

Überlebenszeitanalysen und Logistische Regression

Hartmut Herrmann / Mai 1998

Inhalt

I. Einführung in Überlebenszeitanalysen

- 1.1. Einleitung
- 1.2. Wahrscheinlichkeitsmodell
- 1.3. Die vier Programmaufrufe

II. Sterbetafeln

- 2.1. Das Konzept
- 2.2. Programmgestaltung im Dialogfeld und Beispiel
- 2.3. Ergebnisliste und Graphiken

III. Die Produkt-Limit-Methode

- 3.1. Konzept
- 3.2. Programmgestaltung im Dialogfeld
- 3.3. Gruppenvergleich
- 3.4. Ergebnisliste
- 3.5. Trend- und Schichtanalyse 13

IV. Überlebenszeitanalysen mit Kovariaten, das COX-Modell

- 4.1. Proportional Hazards Regression Model
- 4.2. Programmgestaltung im Dialogfeld und Beispiel
- 4.3. Schätzen von β und Testen von $H_0: \beta = 0$
- 4.4. Kategorielle Kovariate
- 4.5. Schrittweise Analysen
- 4.6. Optionen und Graphiken
- 4.7. Geschichtete Analysen - nicht proportionales Modell

V. Logistische Regression

- 5.1. Einführung
- 5.2. Programmgestaltung im Dialogfeld und Beispiel
- 5.3. Schätzen von β und Testen von $H_0: \beta = 0$
- 5.4. Kategorielle Kovariate
- 5.5. Schrittweise Analysen
- 5.6. Optionen
- 5.7. Zuordnung neuer Fälle
- 5.8. Kreuzvalidierung

Diese Einführung in den Umgang mit Überlebenszeitanalysen und logistischer Regression in SPSS setzt Kenntnisse in der Handhabung von **SPSS for Windows** voraus.

I. Einführung in die Überlebenszeitanalysen

1.1. Einleitung

Die Zielgröße Überlebenszeit meint die Länge eines Intervalls, dessen Anfangspunkt durch ein spezielles Ereignis wie Operation oder Beginn einer Therapie gesetzt wird und dessen Ende das Ereignis Tod - allgemeiner: **Response** - markiert. Das Ziel der Analyse liegt darin,

- die Verteilung der Überlebenszeit bzw. die Überlebenszeitkurve zu ermitteln,
- verschiedene Gruppen von Individuen bezüglich ihrer Überlebenszeitverteilung zu vergleichen,
- zu untersuchen, ob diese Kurve von weiteren Variablen, sog. Kovariaten, beeinflusst wird.

Alle Individuen, die an der Studie für Überlebenszeitanalysen teilnehmen, unterliegen dem Risiko des Responses. Es ist jedoch für ein Individuum möglich, daß es aus der Studie ausscheidet, bevor der Response eingetreten ist, z.B. durch Verweigerung einer Therapie, durch Zusatztherapie mit Einfluß auf die Überlebenszeit, durch Umzug oder durch Beenden der gesamten Studie. Die Kontrolle über das Überleben endet an diesem vorgezogenen Zeitpunkt. Individuen mit solchen unvollständigen Daten bezeichnet man als **zensierte Fälle**.

Wenn z.B. die Gesamtstudiendauer auf zehn Jahre festgelegt ist und sechs Monate vor deren Ende noch ein Patient in die Studie aufgenommen wird und nach einem halben Jahr noch lebt, scheidet er also nach sechs Monaten als zensiert aus der Studie aus.

Als zensiert ist ein Fall auch dann anzusehen, wenn der Tod durch eine Ursache, die nicht Gegenstand der Studie ist, z.B. durch Verkehrsunfall, eintritt.

Der Anlaß, der gegebenenfalls bei einem Patienten zur Zensierung führt, darf nicht mit dem untersuchten Ereignis zusammenhängen. Zu verzerrten Ergebnissen gelangt man z.B., wenn Patienten kurz vor dem sich abzeichnenden Response aus der Studie genommen werden, um sie noch einer anderen Therapie zu unterziehen.

Der Vorteil der Überlebenszeitanalysen liegt darin, daß sie nicht nur die Überlebenszeitintervalle der Verstorbenen, sondern daß sie auch die zensierten Fälle berücksichtigen - jeweils bis zum notierten Zeitpunkt des Verlassens der Studie.

Modell und Analyse der Überlebenszeiten lassen sich auf andere Probleme übertragen: Bei Krankheitsverläufen kann der Response einen Progreß oder die Feststellung der Heilung bedeuten, bei periodisch wiederkehrenden Krankheiten wird die "gesunde Zeit" bis zum Auftreten eines neuen Rezidivs beobachtet, bei stationären Aufenthalten könnten für die Zeit auch die aufgelaufenen Kosten gesetzt werden, im Berufsleben interessiert die Zeit von der Einstellung bis zum Ausscheiden aus einem Betrieb, bei technischen Geräten die Betriebsdauer bis zum ersten Betriebsausfall.

Speziell bei der Auswertung klinischer Studien eignet sich die Überlebenszeitanalyse auch für den Gruppenvergleich von Zeitabschnitten.

1.2. Wahrscheinlichkeitsmodell

Dieser Abschnitt soll dem mathematisch Interessierten zum besseren Verständnis dienen, er kann übergangen werden.

Es sei T das Lebensalter, die Zeit von Geburt bis zum Tod. T ist eine Zufallsgröße mit der Verteilungsfunktion $F(t)$. Dazu bezeichnet man $S(t) = 1 - F(t)$ als Verteilungsfunktion der Überlebenszeit :

$S(t) = \Pr(T > t)$ = Wahrscheinlichkeit, das Alter t zu erleben.

$\Pr(t < T < t+dt | T > t)$ ist die bedingte Wahrscheinlichkeit, unter der Bedingung, daß das Alter t erlebt wird, im kleinen Zeitraum $(t, t+dt)$ zu sterben.

Als Sterbeintensität bzw. **Hazardfunktion** $h(t)$ wird definiert (mit $dt \rightarrow 0$) :

$$h(t) = \Pr(t < T < t+dt | T > t) / dt = \left[\Pr(t < T < t+dt) / dt \right] / \Pr(T > t) = - (dS(t) / dt) / S(t) = - d(\ln S(t)) / dt.$$

Daraus folgt:

$$\ln(S(t)) = - \int_0^t h(v) dv \quad \text{und daraus:} \quad S(t) = \exp \left[- \int_0^t h(v) dv \right]$$

Falls $h(t)$ als konstant gleich h angesehen werden kann, ergibt sich: $S(t) = \exp(- ht)$.

Dann besitzt $F(t) = 1 - S(t) = 1 - \exp(- ht)$ eine Exponentialverteilung.

1.3. Die vier Programmaufrufe

SPSS verteilt die möglichen Analysen, die insbesondere von den Modellannahmen für die Hazardfunktion abhängen, auf vier verschiedene Programmaufrufe. Sie werden angezeigt, wenn man im SPSS-Statistikmenü die **Überlebenszeitanalysen** anklickt:

Sterbetafeln, Kaplan -Meier, COX-Regression und COX mit zeitabhängigen Kovariaten.

Sterbetafeln: Der zu betrachtende Zeitbereich wird in eine Anzahl von z.B. sieben oder zehn gleich lange Intervalle unterteilt. Die Berechnungen, insbesondere Schätzwerte für die Überlebenszeitfunktion, werden dann jeweils zu den Endpunkten der Intervalle durchgeführt. Bezüglich der Überlebenszeitverteilung sind Gruppenvergleiche möglich.

Kaplan-Meier: Die Produkt-Limit-Methode nach Kaplan-Meier erlaubt differenziertere Analysen. Sie stellt Berechnungen zu jedem Zeitpunkt an, an dem ein Response erfolgte. Auch hier sind Gruppenvergleiche, darüber hinaus Trend- und geschichtete Analysen möglich.

COX-Regression: Unter diesem Ansatz wird untersucht, ob die Hazardfunktion und damit die Überlebenszeitfunktion von Kovariaten beeinflusst wird, d.h. ob vermutete Einflußgrößen einen prognostischen Wert für die Überlebenszeitverteilung besitzen.

COX mit zeitabh. Kov.: Es ist möglich, beim COX-Modell die zeitliche Veränderlichkeit der Kovariaten zu berücksichtigen und die mögliche Abhängigkeit der Überlebenszeitverteilung auch von solchen Kovariaten zu untersuchen.

II. Sterbetafeln

(Life Tables, Cutler, Ederer; 1958)

2.1. Das Konzept

Der betrachtete Abschnitt der Zeitachse - er wird zumeist den größten beobachteten Zeitraum umfassen - wird in gleich große Intervalle, z.B. in 10 Intervalle eingeteilt. Für das i -te Intervall wird definiert:

z_i = Länge des Zeitintervalls,

n_i = Anzahl der Lebenden am Anfang des Intervalls,

c_i = Anzahl Zensierter,

die Annahme, daß die Zensierten dieses Intervalls im Mittel bis zur Mitte des Intervalls der Studie angehören, ist dem gleichwertig, daß die Hälfte am Anfang, die übrigen am Ende aus der Studie ausscheiden; daher :

$r_i = n_i - c_i / 2$ = Lebende unter Risiko,

d_i = Anzahl der Gestorbenen (Responses) im Intervall,

$q_i = d_i / r_i$ = bedingte Wahrscheinlichkeit (Schätzwert), im i -ten Intervall zu sterben, unter der Bedingung, daß das Intervall lebend erreicht wird,

$p_i = 1 - q_i$ = bedingte Überlebenswahrscheinlichkeit,

$S_i = (1 - q_1)(1 - q_2) \dots (1 - q_i) = p_1 p_2 \dots p_i$ = Überlebenszeitfunktion, Schätzwert für die Wahrscheinlichkeit, das Ende des i -ten Intervalls, d.h. den Zeitpunkt t_i , zu erleben,

$f_i = (S_i - S_{i+1}) / z_i$ = Dichtefunktion für die Sterbewahrscheinlichkeit

$h_i = (q_i / (1 - q_i / 2)) / z_i = (2q_i / (1 + p_i)) / z_i$ = Hazardfunktion oder Todesrate oder Sterbeintensität, Schätzwert in der Mitte des Intervalls.

Da die Response-Zeitpunkte eines Intervalls nicht alle genau auf das Ende des Intervalls fallen, sondern sich eher gleichmäßig darüber verteilen, erfährt die Überlebenszeitkurve $S(t)$ keine sprunghaften Veränderungen an den Intervallgrenzen t_i , sondern gleichmäßige Veränderungen zwischen den Grenzpunkten (t_i, S_i) , siehe Abb. 2.4. Entsprechend ist der Verlauf von $S(t)$ mit Hilfe der Punkte (t_i, S_i) als Polygonzug darzustellen. Nicht ganz angemessen ist also die in der SPSS-Graphik gezeigte Treppenfunktion.

2.2. Programmgestaltung im Dialogfeld und Beispiel

Die verwendeten Daten entstammen einer kontrollierten Studie 'Therapievergleich' mit einer Test- und einer Standardgruppe von je 38 Patienten. Für diese wurden der beobachtete Zeitraum (Variable *zeit*), der Status am Ende dieses Zeitraums ($tot = 1$, zensiert = 2 in der Variablen *state*) und die Therapieart notiert, darüber hinaus das Geschlecht und die Diagnose (vier verschiedene), die in beiden Therapiegruppen in etwa gleichverteilt vorkommen. Da die Lifetable nur für die ersten 600 Tage interessiert, gehen alle länger Lebenden als zensierte Fälle ein. Die Werte der Sterbetafel sollen zu Zeitpunkten in Abständen von je 100 Tagen berechnet werden.

Das Anklicken des Menüunterpunktes 'Sterbetafeln' eröffnet das zugehörige Dialogfeld, siehe **Abb. 2.1**. Das Dialogfeld zeigt zu Beginn im linken Rechteck alle definierten Variablen der Datenbank. Die für die Analyse benötigten Variablen sind in die entsprechenden Felder des mittleren Bereichs hinüberzuschieben, wie in Abb. 2.1 zu unserem Beispiel bereits geschehen:

Oben die Variable *zeit*, die den Zeitraum in Tagen angibt; dazu sind näher zu bestimmen: der Zeitbereich (0 bis ...), der insgesamt berücksichtigt werden soll, hier also bis 600 Tage, und die Schrittlänge für die Unterteilung der Gesamtzeit in Teilintervalle, hier 100 Tage.

Notwendig muß auch die Statusvariable eingefügt und - nach Anklicken von "**Ereignis definieren**" - der oder die Codes, die einen Response kennzeichnen,

benannt werden, hier die Zahl 1.

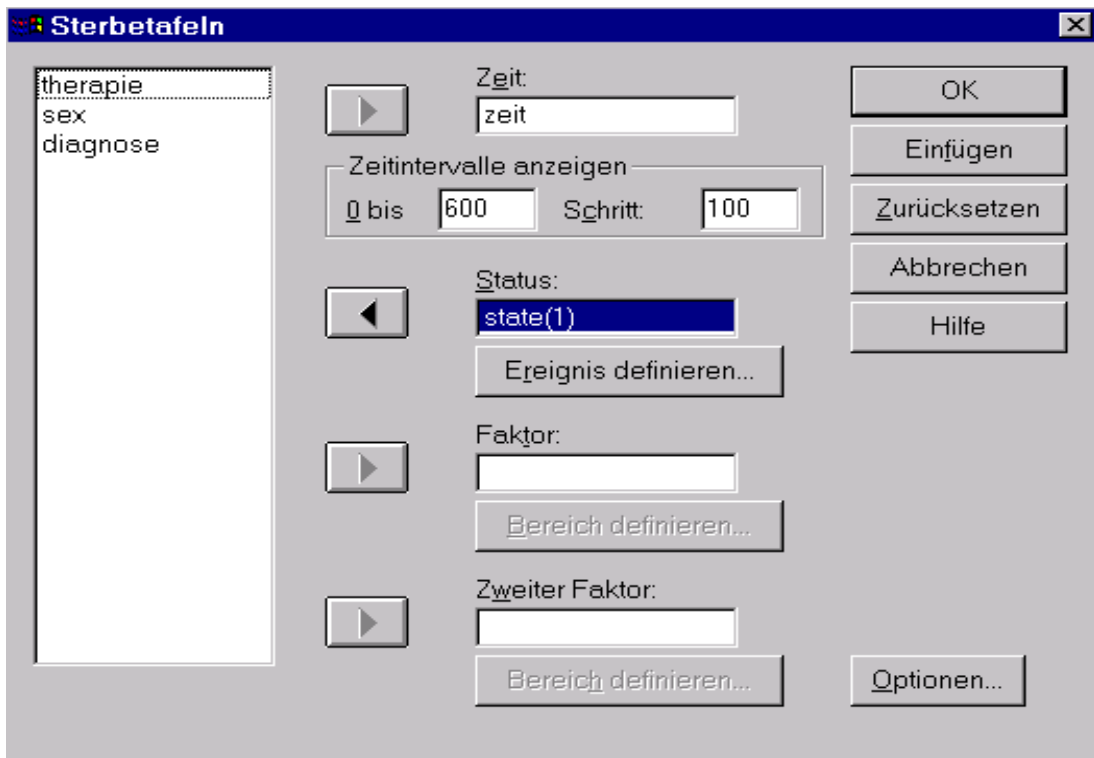


Abb. 2.1 Dialogfeld 'Sterbetafel'

Die mögliche Angabe von bis zu zwei **Faktoren** erlaubt die Bildung von Untergruppen; durch einen Faktor werden z.B. Diagnose-, Geschlechts- oder Therapiegruppen definiert. Je Faktor sind die Gruppencodes, die berücksichtigt werden sollen (kleinster und größter Wert), im Untermenü "**Bereich definieren**" festzulegen. Per Default wird dann für jede Stufe des einen Faktors, falls nur einer angegeben ist, oder für jede Stufenkombination der beiden Faktoren je eine eigene Sterbetafel berechnet und ausgedruckt.

Auswahlmöglichkeiten für **Graphiken** und **statistische Tests** erreicht man über das Anklicken von "**Optionen**". Es erscheint das Feld der **Abb. 2.2** mit folgenden Wahlmöglichkeiten:

1. Der Ausdruck der Sterbetafeln kann gewählt oder unterdrückt werden.
2. Es können verschiedene Graphiktypen angefordert werden, und zwar:

Überlebenszeitkurve, logarithmierte Überlebenszeitkurve, Kurve der Hazardfunktion, Dichte der Überlebenszeit.

Falls ein Gruppenfaktor angegeben ist, berechnet SPSS die Kurvenverläufe getrennt für jede Gruppe .

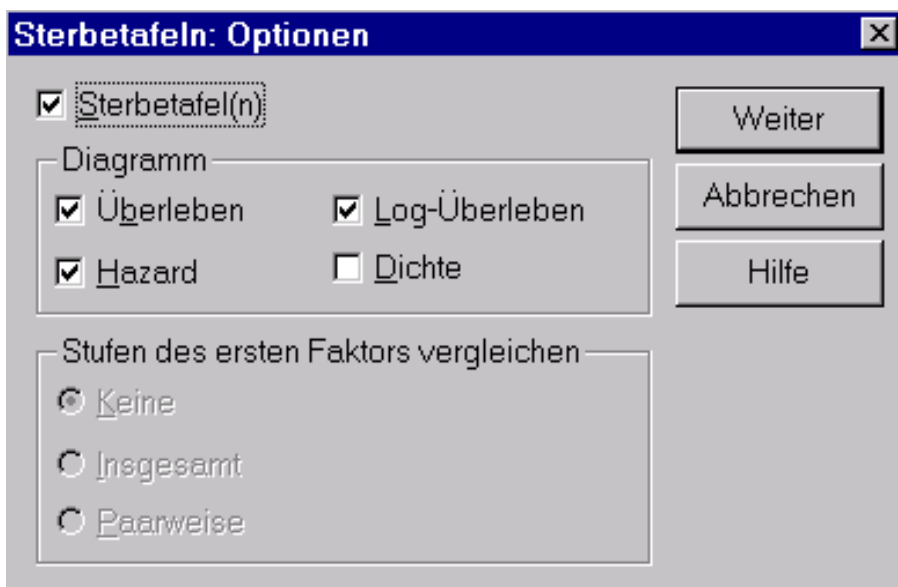


Abb. 2.2 Dialogfeld 'Sterbetafel: Optionen'

Zu den Graphiken: Sind die Überlebenszeiten exponentialverteilt, verläuft die Hazardfunktion konstant und die logarithmierte Überlebenszeitfunktion linear.

3. Vergleich der durch den ersten Faktor gebildeten Untergruppen hinsichtlich ihrer Überlebenszeitverteilung: Nullhypothese: die

Überlebenszeitverteilungen in den Gruppen sind alle gleich. Sofern durch diesen Faktor mehr als zwei Untergruppen definiert sind, wird der paarweise Vergleich **nicht** im Rahmen einer multiplen Testprozedur durchgeführt.

Ist auch ein zweiter Gruppenfaktor angegeben, so werden für jeden Level dieses Faktors je getrennt die Vergleiche zum ersten Faktor gerechnet. Siehe dazu die Möglichkeiten der geschichteten Analyse in Kapitel III.

2.3. Ergebnisliste und Graphiken

Zum Beispiel von Abschnitt 2.2 klickt man im Dialogfeld der Abb. 2.1 das Feld **OK** an. Daraufhin produziert SPSS die Ergebnisliste der **Abb. 2.3**. Die Liste wiederholt, falls SPSS so eingestellt ist, zunächst die Folge der auszuführenden Programmbefehle, die - in Umgehung des Dialogfeldes - auch über das Syntaxfenster hätten eingegeben werden können.

Nach der Anzahl der gültigen Fälle folgen die Berechnungen zur Sterbetafel. In die Berechnung dieser einen Sterbetafel gehen alle Patienten ein; denn ein Faktor für gruppenweise Berechnung wurde nicht angegeben.

This subfile contains: 76 observations

Life Table

Survival Variable ZEIT

Inter- val	Number Entrng this Intrvl	Number Wdrawn During Intrvl	Number Exposd to Risk	Number of Termnl Events	Cumul Propn Termin- natng	Propn Surv- living	Propn Surv at End	Hazard Rate	SE of Cumul Surv- living	SE of Hazard Rate
0	76,0	4,0	74,0	28,0	,3784	,6216	,6216	,0047	,0564	,0009
100	44,0	3,0	42,5	11,0	,2588	,7412	,4607	,0030	,0591	,0009
200	30,0	2,0	29,0	6,0	,2069	,7931	,3654	,0023	,0583	,0009
300	22,0	2,0	21,0	4,0	,1905	,8095	,2958	,0021	,0566	,0010
400	16,0	2,0	15,0	1,0	,0667	,9333	,2761	,0007	,0562	,0007
500	13,0	,0	13,0	2,0	,1538	,8462	,2336	,0017	,0550	,0012
600+	11,0	11,0	5,5	,0	,0000	1,0000	,2336	**	,0550	**

** These calculations for the last interval are meaningless.

The median survival time for these data is 175,59

Die Sterbetafel führt in der ersten Spalte die Startzeitpunkte der Intervalle auf, anschließend die in Abschnitt 2.2 definierten Größen n_i , c_i , r_i , d_i , q_i , p_i , S_i , f_i und h_i ; darunter wird für die Schätzwerte S_i , f_i und h_i jeweils auch der Standard Error (SE) mit ausgegeben. Die letzte Zeile der Tabelle ergibt bis auf die Anzahl n_i , Spalte 2, kaum verwertbare Zahlen.

Die Werte h_i der Hazardfunktion ganz rechts weisen erkennbare, jedoch keine groben Abweichungen vom konstanten Verlauf auf. Ihrer abfallenden Tendenz entsprechen zunehmende Werte p_i in der viertletzten Spalte.

Mit den Werten S_i , $\ln(S_i)$ und h_i sind die Graphiken der **Abb. 2.4 - 2.6** erstellt. Die Abweichung von der Konstanz in Abb.2.6 bedeutet: die Überlebenserwartung steigt, wenn eine schwierige Anfangsphase überstanden ist.

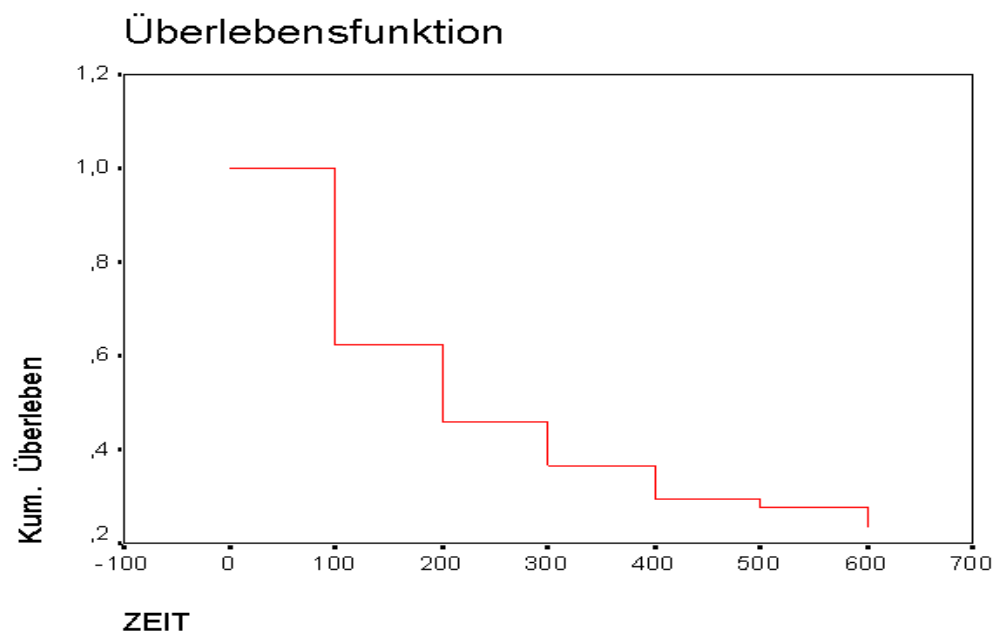


Abb. 2.4. Überlebensfunktion , von SPSS ausgegebene Treppenfunktion, richtiger wäre ein die Eckpunkte verbindender Polygonzug

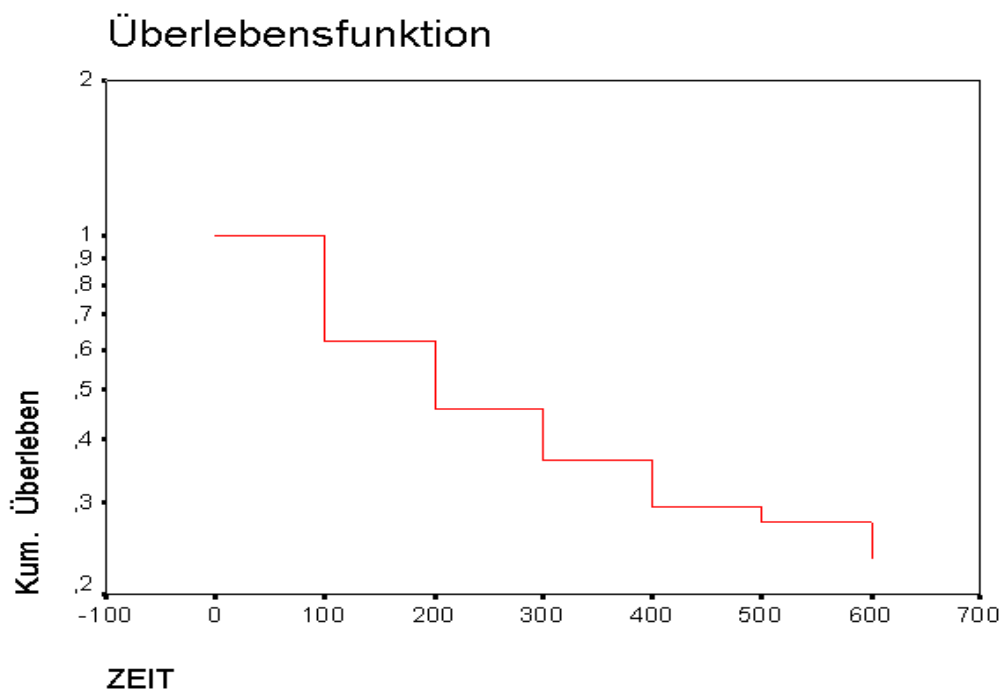


Abb. 2.5. Log-Überlebensfunktion, die Bezeichnung 'Überlebensfunktion' und 'Kum. Überleben' setzt SPSS falsch. Richtig ist statt dessen 'Log-Überlebensfunktion' und 'Log kum. Überleben'.

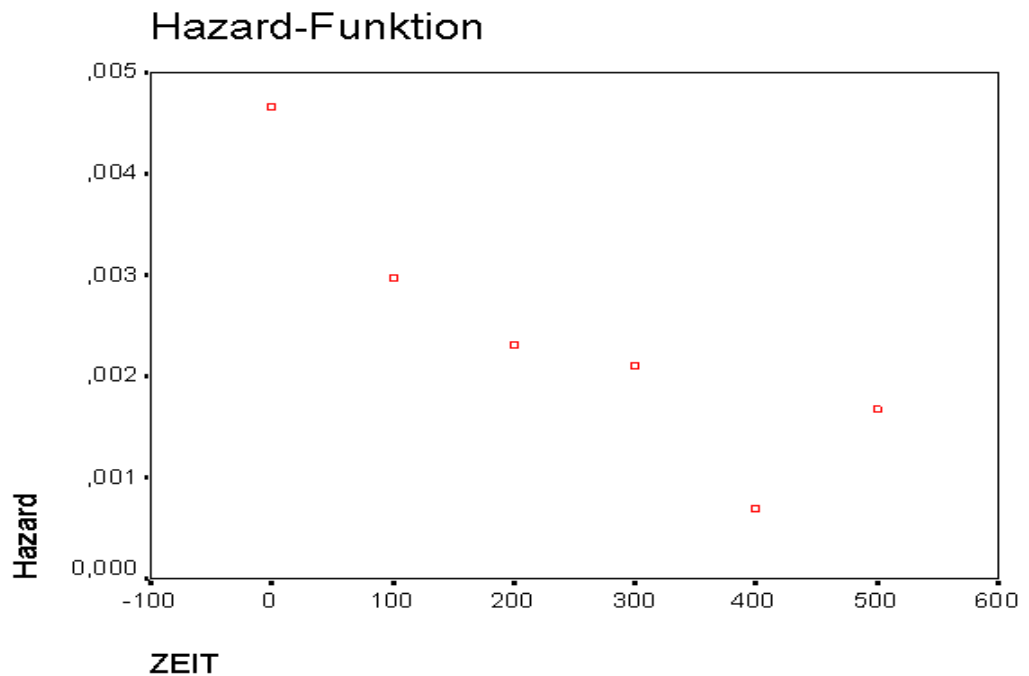


Abb. 2.6. Hazard-Funktion, die Abweichung von einem konstanten Verlauf tritt hier deutlicher hervor als die Abweichung von der Linearität in Abb.2.5

III. Die Produkt-Limit-Methode

(Kaplan, Meier; 1958)

3.1. Konzept

Für jeden der zu Beginn vorhandenen N Fälle gibt es einen Endzeitpunkt t der Beobachtung. Es seien $t_1 < t_2 < \dots < t_N$ die geordneten Endzeitpunkte. Dann werden Schätzwerte der Überlebenszeitfunktion mit folgender Gleichung ermittelt:

$$S(t) = \prod_{i: t_{i-1} < t_i < t} \left[\frac{\delta_i}{N - i + 1} \right], \quad \text{Produkt über alle } t_i < t$$

1, falls t_i Response-Zeitpunkt

mit $\delta_i =$

0, falls t_i Zeitpunkt einer 'zensierten Beobachtung'

Hier wird der geschätzte Verlauf der Überlebenszeitkurve durch eine fallende Treppenfunktion dargestellt. Die Sprunghöhe $S(t_{i-1}) - S(t_i)$ zum Zeitpunkt t_i gibt die bedingte Sterbewahrscheinlichkeit für das Intervall (t_{i-1}, t_i) unter der Bedingung, daß der Zeitpunkt t_{i-1} lebend erreicht wurde, dividiert durch $S(t_{i-1})$, an.

Der Standard Error se zum Schätzwert $S(t)$ ist

$$se = S(t) \sqrt{\sum_{i: t_{i-1} < t_i < t} \frac{\delta_i}{(N - i)(N - i + 1)}}, \quad \text{Summe über alle } t_i < t$$

Die Rechnung unterscheidet nicht zwischen verschiedenen Ursachen der Zensierung.

3.2. Programmgestaltung im Dialogfeld

Das Dialogfeld, **Abb. 3.1**, ist ähnlich dem in Abb. 2.1 aufgebaut. Erforderlich sind eine Variable für das Zeitintervall und eine, die den Status am Ende des Zeitintervalls angibt inklusive des oder der Codes für das Response-Ereignis.

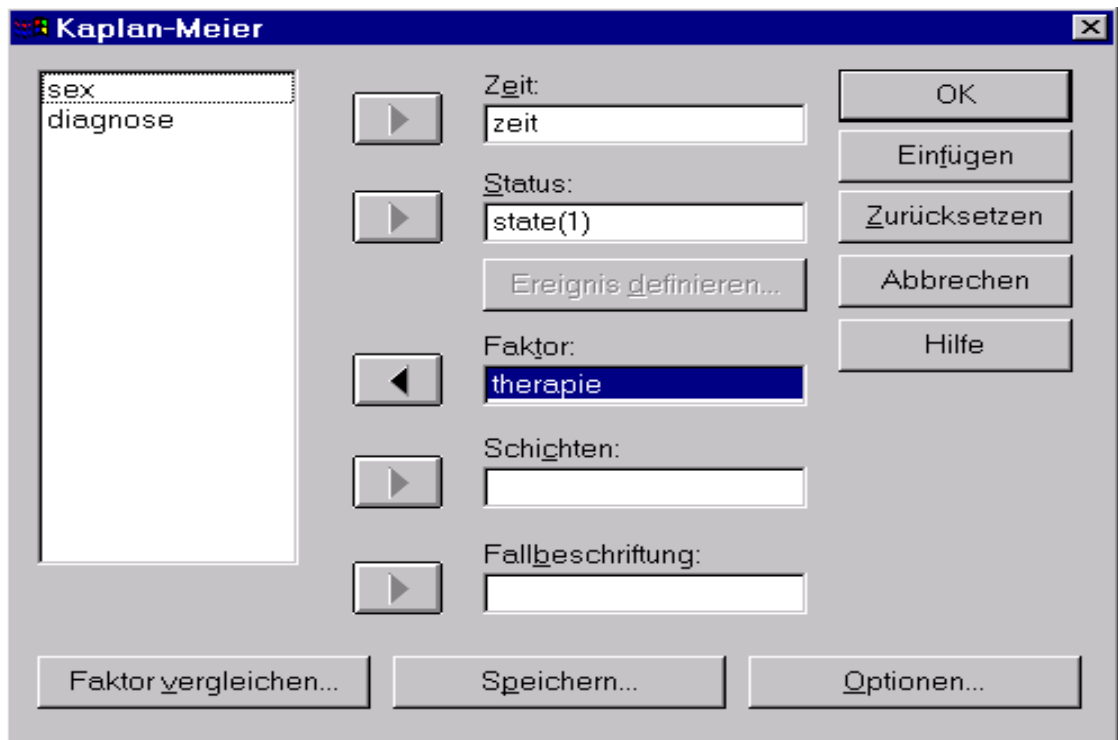


Abb. 3.1 Dialogfeld 'Kaplan-Meier'

Abb. 3.1 zeigt das Beispiel von Abschnitt 2.2, wobei hier noch der Faktor *therapie* berücksichtigt werden soll.

Wird ein Faktor angegeben, können die durch den Faktor definierten Gruppen bezüglich der Überlebenszeitverteilung verglichen werden. Auch hier ist es möglich, einen zweiten Faktor anzugeben, und zwar in das Feld, das mit 'Schichten' überschrieben ist. Die durch diesen Faktor definierten Gruppen werden entsprechend der anschließenden Analyse als Schichten bezeichnet, siehe Abschnitt 3.5.

Über das Optionsfeld 'Optionen' wird der Umfang der Ergebnisliste und werden die gewünschten Graphiken festgelegt, siehe Abb. 3.2. Vorgesehen sind per Default der Ausdruck der Survival-Tabelle(n), Mittelwert(e) und Median(e). Durch das Anklicken einer der Zeilen kann das Gemeinte angefordert oder auch unterdrückt werden.

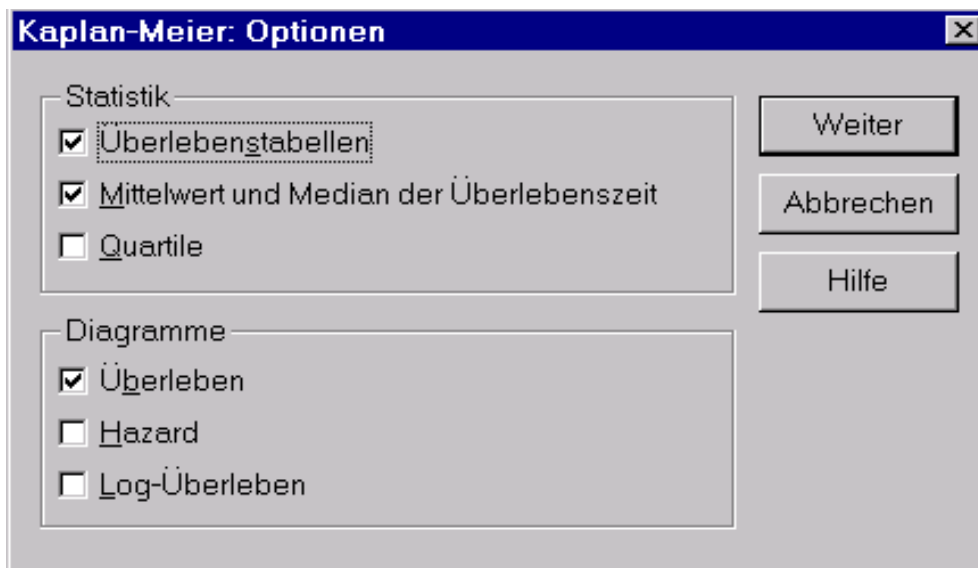


Abb. 3.2 Dialogfeld 'Kaplan-Meier: Optionen'

Zu den Graphiken: Wenn das einfache Modell der exponential verteilten Überlebenszeiten gilt, verläuft die kumulierte Hazard-Funktion linear steigend und die Log-Survival-Funktion linear abfallend, vergl. Abb. 3.5.

3.3. Gruppenvergleich

Um die durch den Faktor definierten Gruppen zu vergleichen, klickt man in der untersten Zeile des Kaplan-Meier- Dialogfeldes das Optionsfeld 'Faktor vergleichen' an. Daraufhin wird das Dialogfeld der Abb. 3.3 eingeblendet.

Geprüft wird die **Nullhypothese**: die wahren Überlebenszeitkurven der Gruppen haben den gleichen Verlauf.

Demgegenüber steht die **Alternativhypothese**: die Verläufe unterscheiden sich.

Es können bis zu drei verschiedene Teststatistiken angegeben werden. Bei einem Vergleich von k Gruppen besitzen die Teststatistiken näherungsweise eine Chi-Quadrat-Verteilung mit $k-1$ Freiheitsgraden. Die Qualität der Näherung wie auch die Power der Tests ist an den Gesamtstichprobenumfang und an die Anzahl der Response-Fälle gebunden.

Um ihre Unterschiede zu verdeutlichen, seien die Teststatistiken kurz skizziert. Zunächst die annähernd chi-Quadrat-verteilte Log-Rank-Statistik: In allen k Gruppen zusammen gebe es m Responsezeitpunkte $t_i, i=1,2,\dots,m$. Zu jedem Zeitpunkt wird in jeder Gruppe $j, j=1,2,\dots,k$, eine Differenz d_{ij} gebildet: die Anzahl der Responsefälle zu dem Zeitpunkt minus dem geschätzten Erwartungswert dieser Anzahl. Die Differenzen werden je Gruppe über alle Zeitpunkte aufaddiert, die Summe wird quadriert und durch den Erwartungswert aller Responsefälle der Gruppe dividiert; so ergeben sich die Chi-Quadratanteile je Gruppe.

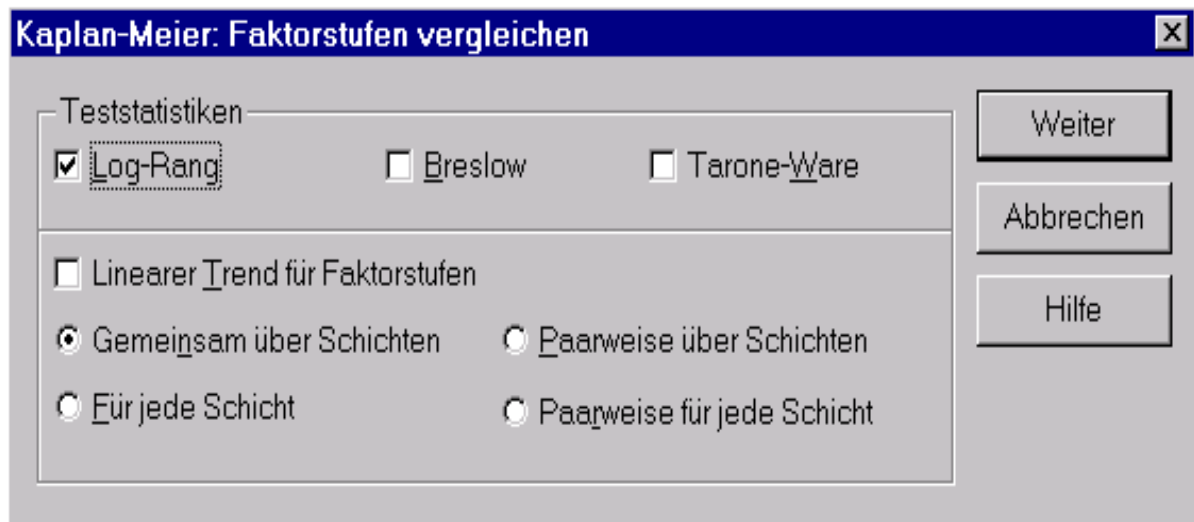


Abb. 3.3 Dialogfeld zum Gruppenvergleich

Die zu jedem Zeitpunkt $i, i=1,2,\dots,m$, gebildeten Differenzen d_{ij} können noch mit einem **Gewichtsfaktor w_i** beeinflusst werden. Die oben genannten drei Teststatistiken unterscheiden sich in der Wahl dieser Gewichtsfaktoren. Ist M_i die Anzahl der unter Responderisiko stehenden Fälle zum i -ten Responsezeitpunkt, so gilt:

$w_i = 1$ für die Log-Rank-Teststatistik

$w_i = M_i$ für die BRESLOW-Teststatistik,

$w_i = \sqrt{M_i}$ für die TARONE-WARE-Teststatistik,

Dies bedeutet: verglichen mit dem Log-Rank-Test gibt der Test nach BRESLOW den Beobachtungen zu frühen Zeitpunkten ein größeres Gewicht als denen zu späteren Zeitpunkten. Die dritte Teststatistik liegt in der Gewichtung zwischen den ersten beiden, entsprechend liegt ihr p -Wert zwischen den p -Werten der beiden anderen Tests. Siehe die Ergebnisliste im folgenden Abschnitt. Im Zweifelsfall ist es sinnvoll, sich für den TARONE-WARE-Test zu entscheiden.

Umfang und Verteilung der zensierten Fälle kann die Teststatistiken erheblich beeinflussen. Eine größere Anzahl von Zensierungen zu frühen Zeitpunkten wirkt sich in größeren Unterschieden zwischen den Teststatistiken aus. Unterscheiden sich die Gruppen erheblich bezüglich Umfang und Verteilung der zensierten Fälle, so sind die Ursachen sorgfältig zu ergründen; denn nach Voraussetzung darf ein Zusammenhang zwischen Zensierung und Überlebenszeit nicht bestehen.

3.4. Ergebnisliste

Wir zeigen und kommentieren die Ergebnisliste zum Beispiel 'Therapievergleich' auf den beiden folgenden Seiten. Darin werden die Patientengruppen mit den Therapien 'Test' bzw. 'Standard' gegenübergestellt.

Die Liste zeigt zunächst die Programminstruktionen in der SPSS-Befehlssprache. d.h. die verwendeten Variablen und Parameter und die angeforderten Optionen. Es folgt, getrennt nach Gruppen, je eine der Sterbetafel entsprechende Tabelle, je Fall eine Zeile, geordnet nach zunehmender Länge des Zeitintervalls. Die Patienten, die während des Studienablaufs die Studie verlassen haben, und diejenigen, die am Ende der Studie bzw. am Ende der maximalen Beobachtungszeit noch leben, gelten in gleicher Weise als zensiert. Die geschätzte kumulative Survivalfunktion verläuft wie eine Treppenfunktion: Änderungen erfährt sie nur zu Zeitpunkten von Responsefällen, nicht jedoch zu Zeitpunkten, an denen zensierte Fälle aus der Studie ausscheiden. Die zeitliche Änderung des Standard Errors se der Funktion läßt sich anhand der Formel für se in Abschnitt 3.1 erklären.

Der je Gruppe ermittelte Median als Zeitpunkt für den 50 %-Wert der kumulativen Überlebenszeit liegt in der ersten, der Testgruppe, am Rand der Responsezeitpunkte und läßt sich darum nur sehr unzuverlässig schätzen, das Konfidenzintervall ist groß; vergleiche dazu die entsprechenden Werte in der zweiten Gruppe.

Die Teststatistiken zeigen einen signifikanten Unterschied zwischen den Überlebenszeitverteilungen der beiden Gruppen. Die Unterschiede sind im Anfangsbereich geringer als bei höheren Zeitwerten, so daß der p -Wert zur BRESLOW-Statistik, die die Anfangswerte stärker berücksichtigt, größer ausfällt als der p -Wert des ungewichteten Log-Rank-Tests; die TARONE-WARE-Statistik liegt erwartungsgemäß dazwischen.

KM zeit BY therapie /STATUS=state(1) .

/PRINT TABLE MEAN

/PLOT SURVIVAL HAZARD LOGSURV

/TEST LOGRANK BRESLOW TARONE

/COMPARE OVERALL POOLED .

Abb. 3.4 Ergebnisliste zum Therapievergleich

Survival Analysis for ZEIT .

Factor THERAPIE = Test

Time	Status	Cumulative Survival	Standard Error	Cumulative Events	Number Remaining
,00	gestorben	,9737	,0260	1	37
3,00	gestorben	,9474	,0362	2	36
8,00	gestorben			3	35
8,00	gestorben	,8947	,0498	4	34
11,00	gestorben	,8684	,0548	5	33
12,00	gestorben	,8421	,0592	6	32
25,00	verloren			6	31
35,00	gestorben	,8149	,0632	7	30
42,00	gestorben	,7878	,0667	8	29
72,00	gestorben	,7606	,0697	9	28
82,00	gestorben	,7334	,0723	10	27
92,00	verloren			10	26
95,00	verloren			10	25
100,00	verloren			10	24
110,00	gestorben	,7029	,0755	11	23
117,00	gestorben	,6723	,0781	12	22
118,00	gestorben	,6418	,0803	13	21
126,00	gestorben	,6112	,0821	14	20
132,00	gestorben	,5806	,0835	15	19
144,00	verloren			15	18
162,00	gestorben	,5484	,0849	16	17
219,00	verloren			16	16
228,00	gestorben	,5141	,0862	17	15
314,00	gestorben	,4798	,0870	18	14
332,00	verloren			18	13
378,00	verloren			18	12
411,00	verloren			18	11
436,00	verloren			18	10
600,00	zensiert			18	9
600,00	zensiert			18	8
600,00	zensiert			18	7
600,00	zensiert			18	6
600,00	zensiert			18	5
600,00	zensiert			18	4
600,00	zensiert			18	3
600,00	zensiert			18	2
600,00	zensiert			18	1
600,00	zensiert			18	0

Number of Cases: 38 Censored: 20 (52,63%) Events: 18 .

	Survival Time	Standard Error	95% Confidence Interval
Mean:	337,51	44,25	(250,78; 424,25)
(Limited to 600,00)			
Median:	314,00	,	(, ; ,)

Factor THERAPIE = Standard

Time	Status	Cumulative Survival	Standard Error	Cumulative Events	Number Remaining
------	--------	---------------------	----------------	-------------------	------------------

Überlebenszeitanalysen und Logistische Regression

1,00 gestorben	,9737	,0260	1	37
3,00 gestorben	,9474	,0362	2	36
7,00 gestorben	,9211	,0437	3	35
8,00 gestorben	,8947	,0498	4	34
15,00 gestorben	,8684	,0548	5	33
18,00 gestorben	,8421	,0592	6	32
24,00 gestorben	,8158	,0629	7	31
25,00 gestorben	,7895	,0661	8	30
30,00 gestorben	,7632	,0690	9	29
33,00 gestorben	,7368	,0714	10	28
36,00 gestorben	,7105	,0736	11	27
44,00 gestorben	,6842	,0754	12	26
45,00 gestorben	,6579	,0770	13	25
48,00 gestorben	,6316	,0783	14	24
51,00 gestorben	,6053	,0793	15	23
73,00 gestorben	,5789	,0801	16	22
84,00 gestorben	,5526	,0807	17	21
87,00 verloren			17	20
90,00 gestorben	,5250	,0812	18	19
111,00 gestorben	,4974	,0815	19	18
112,00 gestorben	,4697	,0815	20	17
140,00 gestorben	,4421	,0813	21	16
151,00 gestorben	,4145	,0808	22	15
183,00 verloren			22	14
186,00 gestorben	,3849	,0802	23	13
201,00 gestorben	,3553	,0793	24	12
229,00 gestorben	,3257	,0781	25	11
231,00 gestorben	,2961	,0764	26	10
231,00 verloren			26	9
242,00 gestorben	,2632	,0746	27	8
283,00 gestorben	,2303	,0722	28	7
357,00 gestorben	,1974	,0690	29	6
380,00 gestorben	,1645	,0648	30	5
389,00 gestorben	,1316	,0596	31	4
467,00 gestorben	,0987	,0530	32	3
567,00 gestorben	,0658	,0444	33	2
587,00 gestorben	,0329	,0322	34	1
600,00 zensiert			34	0

Number of Cases: 38 Censored: 4 (10,53%) Events: 34 .

	Survival Time	Standard Error	95% Confidence Interval
Mean:	183,48	30,72	(123,27 ; 243,68)
(Limited to	600,00)		
Median:	111,00	41,30	(30,05 ; 191,95)

Survival Analysis for ZEIT .

	Total	Number Events	Number Censored	Percent Censored
THERAPIE Test	38	18	20	52,63
THERAPIE Standard	38	34	4	10,53
Overall	76	52	24	31,58

Test Statistics for Equality of Survival Distributions for THERAPIE .

	Statistic	df	Significance
Log Rank	9,90	1	,0017
Breslow	4,29	1	,0384
Tarone-Ware	6,44	1	,0112

Überlebensfunktionen

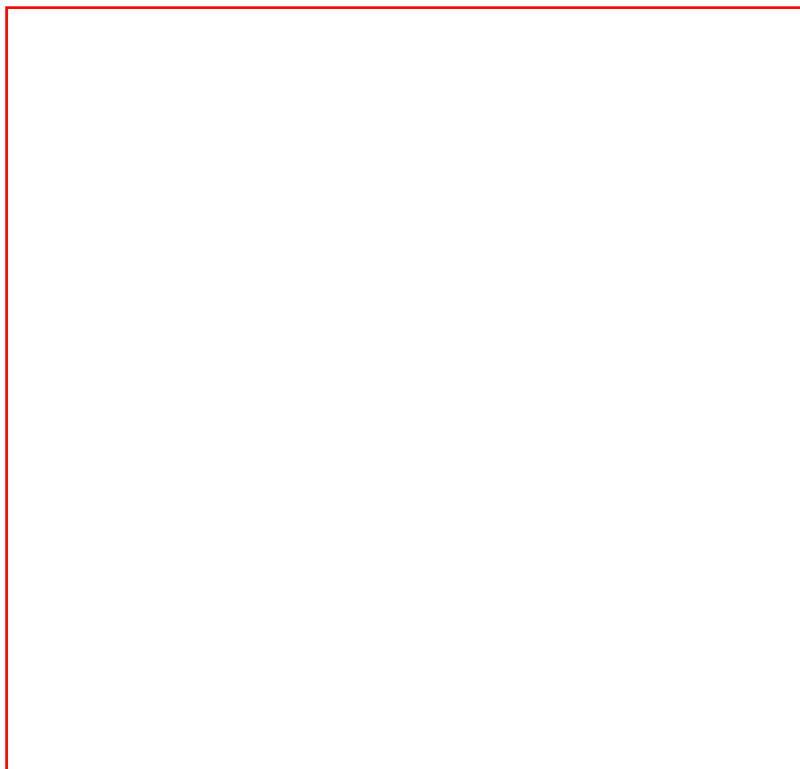
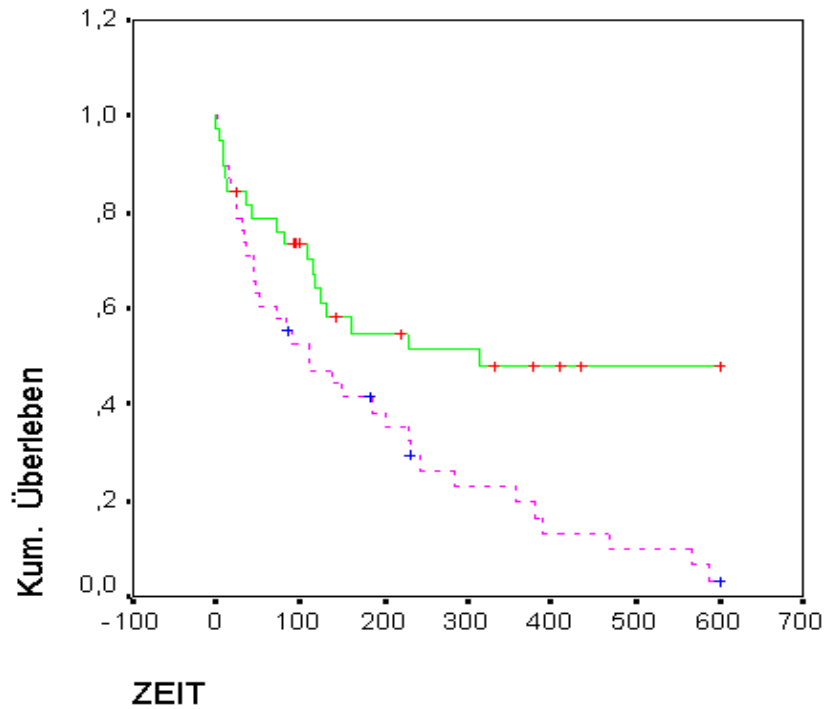


Abb. 3.5.

Die Überlebenszeitfunktion (links) und die logarithmierte Überlebenszeitfunktion (rechts) zu den Therapiegruppen Standard (- - -) und Test (- . . .), + markiert Endzeitpunkt eines zensierten Falles.

Abb. 3.6.

3.5 Trend- und Schichtanalyse

Es gibt Situationen, in denen die k Stufen der Gruppenvariable durch die Klassenbildung eines quantitativen Merkmals entstehen und man vermutet, daß sich diese ordinale Struktur auch zwischen den k Überlebenszeitkurven zeigt.

Als die k Gruppen können z.B. die Risikogruppen von Patienten auftreten, eingeteilt nach den Rauchergewohnheiten 'nicht', 'leicht', 'mittel' und 'stark', oder eingeteilt in Übergewichtsklassen. Oder es werden Behandlungen mit k verschiedenen Dosierungen eines Pharmakons vorgenommen, und man fragt nach der Abhängigkeit der Heilungszeit von der Dosierung.

In diesen Fällen können die Trendversionen des Log-Rank-, BRESLOW- und/oder TARONE-WARE-Tests angefordert werden, indem man im Dialogfeld der Abb. 3.3 für den Faktorvergleich '**Linearer Trend für Faktorstufen**' markiert. Die Trendversionen eignen sich besser, ordinale Effekte als signifikant zu erkennen, als es die allgemeinen Testversionen vermögen. Bei einer derartigen Trendanalyse wird angenommen, daß die

Abstände zwischen je zwei benachbarten Gruppen annähernd gleich groß ausfallen, mit anderen Worten, daß die Unterschiede zwischen benachbarten Gruppen gleichgewichtig sind.

Es ist möglich, daß beim Gruppenvergleich neben dem Gruppenfaktor noch andere kontrollierte Merkmale auftreten: sog. Kovariate, die die Überlebenszeitkurven beeinflussen und so zu verzerrten Testergebnissen führen. Eine kategorielle Kovariate bezeichnet man auch als **Schichtvariable**, die Kategorien als Schichten; z.B. kann das Geschlecht solch eine Schichtvariable sein. Es könnte dann sinnvoll scheinen, den Vergleich nach den Schichten getrennt durchzuführen - bei entsprechend geringeren Stichprobenumfängen. Dazu wäre im Dialogfeld für den Faktorvergleich **'Für jede Schicht'** zu markieren.

Mit dem Ziel, verzerrte Testergebnisse, die durch Kovariate hervorgerufen werden, zu vermeiden und durch größeren Stichprobenumfang die Aussagekraft der approximativen Testergebnisse zu sichern, bietet SPSS die geschichtete Analyse an. Man fordert sie an, indem man im Dialogfeld für den Faktorvergleich **'Gemeinsam über Schichten'** markiert. Um bei dieser Analyse die Teststatistik zu ermitteln, werden die Chi-Quadratanteile für jede Schicht getrennt berechnet, aber dann über alle Schichten aufaddiert.

Ein paarweiser Vergleich der Gruppen ist möglich, siehe Abb. 3.3, indem für jedes Gruppenpaar die gewünschte Teststatistik und der zugehörige p-Wert berechnet wird. Auch hier kann der Vergleich gemeinsam über die Schichten oder getrennt für jede Schicht durchgeführt werden.

Bei der graphischen Darstellung gibt SPSS die angeforderten Kurven getrennt für jede Schicht aus, falls eine Schichtvariable angegeben wurde.

IV. Überlebenszeitanalysen mit Kovariaten, das COX-Modell

Die Überlebenszeitanalysen betrachten Zeitintervalle, siehe Abschnitt 1.1, z.B. die Zeit von Therapiebeginn bis zur Heilung, von der Krebsoperation bis zum Tod, von der Transplantation bis zur Abstoßungsreaktion. Die Länge dieser Intervalle wird in der Regel von mehreren Faktoren erheblich beeinflußt: von begleitenden Therapiebedingungen und von Faktoren, die die Konstitution der Patienten kennzeichnen, sog. Risiko- oder prognostische Faktoren.

Ein von COX (1972/75) konzipiertes Modell erlaubt es zu untersuchen, welche der vermuteten Einflußgrößen, die bei den Patienten einer Studie beobachtet wurden, sich erheblich auf die 'Überlebenszeit' auswirken und welche davon genügen, um das prognostische Potential aller dieser Kovariaten auszuschöpfen.

4.1. Proportional Hazards Regression Model

Das Modell wird kurz als COX-Modell bezeichnet. Wir knüpfen an die Einführung in Kapitel I an.

Mit $P(t)$ als Überlebenszeitverteilung und $h(t)$ als Hazardfunktion oder Sterbeintensität gilt die Beziehung

$$S(t) = \exp \left[- \int_0^t h(v) dv \right]$$

Es sei $\mathbf{z}' = (z_1 \ z_2 \ \dots \ z_K)$ ein Vektor von K - zunächst zeitlich konstanten - Kovariaten, also von Einflußgrößen oder prognostischen Faktoren zur Kennzeichnung individueller Eigenschaften oder Randbedingungen, z.B. Alter, Geschlecht, Behandlungsgruppen, Risikofaktoren, . . .

Die Einflüsse der Kovariaten auf die Überlebenszeiten werden im COX-Modell durch die spezielle Form der nun auch von \mathbf{z} abhängigen Hazardfunktion $h(t, \mathbf{z})$ erfaßt:

$$h(t, \mathbf{z}) = h_0(t) \exp(\beta_1 z_1 + \beta_2 z_2 + \dots + \beta_K z_K) = h_0(t) \exp(\mathbf{\beta}'\mathbf{z}).$$

Hierin bedeutet $h_0(t)$ eine unbekannte Hazardfunktion, die für ein Individuum mit $\mathbf{z} = \mathbf{0}$ gilt (Hazardbasisfunktion), und $\mathbf{\beta}$ einen Vektor unbekannter Regressionskoeffizienten.

Dieses Proportional Hazards Regression Model enthält zwei Annahmen:

- Die Proportionalität: die Hazardfunktion läßt sich als Produkt einer nur von der Zeit abhängigen Basisfunktion h_0 und einem nur von den Kovariaten abhängigen zweiten Faktor, der Risikofunktion, darstellen.
- Die Form der Risikofunktion: sie ist als loglineare Funktion der Kovariaten angesetzt.

Für die Darstellung der Überlebenszeitfunktion ergibt sich:

$$\begin{aligned} S(t) &= \exp \left[- \int_0^t h(v, \mathbf{z}) dv \right] = \exp \left[- \int_0^t h_0(v) \exp(\mathbf{\beta}'\mathbf{z}) dv \right] \\ &= \left\{ \exp \left[- \int_0^t h_0(v) dv \right] \right\} \exp(\mathbf{\beta}'\mathbf{z}) = \left[S_0(t) \right] \exp(\mathbf{\beta}'\mathbf{z}) \end{aligned}$$

mit $So(t)$ als 'Baseline-Survival'.

Unter dem COX-Modell besitzen zwei Individuen mit je eigenem Vektor \mathbf{z}_1 und \mathbf{z}_2 von Kovariaten eine je eigene Hazardfunktion $h_1(t) = h(t, \mathbf{z}_1)$ und $h_2(t) = h(t, \mathbf{z}_2)$ und entsprechend eine **je eigene Überlebenszeitfunktion** $S_1(t)$ und $S_2(t)$. Diese beiden Funktionen zu zwei Vektoren \mathbf{z}_1 und \mathbf{z}_2 unterscheiden sich also durch unterschiedliche Potenzen der 'Baseline-Survival' $So(t)$. Daher können sich die Kurvenverläufe dieser beiden Funktionen bei Gültigkeit des COX-Modells **nicht überschneiden**.

Bei einem positiven Koeffizienten β_i wird $h(t)$ größer mit wachsendem z_i . Das bedeutet einen ungünstigen Einfluß auf die Überlebenszeit, denn um so eher nimmt $P(t)$ kleine Werte an. Entsprechend bedeutet ein negativer Koeffizient β_j , daß von der j -ten Kovariaten bei wachsenden Werten ein günstiger Einfluß auf die Überlebenszeit ausgeht.

Weiter gilt: Der Quotient bzw. der Proportionalitätsfaktor der beiden Hazardfunktionen $h_1(t)$ und $h_2(t)$ ist unabhängig von der Zeit t :

$$\frac{h_1(t)}{h_2(t)} = \frac{h_0(t) \exp(\beta'z_1)}{h_0(t) \exp(\beta'z_2)} = \frac{\exp(\beta'z_1)}{\exp(\beta'z_2)}$$

4.2. Programmgestaltung im Dialogfeld und Beispiel

Als Beispiel dient eine Untersuchung an 117 Patienten mit Lebertransplantationen. Dabei geht es um die Bewertung bekannter und vermuteter prognostischer Faktoren für die Überlebenschancen der Patienten. Während der prognostische Wert von Bilirubin (*bili*) und Indocyanin grün (*icgt*) bekannt ist, soll die Bedeutung von Monoethylglycinexylid (*megx*) noch ermittelt werden.

Die Kontrolle der Patienten erstreckt sich von der Transplantation an über maximal 120 Tage. Der Status am Ende des Zeitintervalls hat den Code 2, wenn der Patient in diesem Zeitraum von 120 Tagen verstorben ist.

Das Dialogfeld 'Cox Regression', **Abb. 4.1**, erscheint in ähnlicher Form wie die Dialogfelder 'Sterbetafel' und 'Kaplan Meier'. Notwendig ist die Angabe einer Variablen für das Zeitintervall sowie der Statusvariablen inklusive der Definition der Response-Codes, wie in Abb. 3.1 bereits geschehen. Darüber hinaus muß wenigstens eine Kovariante benannt werden.

Wird vermutet, daß zwischen zwei Variablen Wechselwirkungen bestehen, die die Überlebenszeit beeinflussen könnten, kann auch das Produkt der beiden Variablen als Kovariante aufgenommen werden. Dazu sind beide Variablen im linken Variablenfeld zu markieren und dann **>a*b>** anzuklicken. Während der Benennung der Kovariaten kann die darüberstehende Blocknummer schrittweise erhöht werden, um Variablenblocks zu bilden. Stehen z.B. drei Variablen im Kovariatenfeld, und klickt man den darüber stehenden Schalter **Weiter** an, werden diese Variablen zu einem Block 1 zusammengefaßt und das Feld der Kovariaten für die Benennung von Variablen für einen Block 2 freigemacht, usw. Durch das Anklicken der Schalter **Zurück** oder **Weiter** können die Inhalte verschiedener Blöcke im Kovariatenfeld angezeigt werden. Über die Verwendungsmöglichkeiten von Blöcken am Ende von Abschnitt 4.5.

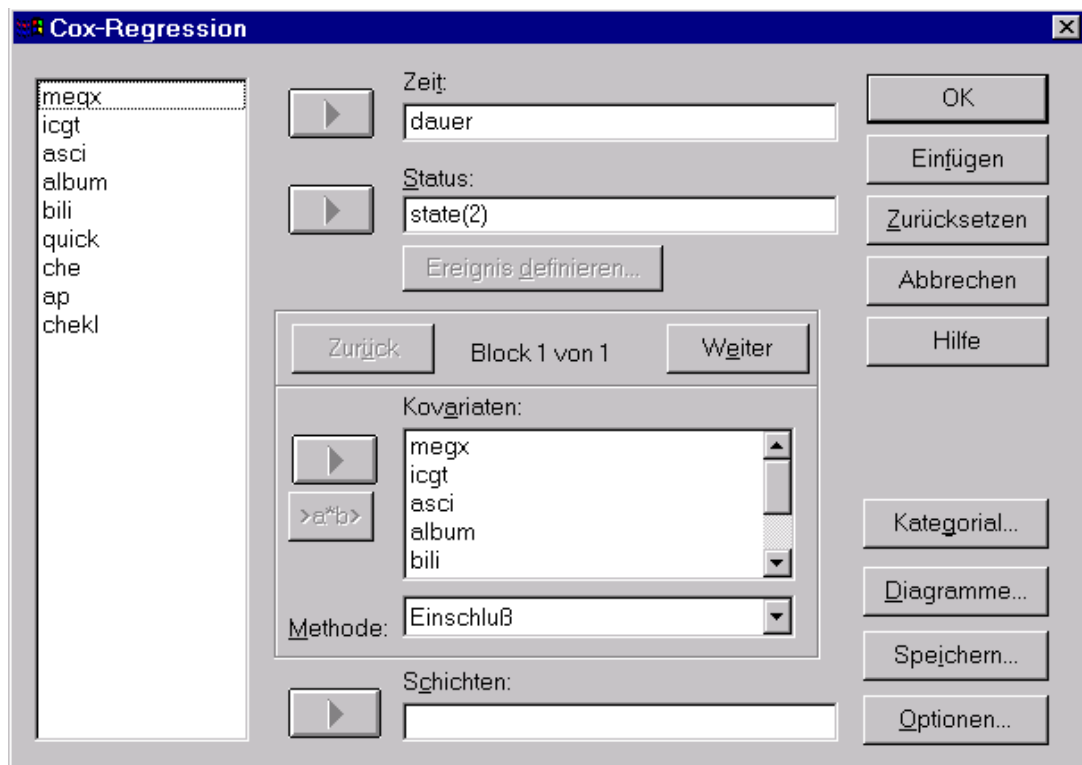


Abb. 4.1 Dialogfeld 'COX-Regression'

Als Auswertungsmethode wird zunächst **Einschluß** vorgeschlagen. Damit gehen alle Kovariate zugleich in die Regressionsgleichung ein. Alternativen

dazu in Abschnitt 4.5.

Der in Abb. 4.1 festgelegte Programmaufruf führt zur Ergebnisliste der **Abb. 4.2**. Man sollte die Grundeinstellung von SPSS so wählen, daß die Ergebnisliste den Programmaufruf mit enthält, um besser nachprüfen zu können, was das Programm gerechnet hat. Insbesondere werden dabei auch unveränderte Voreinstellungen wie hier die p-Werte für Variableneinschluß oder -ausschluß sowie die maximale Anzahl der Einzelschritte bei den iterativen Berechnungen angezeigt.

Wichtig für den Anwender ist anschließend die Möglichkeit einer Kontrolle der gültigen und ungültigen Fälle. Im Weiteren umfaßt die Ausgabeliste eine Prüfung des Modells - hier mit vier Einflußgrößen - sowie eine Kennzeichnung der eingehenden Variablen.

4.3. Schätzen von β und Testen von $H_0: \beta = 0$

Zum vorliegenden Modell und den Werten der Kovariaten und den beobachteten Response-Zeitpunkten bedeutet die Likelihoodfunktion $L(\mathbf{b})$ die Wahrscheinlichkeitsdichte für das Auftreten der Responsefälle zu den beobachteten Zeitpunkten. Der Vektor \mathbf{b} steht für den noch zu schätzenden Vektor β . Die Schätzwerte B_i für β_i werden so bestimmt, daß $L(\mathbf{b})$ maximal wird, oder, was auf dasselbe Ergebnis führt, aber formal einfacher zu handhaben ist, daß $\ln(L(\mathbf{b}))$ maximal wird. Dieser maximale Wert wird auch kurz mit LL bezeichnet. Die Lösung ist über ein nichtlineares Gleichungssystem nur iterativ, d.h. mit höherem Rechenaufwand zu erhalten.

Man kann das Modell dadurch ändern, daß man nur einige oder gar keine der Kovariaten in die Modellgleichung aufnimmt. Entsprechend ändert sich der Wert LL. Ist M1 ein Modell mit keiner oder nur wenigen Kovariaten und M2 ein Modell, daß gegenüber M1 um wenigstens eine Kovariate erweitert und sonst gleich ist, so gilt für die zugehörigen log Likelihood-Werte LL1 und LL2: $-2(LL1 - LL2)$ ist annähernd Chi2-verteilt und ist die Teststatistik im sog. **Likelihood-Ratio-Test** (LR-Test). Die Anzahl df der Freiheitsgrade ist gleich der Anzahl der Kovariaten, um die M2 gegenüber M1 erweitert ist. Der LR-Test eignet sich zum Prüfen der **globalen Nullhypothese $H_0: \beta = 0$** .

Für dieses Prüfen läßt sich ebenso der sog. **Skore-Test** einsetzen. Hier wird die Teststatistik anhand erster und zweiter Ableitungen der LL-Funktion an der Stelle $\beta = 0$ berechnet. Der Skore-Test läßt sich mit vergleichsweise geringem rechnerischen Aufwand durchführen, ist dafür aber nur als Näherung für den LR-Test anzusehen.

COXREG .

```
dauer /STATUS=state(2)
/METHOD=ENTER megx icgt bili icgt*megx
/CRITERIA=PIN(.05) POUT(.10) ITERATE(20) .
```

Abb. 4.2. Ergebnisliste zum
Programmaufruf
der Abb.4.1.

```
117 Total cases read .
    0 Cases with missing values
    0 Valid cases with non-positive times
    0 Censored cases before the earliest event in a stratum
    0 Total cases dropped
117 Cases available for the analysis
```

Dependent Variable: DAUER .

Events	Censored
46	71 (60,7%)

Beginning Block Number 0 . Initial Log Likelihood Function

```
-2 Log Likelihood      417,759
```

Beginning Block Number 1 . Method: Enter

Variable(s) Entered at Step Number 1. .

```
MEGX
ICGT
BILI
ICGT*MEGX
```

Log likelihood converged after 6 iterations .

-2 Log Likelihood	345,462		
	Chi-Square	df	Sig
Overall (score)	60,710	4	,0000
Change (-2LL) from			
Previous Block	72,297	4	,0000
Previous Step	72,297	4	,0000

----- Variables in the Equation ----- .

Variable	B	S.E.	Wald	df	Sig	R	Exp(B)
MEGX	-,1163	,0428	7,3617	1	,0067	-,1133	,8902
ICGT	,0409	,0264	2,3940	1	,1218	,0307	1,0417
BILI	-2,27E-04	,0015	,0237	1	,8778	,0000	,9998
ICGT*MEGX	9,281E-05	,0021	,0019	1	,9656	,0000	1,0001

Covariate Means .

Variable	Mean
MEGX	22,2308
ICGT	15,9769
BILI	83,1197
ICGT*MEGX	275,7214

Die Liste in Abb. 4.2 weist die Prüfung der globalen (overall) Nullhypothese mit dem Skore-Test auf und anschließend die Prüfung der Modellveränderung gegenüber dem vorausgehenden Modell ohne Kovariate (Previous Step) mit dem LR-Test (-2LL), womit hier ebenfalls die globale Nullhypothese geprüft wird. Weiter enthält die Liste für jede Kovariate den geschätzten Koeffizienten **B** und die zugehörige Standardabweichung **S.E.** Daraus wird der Quotient $B/S.E.$ gebildet und dann quadriert. Die Zahl $(B/S.E.)^2$ ist in der mit **Wald** überschriebenen Spalte ausgedruckt und bedeutet die Statistik des Wald-Tests. Sie ist Chi²-verteilt und hat einem Freiheitsgrad für eine quantitative Variable und k-1 Freiheitsgrade für eine kategorielle Variable, die k Stufen aufweist.

R bedeutet die partielle Korrelation der unabhängigen Variablen mit der Hazardfunktion; jedoch wird $R = 0$ gesetzt, falls der Wert der Wald-Statistik $< 2*df$.

Zur Kovariaten z gibt **Exp(B)** (= Hodschkin-Lehmann-Schätzer bzw. Hazard-Ratio) an, um das wievielfache die Hazardfunktion ansteigt bzw. abfällt, wenn man z um eine Einheit vergrößert. Mit graphischen Darstellungen zu verschiedenen Wertekombinationen der Kovariaten läßt sich solch ein Effekt auf die Survivalfunktion veranschaulichen, siehe Abschnitt 4.6. Wenn z 0-1-verteilt ist, stellt Exp(B) das relative Risiko dar, d.h. das Verhältnis der geschätzten Hazard unter $z = 1$ zur geschätzten Hazard unter $z = 0$. Zum Konfidenzintervall von Exp(B) s. Abb.4.6.

4.4. Kategorielle Kovariate

Treten kategorielle Variable auf, also solche mit 2, 3 oder 4 Stufen (möglichst nicht mehr), sind diese zu kenn-zeichnen. Dazu klickt man im Dialogfeld Abb. 4.1 rechts **Kategorial...** an. In dem dann eingeblendeten Dialogfeld, siehe **Abb. 4.3**, schiebt man von den links aufgeführten Kovariaten die kategoriellen in das mittlere Feld.

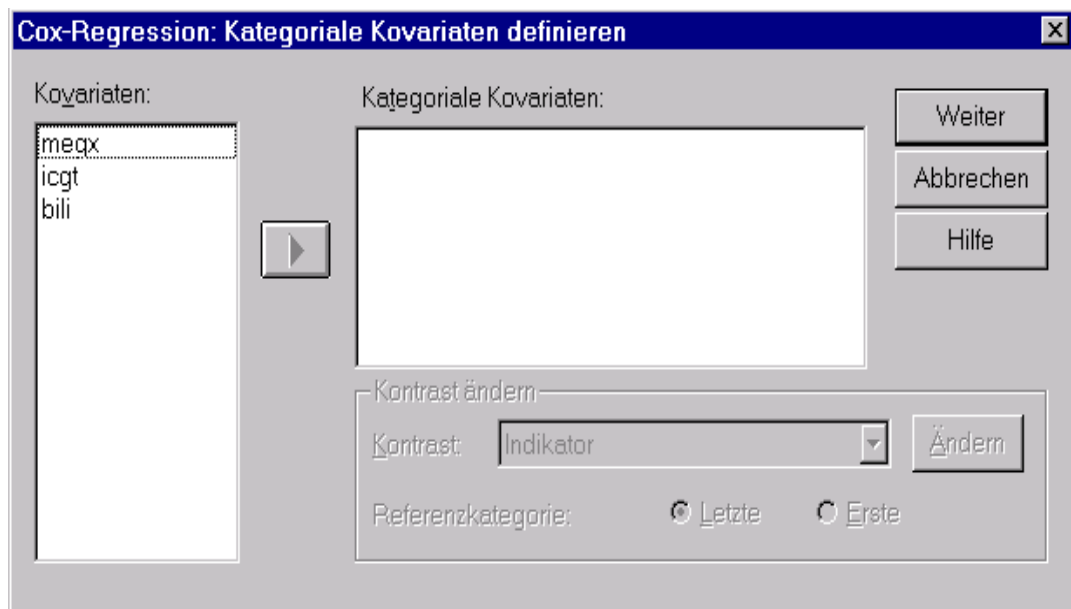


Abb. 4.3. Definition kategorieller Variablen

Für $k > 2$ gilt: Zu den k Stufen werden $k-1$ Kontraste in Form von $k-1$ Designvariablen gebildet. Diese Designvariablen können in die Modellgleichung eingehen und für sie werden die Koeffizienten berechnet. Chiquadrat- und p-Wert werden für sie, aber auch für den Haupt- oder Gesamteffekt der kategoriellen Variablen ermittelt. Bei der schrittweisen Analyse, s.u., hängt es von diesem p-Wert zum Gesamteffekt ab, ob die zugehörigen Designvariablen in die Modellgleichung eingehen - und zwar alle gemeinsam - oder alle nicht.

Die Form von Kontrasten ist für jede kategorielle Variable in verschiedener Weise möglich. Standardmäßig wird von SPSS die Form **Indikator** eingesetzt. Hier wird für die 1-te bis $(k-1)$ -te Kategorie der Variablen, wenn die Referenzkategorie die k -te Kategorie ist, oder für die 2-te bis k -te Kategorie, wenn die erste Kategorie die Referenzkategorie ist, angegeben, ob sie zutrifft (Wert = 1) oder nicht (Wert = 0). Die Referenzkategorie trifft

zu, wenn alle anderen Kategorien nicht zutreffen.

Eine andere gebräuchliche ist die **einfache** Differenz. Hier wird zur Referenzkategorie für jede der übrigen $k-1$ Kategorien die Differenz zur Referenzkategorie gebildet..

Bei den **Differenz**-Kontrasten bzw. den **Helmert**-Kontrasten wird der Effekt jeder Kategorie - bis auf die erste bzw. bis auf die letzte - mit dem mittleren Effekt der vorausgehenden bzw. der nachfolgenden Kategorie gebildet.

Bei ordinalskalierten Variablen wie *index* in Abb. 5.3 kann man **Polynomiale** Kontraste benutzen, um mit dem ersten Kontrast den linearen Effekt, mit dem zweiten den quadratischen usw. zu prüfen. Hier kann durch geeignete Wahl von Kodierungen (z.B. 1, 3, 4, 7 statt 1, 2, 3, 4) die unterschiedliche Nähe benachbarter Kategorien berücksichtigt werden.

Siehe als Beispiel die Variable *chekl* in der Ergebnisliste der Abb. 4.5

4.5. Schrittweise Analysen

Da die Kovariaten in der Regel untereinander abhängig sind, genügen meist wenige von ihnen, um schon den Gesamteinfluß auf die Überlebenszeitfunktion auszuschöpfen. Das Ziel einer schrittweisen Analyse liegt darin, diese möglichst wenigen, alle übrigen ausschöpfenden Merkmale herauszufinden. Die verschiedenen Alternativen der schrittweisen Verfahren werden gezeigt, wenn man im Dialogfeld der Abb. 4.1 in der Zeile **Methode** den rechten nach unten weisenden Pfeil anklickt, siehe **Abb.4.4**.

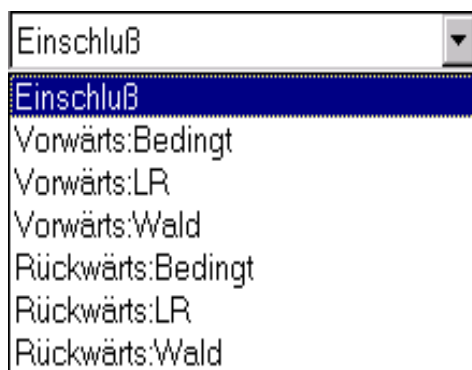


Abb. 4.4 Alternativen der Auswertungsmethode

Die Methode **Einschluß** nimmt alle Variablen in die Modellgleichung auf.

Die **vorwärts** gerichteten Methoden beginnen mit einem Modell, das nur eine Konstante, aber keine Kovariate enthält. Im ersten Schritt wird für jede Kovariate geprüft - hierfür setzt SPSS leider nur den Skore-Test ein -, ob sie einen signifikanten Einfluß ausübt. Dazu wird ein χ^2 -Wert und der zugehörige p-Wert ausgegeben. Die Kovariate mit dem kleinsten p-Wert wird in die Modellgleichung aufgenommen, falls der p-Wert kleiner ist als der vorgegebene Grenzwert, in der Regel 0.05. Gibt es keine Kovariate mit $p < 0.05$, endet die Analyse. Andernfalls wird im folgenden Schritt die Verbesserung, die die ausgewählte Variable bringt, mit Hilfe des LR-Tests (Previous Step) noch einmal beschrieben, wird zum aktuellen Modell die Klassifikationstabelle ausgegeben und werden die Eigenschaften der 'Variables in the Equation' gekennzeichnet, siehe die Erläuterungen zum nachfolgenden Datenbeispiel. Am Ende des Schritts wird geprüft, welche der restlichen Kovariaten eine zusätzliche signifikante Verbesserung bringt. Die Variable mit dem nun kleinsten p-Wert, falls kleiner 0.05, geht als zweite Variable in die Modellgleichung ein - und so fort.

In jeden Schritt eingefügt ist die Tabelle 'Model if Term Removed'. Hier wird für jede Kovariate in der Gleichung geprüft, ob ihr Entfernen (remove) aus der Gleichung eine signifikante Verschlechterung des Modells bewirkt. Je nach Wahl der Methode kann hierfür der Likelihood-Ratio-Test aufgrund von **bedingten** Schätzwerten, der LR-Test (**LR**) oder der **Wald**-Test eingesetzt werden. Es wird die Kovariate wieder aus der Modellgleichung entfernt, deren p-Wert am größten ist, falls er größer als der Grenzwert für den Ausschluß ausfällt (zur Festlegung dieses p-Wertes siehe Abschnitt 5.6) und es würde dann im folgenden Schritt zunächst das in der Weise reduzierte Modell gerechnet. Um der Stabilität des Verfahrens willen muß dieser Grenzwert größer als der Aufnahmegrenzwert sein.

Umgekehrt wird bei den **rückwärts** gerichteten Methoden vorgegangen. Zunächst werden alle Variablen in die Gleichung aufgenommen und dann nach obigem Muster schrittweise die unergiebigsten Variablen aus der Gleichung entfernt.

Falls mehr als ein **Variablenblock** definiert ist, kann für jeden Block eine eigene Methode festgelegt werden. Sollen z.B. bei einer schrittweisen Analyse einige Variablen in jedem Fall in die Modellgleichung eingehen, werden diese in den Block 1 zusammengefaßt und mit der Methode **Einschluß** bearbeitet, während die restlichen Variablen in den Block 2 gegeben und mit einer schrittweisen, z.B. vorwärts gerichteten Methode auf ihre prognostische Bedeutung hin untersucht werden. Hierbei ist die Reihenfolge der beiden Blöcke bewußt so gewählt. Denn die Variablenblöcke werden in aufsteigender Reihenfolge ihrer Nummern bearbeitet. Die Variablen des Blocks 1 werden allerdings nur dann geschlossen in die Modellgleichung übernommen, wenn der entsprechende Test (Change from Previous Block) es zuläßt.

Wir verwenden das Datenbeispiel von Abschnitt 4.2, bei dem nun auf die acht Kovariaten *mexg*, *icgt*, *asci*, *album*, *bili*, *quick*, *ap* und *chekl* die Vorwärtsanalyse (**LR**) angewandt werden soll. Von den acht Kovariaten ist *chekl* als kategoriell zu kennzeichnen. Der Kontrast **Indikator** wird beibehalten. Dazu die nachfolgende Ergebnisliste

COXREG

```
dauer /STATUS=state(2)
/CONTRAST (chekl)=Indicator
/METHOD=FSSTEP(LR) mexg icgt asci album bili quick ap chekl
```

/PLOT SURVIVAL

/CRITERIA=PIN(.05) POUT(.06) ITERATE(20) .

Indicator Parameter Coding

	Value	Freq	(1)	(2)
CHEKL	1,00	26	1,000	,000
	2,00	43	,000	1,000
	3,00	48	,000	,000

Abb. 4.5. Ergebnisliste zum COX-Modell mit 8 Kovariaten

Dependent Variable: DAUER .

Events	Censored
46	71 (60,7%)

Beginning Block Number 0. Initial Log Likelihood Function

-2 Log Likelihood 417,759

----- Variables not in the Equation ----- .

Residual Chi Square = 60,57 with 9 df Sig = ,0000

Variable	Score	df	Sig	R ,
MEGX	36,0496	1	,0000	,2855
ICGT	30,3351	1	,0000	,2604
ASCI	5,4847	1	,0192	,0913
ALBUM	1,8294	1	,1762	,0000
BILI	27,8962	1	,0000	,2490
QUICK	8,7762	1	,0031	,1274
AP	1,4992	1	,2208	,0000
CHEKL	16,9266	2	,0002	,1759
CHEKL(1)	15,8354	1	,0001	,1820
CHEKL(2)	,1913	1	,6619	,0000

Beginning Block Number 1. Method: Forward Stepwise (Likelihood Ratio)

Variable(s) Entered at Step Number 1.. MEGX

-2 Log Likelihood 351,708

	Chi-Square	df	Sig
Overall (score)	36,050	1	,0000
Change (-2LL) from			
Previous Block	66,051	1	,0000
Previous Step	66,051	1	,0000

----- Variables in the Equation -----

Variable	B	S.E.	Wald	df	Sig	R	Exp(B)
MEGX	-,1195	,0198	36,4964	1	,0000	-,2874	,8874

----- Variables not in the Equation ----- .

Residual Chi Square = 19,62 with 8 df Sig = ,0119

Variable	Score	df	Sig	R ,
ICGT	7,1709	1	,0074	,1113
ASCI	6,5888	1	,0103	,1048
ALBUM	,5931	1	,4412	,0000
BILI	,6186	1	,4316	,0000
QUICK	1,5661	1	,2108	,0000
AP	1,3868	1	,2390	,0000
CHEKL	10,4134	2	,0055	,1239
CHEKL(1)	9,8370	1	,0017	,1370
CHEKL(2)	5,0526	1	,0246	,0855

-----Model if Term Removed-----

Term	Loss		
Removed	Chi-square	df	Sig
MEGX	66,0506	1	,0000

Variable(s) Entered at Step Number 2.. CHEKL

Dieser Schritt ist hier ausgelassen

Variable(s) Entered at Step Number 3.. ICGT

Log likelihood converged after 5 iterations.

-2 Log Likelihood 338,247

	Chi-Square	df	Sig
Overall (score)	55,273	4	,0000
Change (-2LL) from			
Previous Block	79,512	4	,0000
Previous Step	4,263	1	,0390

----- Variables in the Equation -----

Variable	B	S.E.	Wald	df	Sig	R	Exp(B)
MEGX	-,1225	,0217	31,9083	1	,0000	-,2676	,8847
ICGT	,0335	,0155	4,6765	1	,0306	,0800	1,0341
CHEKL			7,3150	2	,0258	,0891	
CHEKL(1)	,5464	,3912	1,9512	1	,1625	,0000	1,7271
CHEKL(2)	-,4334	,3872	1,2530	1	,2630	,0000	,6483

----- Variables not in the Equation -----

Residual Chi Square = 4,713 with 5 df Sig = ,4519

Variable	Score	df	Sig	R
ASCI	2,4976	1	,1140	,0345
ALBUM	2,2185	1	,1364	,0229
BILI	,0743	1	,7852	,0000
QUICK	,0327	1	,8566	,0000
AP	,4194	1	,5172	,0000

-----Model if Term Removed-----

Term	Loss		
Removed	Chi-square	df	Sig
MEGX	50,2685	1	,0000
ICGT	4,2629	1	,0390
CHEKL	7,2445	2	,0267

No more variables can be added or deleted.

Covariate Means

Variable	Mean
MEGX	22,2308
ICGT	15,9769
ASCI	1,9915
ALBUM	36,4103
BILI	83,1197
QUICK	63,6325
AP	475,7607
CHEKL(1)	,2222
CHEKL(2)	,3675

Die vorausgehende Ergebnisliste zeigt oben den Programmaufruf - die Befehle PLOT und CRITERIA werden im Abschnitt 4.6 erläutert -, anschließend die Definition der zwei Designvariablen zur Variablen *chekl*: zu den Stufen 1, 2 und 3 sind in den Spalten (1) und (2) die entsprechenden

Werte der Designvariablen angegeben; diese Werte werden mit dem gewählten Kontrast festgelegt, d.h. würde ein anderer Kontrast gewählt, würden hier andere Werte stehen.

Im Weiteren läßt die Ergebnisliste die schrittweise Analyse erkennen. Nach jedem Schritt wird die Modellgleichung in der Zusammensetzung der Kovariaten geändert oder das Verfahren beendet. Der nullte Schritt (Block Number 0) dient dazu, den Einfluß jeder Kovariaten unabhängig von den übrigen zu prüfen und mit dem p-Wert zu bewerten. Die Kovariate, die den größten Chiquadratwert bzw. den kleinsten p-Wert ergibt - das ist hier *mex* - wird in die Modellgleichung des 'Step Number 1' aufgenommen, da die Bedingung $p < 0.05$ erfüllt ist.

Zu Beginn eines Schrittes wird mit dem Overall-Score-Test für alle Variablen in der Gleichung die Hypothese geprüft, ob ihre β -Koeffizienten Null sind. Die Verbesserung gegenüber dem vorausgehenden Schritt, d.h. die Hypothese, daß der Koeffizient nur der aufgenommenen Variablen Null ist, prüft der LR-Test. Im ersten Schritt also prüfen beide Tests dieselbe Hypothese; dennoch führen sie auf einen erheblichen Unterschied im Chiquadratwert und verdeutlichen damit den groben Näherungscharakter des Skore-Tests.

Es folgt die in Abschnitt 4.3 beschriebene Kennzeichnung der Kovariaten, mit der die Modellgleichung gebildet ist.

In der anschließenden Tabelle wird zu jeder Kovariaten, die noch nicht in der Gleichung enthalten ist, mit dem Skore-Test geprüft, ob ihr Hineinnehmen eine Verbesserung der Vorhersagegenauigkeit erbringt, und es wird die Kovariate mit kleinstem p-Wert, falls < 0.05 , in die Gleichung aufgenommen. Zu jeder Kovariaten, die bereits in der Gleichung enthalten ist, wird geprüft (LOSS CHI-SQUARE), ob ein Entfernen aus der Gleichung eine Verschlechterung der Vorhersagegenauigkeit bewirkt, und es wird diejenige Kovariate wieder herausgenommen, deren p-Wert am größten ist, falls > 0.06 .

Das Ergebnis der Analyse besagt, daß neben dem starken Einfluß von *mex* für sich genommen auch andere Merkmale einen erheblichen prognostischen Wert haben, insbesondere *icgt*, *billi* und *chekl*, der jedoch von der Variablen *mex* weitgehend schon ausgeschöpft wird. Über *mex* hinaus haben nur die Kovariate *chekl* und *icgt* einen signifikanten, wenn auch geringen prognostischen Wert.

4.6. Optionen und Graphiken

Mit Hilfe der Optionen kann man, wie **Abb. 4.6** erkennen läßt, den Inhalt der Ergebnisliste und die schrittweise Analyse steuern.

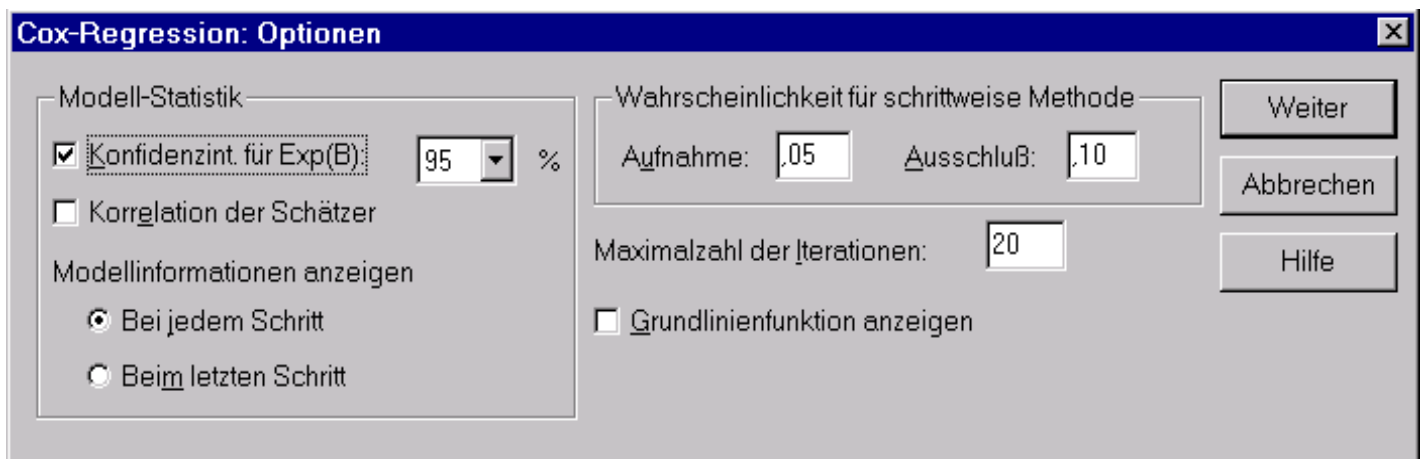


Abb. 4.6 Optionen bei der Anwendung des COX-Modells

Die p-Grenzwerte für Aufnahme (Enter) und Ausschluß (Remove) einer Kovariaten in die bzw. aus der Modellgleichung können nur auf 2 Dezimalstellen genau angegeben werden, der erste wird in der Regel auf 0.05 festgesetzt, der zweite sollte - z.B. mit 0.06 oder 0.10 - größer als der erste ausfallen. Die maximale Anzahl der Iterationen bei der Berechnung der Log Likelihood ist mit 20 schon hoch angesetzt. Diese Einstellungen werden im Programmaufruf zu Beginn der Ergebnisliste mit aufgeführt.

Die Grundlinienfunktion $S_0(t)$ ist in Abschnitt 4.1 definiert: sie liegt vor, wenn in der Modellgleichung die Konstante und alle Koeffizienten β den Wert Null annehmen. Sie wird in Form einer Tabelle ausgegeben.

Im Arbeitsfeld für Graphiken, siehe **Abb. 4.7**, können drei verschiedene Graphiktypen angefordert werden, die Survival-, die Hazard- und die Log-minus-Log-Survivalfunktion.

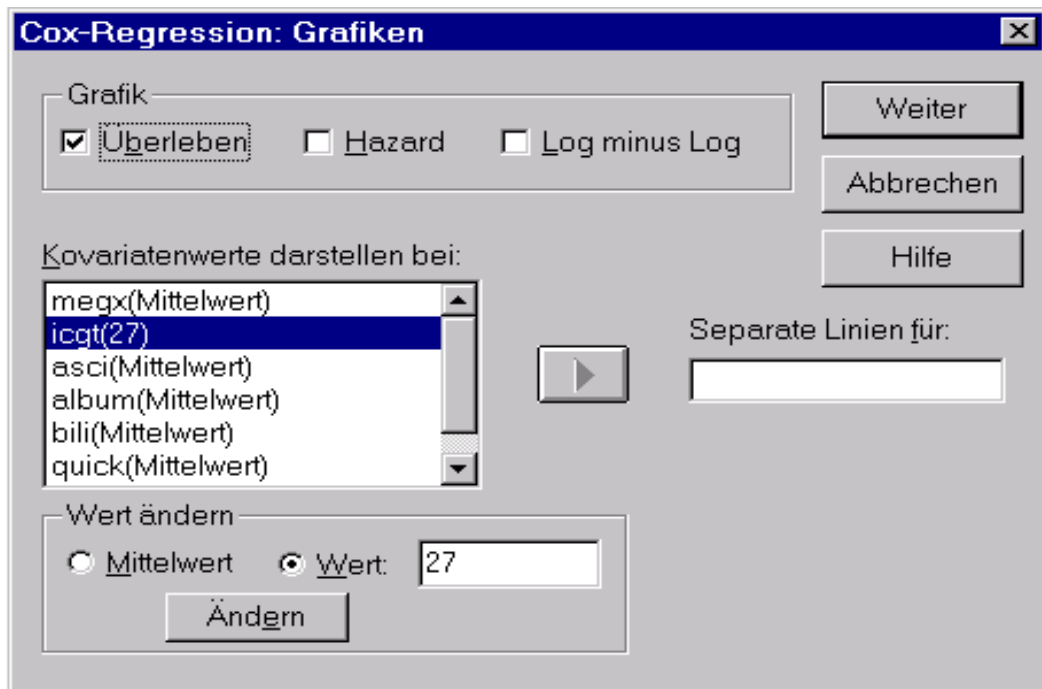


Abb. 4.7. Dialogfeld 'Graphiken'

Jede Kovariate kann verschiedene Werte annehmen. Zu jeder Wertekombination (Muster) der Kovariaten gibt es eine je eigene Survivalfunktion. Wenn Graphiken angefordert werden, wird standardmäßig die Funktion, die zu den Mittelwerten der Kovariaten gehört, als Graphik erstellt. Zusätzlich können Graphiken für ein beliebig anderes Wertemuster erzeugt werden. Die gewünschten Werte werden in das Dialogfeld 'Graphiken' eingegeben. Um z.B. bei der Variablen *icgt* den Mittelwert durch den Wert 27 zu ersetzen, wird die Variable *icgt* markiert, dann im Feld darunter 'Wert' angeklickt, die Zahl 27 eingegeben und 'Ändern' bestätigt.

In der Abb.4.8 sind die Survival- und die Log minus Log-Survival-Funktion dargestellt.

LML-Funktion bei Mittelwert der Kovariaten

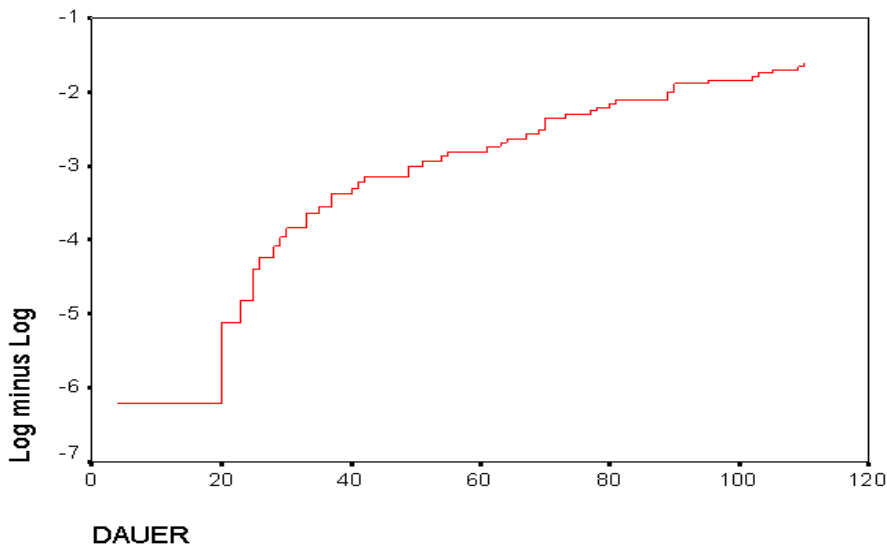
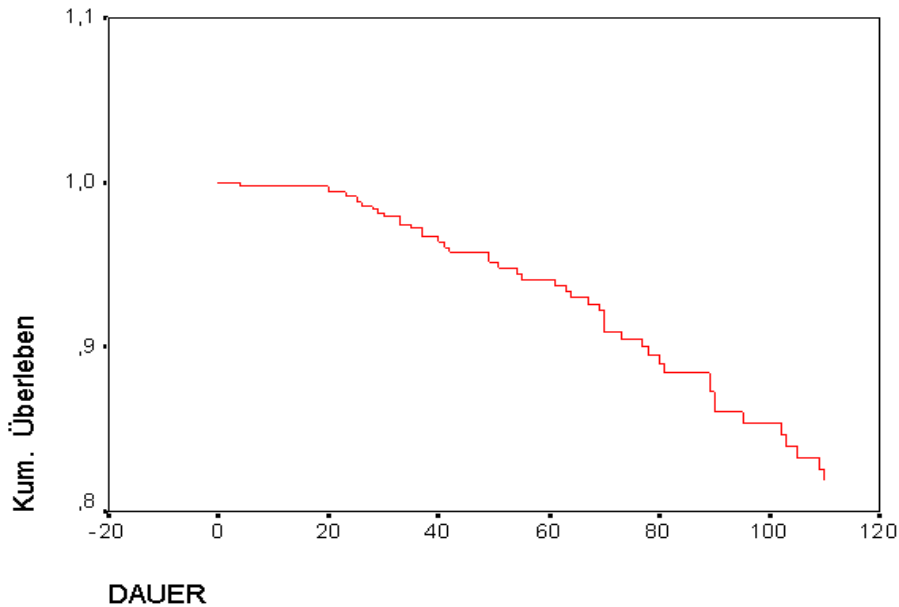


Abb. 4.8. Graphiken zum Beispiel aus 4.5

Überlebensfunktion bei Mittelwert der Kovariaten



4.7. Geschichtete Analysen - nicht proportionales Modell

Eine geschichtete Analyse gibt die Möglichkeit, die Voraussetzung der Proportionalität abzuschwächen, indem man zuläßt, daß für die in Gruppen (Schichten) aufgeteilten Patienten die Proportionalität nur innerhalb der Gruppen gilt - mit je eigener beliebiger Grundlinienfunktion, wogegen der eine Parametervektor β für alle Gruppen gemeinsam geschätzt wird. Die Variable zur Festlegung der **Schichten** wird unten im Dialogfeld der Abb. 4.1 angegeben.

Wird der Ausdruck einer Grundlinienfunktion gewünscht, gibt SPSS je Schicht eine eigene Funktionstabelle aus.

Fordert man zur geschichteten Analyse Graphiken an, wird die Kurve eines Typs je Schicht gesondert berechnet und werden dann alle Kurven dieses Typs in einer Graphik zusammengeführt. So ist ein Vergleich dieser Kurven untereinander in einfacher Weise möglich.

Wie weit das Modell der geschichteten Analyse erforderlich war, erkennt man, wenn man die Log-minus-Log-Survivalkurven der verschiedenen Schichten miteinander vergleicht:

Das proportionale (d.h. ungeschichtete) Modell gilt, wenn die Kurven parallel verlaufen und sich über den ganzen Bereich nur durch eine Konstante unterscheiden. Dann ist es sinnvoll, die Schichtvariable in die Gruppe der Kovariaten aufzunehmen.

V. Logistische Regression

5.1. Einführung

In der logistischen Regressionsanalyse wird eine Relation zwischen einer binären abhängigen Variablen Y und einer Menge von qualitativen (kategoriellen) und/oder quantitativen (intervallskalierten) Variablen gesetzt und untersucht. An die Verteilungen dieser Variablen werden keine Bedingungen geknüpft. Die Codes der Ergebnisvariablen Y sind :

0 für Mißerfolg, Fehl, negatives Testergebnis,

1 für Erfolg, positives Testergebnis.

Das Wahrscheinlichkeits- oder Regressionsmodell lautet:

$$P(Y = 1) = P = \exp(u) / [1 + \exp(u)] \quad P(Y = 0) = 1 - P$$

Darin stellt $u = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$

eine lineare Funktion der p unabhängigen Einflußgrößen X_1, X_2, \dots, X_p und der Konstanten dar, d.h. eines Vektors \mathbf{X} mit $\mathbf{X}' = (1 \ X_1 \ X_2 \dots \ X_p)$. Die Einflußgrößen werden auch **prognostische Faktoren** oder **Kovariate** genannt.

$\beta' = (\beta_0 \ \beta_1 \ \dots \ \beta_p)$ bedeutet einen unbekanntem Parametervektor; er wird nach der Maximum-Likelihood-Methode geschätzt.

Man bezeichnet den Quotienten $P/(1 - P)$ als Odds und den Logarithmus davon als die Log Odds oder Logit der Wahrscheinlichkeiten P und $1 - P$, und es gilt mit obigem Modell:

$$\ln [P / (1 - P)] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p = \beta' \mathbf{X} .$$

Man kann vorgeben, daß alle Kovariate in die Regressionsgleichung aufgenommen werden. Häufiger wird eine **schrittweise Analyse** gewählt. Hier werden, wenn der Benutzers nicht anderes vorgibt, in die endgültige Regressionsgleichung nur solche Faktoren aufgenommen, deren prognostischer Beitrag signifikant ausfällt. Da die prognostischen Faktoren in der Regel voneinander abhängen, kann der Einfluß des einen oder anderen Faktors schon von denjenigen Faktoren erfaßt werden, die bereits in der Gleichung stehen. Das **Ziel** der schrittweisen Analyse liegt darin, für die Modellgleichung die prognostischen Faktoren und nur die auszuwählen, die nötig sind, um den Einfluß aller Faktoren zu erfassen. Die Einwirkung des Benutzers kann z.B. darin liegen, daß er Faktoren auswählt, die in jedem Fall in die Gleichung eingehen sollen. - Neben den (Haupt-) Effekten der Faktoren lassen sich bei quantitativen Variablen auch Wechselwirkungseffekte, d.h. Effekte der Produkte zweier Faktoren, prüfen. Ein kategorieller prognostischer Faktor mit k Ausprägungen wird durch eine Menge von k-1 dichotomen Design-variablen ersetzt. Diese werden aber stets als zusammenhängend behandelt, d.h. die k-1 Designvariablen sind gemeinsam entweder in der Gleichung oder nicht darin. Die Ergebnisvariable läßt sich als eine Variable ansehen, die die Beobachtungseinheiten (Patienten) in zwei Gruppen einteilt. Man untersucht also in der logistischen Regressionsanalyse, wie weit die prognostischen Faktoren die Wahrscheinlichkeit dieser Gruppenzugehörigkeit beeinflussen. Es sei darauf hingewiesen, daß ein ähnliches Ziel auch von der schrittweisen Diskriminanzanalyse angestrebt wird, jedoch unter der Voraussetzung normalverteilter Faktoren. Diese Voraussetzung ist nicht als sehr streng anzusehen. Ist sie jedoch erfüllt, sollte die Diskriminanzanalyse als wirkungsvolleres Verfahren vorgezogen werden. Konzept und Handhabung der kategoriellen Variablen, der Tests und der schrittweisen Verfahren gleichen denen zum COX-Modell in der Überlebenszeitanalyse.

5.2. Programmgestaltung im Dialogfeld und Beispiel

Als Beispiel dient eine Untersuchung an 89 Patienten einer Therapiestudie mit zwei Diagnosen: akute D. und chronische D. Dabei geht es um die Bewertung vermuteter prognostischer Faktoren für die Heilungschancen der Patienten unter Therapie. Als Einflußfaktoren sind vorgesehen: die Diagnose, ein Bewertungsindex der Patienten mit drei Stufen sowie drei quantitative Merkmale, die die Eigenschaften von Leber, Blut und Urin kennzeichnen.

Zum Programmaufruf wird unter den Statistiken die Regression angewählt und dann unter den Regressionen die **logistische Regression**. Das Dialogfeld der folgenden **Abb. 5.1**, zeigt auf der Linken Seite ein Fenster (Rechteck),



Abb. 5.1. 'Logistische

Regression'

in dem alle definierten Variablen der Datenbank angesprochen werden können. Markierte Variable werden in das Feld für die abhängige Variable bzw. für die Kovariaten übertragen. Wird bei quantitativen Variablen vermutet, daß zwischen zwei Variablen bezüglich der abhängigen Variablen Wechselwirkungen bestehen, kann auch das Produkt der beiden Variablen als Kovariate aufgenommen werden. Dazu sind beide Variablen im linken Variablenfeld zu markieren und dann **>a*b>** anzuklicken.

Während der Benennung der Kovariaten kann die darüberstehende Blocknummer schrittweise erhöht werden. So lassen sich Variablenblocks bilden. Die Auswertungsmethode, siehe Abschnitt 5.4, wird je Block festgelegt. Als Auswertungsmethode wird standardmäßig **Einschluß** vorgeschlagen. Damit gehen alle Kovariate zugleich in die Regressionsgleichung ein. Alternativen dazu in Abschnitt 5.4.

Der mit Abb. 5.1 festgelegte Programmaufruf führt zur Ergebnisliste der **Abb. 5.2**. Man sollte die SPSS-Grundeinstellung immer so wählen, daß die Ergebnisliste den Programmaufruf mit enthält, um besser nachprüfen zu können, was das Programm gerechnet hat. Insbesondere werden dabei auch unveränderte Voreinstellungen wie hier die p-Werte für Variableneinschluß oder -ausschluß sowie die maximale Anzahl der Einzelschritte bei den iterativen Berechnungen angezeigt. Anschließend werden die Anzahlen der gültigen und ungültigen Fälle ausgegeben. Sie sollten stets auf ihre Richtigkeit hin kontrolliert werden.

Weiter umfaßt die Ergebnisliste eine Prüfung des Modells - hier mit drei Einflußgrößen -, eine Klassifikationstabelle - vergl. dazu Abschnitt 5.8 -, sowie Charakteristika der eingehenden Variablen, die im folgenden Abschnitt näher besprochen werden.

5.3. Schätzen von β und Testen von $H_0: \beta = 0$

Zum vorliegenden Modell und zu den Realisationen (y_i, \mathbf{x}_i) , $i = 1, 2, \dots, n$, des Vektors (Y, \mathbf{X}) von Ergebnisvariable und Kovariaten bedeutet die Likelihoodfunktion

$$L(\mathbf{b}) = \prod_{i=1}^n \left\{ \frac{\exp(y_i \mathbf{b}' \mathbf{x}_i)}{1 + \exp(\mathbf{b}' \mathbf{x}_i)} \right\}$$

die Wahrscheinlichkeit für das Auftreten der beobachteten Erfolge und Mißerfolge. Darin meint $\mathbf{b}' \mathbf{x}_i$ die Summe $b_0 + b_1 x_{i1} + \dots + b_p x_{ip}$. Der Vektor \mathbf{b} steht für den noch zu schätzenden Vektor β . Die Schätzwerte B_i für β_i werden so bestimmt, daß $L(\mathbf{b})$ maximal wird. Die Lösung ist über ein nicht lineares Gleichungssystem nur iterativ, d.h. mit höherem Rechenaufwand zu erhalten. Der Logarithmus des Maximums (Log Likelihood) wird mit LL bezeichnet.

Man kann das Modell dadurch ändern, daß man nur einige oder gar keine der Kovariaten in die Modellgleichung aufnimmt. Entsprechend ändert sich der Wert LL. Ist M1 ein Modell mit keiner oder nur wenigen Kovariaten und M2 ein Modell, daß gegenüber M1 um wenigstens eine Kovariate erweitert und sonst gleich ist, so gilt für die zugehörigen Log-Likelihood-Werte LL1 und LL2: $-2(LL1 - LL2)$ ist annähernd Chi2-verteilt und ist die Teststatistik im sog. **Maximum-Likelihood-Ratio-Test** (MLR-Test). Die Anzahl df der Freiheitsgrade ist gleich der Anzahl der Kovariaten, um die M2 gegenüber M1 erweitert ist.

Im MLR-Test wird der Model Chi-Square-Wert zum Prüfen der **globalen Nullhypothese** $H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$ eingesetzt, bei der schrittweisen Analyse der Step-Chi-Square zur Prüfung der Modellverbesserungen gegenüber dem vorausgehend Schritt. Wenn die Variablen in Blöcken zusammengefaßt wurden, gibt der Block-Chi-Square die Modellverbesserungen gegenüber dem vorausgehend verwendeten Block an.

Die beiden R^2 -Werte nach *Cox & Snell* und nach *Nagelkerke* haben den Charakter eines Bestimmtheitsmaßes. Sie geben ähnlich wie in der linearen Regression den Anteil der erklärten Variation der Zielgröße an. Während der erste Wert stets kleiner 1 ist, ist der zweite die Normierung des ersten mit $0 \leq R^2 \leq 1$.

In der schrittweisen Analyse verwendet SPSS für das Prüfen der Modellverbesserungen durch einzelne Variablen und für ihre Auswahl den sog. **Skore-Test**. Hier wird die Chi2-Teststatistik anhand erster und zweiter Ableitungen der Maximum-Likelihood-Funktion berechnet. Der Skore-Test läßt sich mit vergleichsweise geringem rechnerischen Aufwand durchführen, ist dafür aber als gröbere und damit vorläufige Näherung anzusehen. Siehe auch Abschnitt 5.5.

Als dritten Test zur Bewertung von Kovariaten verwendet SPSS den **Wald-Test**: Zu seiner Erläuterung wird die Ergebnisliste der **Abb. 5.2** betrachtet. Die Liste enthält unten für jede Kovariate (Variables in the Equation) den geschätzten Koeffizienten B und die zugehörige Standardabweichung S.E. Daraus wird der Quotient B/S.E. gebildet und dann quadriert. Die Zahl $(B/S.E.)^2$ ist in der mit **Wald** überschriebenen Spalte ausgedruckt und bedeutet die Statistik des Wald-Tests. Sie ist näherungsweise Chi²-verteilt mit einem Freiheitsgrad für eine quantitative Variable und mit k-1 Freiheitsgraden für eine kategorielle Variable mit k Stufen (siehe Abschnitt 5.4).

R gibt die partielle Korrelation der unabhängigen Variablen mit der abhängigen, den Log Odds, an; jedoch wird $R = 0$ gesetzt, falls der Wert der Wald-Statistik $< 2 \cdot df$.

Exp(B) gibt an, um das wievielfache die Odds ansteigen bzw. abfallen, wenn man die betreffende Kovariate um eine Einheit vergrößert und die übrigen Kovariaten konstant hält.

```
LOGISTIC REGRESSION VAR=erfolg
  /METHOD=ENTER leber blut urin
  /CRITERIA PIN(.05) POUT(.10) ITERATE(20) CUT(.5) .
```

```
Total number of cases: 89 (Unweighted)
```

```
Number of selected cases: 89
```

```
Number of unselected cases: 0
```

```
Number of selected cases: 89
```

```
Number rejected because of missing data: 0
```

```
Number of cases included in the analysis: 89
```

Abb. 5.2. Ergebnisliste zum

Programmaufruf

der Abb.5.1.

```
Dependent Variable.. ERFOLG
```

```
Beginning Block Number 0. Initial Log Likelihood Function
```

```
-2 Log Likelihood 109 , 23742
```

```
* Constant is included in the model .
```

```
Beginning Block Number 1. Method: Enter .
```

```
Variable(s) Entered on Step Number 1.
```

```
LEBER
```

```
BLUT
```

```
URIN
```

-2 Log Likelihood	86,449		
Goodness of Fit	74,638		
Cox & Snell - R ²	,226		
Nagelkerke - R ²	,320		
	Chi-Square	df	Significance
Model	22,788	3	,0000
Block	22,788	3	,0000
Step	22,788	3	,0000

Classification Table for ERFOLG .

The Cut Value is ,50

Predicted

Observed	Heilung		Abstoss		Percent Correct
	H	I	A	I	
Heilung	9	18			33,33%
Abstoss	8	54			87,10%
	Overall				70,79%

----- Variables in the Equation ----- .

Variable	B	S.E.	Wald	df	Sig	R	Exp(B)
LEBER	-,0388	,0248	2,4508	1	,1175	-,0642	,9619
BLUT	-,0305	,0151	4,0524	1	,0441	-,1371	,9700
URIN	,0340	,0316	1,1539	1	,2827	,0000	1,0346
Constant	3,5994	1,0617	11,4937	1	,0007		

Bildet man mit den geschätzten Koeffizienten B die Modellgleichung und setzt die beobachteten Werte der Kovariaten *leber*, *blut* und *urin* ein, ergibt dies eine Zuordnungswahrscheinlichkeit P zur Gruppe A der Patienten mit Abstoßung und eine Zuordnungswahrscheinlichkeit 1-P zur Gruppe H der

Patienten mit Heilung. Ein Patient wird dann aufgrund der Meßwerte der Gruppe A zugeordnet (predicted), wenn $P \geq 0.5$ (Cut Value) ist, andernfalls wird er der Gruppe H zugeordnet. Zur Änderung des Cut Values siehe Abschnitt 5.6. Zum Modell mit den geschätzten Koeffizienten und dem festgelegten Cut Value gibt die **Klassifikationstabelle** für die Patienten mit Abstoßung bzw. mit Heilung an, wie viele von diesen richtig der Gruppe A bzw. der Gruppe H zugeordnet würden, wenn nur die Kenntnis der Kovariaten genutzt würde. Mehr dazu unter dem Begriff **Kreuzvalidierung** in Abschnitt 5.8.

5.4. Kategoriale Kovariate

Treten kategoriale Variable auf, also solche mit $k = 2, 3$ oder 4 Stufen (möglichst nicht mehr), sind diese als solche zu kennzeichnen. Dazu klickt man im Dialogfeld der Abb. 5.1 unten **Kategorial...** an. In dem dann eingeblendeten Dialogfeld, siehe **Abb. 5.3**, schiebt man von den links aufgeführten Kovariaten die kategoriellen in das mittlere Feld.

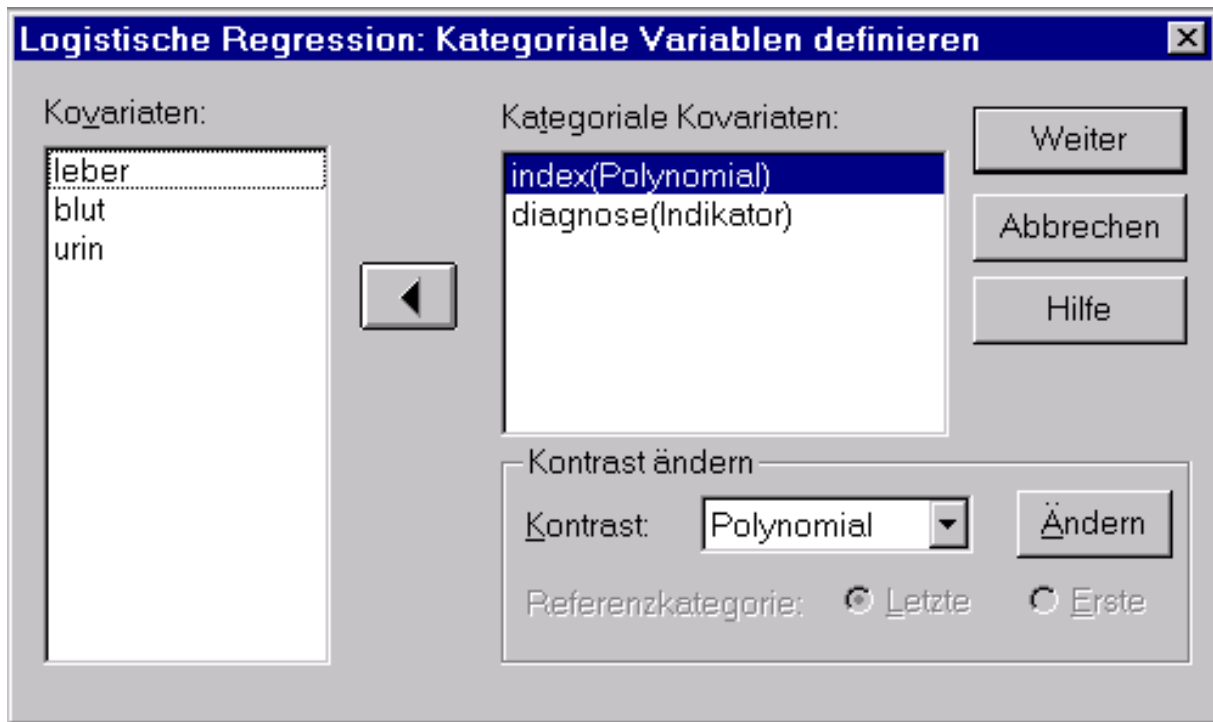


Abb. 5.3. Definition kategorieller Variablen

Für $k > 2$ gilt: Zu den k Stufen werden $k-1$ Kontraste in Form von $k-1$ Designvariablen gebildet. Diese Designvariablen können in die Modellgleichung eingehen und für sie werden die Koeffizienten berechnet. Chi-Quadrat- und p-Wert werden für sie, aber auch für den Haupt- oder Gesamteffekt der kategoriellen Variablen ermittelt. Bei der schrittweisen Analyse, s.u., hängt es von diesem p-Wert zum Gesamteffekt ab, ob die zugehörigen Designvariablen in die Modellgleichung eingehen - und zwar alle gemeinsam - oder alle nicht.

Die Form von Kontrasten ist für jede kategorielle Variable in verschiedener Weise möglich. Vom Programm wird die Form **Indikator** vorgeschlagen. Hier wird für die 1-te bis $(k-1)$ -te Kategorie der Variablen, wenn die Referenzkategorie die k -te Kategorie ist, oder für die 2-te bis k -te Kategorie, wenn die erste Kategorie die Referenzkategorie ist, angegeben, ob sie zutrifft (Wert = 1) oder nicht (Wert = 0). Die Referenzkategorie trifft zu, wenn alle anderen Kategorien nicht zutreffen.

Eine andere gebräuchliche ist die **einfache Differenz**. Hier wird zur Referenzkategorie für jede der übrigen $k-1$ Kategorien die Differenz zur Referenzkategorie gebildet.

Bei den **Differenz**-Kontrasten bzw. den **Helmert**-Kontrasten wird der Effekt jeder Kategorie - bis auf die erste bzw. bis auf die letzte - mit dem mittleren Effekt der vorausgehenden bzw. der nachfolgenden Kategorie gebildet.

Bei ordinalskalierten Variablen wie *index* in Abb. 5.3 kann man **Polynomiale** Kontraste benutzen, um mit dem ersten Kontrast den linearen Effekt, mit dem zweiten den quadratischen usw. zu prüfen. Hier kann durch geeignete Wahl von Kodierungen (z.B. 1, 3, 4, 7 statt 1, 2, 3, 4) die unterschiedliche Nähe benachbarter Kategorien berücksichtigt werden.

Beim Kontrast **Abweichung** wird für jede Kategorie der Variablen die Differenz zum Gesamteffekt gebildet; ausgenommen davon ist die Referenzkategorie, für diese kann der Koeffizient B aus den übrigen Koeffizienten berechnet werden, denn die Summe aller k Koeffizienten ist Null. Siehe als Beispiel die Variable *index* mit 3 Kategorien in der Ergebnisliste der Abb. 5.5, wie zum polynomialen Kontrast die Kodierung der Designvariablen festgelegt ist.

5.5. Schrittweise Analysen

Da die Kovariaten in der Regel untereinander abhängig sind, genügen meist wenige von ihnen, um schon den Gesamteinfluß auf die Odds, den Quotienten $P/(1-P)$, auszuschöpfen. Das Ziel einer schrittweisen Analyse liegt darin, diese meist wenigen, alle übrigen mit ausschöpfenden Merkmale herauszufinden. Die verschiedenen Alternativen der schrittweisen Verfahren werden gezeigt, wenn man im Dialogfeld der Abb. 5.1 in der Zeile **Methode** den rechten nach unten weisenden Pfeil anklickt, siehe **Abb.5.4**.

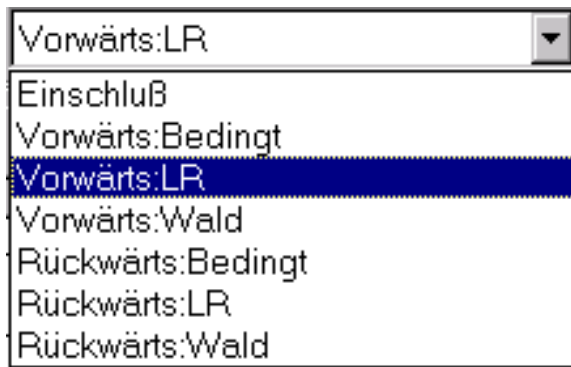


Abb. 5.4. Alternativen der Auswertungsmethode

Die **vorwärts** gerichteten Methoden beginnen mit einem Modell, das nur eine Konstante, aber keine Kovariate enthält. Im ersten Schritt wird für jede Kovariate mit dem Skore-Test geprüft, ob sie einen signifikanten Einfluß ausübt. Dazu wird der Chi²-Wert und der zugehörige p-Wert ausgegeben. Die Kovariate mit dem kleinsten p-Wert wird in die Modellgleichung aufgenommen, falls der p-Wert kleiner ist als der vorgegebene Grenzwert, in der Regel 0.05. Gibt es keine Kovariate mit $p < 0.05$, endet die Analyse. Andernfalls wird im folgenden Schritt die Verbesserung, die die ausgewählte Variable bringt, mit Hilfe des MLR-Tests (Step) noch einmal beschrieben, und es werden zum aktuellen Modell die Klassifikationstabelle ausgegeben und die Eigenschaften der 'Variables in the Equation' gekennzeichnet, siehe die Erläuterungen zum nachfolgenden Datenbeispiel. Am Ende des Schritts wird geprüft, welche der restlichen Kovariaten eine zusätzliche signifikante Verbesserung bringt. Die Variable mit dem nun kleinsten p-Wert, falls kleiner 0.05, geht als zweite Variable in die Modellgleichung ein - und so fort.

In jeden Schritt eingefügt ist die Tabelle 'Model if Term Removed'. Hier wird für jede Kovariate in der Gleichung geprüft, ob ihr Entfernen (remove) aus der Gleichung eine signifikante Verschlechterung des Modells bewirkt. Je nach Wahl der Methode kann hierfür der Likelihood-Ratio-Test aufgrund von **bedingten** Schätzwerten, der MLR-Test (**LR**) oder der **Wald**-Test eingesetzt werden. Es wird die Kovariate wieder aus der Modellgleichung entfernt, deren p-Wert am größten ist, falls er größer als der Grenzwert für den Ausschluß ausfällt (zur Festlegung dieses p-Wertes siehe Abschnitt 5.6) und es würde dann im folgenden Schritt zunächst das in der Weise reduzierte Modell gerechnet. Um der Stabilität des Verfahrens willen muß dieser Grenzwert größer als der Aufnahmegrenzwert sein.

Umgekehrt wird bei **rückwärts** gerichteten Methoden vorgegangen. Zuerst werden alle Variablen in die Gleichung aufgenommen und dann nach obigem Muster schrittweise die unergiebigsten Variablen aus der Gleichung entfernt.

Wir verwenden das Datenbeispiel von Abschnitt 5.2, bei dem nun auf die fünf Kovariaten *index*, *diagnose*, *leber*, *blut* und *urin* die Vorwärtsanalyse (LR) angewandt werden soll.

Von den fünf Kovariaten sind *index* und *diagnose* als kategoriell zu kennzeichnen. Für die ordinale Kovariate *index* wählen wir den Kontrast **Polynomial**, für *diagnose* den Kontrast **Indikator**. Dazu die nachfolgende Ergebnisliste in **Abb. 5.5**. Sie ist eingeschränkt auf die bedeutsamen Ausschnitte und enthält insbesondere nicht den Step Number 2 mit der zusätzlichen Variablen *blut*.

Die Liste zeigt oben den Programmaufruf - der Befehl CRITERIA wird im Abschnitt 5.6 erläutert -, anschließend zur kategoriellen Kovariaten *index* die Definition der Designvariablen: Zu den Stufen 1, 2 und 3 sind in den Spalten (1) und (2) die entsprechenden Werte der Designvariablen angegeben; diese speziellen Werte werden mit dem gewählten Kontrast - hier Polynomial - festgelegt.

Im Weiteren läßt die Ergebnisliste die schrittweise Analyse erkennen. Variablenblock 0 enthält nur die Konstante, Variablenblock 1 alle Kovariaten. Hier wird nach jedem Schritt die Modellgleichung in der Zusammensetzung der Kovariaten geändert oder das Verfahren beendet. Vorausgehend (vor Step Nummer 1) wird der Einfluß jeder Kovariaten unabhängig von den übrigen mit dem Skore-Test geprüft und mit dem p-Wert bewertet. Die Kovariate, die den größten Chi-Quadratwert bzw. den kleinsten p-Wert ergibt - das ist hier *index* - wird in die Modellgleichung des 'Step Number 1' aufgenommen, da die Bedingung $p < 0.05$ erfüllt ist.

```
LOGISTIC REGRESSION VAR=erfolg
  /METHOD=FSSTEP(LR) index diagnose leber blut urin
  /CONTRAST (index)=Polynomial /CONTRAST (diagnose)=Indicator
  /CRITERIA PIN(.05) POUT(.06) ITERATE(20) CUT(.5) .
```

Parameter

	Value	Freq	Coding	
			(1)	(2)
INDEX				
leicht	1	43	-,707	,408
mittel	2	22	,000	-,816
schwer	3	24	,707	,408
DIAGNOSE				
AA	1	62	1,000	
BB	2	27	,000	

Abb. 5.5. Ergebnisliste zur logistischen Regression mit 5 Kovariaten

Dependent Variable.. ERFOLG

Beginning **Block Number 0**. Initial Log Likelihood Function

-2 Log Likelihood 109,23742

----- Variables in the Equation -----

Variable	B	S.E.	Wald	df	Sig	R	Exp(B)
Constant	,8313	,2306	12,9981	1	,0003		

Beginning Block Number 1. Method: Forward Stepwise (LR)

----- Variables not in the Equation -----

Residual Chi Square 39,185 with 6 df Sig = ,0000

Variable	Score	df	Sig.	R
INDEX	<u>29,7498</u>	2	,0000	,4855
INDEX(1)	29,3645	1	,0000	,5005
INDEX(2)	,0303	1	,8617	,0000
DIAGNOSE(1)	11,6634	1	,0006	,2974
LEBER	6,0350	1	,0140	,1922
BLUT	16,2506	1	,0001	,3612
URIN	9,9126	1	,0016	,2691

Variable(s) Entered on Step Number 1.. INDEX

-2 Log Likelihood	78,257
Goodness of Fit	88,993
Cox & Snell - R ²	,294
Nagelkerke - R ²	,416

	Chi-Square	df	Significance
Model	30,980	2	,0000
Block	30,980	2	,0000
Step	<u>30,980</u>	2	,0000

Classification Table for ERFOLG

The Cut Value is ,50

Observed	Predicted						Percent Correct
	Heilung		Abstoss				
	H	I	I	A			
Heilung	H	I	17	I	10	I	62,96%
Abstoss	A	I	7	I	55	I	88,71%
Overall							80,90%

----- Variables in the Equation -----

Variable	B	S.E.	Wald	df	Sig	R	Exp(B)
INDEX			<u>22,0604</u>	2	,0000	,4066	
INDEX(1)	-2,4589	,5291	21,5949	1	,0000	-,4235	,0855
INDEX(2)	,0729	,4827	,0228	1	,8800	,0000	1,0756
Constant	,8216	,2924	7,8960	1	,0050		

----- Model if Term Removed -----

Term	Log Likelihood	-2 Log LR	df	Significance of Log LR
INDEX	-54,619	30,980	2	,0000

----- Variables not in the Equation -----

Variable	Score	df	Sig	R

DIAGNOSE(1)	,4213	1	,5163	,0000
LEBER	6,5149	1	,0107	,2033
BLUT	10,4479	1	,0012	,2781
URIN	4,6739	1	,0306	,1565

Variable(s) Entered on Step Number 2.. BLUT Dieser Schritt ist ausgelassen.

Variable(s) Entered on Step Number 3.. LEBER

-2 Log Likelihood	62,760
Goodness of Fit	61,401
Cox & Snell - R ²	,407
Nagelkerke - R ²	,575
	Chi-Square df Significance
Model	46,478 4 ,0000
Step	4,755 1 ,0292

Classification Table for ERFOLG .

The Cut Value is ,50

Observed	Predicted			Percent Correct
	Heilung Abstoss			
	H	I	A	
Heilung	H	I	20 I 7 I	74,07%
Abstoss	A	I	8 I 54 I	87,10%
Overall				83,15%

----- Variables in the Equation -----

Variable	B	S.E.	Wald	df	Sig	R	Exp(B)
INDEX			17,5189	2	,0002	,3518	
INDEX(1)	-2,5324	,6092	17,2799	1	,0000	-,3740	,0795
INDEX(2)	,0415	,5759	,0052	1	,9426	,0000	1,0423
LEBER	-,0622	,0301	4,2864	1	,0384	-,1447	,9396
BLUT	-,0155	,0058	7,2473	1	,0071	-,2192	,9846
Constant	4,7298	1,3354	12,5446	1	,0004		

----- Model if Term Removed -----

Term	Log Likelihood	-2 Log LR	df	Significance of Log LR
INDEX	-43,924	25,088	2	,0000
LEBER	-33,758	4,755	1	,0292
BLUT	-35,790	8,820	1	,0030

----- Variables not in the Equation -----

Variable	Score	df	Sig.	R
DIAGNOSE(1)	,1576	1	,6914	,0000
URIN	,8246	1	,3639	,0000

No more variables can be deleted or added .

Zu Beginn eines Schrittes wird mit dem MLR-Test (Zeile "Model ...") die Hypothese geprüft, daß die β -Koeffizienten für alle Variablen in der Gleichung Null sind. Die Verbesserung gegenüber dem vorausgehenden Schritt, d.h. die Hypothese, daß der Koeffizient nur der aufgenommenen Variablen Null ist, prüft der MLR-Test (Zeile "Step ..."). Im ersten Schritt also prüfen beide Tests dieselbe Hypothese. Es folgen die Klassifikationstabelle, die den Anteil richtig klassierter Fälle mit 81% ausweist, und danach die in Abschnitt 5.3 beschriebene Kennzeichnung der aktuellen Modellvariablen mit Bewertung durch den Wald-Test. Zu einer in einem Schritt aufgenommenen Variablen können zwischen dem χ^2 des Skore-Tests (in der Liste 'Variables not in the equation' des vorausgehenden Schritts), dem χ^2 des MLR-Tests (Step) und dem χ^2 des Wald-Tests in der Tabelle 'Variables in the Equation' erhebliche

Unterschiede auftreten. Dies wird z.B. bei der ersten ausgewählten Kovariaten *index* mit den Chi²-Werten 29,75 im Score-Test, 30,98 im MLR-Test (Step) und 22,06 im Wald-Test deutlich. Die Unterschiede verdeutlichen - bezogen auf den MLR-Test - die Ungenauigkeit des Wald-Tests. Diese Ungenauigkeit gilt im allgemeinen ebenso auch für den Skore-Test, mit dem die Variablenauswahl vorgenommen wird.

Das Prüfen der Kovariaten in der Tabelle 'Model if Term Removed' mit dem angeforderten MLR-Test wird erst interessant, wenn mehr als eine Variable in der Modellgleichung stehen.

In der Tabelle 'Variables not in the Equation' wird zu jeder dieser Variablen mit dem Skore-Test geprüft, ob ihr Aufnahme in die Modellgleichung eine Verbesserung der Vorhersagegenauigkeit erbringt, und es wird die Kovariate mit kleinstem p-Wert, das ist hier *blut* mit $p = 0.0012$, in die Gleichung aufgenommen. Am p-Wert von 0.5164 zur Variablen *diagnose* wird deutlich, daß ihr ursprünglich starker Einfluß ($p = 0.0006$) nahezu vollständig durch die Variable *index* ausgeschöpft wird.

Die weiteren analogen Ergebnisse im 3. Schritt zeigen, daß das Hineinnehmen der Variablen *blut* und *leber* in die Modellgleichung die Klassifikationssicherheit - trotz Signifikanz - nur von 81% auf 83% erhöht. Das Bestimmtheitsmaß nach Nagelkerke hat den Wert $R^2 = 0,575$.

5.6. Optionen

Das Anklicken von **Optionen** im Dialogfeld der Abb. 5.1 unten rechts blendet das Dialogfeld der **Abb. 5.6** ein.

Im unteren Feldbereich können die Kriterien der Analyse, insbesondere die Grenzen der p-Werte für Aufnahme und Ausschluß von Variablen eingesehen und geändert werden. Der p-Wert für den Ausschluß soll stets größer als der für den Einschluß gesetzt werden. Bei einem großen Stichprobenumfang sind zwei kleinere p-Werte wie 0.01 und 0.02 vorzuziehen, da sonst auch eine fast bedeutungslose Variable einen statistisch signifikanten Einfluß aufweisen und damit die Anzahl von Variablen in der Modellgleichung zu groß werden kann.

Abb. 5.6 Dialogfeld 'Optionen'

In der Zeile über den p-Werten besteht zu einer schrittweisen Analyse die Möglichkeit zu wählen, ob Zwischenergebnisse nach jedem Schritt oder nur zusammenfassende Ergebnisse nach dem letzten Schritt ausgegeben werden sollen. Solch eine Zusammenfassung zu den schrittweisen Ergebnissen der Abb. 5.5 zeigt die abschließende **Abb. 5.7**. Sie enthält Angaben nur zu denjenigen Kovariaten, die in den einzelnen Schritten für das Modell ausgewählt wurden; insbesondere erfährt man nicht, wie groß der Einfluß jeder einzelnen Kovariaten, für sich betrachtet, ausfällt.

Das Risiko der Fehlklassierung (Fehlerrate) läßt sich durch Veränderung des **Cut Values** (Klassifikationsschwellenwert) zu Gunsten des einen oder des anderen Ergebniswertes steuern. In Abb. 5.6 ist er auf 0.35 gesetzt. Das führt in der Klassifikationstabelle der Abb. 5.7 für die Abstoßung zu einer Fehlerrate von 6% ($100\% - 93,55\% = 6,45\%$) statt der 13% in Abb. 5.5. Gleichzeitig erhöht sich die Fehlerrate für die Heilung. Siehe dazu auch den Abschnitt 5.8 zur Kreuzvalidierung.

Hilfreich kann es sein, **Klassifikationsdiagramme** anzufordern (im Dialogfeld oben). Solch ein Diagramm zeigt für jeden Fall die Zuordnungswahrscheinlichkeit zu der mit 1 codierten Gruppe, hier die Patienten mit Abstoßung, siehe die Ergebnisliste der **Abb. 5.7**. Mit dem Schwellenwert von 0.5 bzw. 0.35 auf der Abszisse kann der Leser das Ergebnis der Klassifikationstabelle in Abb. 5.5, S. 28, bzw. Abb. 5.7 selbst leicht nachvollziehen. Durch Verschieben dieses Schwellenwertes lassen sich andere Risikoverteilungen durchspielen. Man kann auch einen **Unsicherheitsbereich** definieren, z.B. in den Grenzen von 0.25 bis 0.75, der weitere Untersuchungen für die in diesen Bereich fallenden Patienten erfordert, und fällt für Patienten in den Randbereichen Entscheidungen mit hoher Zuverlässigkeit. Mitunter ist es wünschenswert zu wissen, welche der Zuordnungswahrscheinlichkeiten, die im Klassifikationsdiagramm abzulesen sind, zu welchen Patienten gehören, d.h. wie sicher, unsicher oder falsch der einzelne Fall einer Gruppe zugeordnet wird. Man erfährt dies, wenn man die fallweise Auflistung der Residuen anfordert.

Bei einer schrittweisen Analyse werden die Residuen am Ende der Ergebnisliste aufgeführt, Klassifikationstabelle und, falls angefordert,

Klassifikationsdiagramm zählen zu den Zwischenergebnissen nach jedem Schritt.

Abbildung 5.7 enthält darüber hinaus die Tabelle der Korrelationen, die zwischen den einzelnen Parameterschätzungen bestehen. Auch diese Tabelle, falls angefordert, zählt zum Bestandteil der Zwischenergebnisse.

Beginning Block Number 1. Method: Forward Stepwise (LR)

Step	Improv.			Model			Correct	
	Chi-Sq.	df	sig	Chi-Sq.	df	sig	Class	% Variable
1	30,980	2	,000	30,980	2	,000	80,90	IN: INDEX
2	10,742	1	,001	41,722	3	,000	78,65	IN: BLUT
3	4,755	1	,029	46,478	4	,000	79,78	IN: LEBER

No more variables can be deleted or added.

-2 Log Likelihood 62,760

Goodness of Fit 61,401

	Chi-Square	df	Significance	
Model	46,478	3	,0000	Abb. 5.7. Ausgabe zur schrittweisen
Block	46,478	4	,0000	Analyse ohne Zwischen-
Step	4,755	1	,0292	ergebnisse

Classification Table for ERFOLG

The Cut Value is ,35

Observed	Predicted						Percent Correct
	Heilung			Abstoss			
	H	I	A	H	I	A	
Heilung	H	I	13	I	14	I	48,15%
Abstoss	A	I	4	I	58	I	93,55%
Overall							79,78%

Variables in the Equation							
Variable	B	S.E.	Wald	df	Sig	R	Exp(B)
INDEX			17,5189	2	,0002	,3518	
INDEX(1)	-2,5324	,6092	17,2799	1	,0000	-,3740	,0795
INDEX(2)	,0415	,5759	,0052	1	,9426	,0000	1,0423
LEBER	-,0622	,0301	4,2864	1	,0384	-,1447	,9396
BLUT	-,0155	,0058	7,2473	1	,0071	-,2192	,9846
Constant	4,7298	1,3354	12,5446	1	,0004		

Correlation Matrix:

	Constant	INDEX(1)	INDEX(2)	LEBER	BLUT,
Constant	1,00000	-,36989	-,01261	-,70099	-,72737
INDEX(1)	-,36989	1,00000	-,13385	,25088	,23187
INDEX(2)	-,01261	-,13385	1,00000	,21728	-,16802
LEBER	-,70099	,25088	,21728	1,00000	,08496
BLUT	-,72737	,23187	-,16802	,08496	1,00000

Observed Groups and Predicted Probabilities

16 +	A+
I	AI

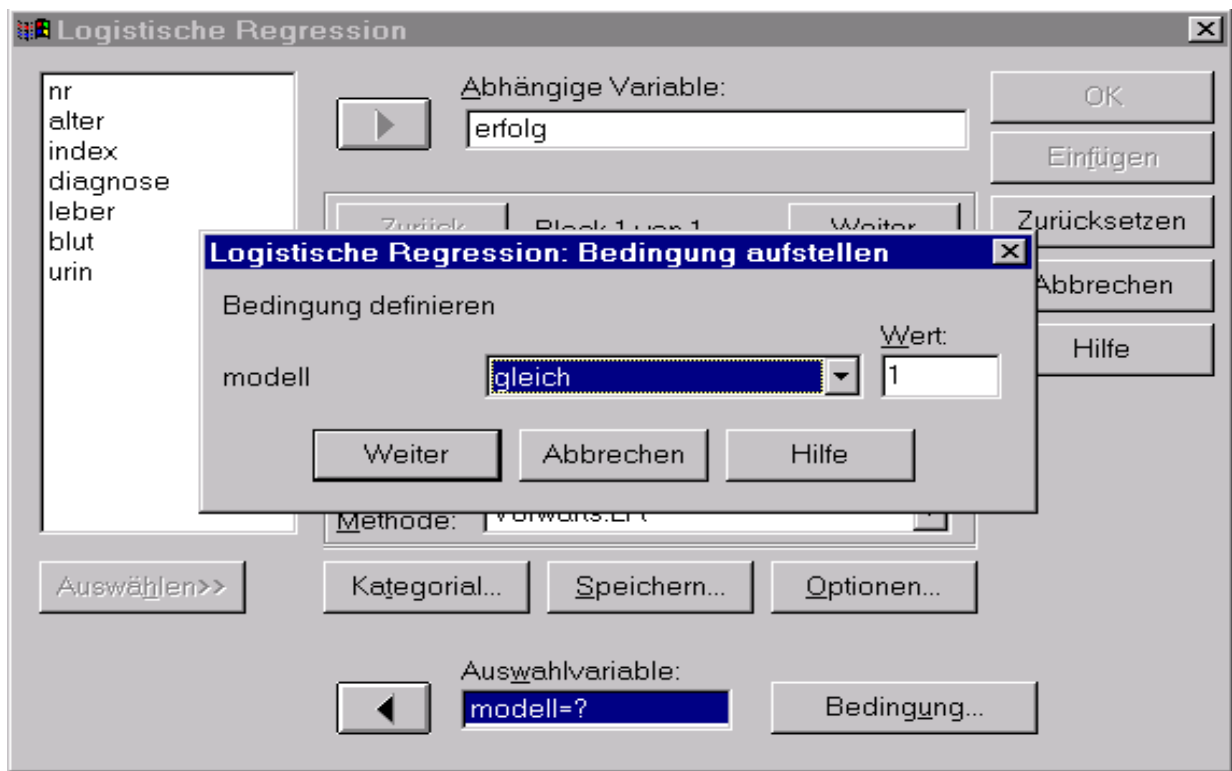


Abb. 5.8. Modellbildung mit ausgewählten Fällen

Die Programmdurchführung liefert neben der üblichen Ergebnisliste die Zuordnungswahrscheinlichkeiten und die vorhergesagten Gruppenzugehörigkeiten zu **allen** Fällen. Sie sind unter den Variablenamen *pre_1* und *pgr_1* in der Datendatei zu finden. Die Klassierung hängt dabei vom vorgegebenen Schwellenwert (Cut Value) ab. Die Rechnung kann für verschiedene Schwellenwerte wiederholt und die Gruppenzugehörigkeit jeweils unter einer neuen Variablen (*pgr_2*, *pgr_3*...) *abgespeichert* werden.

Andere Zuordnungen: Durch eine Rekodierung der Variablen *pre_1* kann die Gruppenzugehörigkeit und damit das Risiko einer Fehlklassierung im nachhinein auch anders definiert werden; z.B. ($p \leq 0.3$) = 1, ($0.3 < p < 0.7$) = 2, ($p \geq 0.7$) = 3 weist die Codes 1 und 3 als 'sichere' Klassierung und Code 2 als unsichere Klassierung aus.

Simulationen: Verwendet man als neue Fälle spezielle Wertekombinationen der Kovariaten (Simulationen) und ihre gezielten Veränderungen, lassen sich die Auswirkungen solcher Veränderungen auf die Zuordnungswahrscheinlichkeiten unmittelbar ablesen.

5.8. Kreuzvalidierung

Die Lernstichprobe (y_i, \mathbf{x}_i) , $i=1,2,\dots,n$, sei mit Z bezeichnet; darin sind y_i der Ergebniswert und \mathbf{x}_i eine Realisation des Vektors der Kovariaten. Aus Z wird das spezielle Funktionsmodell der logistischen Regression abgeleitet und werden die Koeffizienten geschätzt. Damit kann für ein \mathbf{x}_0 aus der Grundgesamtheit die Wahrscheinlichkeit dafür berechnet werden, daß das Ergebnis $Y = 1$ ist. Zusammen mit dem Schwellenwert s kann so eine Klassifikations- oder **Vorhersageregeln** $\eta = \eta_{Z,s}$ festgelegt werden, die für dieses \mathbf{x}_0 das Ergebnis $Y = 1$ vorhersagt, wenn die berechnete Wahrscheinlichkeit größer oder gleich dem Schwellenwert ist.

Als Maß für die Zuverlässigkeit der Vorhersageregeln und zugleich als Maß für die Gültigkeit des Modells (Validierung) wird die **Fehlerrate** $\text{Err}(Z,s)$ verwendet, oder auch die **Sensitivität** als Anteil der richtigen Vorhersagen für das Ergebnis $Y = 1$ und die **Spezifität** als Anteil der richtigen Vorhersagen für das Ergebnis $Y = 0$. Fehlerrate, Sensitivität und Spezifität sind nicht bekannt und müssen geschätzt werden.

In den Klassifikationstabellen der Abb. 5.1, 5.5 und 5.7 wird der Anteil $R = R(Z,s)$ der insgesamt (overall) richtig klassierten Fälle der Lernstichprobe Z in % angegeben. Entsprechend gilt für den Anteil err der falsch klassierten Fälle:

$$\text{err} = \text{err}(Z,s) = 100 - R (\%).$$

Diese Fehlerrate bezeichnet man jedoch als **scheinbare Fehlerrate**; denn als Schätzwert für die wahre Fehlerrate, die sich bei Anwendung auf Teststichproben ergibt, ist err zu optimistisch, da Z zuvor auch schon für die Erzeugung der Vorhersageregeln η verwandt wurde.

Den Bedarf nach einem verzerrungsfreien Schätzwert für Err zumindest bei nicht sehr großen Stichproben erfüllt die **Kreuzvalidierung**. Dazu wird zu jedem einzelnen Element (y_i, \mathbf{x}_i) der Lernstichprobe Z eine Lernstichprobe $Z_{(i)}$ gebildet, die bis auf (y_i, \mathbf{x}_i) alle Elemente von Z enthält (Methode des 'leaving one out'). Die Koeffizienten der logistischen Regression werden dann anhand dieser reduzierten Lernstichprobe $Z_{(i)}$ bestimmt. Damit kann für jedes Element (y_i, \mathbf{x}_i) eine eigene Vorhersageregeln $\eta_{Z_{(i)},s}$ gebildet und auf dieses Element selber angewendet werden, ohne daß diese Vorhersageregeln vom Element (y_i, \mathbf{x}_i) selbst beeinflusst wird. Die in der Weise zu Z ermittelte Klassifikationstabelle (Kreuzvalidierung) führt auf einen verzerrungsfreien Schätzwert für Err .

Diese Methode des 'leaving one out' stellt einen Spezialfall der Kreuzvalidierung dar; der allgemeine Fall ist der des 'leaving k out'.

SPSS kann den Schritt der Kreuzvalidierung nicht ausführen, jedoch kann man anfordern, daß zu jedem Fall i mit den Daten (y_i, \mathbf{x}_i) die Regressions-

koeffizienten anhand von $Z_{(i)}$ bestimmt und in die Datendatei geschrieben werden. Daraus kann man zu jedem Fall die Zuordnungswahrscheinlichkeiten zum Ergebnis (0 oder 1) berechnen, mit dem Schwellenwert s die Zuordnung festlegen, dann in einer Kreuztabelle dem tatsächlichen Ergebniswert y gegenüberstellen und daraus die Fehlerrate bestimmen. Dies soll mit dem obigen Datenbeispiel vorgeführt werden. Im Dialogfeld der Abb. 5.1 klickt man 'Speichern' an und erhält folgendes Dialogfeld:



Abb. 5.9. Zum Speichern der fallweise berechneten Koeffizienten

Mit **DfBeta** wird zu jeder Variablen des letzten ausgewählten Modells und zur Konstanten ein neuer Schätzwert des Regressionskoeffizienten berechnet, und zwar für den Fall i der Schätzwert, den man erhält, wenn man diesen Fall aus der Lernstichprobe Z fortläßt. Genauer: es werden zur Liste der Regressionskoeffizienten und der Konstanten neue Variablen in der Datei erzeugt, die für jeden Fall die **Abweichung** der Regressionskoeffizienten gegenüber der Ausgangsschätzung enthalten.

So kann für jeden Fall mit diesen Parameterwerten und den Variablenwerten die Wahrscheinlichkeit für einen Erfolg berechnet und daraus die Zuordnung zu Erfolg oder Mißerfolg bestimmt werden.

Im vorliegenden Beispiel verwendet SPSS die Variablennamen $dfb0_1$ für die Konstante, $dfb1_1$ zu $index(1)$, $dfb2_1$ zu $index(2)$, $dfb3_1$ zu $leber$ und $dfb4_1$ zu $blut$. Der neue Schätzwert ist gleich dem alten Schätzwert minus der Abweichung. Dafür vergeben wir die Variablennamen $Kval0, \dots, Kval4$. Mit der Variablen $prob$ werden die Zuordnungswahrscheinlichkeiten zum Ergebnis '1' berechnet und in $probkl$ die Zuordnungsclassen '0' oder '1' zum Schwellenwert von 0,5 festgelegt. Für die Berechnung von $prob$ sind zuvor die zur Kovariaten $index$ gehörigen zwei Designvariablen $index1$ und $index2$ in die Datendatei einzufügen und mit den Kodierungen, wie sie aus der Ergebnisliste von Abb. 5.5 entnommen werden können, zu versehen. z ist eine Zwischengröße.

```

COMPUTE kval0 = 4.7298 - dfb0_1.
COMPUTE kval1 = -2.5324 - dfb1_1.
COMPUTE kval2 = 0.0415 - dfb2_1.
COMPUTE kval3 = -0.0622 - dfb3_1.
COMPUTE kval4 = -0.0155 - dfb4_1.

COMPUTE index1 = index .
COMPUTE index2 = index .
RECODE index1 (1=-0.707) (2=0) (3=0.707).
RECODE index2 (1=0.408) (2=-0.816) (3=0.408).

COMPUTE z = kval0 + kval1*index1 + kval2*index2 + kval3*leber
           + kval4*blut.
COMPUTE prob = EXP(z)/(1 + EXP(z)).
COMPUTE probkl = prob.
RECODE probkl (Lowest thru 0.499999=0) (0.500 thru Highest=1).

CROSSTABS /TABLES = ERFOLG BY PROBKL ...

```

ERFOLG * PROBKL Kreuztabelle

Anzahl		PROBKL		Gesamt
		,00	1,00	
ERFOLG	Heilung	18	9	27
	Abstoss	8	54	62
Gesamt		26	63	89

Fehlerrate für Heilung 33,3 %,
für Abstoßung 12,9 %,
insgesamt 19,1 %.

Abb. 5.10 Rechenschritte und Ergebnis der Kreuzvalidierung

Von 89 Fällen insgesamt 17 Fehlklassifikationen, das bedeutet für die wahre Fehlerrate einen Schätzwert von

$$\mathbf{Err(Z, 0,5) = 19 \%}$$

gegenüber der scheinbaren Fehlerrate von $\mathbf{err(Z, 0,5) = 17 \%}$.

Die Größe der Differenz dieser beiden Schätzwerte (der Optimismus der scheinbaren Fehlerrate) nimmt ab mit wachsendem Stichprobenumfang. Ebenfalls mit der Kreuzvalidierung erhält man eine verzerrungsfreie Schätzung der Sensitivität - hier von $54/62 = 87 \%$ - und der Spezifität - hier von $18/27 = 67 \%$.