20	19	18	17	16	14	13	12	10	8	7	6	5	4	2

Formeln und Ergebnisse:

$$e(a) = n(a) \cdot d(\text{ges}) / n(\text{ges})$$

$$D_a = \text{SUMME}(d(a)) = 7$$

$$D_b = \text{SUMME}(d(b)) = 4$$

$$e(b) = n(b) \cdot d(\text{ges}) / n(\text{ges})$$

$$E_a = \text{SUMME}(e(a)) = 5.25$$

$$E_b = \text{SUMME}(e(b)) = 5.75$$

Sterberatenverhältnis:

$$\text{Definition: } R = (D_a/E_a) / (D_b/E_b)$$

$$\text{Ergebnis: } R = 1.92$$

Logranktest:

$$X^2 = \frac{(D_a - E_a)^2}{E_a} + \frac{(D_b - E_b)^2}{E_b}$$

$$\text{Ergebnis des Tests: } X = 1.1, \quad df = 1; \quad \text{nicht signifikant}$$

Nachtrag - Überlebenskurve für Gruppe B:

t	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
d	0	0	0	1	0	0	1	0	1	0	0	0	0	1	0
z	0	0	0	0	1	1	0	1	0	1	0	0	1	0	1
n(b)	10	10	10	10	9	8	7	6	5	4	3	3	3	2	1
1-λ	1	1	1	.90	1	1	.89	1	.80	1	1	1	1	.50	1
S	1	1	1	.90			.77		.62					.31	.31

So ein Ergebnis nennt man signifikant auf dem 5%-Niveau. Wäre X-Quadrat größer als 6.64, so würde das Ergebnis als auf dem 1%-Niveau gesichert gelten.

Anmerkung: Es handelt sich um einen zweiseitigen Test.

Die kritischen Werte 3.84 und 6.64 gelten für den hier beschriebenen Vergleich zweier Gruppen. Der Vergleich mittels X-Quadrat ist auch bei mehreren Gruppen möglich.

Die Formel für X-Quadrat wird dazu um die entsprechenden Summanden erweitert und die kritischen Werte finden sich in Tabellen der Chi-Quadrat Verteilung. Ist ein Ergebnis auf einem vorher gewählten Niveau signifikant, dann heißt das, ein so großer oder größerer Unterschied wäre kaum durch zufällige Unterschiede zwischen den beiden Gruppen zustande gekommen.

Das legt die Vermutung nahe, daß ein beobachteter signifikanter Unterschied nicht durch Zufall, sondern durch inhaltliche Verschiedenheiten der Gruppen entstanden ist. In unserem Fall könnten das unterschiedliche Therapieformen sein (wenn der Unterschied signifikant wäre!).

Unser X-Quadrat beträgt 1.1 und ist damit gar nicht signifikant. Der beeindruckende Unterschied kann ganz gut rein zufällig aufgetreten sein. Er ist ungefähr so groß, wie er durchschnittlich zufallsbedingt zu erwarten ist.

Das bedeutet nicht, daß der Unterschied an sich klein ist, sondern, daß man viel mehr Patienten beobachten müßte, um dann einen gleich großen Unterschied als signifikant ausweisen zu können.

Würde man weiterhin Personen in die Studie aufnehmen und so lange beobachten, bis man rund 50 Todesfälle verzeichnet, hätte man eine Chance von 50%, einen derartigen Unterschied als signifikant ausweisen zu können, wenn der Unterschied zwischen den beiden Gruppen in Wahrheit einem Sterberatenverhältnis von 1:2 entspräche.

Besteht hingegen in Wahrheit kein Unterschied zwischen den Patientengruppen, ist zu erwarten, daß sich das Sterberatenverhältnis mit zunehmender Patientenzahl dem Wert 1 nähert.

Das prinzipielle Konzept eines Signifikanztests ist problematisch, und eher den Bedürfnissen einer "reinen Wissenschaft" angepaßt.

Dort kann man den Standpunkt vertreten, es sei besser etwas nicht zu wissen, als es irrtümlich für wahr zu halten. Die Therapieforchung dient aber mehr der Verbesserung von Entscheidungen, die auf jeden Fall getroffen werden müssen.

Es ist nicht möglich, sich wegen Nichtwissens nicht zu entscheiden, da jedes therapeutische Vorgehen oder Nichtvorgehen implizit eine Entscheidung darstellt.

Daher ist es zuerst wichtig, beispielsweise unterschiedliche Wirksamkeit von Therapien abzuschätzen, und die Genauigkeit der Schätzung in einem entscheidungstheoretischen Gesamtzusammenhang zu sehen.

Hier führt es zu weit, diese Fragen zu behandeln, und so sei das oben Gesagte nur als Andeutung gesehen.

Anmerkung: Es ist wichtig, die unterschiedlichen Fragestellungen, die dem Sterberatenverhältnis und dem Logranktest zugrundeliegen, auseinanderzuhalten.

Das Sterberatenverhältnis ist ein Maß für den Unterschied im Absterbevorgang zwischen zwei Gruppen. Es beschreibt, wie sich zwei tatsächlich beobachtete Gruppen voneinander unterscheiden. Damit liefert es auch eine Schätzung, wie sich zwei derartige Gruppen im allgemeinen unterscheiden würden.

Der Logranktest untersucht, ob ein festgestelltes Sterberatenverhältnis mit großer Wahrscheinlichkeit durch Zufallsschwankungen zustande gekommen sein könnte oder ob relativ sicher auch bei weiteren Untersuchungen unter gleichen Bedingungen zwischen zwei derartigen Gruppen ein von 1 verschiedenes Sterberatenverhältnis zu beobachten sein würde. Man erinnere sich daran, daß bei einem Sterberatenverhältnis ein Wert von 1 !! Unterschiedslosigkeit bedeutet. Das hieße z.B., eine Therapie, die zur anderen ein Sterberatenverhältnis von 1:1 bewirkt, verändert die Sterbechancen der Patienten **nicht**.

Periodenvergleich

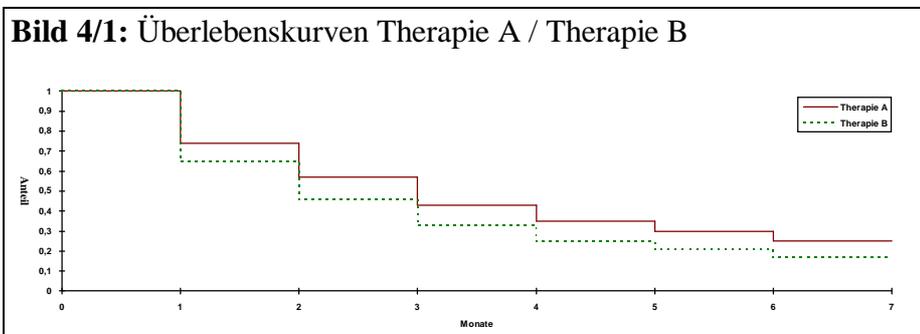
Neben dem Vergleich zweier oder mehrerer Gruppen über die gesamte Beobachtungszeit hinweg, kann man sich auch für Unterschiede zwischen diesen Gruppen in bestimmten Perioden interessieren. Sowohl das Sterberatenverhältnis als auch der Logranktest lassen sich auf den Vergleich innerhalb einer bestimmten Periode beschränken. Will man z.B. die unterschiedliche Wirksamkeit zweier Therapien im ersten Behandlungsjahr vergleichen, so errechnet man sowohl die D's als auch die E's nur für das erste Jahr und bildet daraus das Sterberatenverhältnis R und X-Quadrat. Stehen die Perioden, für die man Vergleiche durchführen will, nicht bereits vor Kenntnis der Daten fest, sollte man bei der Interpretation etwa gefundener großer, bzw. signifikanter Unterschiede sehr vorsichtig sein.

Der geschichtete Gruppenvergleich

Wir haben im Zusammenhang mit Beispiel 3 den Vergleich zweier Gruppen dargestellt. Nun sollen zwei Therapiearme in Beispiel 4 untersucht werden.

Beispiel 4 ist frei erfunden. Dabei wurde nicht auf Plausibilität geachtet. Es wurde so konstruiert, daß Konzepte einer multivariaten Auswertung leicht darstellbar sind.

Tabelle 4/1 und Bild 4/1 zeigen den Vergleich zwischen den Therapien A und B.



4500 Personen wurden mit A behandelt und 3000 mit B. Nach 6 Monaten wurde ausgewertet. Das Sterberatenverhältnis beträgt .79 oder 1/1.26 .

Anmerkung: Es ist sinnvoll, bei einem Verhältnis immer auch den Kehrwert anzugeben, also .79 und 1/1.26 ,da das Ergebnis nur davon abhängt, ob man A zu B oder B zu A in Verhältnis setzt.

Die Sterberaten in A und B verhalten sich wie .79 zu 1. Das ist gleichbedeutend damit, daß sich die Sterberaten von B zu A wie 1.26 zu 1 verhalten.

Tabelle 4/1: Vergleich zweier Therapiegruppen - Sterberatenverhältnis und Logranktest für Therapie A zu Therapie B des Beispiel 4

Zeitpunkt -->	1	2	3	4	5	6	SUMMEN
GRUPPE A: 00+01 = Therapie A							
d(a)..Tote	1190	743	618	357	231	245	3384
z(a)..Zensierte	0	0	0	0	0	1116	1116
n(a)..P unter Risiko	4500	3310	2567	1949	1592	1361	4500
sr(a)..lokale Überlebensrate	.74	.78	.76	.82	.85	.82	
S(a)..kumulative Überlebensrate	.74	.57	.43	.35	.30	.25	
e(a)..erw Tote	1344	819.95	670.13	380.60	244.25	256.21	3715.14
GRUPPE B: 10+11 = Therapie B							
d(b)..Tote	1050	560	415	214	130	130	2499
z(b)..Zensierte	0	0	0	0	0	501	501
n(b)..P unter Risiko	3000	1950	1390	975	761	631	3000
sr(b)..lokale Überlebensrate	.65	.71	.70	.78	.83	.79	
S(b)..kumulative Überlebensrate	.65	.46	.33	.25	.21	.17	
e(b)..erw Tote	896	483.05	362.87	190.40	116.75	118.79	2167.86
d(gesamt)	2240	1303	1033	571	361	375	
n(gesamt)	7500	5260	3957	2924	2353	1992	

Sterberatenverhältnis = $.79 = 1/1.26$; X-Quadrat = 80.10, df = 1

Mit dem Ergebnis könnten wir zufrieden sein, aber nun fällt unser Blick auf Tabelle 4/2. Wir sehen, daß 500 junge und 4000 alte Patienten Therapie A erhalten haben sowie 1000 junge Patienten und 2000 alte Patienten Therapie B. (Wann die Jugend aufhört und das Alter beginnt, ist hier belanglos.)

Tabelle 4/2: Vergleich der Altersstruktur in den beiden Therapiegruppen

Zahl der Patienten	Junge	Alte	Verhältnis
Therapie- gruppe A	500	4000	1 : 8
B	1000	2000	1 : 2

In Therapiegruppe A gibt es also 8 mal so viele Alte wie Junge, in der Gruppe B nur doppelt so viele. Das gibt zu denken, da Junge oft bessere Heilungschancen haben als Alte.

Das Sterberatenverhältnis von $.79$ bedeutet, daß unter A weniger sterben als unter B. Jetzt, da wir wissen, daß in Therapiegruppe A wesentlich mehr alte Patienten waren als in Gruppe B, vermuten wir, daß der Vergleich von Therapie A und B noch deutlicher für A ausgefallen wäre, wenn wir die ungleiche Altersstruktur berücksichtigt hätten. Das wollen wir nachholen.

Wir vergleichen dazu Therapie A mit B zuerst nur bei jungen Patienten (Tabelle 4/3), und dann nur bei alten Patienten (Tabelle 4/4).

In beiden Fällen erhalten wir ein Sterberatenverhältnis von $.67$ oder $1/1.49$. Die beiden Therapien unterscheiden sich in den Untergruppen (oder Schichten, daher die Bezeichnung "geschichteter Vergleich") deutlicher, als nach dem Gesamtvergleich zu erwarten war.

Hat man ein Ergebnis, wie in diesem konstruierten Beispiel, wird man natürlich schließen, daß $.67$ das unverzerrte Verhältnis von A zu B ist, während das erste Ergebnis, nämlich $.79$, durch die Unterschiede in der Altersstruktur verzerrt war.

Im allgemeinen erhält man aber in den einzelnen Untergruppen nicht genau das gleiche Ergebnis und man kombiniert dann die Einzelergebnisse nach der Formel in Tabelle 4/5. Im Beispiel erhalten wir dann $.68$ oder $1/1.45$ also etwa dasselbe Ergebnis wie in den einzelnen Altersschichten.

Tabelle 4/3: Vergleich der Therapieformen A und B nur für junge Patienten

Zeitpunkt -->	1	2	3	4	5	6	SUMMEN
GRUPPE A: 00 = Therapie A;Junge							
d(a)..Tote	70	52	49	33	24	27	255
z(a)..Zensierte	0	0	0	0	0	245	245
n(a)..P u Risiko	500	430	378	329	296	272	500
sr(a)..lok Üblrat	.86	.88	.87	.90	.92	.90	
S(a)..kum Üblrat	.86	.76	.66	.59	.54	.49	
e(a)..erw Tote	93.33	68.38	64.47	42.91	31.20	34.92	335.22
GRUPPE B: 10 = Therapie B;Junge							
d(b)..Tote	210	142	126	78	54	58	668
z(b)..Zensierte	0	0	0	0	0	332	332
n(b)..P u Risiko	1000	790	648	522	444	390	1000
sr(b)..lok Üblrat	.79	.82	.81	.85	.88	.85	
S(b)..kum Üblrat	.79	.65	.52	.44	.39	.33	
e(b)..erw Tote	186.67	125.62	110.53	68.09	46.80	50.08	587.78
d(gesamt)	280	194	175	111	78	85	
n(gesamt)	1500	1220	1026	851	740	662	

Sterberatenverhältnis = $.67 = 1/1.49$; X-Quadrat = 30.15, df = 1

Tabelle 4/4: Vergleich der Therapieformen A und B nur für alte Patienten

Zeitpunkt -->	1	2	3	4	5	6	SUMMEN
GRUPPE A: 01 = Therapie A;Alte							
d(a)..Tote	1120	691	569	324	207	218	3129
z(a)..Zensierte	0	0	0	0	0	871	871
n(a)..P u Risiko	4000	2880	2189	1620	1296	1089	4000
sr(a)..lok Üblrat	.72	.76	.74	.80	.84	.80	
S(a)..kum Üblrat	.72	.55	.41	.32	.27	.22	
e(a)..erw Tote	1306.67	790.57	640.79	359.48	227.38	237.45	3562.35
GRUPPE B: 11 = Therapie B;Alte							
d(b)..Tote	840	418	289	136	76	72	1831
z(b)..Zensierte	0	0	0	0	0	169	169
n(b)..P u Risiko	2000	1160	742	453	317	241	2000
sr(b)..lok Üblrat	.58	.64	.61	.70	.76	.70	
S(b)..kum Üblrat	.58	.37	.23	.16	.12	.08	
e(b)..erw Tote	653.33	318.43	217.21	100.52	55.62	52.55	1397.65
d(gesamt)	1960	1109	858	460	283	290	
n(gesamt)	6000	4040	2931	2073	1613	1330	

Sterberatenverhältnis = $.67 = 1/1.49$ X-Quadrat = 187.07, df = 1

Tabelle 4/5: Geschichteter Vergleich der Therapieformen A und B - Schichtungsvariable ist das Alter in zwei Gruppen (zu Beispiel 4)

$$Da = 255 + 3129 = 3384$$

$$Ea = 335.22 + 3562.35 = 3897.57$$

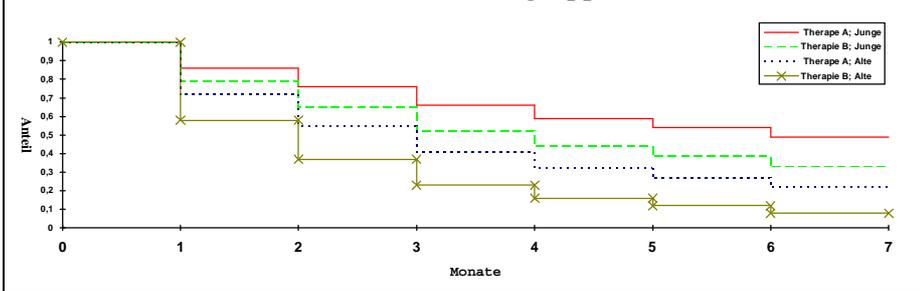
$$Db = 668 + 1831 = 2499$$

$$Eb = 587.78 + 1397.65 = 1985.43$$

$$\text{Sterberatenverhältnis} = \frac{Da \cdot Eb}{Db \cdot Ea} = \frac{3384 \cdot 1985.43}{2499 \cdot 3897.57} = .689 = 1/1.45$$

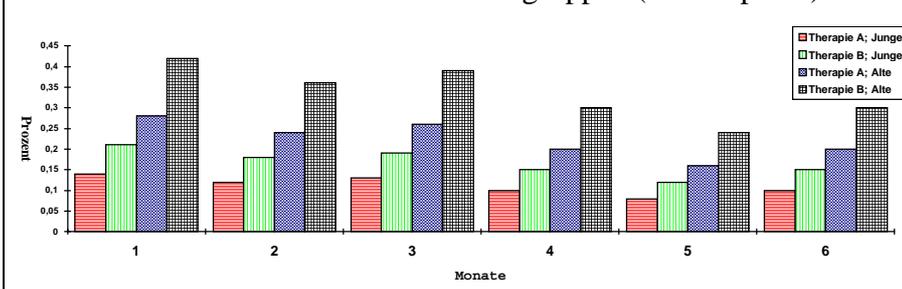
Sehen wir uns die zugehörigen Grafiken an. Bild 4/1 zeigt die Überlebenskurven von Therapie A und Therapie B, Bild 4/2 zeigt die Kurven für die einzelnen Untergruppen. Die Hazardfunktionen der vier Untergruppen sind in Bild 4/3 dargestellt. Ihnen entnimmt man, nach welchen Vorgaben das Beispiel konstruiert wurde.

Betrachten wir zuerst die Patienten, die mit Therapie A behandelt wurden und jung sind.

Bild 4/2: Überlebenskurven der 4 Untergruppen

zweiten Monat 13 Prozent sterben und so weiter.

Sie haben im ersten Monat einen Hazard von .14, im 2. Monat einen von .12, in den folgenden .13, .10, .08, .10. Das bedeutet, wie anhand des ersten Beispiels erklärt, daß von allen 500 Patienten im ersten Monat 14 Prozent sterben, daß von denen, die den Beginn des zweiten Monats erleben, im

Bild 4/3: Hazardfunktionen der 4 Untergruppen (zu Beispiel 4)

Vergleicht man die Hazardfunktion dieser Untergruppe mit jener der jungen Patienten unter Therapie B, dann sieht man, daß sich die Hazardfunktion der Untergruppe B; Jung aus der Funktion der ersten Untergruppe errechnen läßt.

Multipliziert man den Hazard für

den 4. Monat (.10) mit 1.5 erhält man den Hazard der Untergruppe B; Jung im 4. Monat nämlich .15. Das gilt für jeden Monat. Die beiden Hazardfunktionen unterscheiden sich nur um einen Proportionalitätsfaktor von 1.5. Einen entsprechenden Vergleich kann man auch zwischen Therapie A und B für die alten Patienten anstellen. Er ergibt denselben Faktor, obwohl der Hazard für alte Patienten an sich höher als für junge ist.

Der Unterschied zwischen jungen und alten Patienten läßt sich analog durch einen Proportionalitätsfaktor darstellen. Vergleicht man nach Therapien getrennt junge und alte Patienten, dann bekommt man jedesmal einen Faktor 2.

Alte Patienten haben im Beispiel 4 ein doppelt so großes Sterberisiko wie junge Patienten unter der jeweils gleichen Therapie.

In diesem Beispiel hat der Altersunterschied eine größere Bedeutung als die verschiedenen Therapieformen.

Proportionaler Hazard besteht dann, wenn eine Gruppe eine beliebige Hazardfunktion hat, von der sich die Hazardfunktion einer anderen Gruppe nur mehr um einen Proportionalitätsfaktor unterscheidet.

Die Berechnungen zum Beispiel 4 haben ein geschichtetes Sterberatenverhältnis zwischen den Therapien von .689 oder 1/1.45 ergeben. Das stimmt recht gut mit dem Proportionalitätsfaktor der vorgegebenen Hazardfunktionen von 1.5 überein.

- Anmerkung:
1. Breslow (1975) behandelt eingehend die Frage, wieweit sich das Sterberatenverhältnis als Schätzer für den Proportionalitätsfaktor eignet.
 2. Im Beispiel 4 wurden die Überlebenszeiten in Monaten angegeben und dadurch die Todesfälle gruppiert. Die Zeit wurde also als diskrete Variable behandelt und nicht, wie die Modelle voraussetzen, als kontinuierliche. Das bedingt kleine Unterschiede der Ergebnisse, die aber für das Verständnis der Modellvorstellung eines Proportional-Hazard-Modells nicht bedeutend sind.
 3. Man kann zwar ein Sterberatenverhältnis auch errechnen, wenn die Situation proportionaler Hazards nicht gegeben ist, es läßt sich dann jedoch nicht leicht interpretieren und ist systematisch von der Beobachtungsdauer abhängig.
Einem Mittelwerts- oder Medianvergleich ist es unter Umständen auch dann noch vorzuziehen.

geschichteter Logranktest

Mit den Werten für D_a , E_a , D_b und E_b , die man zur Berechnung des geschichteten Sterberatenverhältnisses verwendet hat, kann man auch einen geschichteten Logranktest rechnen. Man setzt dazu nur diese Werte in die Formel für X-Quadrat ein und vergleicht das Ergebnis mit denselben kritischen Werten, wie beim Gesamtgruppenvergleich (siehe Mantel, Haenszel 1959 und Mantel 1963).

ein Proportional-Hazard-Model (siehe Cox 1972, Breslow 1975)

Wir haben vorerst das geschichtete Sterberatenverhältnis zwischen den beiden Therapieformen in Beispiel 4 ausgerechnet und dann die Hazardfunktionen der einzelnen Untergruppen betrachtet.

Da das Beispiel rigoros unrealistisch konstruiert ist, lassen sich die Proportionalitätsfaktoren (die, wie oben erwähnt durch das Sterberatenverhältnis geschätzt werden können) für Therapie A zu Therapie B und für Jung zu Alt aus Bild 4/3 gut bestimmen. Die Hazardfunktionen von Therapie A zu B verhalten sich wie 1 zu 1.5 und von Jung zu Alt wie 1 zu 2.

Betrachtet man den Hazard der Untergruppe Therapie A; Jung in einem bestimmten Monat als den Basis-Hazard, dann erhält man den Hazard einer Person in einer anderen Untergruppe, indem man den Basis-Hazard bei Anwendung der Therapie B mit 1.5 multipliziert oder, wenn die Person alt ist, mit 2.

Hat eine junge Person mit Therapie A im 5. Monat einen Hazard von .08 so hat eine junge Person mit Therapie B im selben Monat einen Hazard von $.08 * 1.5 = .12$ oder eine alte Person mit Therapie A einen Hazard von $.08 * 2 = .16$. Eine Person, die Therapie B erhält und alt ist, hat dann einen Hazard von $.08 * 1.5 * 2 = .24$.

Die Einflüsse der einzelnen Faktoren - Therapie und Alter - werden, wie zu sehen, miteinander multipliziert.

Das P-H-Modell macht also zwei wichtige Annahmen:

1. die Annahme der proportionalen Hazardfunktionen.
2. die Annahme des multiplikativen Zusammenhanges der einzelnen Faktoren.

Im Beispiel 4 sind beide Annahmen erfüllt. Wir können daher die oben beschriebenen Zusammenhänge formal wie in Tabelle 4/6 darstellen.

Tabelle 4/6: Darstellung der Zusammenhänge in Beispiel 4
Ein einfaches Proportional-Hazard-Model nach Breslow (1975)

Kodierung der Kovariablen (z_1, z_2):

z_1 ... Therapie: A = 0, B = 1

z_2 ... Altersgruppe: Jung = 0, Alt = 1

Hazard einer Person mit den Kovariablenwerten z_1, z_2 zum Zeitpunkt t :

$$\lambda(t; z_1, z_2) = \lambda_0(t) * \exp(\beta_1 * z_1 + \beta_2 * z_2)$$

$$\lambda(t; z_1, z_2) = \lambda_0(t) * R_1^{z_1} * R_2^{z_2} \quad \text{mit } R_i = \exp(\beta_i)$$

R_i : Sterberatenverhältnis zwischen Gruppe 1 und Gruppe 2 der Kovariablen i .

Prognoseindex: $PI = \beta_1 * z_1 + \beta_2 * z_2$ $\lambda(t; PI) = \lambda_0(t) * \exp(PI)$

Tabelle 4/7: Zusammenfassung der Sterberatenverhältnisse in den verschiedenen Untergruppen

Vergleich	Sterberatenverhältnis	siehe
Therapie A/Therapie B einfach	.79 = 1/1.26	Tab 4/1
Therapie A/Therapie B nur junge Patienten	.67 = 1/1.49	Tab 4/3
Therapie A/Therapie B nur alte Patienten	.67 = 1/1.49	Tab 4/4
Therapie A/Therapie B geschichtet nach jung/alt	.69 = 1/1.45	Tab 4/5
junge Pat / alte Pat einfach	.58 = 1/1.72	
junge Pat / alte Pat nur Therapie A	.50 = 1/2.00	
junge Pat / alte Pat nur Therapie B	.51 = 1/1.97	
junge Pat / alte Pat geschichtet nach Therapie A/Therapie B	.54 = 1/1.86	

Dabei werden die Einflußfaktoren Therapie und Alter, wie in der Literatur üblich, Kovariablen genannt. Eine einfache Form, die Zugehörigkeit zu einer Untergruppe zu beschreiben, besteht darin, für jede Person in einer Therapievariablen die angewandte Therapieform mit 0 bzw. 1 zu vermerken, entsprechend Therapie A bzw. B. Die Altersgruppe Jung bzw. Alt wird ebenso vermerkt.

Nach dieser Kodierung braucht man nur mehr in die Formel einzusetzen:

$$\lambda(t; z_1, z_2) = \lambda_0(t) * R_1^{z_1} * R_2^{z_2}$$

Berechnen wir als Beispiel den Hazard einer alten Person unter Therapie A im 2. Monat: $\lambda_0(2)$, der Basishazard im 2. Monat beträgt .12 .

Das ist der Hazard der Jungen unter Therapie A oder allgemein, der Hazard jener Personen, die in allen Kovariablen den Wert 0 haben.

R_1 ist der Proportionalitätsfaktor der Kovariable Therapie (1.5) und R_2 der Faktor der Kovariable Alter (2.0), z_1 und z_2 sind die Werte der Beispielperson in den Kovariablen. Eine Person unter Therapie A hat $z_1=0$ und, wenn sie alt ist $z_2=1$.

so ergibt sich der Hazard mit

$$\lambda(2; 0, 1) = 0.12 * 1.5^0 * 2.0^1$$

$$\lambda(2; 0, 1) = 0.12 * 1 * 2.0$$

$$\lambda(2; 0, 1) = 0.24$$

Anmerkung: Weiters gilt, wenn der Hazard $\lambda(t; z_1, z_2, z_3, \dots) = \lambda_0(t) * prop$
dann ist die Überlebenskurve $S(t; z_1, z_2, z_3, \dots) = S_0(t)^{prop}$.
prop sei der Faktor, der sich aus Berücksichtigung aller Kovariablen ergibt.

In der Literatur ist folgende äquivalente Darstellung üblich:

$$\lambda(t; z_1, z_2, \dots) = \lambda_0(t) * \exp(\beta_1 * z_1 + \beta_2 * z_2 + \dots) \text{ mit } R_i = \exp(\beta_i) \text{ oder } \beta_i = \ln(R_i)$$

Das läßt sich umformen in $\ln(\lambda(t; z_1, z_2, \dots) / \lambda_0(t)) = \beta_1 * z_1 + \beta_2 * z_2 + \dots$

Danach definiert man den Prognoseindex: $PI = \beta_1 * z_1 + \beta_2 * z_2 + \dots$

und kann weiter schreiben: $\lambda(t; z_1, z_2, \dots) = \lambda_0(t) * \exp(PI)$.

Das Modell ist nicht, wie in Beispiel 4 auf zwei Kovariablen beschränkt. Es lassen sich neben dichotomen auch mehrkategoriale und kontinuierliche Variablen einbeziehen.

Prüfung der Annahmen:

Weiter oben haben wir die zwei wesentlichen Annahmen des P-H- Modells erwähnt. Praktisch hängt es vom verwendeten Computerprogramm ab, welche Möglichkeiten man hat, die Annahmen zu prüfen. Theoretisch mögliche Prüfungen werden in der Literatur dargestellt (siehe Andersen 1982, Cox 1979, Elashoff 1983).

Interpretation:

Die Analyse im Rahmen des P-H-Modells kann mit einer schrittweisen logistischen Regression erfolgen. Es werden Proportionalitätsfaktoren für ein Set von Kovariablen geschätzt.

Dieses Vorgehen ist für die Interpretation der Ergebnisse von wesentlicher Bedeutung. Es führt aber zu weit, es in diesem Rahmen darzustellen.

Setzen wir voraus, bereits vor der Auswertung einige wenige Kovariablen zu kennen, und deren Einfluß auf die Überlebenszeit abschätzen zu wollen.

In dieser Situation schätzen wir mit einem Computerprogramm die Proportionalitätsfaktoren der einzelnen Kovariablen. Im allgemeinen ist damit auch eine Signifikanzprüfung der einzelnen Einflüsse verbunden. Das Ergebnis für Beispiel 4 sieht dann wie in Tabelle 4/8 aus.

Tabelle 4/8: Ergebnisse einer Analyse der Daten aus Beispiel 4 entsprechend dem Cox-proportional hazard model

Variable	B	S.E.	Wald	df	Sig	R	Exp(B)
Therapie	-.4050	.0275	216.9809	1	.0000	-.0464	.6670
Alter	-.6930	.0374	343.3977	1	.0000	-.0585	.5001

Wir können dann anhand der Größe der Faktoren (in Zusammenhang mit der Kodierung der Kovariablen) die Stärke der Einflüsse der einzelnen Kovariablen beurteilen.

Außerdem läßt sich ein Prognoseindex bilden. Bei seiner Verwendung ist auf zwei wesentliche Punkte zu achten. Erstens darf er nur auf Personen angewandt werden, die mit jenen vergleichbar sind, auf deren Daten die Schätzung beruht.

Zweitens hängt die Brauchbarkeit stark von der Genauigkeit der Schätzung der einzelnen Faktoren und - im Fall einer schrittweisen Selektion - von der Auswahl der Kovariablen ab.

Grundsätzlich ist eine korrekte Interpretation der Ergebnisse einer multivariaten Analyse schwierig. Das gilt nicht nur für das Cox' sche P-H Modell. Die Situation ist jener bei der multiplen Regression oder bei der logistischen Regression vergleichbar.

Besonders in der Onkologie ist die Verwendung des P-H Modelles sehr modern (Stand 1993), aber die Interpretation der Ergebnisse zeigt, daß sich Computerprogramme zur Berechnung eines Modelles wesentlich schneller verbreiten, als das notwendige Verständnis zur entsprechenden Interpretation der Ergebnisse. Im gegebenen Rahmen können auch wir nur warnen.

Weiteres

Neben dem Proportional-Hazard-Model gibt es eine Reihe anderer multivariater Modelle für die Analyse von Überlebensdaten (siehe u.a. Kalbfleisch und Prentice, 1980). Vergleiche der Ergebnisse von Auswertungen mit verschiedenen Modellen findet man u.a. bei Hudec und Platz (1982).

Mir ist wenig weiterführende Literatur bekannt, die ich für ähnlich einfach lesbar halte wie dieses Skriptum.

Empfehlenswert scheint mir der zweiteilige Artikel von Peto et al. (1976, 1977). Er behandelt neben Kaplan-Meier-Verfahren und Logranktest wichtige Punkte der Planung randomisierter Studien. Dort findet man auch eine Tabelle, die die Chance eines signifikanten Ergebnisses unter verschiedenen Bedingungen abschätzen hilft (Tabelle für den Fehler 2. Art).

Der Artikel von Breslow (1975) stellt das P-H-Modell ausführlich in Zusammenhang mit dem geschichteten Logranktest dar. Dieser Darstellung bin ich über weite Strecken gefolgt. Breslow führt weitere Aspekte aus, die hier nicht besprochen werden konnten.

Die im Literaturverzeichnis angegebenen deutschsprachigen Artikel sind i.a. sehr leicht verständlich, gehen aber inhaltlich kaum über den dargestellten Bereich hinaus. Selbstverständlich gibt es aber eine große Zahl guter und tiefergehender englischsprachiger Bücher zum Thema.

Schlußbemerkung

Das Skriptum soll Grundlagen der Überlebenskurvenberechnung ohne Ansprüche an statistische und mathematische Vorbildung darstellen.

Dabei können viele wichtige Themen nicht oder nur unzureichend behandelt werden. Deren Behandlung findet man in Standardlehrbüchern der Statistik.

Text erstellt 1993, geringfügige Korrekturen 2001

Literatur

- Andersen PK.: Testing goodness-of-fit of Cox' s regression and life model. Biometrics 1982;38:67-77.
- Berchtold W.: Klinische Studien:Berechnen und Vergleichen von Überlebenskurven.Schweiz. med. Wschr. 1981;111:128-133.
- Breslow NE.: Contribution to the discussion of the paper by DR Cox. J R Statist Soc B 1972;34:216-7.
- Breslow NE.: Analysis of survival data under the proportional hazard model.Int Stat Rev 1975;1:45-58.
- Cox D.R.: Analysis of binary data. Chapman and Hall, London, 1970.
- Cox D.R.: Regression models and life tables (with discussion). J R Statist Soc B 1972;34:187-220.
- Cox D.R.:A note on the graphical analysis of survival data. Biometrika 1979;66, 188-190.
- Cutler S.J., Ederer F.: Maximum utilization of the life table method in analyzing survival. J Chron Dis. 1958;8:699-712.
- Elashoff J.D.: Surviving Proportional Hazards. Hepatology 1983;Vol 3, No 6:1031-1035.
- Hudec M.: Survival Analysis in Heterogenous Populations - Statistical Models and Concepts. In:Stochastic Modelling of Social Process, 1984.
- Hudec M., Platz H.: Fitting Regression Models to Censored Survival Data. Statistics in Medicine 1982;2:287-293.
- Kalbfleisch JD., Prentice RL.: The statistical Analysis of failure time data.New York:John Wiley and Sons, 1980.
- Kaplan E.L., Meier P.: Nonparametric estimation from incomplete observations. J Amer Statist Ass. 1958;53:457-481.
- Lienert G. A.: Verteilungsfreie Methoden in der Biostatistik, Band I und II. Hain, Meisenheim 1973.
- Mantel N.: Chi-Square Tests with one Degree of Freedom;Extensions of the Mantel-Haenszel Procedure. American Statistical Association Journal 1963;690-700.
- Mantel N., Haenszel W.: Statistical Aspects of the Analysis of Data from Retrospective Studies of Disease. J of the National Cancer Institute 1959;22, 4, 719-748.
- Neiß A., Selbmann H.K., Messerer D.: Analyse von Überlebenskurven. 29.Jahrestagung der Deutschen und Österreichischen Gesellschaften für Hämatologie und Onkologie (1984)
- Peto R., Pike MC., Armitage P., et.al.: Design and Analysis of randomized clinical trials requiring prolonged observation of each Patient.
I Introduction and Design. Brit J Cancer 1976;34:585-612.
II. Analysis and Examples. Brit J Cancer 1977;35;1-39.
- Selbmann, H.K.: Statistische Auswertungsverfahren in der klinisch-therapeutischen Forschung.In:Klinische Pharmakologie Hrsg.:Kuemmerle H.P. (1984)
- TNM-Klassifikation der malignen Tumoren:Statistische Behandlung von Überlebenszeiten.Berlin (1979)
- Tsiatis A.A.:A large sample study of Cox' s regression Model. The Annals of Statistics 1981;9, 1, 93-108.