

Berechnung des LOG-RANK-Tests bei Überlebenskurven

Hans-Dieter Spies
 inVentiv Health Germany GmbH
 Brandenburger Weg 3
 60437 Frankfurt
 hd.spies@t-online.de

Zusammenfassung

Mit Hilfe von Überlebenskurven stellt man die Sterberaten von Populationen über einen Zeitraum/eine Zeitspanne grafisch dar. Das herausragende Ereignis ist hierbei der Tod. Es können aber auch andere Ereignisse definiert werden, was die Anwendung von Überlebenskurven in verschiedenen Bereichen ermöglicht.

SAS stellt hierfür die Prozedur PROC LIFETEST zur Verfügung. Dieser Beitrag erläutert die Kaplan-Meier-Methode, auf der die Prozedur PROC LIFETEST beruht. Dabei wird der Kaplan-Meier-Schätzer für die einzelnen Ereignisse berechnet. Im Weiteren wird der LOG-RANK-Test erklärt, mit dem man die Überlebenskurven von Gruppen (z.B. verschiedene Behandlungsgruppen einer klinischen Studie) vergleichen kann. Der Algorithmus des LOG-RANK-Tests wird zunächst an einem Beispiel mit kleinen Gruppen ausführlich dargestellt.

Schlüsselwörter: Kaplan-Meier-Schätzer, PROC LIFETEST, LOG-RANK-Test

1 Einführung

Die Survival-Analyse untersucht die Verweildauer von Elementen bis zu einem bestimmten Ereignis. Die Elemente können Personen sein, die zwei unterschiedliche Behandlungen bekommen. Das Ereignis kann dabei Tod sein, jedoch auch beliebige andere Endpunkte wie Heilung, Erkrankung oder Eintreten einer Komplikation.

An Hand eines einfachen Beispiels sollen 2 Fragen beantwortet werden:

- 1) Wie groß ist die Wahrscheinlichkeit, dass ein Ereignis zu einer bestimmten Zeit eintritt?
 Der Kaplan-Meier-Schätzer schätzt diese Wahrscheinlichkeit. Das Verfahren wurde 1958 von Edward L. Kaplan und Paul Meier entwickelt. [1]
- 2) Man betrachtet die Überlebenskurven von 2 Gruppen: Wie kann man statistisch signifikante Unterschiede von 2 Überlebenskurven bestimmen?
 Eine Lösung bietet der LOG-RANK -Test. [2] Das ist ein nicht-parametrischer Test, der den Vorteil bietet, dass keine bestimmte Verteilung zugrunde liegen muss.

2 Kaplan–Meier Methode

Beginnen wir mit einem Beispiel:

Gegeben seien 2 Gruppen (A und B) mit jeweils 5 Personen. Es wird ein Ereignis definiert. In Gruppe A tritt dieses Ereignis bei 3 Personen auf, in Gruppe B tritt dieses Ereignis bei allen 5 Personen auf. Wenn bei einer Person das Ereignis nicht auftritt, wird die Dauer angegeben, die diese Person unter Beobachtung ist. In diesem Fall wird ein * (Stern) hinter der Dauer geschrieben (siehe bei Gruppe A). In allen anderen Fällen entspricht die Dauer der Zeitspanne bis zum Zeitpunkt des Eintretens des Ereignisses.

Tabelle 1: Listing der Testdaten

Dauer Gruppe A	Dauer Gruppe B
3	15
6*	22
10	28
15	36
32*	40

Ein großer Vorteil der Kaplan-Meier Methode ist, dass zensierte Daten (es tritt kein Ereignis ein) berücksichtigt werden. Wird bei einer klinischen Studie die Dauer bis zum Auftreten eines bestimmten Ereignisses untersucht, so kann es passieren, dass Patienten vorzeitig ausscheiden, bevor das Ereignis eingetreten ist. In diesen Fällen wird die Dauer bis zum Ausscheiden als zensierte Dauer berücksichtigt. Ebenso Patienten, die bis zum Ende der geplanten Studiendauer ohne Ereignis sind, werden mit der Studiendauer als zensierte Dauer gewertet.

Nun zur ersten Frage: Der Kaplan-Meier-Schätzer (KM-Schätzer), auch als Produkt-Limit-Schätzer bekannt, dient zum Schätzen der Wahrscheinlichkeit, dass bei einem Versuchsobjekt ein bestimmtes Ereignis innerhalb eines Zeitintervalls nicht eintritt.

Der Kaplan–Meier Schätzer $\hat{S}(t)$ ist als Produkt definiert.

$$\hat{S}(t) = \prod_{t(i) \leq t} \frac{n_i - d_i}{n_i}$$

wobei $\hat{S}(0) = 1$

d_i = die Anzahl der Ereignisse, die zu dem Zeitpunkt $t(i)$ eintreten.

n_i = die Anzahl der Personen unter Risiko zu dem Zeitpunkt $t(i)$, das ist die Anzahl der Personen, die bis zu dem Zeitpunkt $t(i)$ noch kein Ereignis hatten und auch noch unter Beobachtung sind.

In Gruppe A sind 2 Personen ohne Ereignis während der Beobachtungszeit. In Gruppe B gibt es zu jeder Person ein Ereignis.

Tabelle 2: Berechnung der Kaplan-Meier-Schätzer

Gruppe	Dauer t(i)	0=Ereignis 1=zensiert	Unter Risiko n _i	Anzahl Ereignis d _i	KM-Schätzer
A	3	0	5	1	$(5-1)/5 = 0,8000$
A	6	1	4	0	$\hat{S}(3) * (4-0)/4 = 0,8000$
A	10	0	3	1	$\hat{S}(6) * (3-1)/3 = 0,5333$
A	15	0	2	1	$\hat{S}(10) * (2-1)/2 = 0,2667$
A	32	1	1	0	$\hat{S}(15) * (1-0)/1 = 0,2667$
B	15	0	5	1	$(5-1)/5 = 0,8000$
B	22	0	4	1	$\hat{S}(15) * (4-1)/4 = 0,6000$
B	28	0	3	1	$\hat{S}(22) * (3-1)/3 = 0,4000$
B	36	0	2	1	$\hat{S}(28) * (2-1)/2 = 0,2000$
B	40	0	1	1	$\hat{S}(36) * (1-1)/1 = 0,0000$

Überlebenskurven werden in der Regel als Treppenkurven dargestellt. Da $\hat{S}(0) = 1$ definiert ist, beginnt die Kurve bei 1 und geht in Stufen runter. Die folgende Abbildung zeigt die Überlebenskurve für die Testdaten.

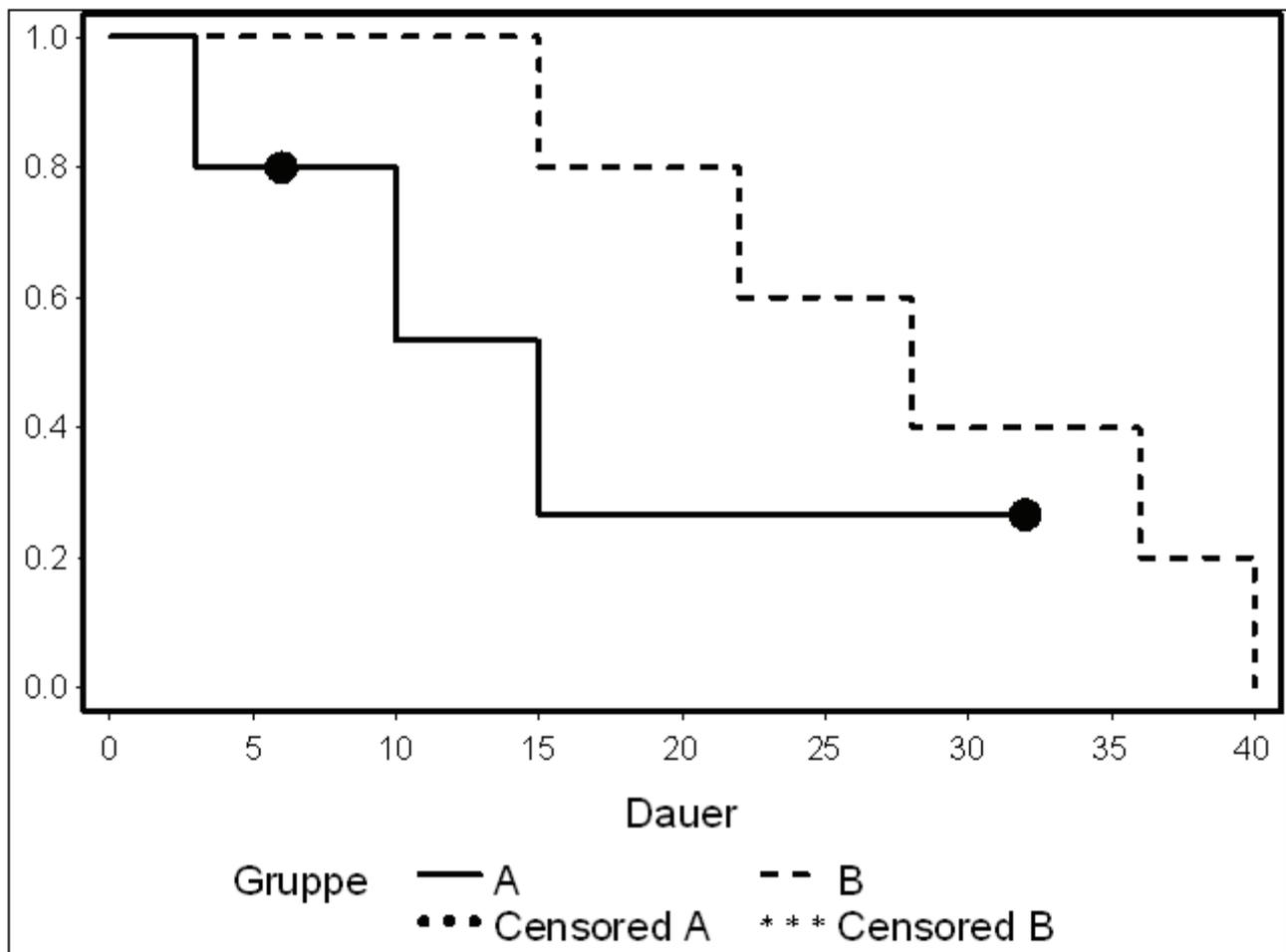


Abbildung 1: Kaplan-Meier Kurve für die Testdaten

Beide Kurven starten bei der Dauer 0 mit dem Wert 1. Da für die Gruppe A bereits nach dem Zeitpunkt $t(1)=3$ das erste Ereignis eintritt, fällt hier die Kurve auf 0.8. Der Zeitpunkt $t(2)=6$ wird durch einen Punkt markiert (diese Person hat kein Ereignis). Die Kurve für die Gruppe A endet zu dem Zeitpunkt $t(5)=32$ bei einem Wert von 0.2667. Die Kurve der Gruppe A geht nicht runter auf 0, da die längste Dauer von einer zensierten Person kommt.

Den Wert für den Kaplan-Meier-Schätzer kann man für jede Dauer an der y-Achse ablesen.

Mit der SAS Prozedur PROC LIFETEST kann man den KM-Schätzer berechnen. Der Programm Code sieht folgendermaßen aus:

```
PROC LIFETEST DATA = testdata;
  STRAT group / test=LOGRANK;
  TEST status;
  TIME duration*status{0};
quit;
```

Im output Fenster von SAS findet man die verschiedenen Kenngrößen wieder:

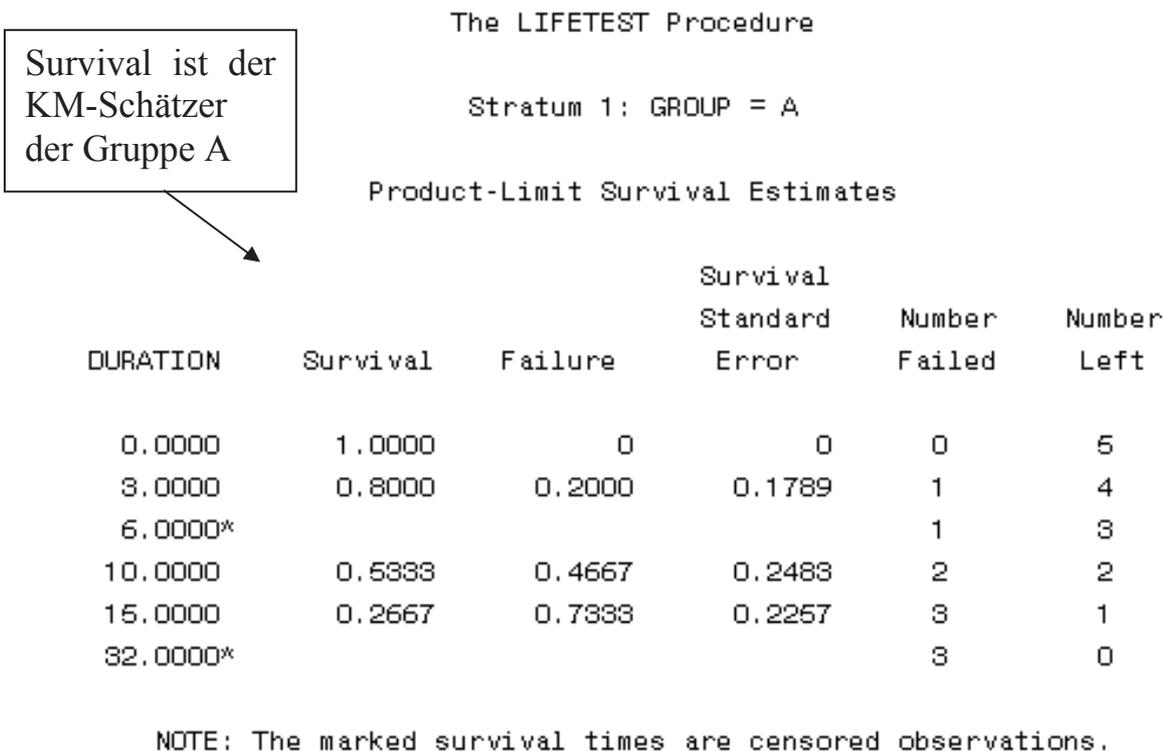


Abbildung 2: SAS Output für PROC LIFETEST

Analog erhält man die Ergebnisse für die Gruppe B.

3 Der LOG-RANK-Test

Der LOG-RANK-Test ist ein in der Statistik häufig verwendeter Test, wenn man Überlebens-Verteilungen von 2 oder mehr Gruppen vergleichen will. In der klinischen Forschung untersucht man in der Regel verschiedene Behandlungsgruppen.

Im Folgenden wird die Konstruktion des Tests für 2 Gruppen ($g=1,2$) dargestellt. Die Dauer aller Ereignisse (aus beiden Gruppen) werde mit $t_1 < t_2 < \dots < t_k$ bezeichnet, wobei sie nach der Länge der Dauer sortiert sind.

Weiter sei

- d_{gj} = Anzahl der Personen in der Gruppe g mit einem Ereignis zum Zeitpunkt t_j ($g = 1,2 ; j = 1,2, \dots ,k$)
- $d_j = d_{1j} + d_{2j}$ = Anzahl aller Ereignisse zum Zeitpunkt t_j
- $n_g(t_j)$ = Anzahl der Personen in der Gruppe g , die zum Zeitpunkt t_j noch unter Risiko sind ($g = 1,2 ; j = 1,2, \dots ,k$)
- $n(t_j) = n_1(t_j) + n_2(t_j)$ = Anzahl aller Personen, die zum Zeitpunkt t_j noch unter Risiko sind ($j = 1,2, \dots ,k$)

Die Information zu jedem Zeitpunkt t_j kann in einer 4 Feldertafel zusammengefasst werden, wobei zu jedem Zeitpunkt t_j das Auftreten eines Ereignisses in Beziehung zu den Personen unter Risiko in jeder Gruppe in Beziehung gesetzt wird.

Betrachtet man im Beispiel die Dauer $t_j = 15$, so tritt nach dieser Dauer in beiden Gruppen ein Ereignis ein.

	Ereignis tritt ein	Personen unter Risiko
Gruppe A	1	2
Gruppe B	1	5

Allgemein ergibt sich zu jedem Zeitpunkt t_j :

	Ereignis tritt ein	Personen unter Risiko
Gruppe 1	d_{1j}	$n_1(t_j)$
Gruppe 2	d_{2j}	$n_2(t_j)$

Dies führt zu der Log-Rank Statistik für 2 Gruppen, die X^2 (Chi-Quadrat) verteilt ist.

$$\frac{(\sum (d_{gj} - (n_i(t_j) * d_{gj})))^2}{\text{Var}(\sum (d_{gj} - (n_i(t_j) * d_{gj})))} \sim X^2$$

wobei $g = 1,2 ; j = 1,2, \dots ,k$ sind.

Im Folgenden wird dem Wert für Chi-Quadrat die Wahrscheinlichkeit des LOG-RANK-Test an Hand der Testdaten nachgerechnet. In der folgenden Tabelle werden die einzelnen Schritte dargestellt. In der oberen Hälfte der Tabelle werden die Ereignisse und die jeweilige Anzahl der Personen unter Risiko dargestellt. Jeweils in der ersten Spalte wird die Dauer gelistet, wann ein Ereignis eintritt.

Bei der Anzahl der Personen unter Risiko müssen zensierte Personen berücksichtigt werden. So ist für den Zeitpunkt $t_j = 10$ in der Gruppe A die Anzahl der Personen unter Risiko =3 (die Personen mit der Dauer =3 und Dauer =6 dürfen nicht mitgezählt werden), während in der Gruppe B zu diesem Zeitpunkt noch alle 5 Personen unter Risiko sind.

Nun wird für Zeitpunkt t_j die erwartete Wahrscheinlichkeit berechnet, ob ein Ereignis eher in Gruppe A oder Gruppe B stattfindet. Die Formel dafür ist: $e_{gj} = d_j * n_{gj} / n_j$. Die Summe der Erwartungswerte für beide Gruppen ergibt jeweils die Anzahl der Ereignisse zu diesem Zeitpunkt. Für die Zeitpunkt $t_j = 15$ ist die Summe gleich 2, sonst ist die Summe jeweils gleich 1.

Die Varianz ist für jeden Zeitpunkt t_j definiert als $(d_j * (n_j - d_j) * n_{1j} * n_{2j}) / (n_j * n_j * (n_{j-1}))$. Die Varianz über alle Zeitpunkte ist die Summe = 1.1720.

Tabelle 3: Berechnung der Varianz

Dauer t_j	Ereignisse zum Zeitpunkt t_j			Personen unter Risiko zum Zeitpunkt t_j		
	A	B	beide	A	B	beide
Gruppe	d_{1j}	d_{2j}	$d_j = d_{1j} + d_{2j}$	n_{1j}	n_{2j}	$n_j = n_{1j} + n_{2j}$
3	1	0	1	5	5	10
10	1	0	1	3	5	8
15	1	1	2	2	5	7
22	0	1	1	1	4	5
28	0	1	1	1	3	4
36	0	1	1	0	2	2
40	0	1	1	0	1	1

Dauer t_j	Erwartete Wahrscheinlichkeit		Beobachtet - Erwartet		Varianz
Gruppe	A	B	A	B	beide
	$e_{1j} = d_j * n_{1j} / n_j$	$e_{2j} = d_j * n_{2j} / n_j$	$d_{1j} - e_{1j}$	$d_{2j} - e_{2j}$	$(d_j * (n_j - d_j) * n_{1j} * n_{2j}) / (n_j * n_j * (n_j - 1))$
3	0.5000	0.5000	0.5000	-0.5000	0.2500
10	0.3750	0.6250	0.6250	-0.6250	0.2344
15	0.5714	1.4286	0.4286	-0.4286	0.3401
22	0.2000	0.8000	-0.2000	0.2000	0.1600
28	0.2500	0.7500	-0.2500	0.2500	0.1875
36	0	1	0	0	0
40	0	1	0	0	0
	1.8964	6.1036	1.1036	-1.1036	1.1720

Nun kann der Wert für Chi-Quadrat berechnet werden. Hierfür kann man entweder die Werte der Gruppe A oder die Werte der Gruppe B zugrunde legen.

Tabelle 4: Berechnung von Chi-Quadrat

	Anzahl der Ereignisse d_g	Summe Erwartungswert e_g	Chi-Quadrat $(d_g - e_g)^2 / \text{var}(d_g - e_g)$
Gruppe A	3	1,8964	1,0391
Gruppe B	5	6,1036	1,0391

$g = \{\text{Gruppe A, Gruppe B}\}$

Die Wahrscheinlichkeit kann mit der SAS Funktion PROBCHI() berechnet werden.

```
data chi;
  x1=1-probchi(1.0391,1);
run;
```

Das Ergebnis ist $x1 = 0.3080308805$

Mit verschiedenen Farben sind die Ergebnisse der einzelnen Schritte eingerahmt.

 Diese Form zeigt die Summe der (Beobachteten – Erwarteten) Werte.

 Diese Form zeigt die Varianz.

 Diese Form zeigt den charakteristischen Wert der Chi-Quadrat Verteilung.

 Diese Form zeigt (1 minus “der Wahrscheinlichkeit von Chi-Quadrat”).

In dem folgenden SAS-Ausdruck der Prozedur LIFETEST findet man die oben berechneten Werte wieder.

The LIFETEST Procedure

Testing Homogeneity of Survival Curves for DURATION over Strata

Rank Statistics

GROUP	Log-Rank
A	1.1036
B	-1.1036

Covariance Matrix for the Log-Rank Statistics

GROUP	A	B
A	1.17201	-1.17201
B	-1.17201	1.17201

Test of Equality over Strata

Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	1.0391	1	0.3080

Abbildung 3: Output von dem SAS Programm mit PROC LIFETEST

In dem SAS output sind alle zuvor berechneten Werte vorhanden und entsprechend markiert.

Das Ergebnis des LOG-RANK-Test mit $p = 0.3080$ besagt, dass sich die 2 Kurven nicht signifikant unterscheiden.

4 Zusammenfassung

Es wurde ein einfaches Beispiel gewählt, um die Rechenschritte der Survival-Analyse zu zeigen. Bei der geringen Fallzahl ist die Aussagekraft des Tests mit Vorsicht zu betrachten, aber bei einer größeren Fallzahl wären die Berechnungen unübersichtlich geworden.

- Berechnung der Kaplan-Meier-Schätzer
Mit Hilfe von Beispieldaten wurde gezeigt, wie die Survival-Analyse zensierte Fälle behandelt.
- Demonstration des LOG-RANK-Tests
Dazu wurde der Wert der Charakteristik der Chi-Quadrat Verteilung berechnet und damit die Wahrscheinlichkeit ermittelt, mit der sich zwei Überlebenskurven unterscheiden.
- Die SAS Prozedur PROC LIFETEST liefert alle gewünschten charakteristischen Werte.
- Ausblick
Neben der Prozedur PROC LIFETEST können mit Hilfe der Prozedur PROC PHREG (Proportional Hazards Regression) Proportional Hazards-Modelle geschätzt werden. Dieser semi-parametrische Ansatz wird verwendet, um den Einfluss erklärender Kovariablen auf die Hazardrate zu untersuchen.

Literatur

- [1] E.L. Kaplan, P. Meier (1958) Nonparametric Estimation from Incomplete Observations. Journal of the American Statistical Association. Vol. 53, 1958, 282
- [2] D. G. Kleinbaum & M. Klein (2005): Survival Analysis. A Self-Learning Text; Springer: USA, 2005 (2.Ed)