SPEZIFIKATIONSTESTS VOM HAUSMANTYP FÜR ÜBERGANGSRATENMODELLE

Dr. Karl AUINGER

Forschungsbericht/
Research Memorandum No. 243

März 1988

Die in diesem Forschungsbericht getroffenen Aussagen liegen im Verantwortungsbereich des Autors und sollen daher nicht als Aussagen des Instituts für Höhere Studien wiedergegeben werden.

INHALT

Zusammenfassung. In dieser Arbeit werden verschiedene Spezifikationstests für das einfache Übergangsratenmodell mit konstanter Hazardrate konstruiert.
Da für Überlebensverteilungen zensierte Daten die Regel sind, wird auf diese Problematik von vorneherein bezug genommen. Es werden die folgenden Zensierungsmuster betrachtet:
(i) single type I-censoring
(ii) type II-censoring
(iii) random-censoring
Für die Fälle (i) und (iii) wird folgende Vorgangsweise gewählt: unter Verwendung einer Funktionalgleichung der Überlebensfunktion ($S(x+y)=S(x)S(y)$) wird für feste x,y der Wert $S(x+y)$ auf zwei verschiedene Arten geschätzt, sodaß die (standardisierte) Differenz asymptotisch normalverteilt ist. Mittels einer konsistenten Schätzung der asymptotischen Varianz werden dann Teststatistiken abgeleitet, die unter der Nullhypothese asymptotisch chi^2-verteilt sind.
Für den Fall (ii) werden mittels Order-Statistiken verschiedene Quantile geschätzt. Hier ist der Quotient zweier derartiger Order-Statistiken (nach geeigneter Standardisierung) asymptotisch normalverteilt mit bekannter Varianz.
Für den Fall (i) wird schließlich noch eine weitere Methode vorgeschlagen: es werden die Maximum-Likelihood Schätzungen des Skalenparameters - bezüglich der gegebenen Zensierungszeit C und einer "künstlichen" Zensierungszeit $C'<C$ gebildet. Ihre standardisierte Differenz ist wieder asymptotisch normalverteilt.
Für alle diese Varianten wird die Menge der Alternativen,gegen die der jeweilige Test konsistent ist, bestimmt.

Summary. In this paper we construct different tests against misspecification for the basic failure time model with a constant hazard function.
Throughout the paper it is taken into account that lifetime data are usually censored. We consider the following censoring mechanisms:
(i) single type I-censoring
(ii) type II-censoring
(iii) random censoring
Using a functional equation of the survival function ($(S(x+y)=S(x)(y))$), we treat the cases (i) and (iii) by the following method: for fixed x,y we estimate the value $S(x+y)$ in two different ways such that the (standardized) difference between these estimates is asymptotically normal. Using a consistent estimate of the asymptotic variance, we derive test-statistics that are - under the null-hypothesis - asymptotically chi-square.
For the case (ii) we use order statistics to estimate different quantiles. Then the ratio of two of these estimates is (when appropriately standardized) asymptotically normal with known variance.
Finally, for the case (i) we propose a further method: estimating the scale parameter by Maximum-Likelihood method once using the given censoring time C, once using an "artificial" censoring time $C'<C$, we obtain two estimates whose standardized difference is asymptotically normal.
For all these possibilities we determine the sets of alternatives against which the tests are consistent.

I.Einleitung

1.Grundbegriffe der Survivalanalyse

Es sollen hier kurz die mathematisch-statistischen Grundlagen der Survivalanalyse skizziert werden.
Survivalanalyse untersucht - ganz allgemein gesprochen - den Übergang von Individuen von einem Stadium (Zustand) in ein anderes, die Dauer des Verweilens in den einzelnen Stadien und die Wahrscheinlichkeit eines Überganges. Wir wollen uns auf den Spezialfall des sogenannten Zwei-Zustands-Modells mit absorbierendem Zielzustand beschränken. Hier befinden sich die interessierenden Objekte zu Beginn der Untersuchung in einem Ausgangszustand, um im Lauf der Zeit in einen anderen Zustand überzugehen, der vom Standpunkt der jeweiligen Untersuchung aus als endgültiger ("Tod") aufgefaßt wird.
Der Zeitpunkt T des Eintritts des Ereignisses - auch Ankunftszeit oder Überlebenszeit genannt - wird als nichtnegative zufällige Veränderliche aufgefaßt. Ganz allgemein besteht die Aufgabe der Survivalanalyse darin, aus dem gegebenen Datenmaterial Aufschlüsse über die Wahrscheinlichkeitsverteilung der zufälligen Variablen T zu gewinnen.
Angenommen, diese Wahrscheinlichkeitsverteilung besitzt eine Dichtefunktion f, dann ist

$$F(t) = P(T \leq t) = \int_0^t f(s)ds$$

die (kumulierte) Verteilungsfunktion der Überlebenszeiten. F(t) gibt für ein Individuum die Wahrscheinlichkeit dafür an, daß sein "Tod" innerhalb des Zeitintervalls [0,t] eintritt. Umgekehrt gibt

die "Überlebensfunktion" S:

$$S(t) = P(T>t) = 1 - F(t) = \int_t^{\infty} f(s)ds$$

die Wahrscheinlichkeit an, mit der ein Individuum den Zeitpunkt t "überlebt",d.h. das besagte Ereignis bis zum Zeitpunkt t noch nicht eingetreten ist.

Die Hazardfunktion r ist definiert als

$$r(t) := \lim_{h->0+} q(t,t+h)/h = \lim_{h->0+} P(t<T\leq t+h|T>t)/h$$

wobei q(t,t+h) die Übergangswahrscheinlichkeit: die Wahrscheinlichkeit des Eintritts des Ereignisses im Intervall (t,t+h], unter der Bedingung, daß es bis zum Zeitpunkt t noch nicht eingetreten ist, ausdrückt. r(t) heißt dann die Hazardrate oder Übergangsrate zum Zeitpunkt t.(Bei Mehrzustandsmodellen sind die Begriffe Hazardrate und Übergangsrate nicht identisch.) Wir wollen voraussetzen, daß die Hazardrate zu jedem Zeitpunkt t>0 definiert ist.

Die Hazardrate ist keine Wahrscheinlichkeit (es können Werte >1 vorkommen). Wohl aber sind für "sehr kleine" Zeitintervalle der Länge h die Hazardrate und die Übergangswahrscheinlichkeit ungefähr proportional:

$$q(t,t+h) \approx r(t)\cdot h$$

Anschaulich heißt dies: je höher die Hazardrate zu einem Zeitpunkt t ist, desto größer ist für ein Individuum das Risiko, zu diesem Zeitpunkt t zu "sterben"; damit ist in der Hazardrate quasi die Intensität des Übergangsprozesses ausgedrückt. In der Hazardfunktion r(.) ist festgehalten, wie sich diese "momentane Todesrate" während der Zeit ändert.

Es gelten folgende Beziehungen:

$$r(t) = f(t)/S(t) = -(dS(t)/dt)/S(t) = -d(\log S(t))/dt$$

$$S(t) = \exp(-\int_0^t r(s)ds)$$

$$f(t) = r(t)\cdot\exp(-\int_0^t r(s)ds)$$

Damit besteht eine bijektive Zuordnung zwischen Ratenfunktion und Wahrscheinlichkeitsverteilung der Ankunftszeiten: erstere ist durch letztere eindeutig bestimmt und umgekehrt.

Die einfachste Form einer Hazardfunktion ist offenbar:

$$r(t) \equiv \alpha > 0$$

Das heißt: die Hazardrate ändert sich nicht während der Zeit. Dem entspricht, daß die Ankunftszeiten exponentialverteilt sind, die Überlebensfunktion also gegeben ist durch

$$S(t) \equiv \exp(-\alpha t)$$

und die Dichtefunktion gegeben ist durch

$$f(t) = \alpha\cdot\exp(-\alpha t).$$

Diese Arbeit stellt sich die Aufgabe, mittels der *Hausman'schen Methode* Tests für die Hypothese

$$H_0: r(t) \text{ ist konstant}$$

oder äquvalent dazu:

$$H_0: S(t) \equiv \exp(-\alpha t)$$

für *zensierte Daten* zu konstruieren.

Ein Überblick über andere Spezifikationstest für Survivalmodelle findet sich z.B. im Buch von Lawless (1982). Weitere Bücher über Survivalanalyse: z.B. Kalbfleisch & Prentice (1980) oder Diekmann & Mitter (1984) und die dort angegebene Literatur.

2.Zensierte Daten

Von zensierten Daten spricht man, wenn die - oder einige der - Realisierungen der zu untersuchenden zufälligen Variablen T_i nicht exakt bekannt, sondern nur obere oder untere Schranken derselben bekannt sind. Da die Realisierungen der Variablen T_i beobachtete Zeiten sind, ist nur der Fall interessant, daß sogenannte Rechtszensierung vorliegt, d.h. daß die - oder einige der - bekannten Daten untere Schranken der tatsächlichen Ankunftszeiten sind. Wir wollen folgende drei Fälle unterscheiden:

i) Single-Type-I-Censoring:

Hier wird die Dauer der Beobachtung (z.B. einer empirischen Untersuchung) im voraus festgelegt, dagegen die Anzahl der tatsächlich beobachteten Ereignisse offengelassen - diese ist dann selbst eine zufällige Variable. Ist C die festgelegte Beobachtungsdauer, dann werden anstatt der N exakten Ankunftszeiten $T_1,\ldots T_N$ Realisierungen folgender Variablen beobachtet:

$$X_i := \min(T_i,C) \text{ und } d_i := I\{T_i \leq C\}$$

Dabei soll $I\{T_i \leq C\}$ die Indikatorfunktion der Menge $\{ :T_i( ) \leq C\}$ bezeichnen. Die zur Verfügung stehenden Daten bestehen dann aus der Menge $\{(X_1( ),d_1( )),\ldots(X_N( ),d_N( ))\}$ von Realisierungen der Paare (X_i,d_i).

ii) Type-II-Censoring:

Hier werden in der Studie N Individuen einer Beobachtung unterzogen, und es wird im voraus festgelegt, die Beobachtung abzubrechen, nachdem das r-te Eintreffen des besagten Ereignisses beobachtet werden konnte ($1<r<N$). Die Anzahl der beobachteten Ereignisse ist also fix, während die Dauer der Studie nicht festgelegt ist.

iii) Random-Censoring:

Im Unterschied zu i) ist hier die *Dauer der Beobachtung* des i-ten Individuums keine Konstante, sondern ebenfalls Realisierung einer zufälligen Variablen: der Zensierungsvariablen C_i. Ähnlich wie unter i) läßt sich hier die Datenstruktur als Realisierung einer Menge von Paaren von Zufallsvariablen (X_i,d_i) darstellen, wobei:

$$X_i := \min(T_i,C_i) \text{ und } d_i := I\{T_i \leq C_i\}$$

bedeutet. Im Unterschied zu i) sind hier die Zensierungszeiten C_i (Realisierungen von) zufälligen Variablen, was die Behandlung dieses Modells erheblich kompliziert macht.

3.Hausman-Tests

Hausman hat in seiner berühmten Arbeit (1978) eine allgemeine Methode vorgeschlagen, Tests auf Fehlspezifikation zu konstruieren: dazu wähle man zwei verschiedene Schätzungen $\hat{\beta}$ und $\tilde{\beta}$, die unter der Nullhypothese konsistent für einen Parameter β des vermuteten Modells sind. Ferner bestimme man die Grenzverteilung von $\sqrt{N}(\hat{\beta}-\tilde{\beta})$. Ist diese eine Normalverteilung mit Mittelwert 0 und Varianz-Kovarianz-Matrix V, dann ist die Statistik

$$H := N(\hat{\beta}-\tilde{\beta})'V^{-1}(\hat{\beta}-\tilde{\beta})$$

asymptotisch chi^2-verteilt mit einer bekannten Anzahl von Freiheitsgraden (meistens dim β). Und dieses asymptotische Verhalten gilt auch dann, wenn statt der (meist unbekannten) Varianz-Kovarianz-Matrix V eine konsistente Schätzung $\hat{V}$ derselben eingesetzt wird. Ferner ist der Test gegen alle Alternativen, für die $\hat{\beta}$ und $\tilde{\beta}$ in Wahrscheinlichkeit gegen verschiedene Grenzwerte konvergieren konsistent.

II. Hausman-Tests auf konstante Hazardfunktion

1.Typ I zensierte Daten

Wir wollen nur den einfachsten Fall, nämlich "single-type-I-censoring" behandeln.
Formalisiert kann dies auf folgende Weise werden: Seien $T_1,...T_N$ unabhängige, identisch verteilte Zufallsvariablen und C eine fix vorgegebene Konstante. Beobachtet werden dann nicht Realisierungen der Variablen T_i, sondern Realisierungen der Variablen

$$X_i := \min(T_i,C) \text{ und } d_i := I\{T_i \leq C\}$$

Die Nullhypothese lautet:

$$H_0: T_i \text{ ist exponentialverteilt}$$

oder äquivalent dazu:

$$r(t) \equiv \alpha$$

Festzuhalten ist noch, daß, wenn nach dem Zeitpunkt C überhaupt keine Beobachtungen mehr gemacht werden können, Verteilungs-, Überlebens-, oder Ratenfunktion nur im Intervall [0,C] geschätzt und getestet werden können. Es könnte dann sein, daß z.B. der tatsächlich zugrundeliegende Prozeß im Intervall [0,C] eine konstante Hazardrate hätte, die erst außerhalb dieses Intervalls eine andere funktionale Form annimmt. Eine derartige Fehlspezifikation kann natürlich bei derartig zensierten Daten durch keinen Test entdeckt werden. Die zu testende Nullhypothese lautet daher exakter:

H_0: T_i hat im Intervall [0,C] eine konstante Hazardfunktion

oder äquivalent dazu:

$$S(t) \equiv \exp(-\alpha t) \text{ auf } [0,C]$$

1.1. Tests mittels Schätzung der Überlebensfunktion durch die empirische Überlebensfunktion

Eine erste Möglichkeit mittels des Hausman-Prinzips einen Spezifikationstest zu konstruieren benützt die Tatsache, daß eine Variable T_i genau dann exponentialverteilt ist, wenn ihre Überlebensfunktion der Funktionalgleichung

$$S(x)S(y) = S(x+y)$$

genügt. Man kann daher $x,y \in (0,C)$ mit $x+y \leq C$ wählen, die Überlebensfunktion an den Stellen x,y und x+y schätzen; trifft die Nullhypothese zu, dann darf die Differenz $|\hat{S}(x+y)-\hat{S}(x)\hat{S}(y)|$ nicht "zu groß" werden.

Exakte Ausführung: Wähle $x,y \in (0,C)$, o.B.d.A. sei $x \leq y$, mit $z := x+y \leq C$; N bezeichne generell die Samplegröße. Für $u \in \{x,y,z\}$ bezeichne N-h(u) die Anzahl der vor dem Zeitpunkt u gemachten Beobachtungen. Dann ist die empirische Überlebensfunktion $\hat{S}$:

$$\hat{S}(u) := h(u)/N$$

eine konsistente, asymptotisch normalverteilte Schätzung für S(u). Seien T_i die unabhängigen, identisch verteilten Variablen der Ankunftszeiten; weiters seien

$$X_i := I\{T_i>x\} \ , \ Y_i := I\{T_i>y\} \ , \ Z_i := I\{T_i>z\}$$

wobei wir für den Fall z = C vereinbaren wollen, daß $T_i>z$ genau dann gilt, wenn $d_i=0$, d.h., wenn es sich um eine zensierte Beobachtung handelt. (Wenn die i-te Beobachtung zensiert wurde, kann angenommen werden, daß die eigentliche i-te Ankunftszeit zumindest um eine sehr kleine Zeitspanne größer ist.)

Die natürlichen Schätzungen für die Überlebensfunktion sind gegeben durch

$$\hat{S}(x) := N^{-1}\Sigma X_i \ , \ \hat{S}(y) := N^{-1}\Sigma Y_i \ , \ \hat{S}(z) := N^{-1}\Sigma Z_i$$

Wir verwenden nun den zentralen Grenzwertsatz in folgender Form:

Satz: Sei (Y_i) eine Folge von unabhängigen, identischverteilten (k-dimensionalen) zufälligen Vektoren mit Erwartungswert $EY_i = m$ und der Varianz-Kovarianz-Matrix $E(Y_i-m)(Y_i-m)' = T$. Dann konvergiert die Variable

$$N^{-\frac{1}{2}}\sum_{i=1}^{N}(Y_i-m)$$

in Verteilung gegen eine multivariate Normalverteilung $N(0,T)$ mit Mittelwert 0 und Varianz-Kovarianz-Matrix T.
(Ein Beweis findet sich z.B. in Anderson (1958),p.74).

Daraus ergibt sich für unsere Schätzungen: die Statistik

$$\sqrt{N}(\hat{S}(x)-S(x),\hat{S}(y)-S(y),\hat{S}(z)-S(z))$$

strebt in Verteilung gegen eine multivariate Normalverteilung $N(0,T)$ mit Mittelwert 0 und Varianz-Kovarianz-Matrix T, die gegeben ist durch

$$T = \mathrm{Cov}(X_i,Y_i,Z_i)$$

Wir müssen also die Kovarianzen $\mathrm{Cov}(U_i,V_i)$ mit $U_i,V_i\in\{X_i,Y_i,Z_i\}$ berechnen. Die U_i sind binomialverteilt mit

$$EU_i = S(u) \text{ und } \mathrm{Var}\ U_i = S(u)F(u)$$

wobei $F(u) = 1 - S(u)$ die Verteilungsfunktion bezeichnet. Weiters gilt:

$$X_iZ_i = Y_iZ_i = Z_i \text{ und } X_iY_i = Y_i$$

da $x\leq y\leq z$. Daher ist $\mathrm{Cov}(X_i,Y_i) = E(X_iY_i) - (EX_i)(EY_i) = EY_i - (EX_i)(EY_i) = S(y)F(x)$. Genauso erhält man $\mathrm{Cov}(X_i,Z_i) = S(z)F(x)$ und $\mathrm{Cov}(Y_i,Z_i) = S(z)F(y)$. Die Matrix T ist daher gegeben durch:

$$T = \begin{bmatrix} S(x)F(x) & S(y)F(x) & S(z)F(x) \\ S(y)F(x) & S(y)F(y) & S(z)F(y) \\ S(z)F(x) & S(z)F(y) & S(z)F(z) \end{bmatrix}$$

Außerdem verwenden wir folgend Satz aus der Theorie der Grenzverteilungen:

Satz: Seien $T_{1N},\ldots T_{kN}$ Statistiken derart, daß die Statistik

$$\sqrt{N}(T_{1N} - m_1,\ldots T_{kN} - m_k)$$

für $N \to \infty$ in Verteilung gegen eine multivariate Normalverteilung $N(0,T)$ konvergiert mit Mittelwert 0 und Varianz-Kovarianz-Matrix T. Weiters sei $g:R^k \to R^n$ eine Abbildung, deren (totale) Ableitung existiere. Dann konvergiert

$$\sqrt{N}(g(T_{1N},\ldots T_{kN}) - g(m_1,\ldots m_k))$$

in Verteilung gegen eine multivariate Normalverteilung $N(0,V)$ mit Mittelwert 0 und Varianz-Kovarianz-Matrix

$$V = G'TG$$

wobei die k x n -Matrix G gegeben ist durch

$$G = Dg(m_1,\ldots m_k).$$

(Ein Beweis ist z.B. in Rao (1965),Ch.6 zu finden.)

Sei nun $g:R^3 \to R$ definiert durch

$$g(u_1,u_2,u_3) = u_1u_2 - u_3.$$

Dann ist nach dem eben zitierten Satz die Statistik

$$\sqrt{N}(g(\hat{S}(x),\hat{S}(y),\hat{S}(z)) - g(S(x),S(y),S(z)))$$

asymptotisch normalverteilt mit Mittelwert 0 und Varianz $\sigma^2 = D'TD$ mit $D = Dg(S(x),S(y),S(z))$. Nun ist $Dg(u_1,u_2,u_3) = (u_2,u_1,-1)'$, daher ist $D' = (S(y),S(x),-1)$. Daraus kann sofort die Varianz σ^2 berechnet werden:

$$\begin{bmatrix} S(x)F(x) & S(y)F(x) & S(z)F(x) \\ S(y)F(x) & S(y)F(y) & S(z)F(y) \\ S(z)F(x) & S(z)F(y) & S(z)F(z) \end{bmatrix} * \begin{bmatrix} S(y) \\ S(x) \\ -1 \end{bmatrix} =$$

$$= \begin{bmatrix} 2S(x)S(y)F(x) - S(z)F(x) \\ S^2(y)F(x) + S(x)S(y)F(y) - S(z)F(y) \\ S(y)S(z)F(x) + S(x)S(z)F(y) - S(z)F(z) \end{bmatrix}$$

Daher ist

$$\sigma^2 = 2S(x)S^2(y)(1-S(x)) - S(y)S(z)(1-S(x)) +$$
$$+ S(x)S^2(y)(1-S(x)) + S^2(x)S(y)(1-S(y)) - S(x)S(z)(1-S(y)) -$$
$$- S(y)S(z)(1-S(x)) - S(x)S(z)(1-S(y)) + S(z)(1-S(z)) =$$
$$= -4S^2(x)S^2(y) + 4S(x)S(y)S(z) - S^2(z) + 3S(x)S^2(y) -$$
$$- 2S(y)S(z) - 2S(x)S(z) + S^2(x)S(y) + S(z).$$

Wenn die Nullhypothese zutrifft, dann gilt $S(z) = S(x)S(y)$, und wir erhalten:

$$\sigma^2 = -S^2(x)S^2(y) + S(x)S^2(y) - S^2(x)S(y) + S(z) =$$
$$= S(z)(1-S(x)S(y)) + S(x)S(y)(S(y)-S(x)) =$$
$$= S(x)S(y)(1-S(z)) + S(z)(S(y)-S(x)).$$

Unter H_0 ist $g(S(x),S(y),S(z)) = 0$, daher ist die Statistik

$$\sqrt{N}(\hat{S}(x)\hat{S}(y) - \hat{S}(z))$$

asymptotisch normalverteilt mit Mittelwert 0 und Varianz σ^2.

Meistens ist der wahre Wert von σ^2 unbekannt; ist $\hat{\sigma}^2$ eine konsistente Schätzung von σ^2, dann haben

$\sqrt{N}(\hat{S}(x)\hat{S}(y) - \hat{S}(z))/\sigma$ und $\sqrt{N}(\hat{S}(x)\hat{S}(y) - \hat{S}(z))/\hat{\sigma}$

die gleiche Grenzverteilung (wie wir später noch sehen werden).

Eine konsistente Schätzung für σ^2 erhält man, wenn man die in den obigen Formeln für σ^2 vorkommenden $S(u)$ durch $\hat{S}(u) = h(u)/N$ ersetzt. Man erhält damit folgende Teststatistiken:

$$H := \frac{\sqrt{N}(h(x)h(y)/N^2 - h(x+y)/N)}{\sqrt{(h(x)h(y)/N^2(1-h(x+y)/N) + h(x+y)(h(y)-h(x))/N^2)}}$$

welche sich noch vereinfachen läßt:

$$H := \sqrt{N}\frac{h(x)h(y)/N - h(x+y)}{\sqrt{(h(x)h(y)(1-h(x+y)/N) + h(x+y)(h(y)-h(x)))}}$$

Wählt man den andere Ausdruck für die Varianz, dann erhält man:

$$H := \frac{\sqrt{N}(h(x)h(y)/N - h(x+y))}{\sqrt{(h(x+y)/N\cdot(N^2-h(x)h(y))/N^2 + h(x+y)(h(y)-h(x)/N^2)}}$$

woraus sich einfacher folgendes ergibt:

$$H := \sqrt{N}\,\frac{h(x)h(y)/N - h(x+y)}{\sqrt{(h(x+y)(N-h(x)h(y)/N) + h(x+y)(h(y)-h(x)))}}$$

Diese Statistiken sind unter der Nullhypothese N(0,1) verteilt. Gilt $S(x)S(y) \neq S(x+y)$ für die gewählten x,y, dann strebt der Nenner von H stochastisch gegen einen endlichen Grenzwert, der Zähler strebt stochastisch gegen $\pm\infty$. Damit ist der beidseitige Test konsistent gegen die Menge aller Alternativen, für die $S(x)S(y) \neq S(x+y)$ gilt.

Eine weitere Möglichkeit, mit dieser Methode einen Test zu konstruieren, benützt die Tatsache, daß unter der Nullhypothese $S(rx) = S^r(x)$ gilt.

Vorgangsweise: man wähle ein $x \leq C$ und ein p mit $0<p<1$. Seien

$$X_i := I\{T_i>x\} \text{ und } Y_i := I\{T_i>px\}$$

Es ist $\hat{S}(x) = N^{-1}\Sigma Xi$ und $\hat{S}(px) = N^{-1}\Sigma Y_i$. Nach der oben zitierten Fassung des zentralen Grenzwertsatzes strebt die Statistik:

$$\sqrt{N}(\hat{S}(x) - S(x),\hat{S}(px) - S(px))$$

in Verteilung gegen eine bivariate Normalverteilung mit Mittelwert 0 und folgender Varianz-Kovarianz-Matrix T:

$\text{Var } X_i = S(x)F(x)$, $\text{Var } Y_i = S(px)F(px)$, $\text{Cov}(X_i,Y_i) = E(X_iY_i) - (EX_i)(EY_i) = EX_i - (EX_i)(EY_i) = S(x)(1-S(px)) = S(x)F(px)$.

Sei nun $g: R^2 \dashrightarrow R$ definiert durch

$$g(u,v) = u - v^{1/p}$$

dann gilt unter der Nullhypothese $g(S(x),S(px)) = 0$. Daher ist

$$\sqrt{N}(\hat{S}(x) - \hat{S}(px)^{1/p})$$

asymptotisch normalverteilt mit Mittelwert 0 und Varianz $\sigma^2 = D'TD$. D bezeichnet dabei die Ableitung von g an der Stelle $(S(x),S(px))$. Daher ist $D' = (1,-1/pS^{(1-p)/p}(px))$. Rechnet man TD

aus, so ergibt sich der Vektor

$$\begin{bmatrix} S(x)F(x) - 1/pS^{(1-p)/p}(px)S(x)F(px) \\ S(x)F(px) - 1/pS^{1/p}(px)F(px) \end{bmatrix}$$

Daher ist

$$D'TD = S(x)F(x) - 1/pS^{(1-p)/p}(px)S(x)F(px) -$$
$$- 1/pS^{(1-p)/p}(px)S(x)F(px) + 1/p^2S^{(2-p)/p}(px)F(px) =$$
$$= S(x) - S^2(x) - 2/pS^{(1-p)/p}(px)S(x) + 2/pS^{1/p}(px)S(x) +$$
$$+ 1/p^2S^{(2-p)/p}(px) - 1/p^2S^{2/p}(px).$$

Unter der Nullhypothese gilt $S(px) = S^p(x)$ und daher auch $S^{1/p}(px) = S(x)$. In diesem Fall gilt also:

$$\sigma^2 = S(x) - S^2(x) - 1/pS^{(1-p)/p}(px)S^{1/p}(px) + 2/pS^{1/p}(px)S^{1/p}(px)$$
$$+ 1/p^2S^{2-p}(x) - 1/p^2S^2(x) =$$
$$= S(x) - S^2(x) - 2/pS^{2-p}(x) + 2/pS^2(x) + 1/p^2S^{2-p}(x) -$$
$$- 1/p^2S^2(x) = S(x) - (1-1/p)^2S^2(x) - 1/p(2-1/p)S^{2-p}(x).$$

Nun ist nach denselben Überlegungen wie vorher die Statistik

$$\sqrt{N}(\hat{S}(x) - \hat{S}^{1/p}(px))/\hat{\sigma}$$

unter H_0 asymptotisch N(0,1)-verteilt, wenn $\hat{\sigma}$ eine konsistente Schätzung für die Streuung ist. Unter der Nullhypothese erhält man eine solche, wenn man in der obigen Formel S durch $\hat{S}$ ersetzt:

$$\hat{\sigma}^2 = h(x)/N - (1-1/p)^2h^2(x)/N^2 - 1/p(2-1/p)h^{2-p}(x)/N^{2-p}.$$

Wie vorher gilt: ist $S^p(x) \neq S(px)$, dann gilt $p \lim H = \pm\infty$, d.h. gegen jede derartige Alternative ist dieser Test konsistent. Für $p = \frac{1}{2}$ wird die Teststatistik am einfachsten: die Varianz wird zu

$$\sigma^2 = S(x) - S^2(x)$$

was wir auch erhalten hätten, wenn wir im vorher behandelten Fall $x = y$ gesetzt hätten. Als konsistente Schätzung für die Varianz σ^2 erhalten wir:

$$\hat{\sigma}^2 = h(x)/N - h^2(x)/N^2$$

Damit ergeben sich folgende Teststatistiken:

$$H := \frac{h(x) - h^2(x/2)/N}{\sqrt{(h(x) - h^2(x)/N)}}$$

welche unter H_0 asymptotisch normalverteilt ist mit Mittelwert 0 und Varianz 1 und

$$H^2 := \frac{(h(x) - h^2(x/2)/N)^2}{|h(x) - h^2(x)/N|}$$

welche unter H_0 asymptotisch chi^2-verteilt ist mit einem Freiheitsgrad.

Fehlspezifikationen, für die für die gewählten Werte von x,y resp. x,p $S(x+y) = S(x)S(y)$ resp. $S(px) = S^p(x)$ gilt, kann ein derartiger Test natürlich nicht entdecken. Es läßt sich aber immerhin eine wichtige Klasse von Alternativen angeben, gegen die die oben erwähnten Tests - unabhängig von der Wahl der x,y resp. x,p - konsistent sind.

Satz: Sei $0<x<C$; ist die Hazardfunktion r(t) auf dem Intervall (0,x) monoton wachsend (resp. monoton fallend) und nicht konstant (d.h. strenge Monotonie wird nicht benötigt), dann gilt für beliebige $u,v\in(0,x)$ mit $u+v=x$: $S(u)S(v)>S(u+v)$ (resp. $S(u)S(v)<S(u+v)$).

Beweis: Sei r(t) monoton wachsend und (o.B.d.A.) $v\leq u$. Es ist

$S(z) = \exp(-\int_0^z r(t)dt)$, daher ist $S(u+v)<S(u)S(v)$ genau dann,

wenn $$\int_0^{u+v} r(t)dt > \int_0^u r(t)dt + \int_0^v r(t)dt$$

gilt,d.h.,wenn

$$\int_u^{u+v} r(t)dt > \int_0^v r(t)dt$$

Nun ist $\int_0^v r(t)dt \leq r(v)\cdot v \leq r(u)\cdot v \leq \int_u^{u+v} r(t)dt$. Würde keine dieser Ungleichungen strikt gelten, dann wäre wegen der Monotonie von r:

$$r(t) \equiv r(u) \text{ für } t\in(0,u+v).$$

Ist r(t) monoton fallend, dann folgt die Behauptung aus der dualen Argumentation.

Insbesondere eignen sich die vorher abgeleiteten Teststatistiken für *einseitige* Tests gegen die Alternative

H_1: r(t) ist monoton wachsend (resp. fallend).

Unter H_1: "r(t) ist monoton wachsend" wäre dann p lim H = $-\infty$, unter H_1: "r(t) ist monoton fallend" wäre dann p lim H = $+\infty$,(oder genau umgekehrt, je nachdem, wie der Zähler in der Teststatistik gewählt wird).

Im folgenden sei noch ein Lemma zitiert, das wir verwendet haben (für einen Beweis siehe z.B. RAO (1965),p.122).

Lemma: Seien (X_n) und (Y_n) Folgen von zufälligen Variablen, sodaß X_n in Verteilung gegen eine Variable X, Y_n in Wahrscheinlichkeit gegen eine Konstante c strebt; dann gelten folgende Beziehungen:

(i) $X_n + Y_n$ strebt in Verteilung gegen X + c

(ii) ist $c \neq 0$, dann strebt X_nY_n in Verteilung gegen cX

ist c = 0, dann strebt X_nY_n in Wahrscheinlichkeit gegen 0.

Daraus folgt erstens etwas, was wir schon verwendet haben:

die Statistiken

$\sqrt{N}(\hat{S}(x)\hat{S}(y)-\hat{S}(z))/\sigma$ und $\sqrt{N}(\hat{S}(x)\hat{S}(y)-\hat{S}(z))/\hat{\sigma}$

respektive

$\sqrt{N}(\hat{S}(x)-\hat{S}^{1/p}(px))/\sigma$ und $\sqrt{N}(\hat{S}(x)-\hat{S}^{1/p}(px))/\hat{\sigma}$

haben asymptotisch die gleichen Verteilungen.
Außerdem läßt sich daraus ein Lemma ableiten, daß wir im folgenden brauchen werden:

Lemma: Sei (d_n) eine Folge von zufälligen Vektoren und A eine Matrix. (B_n) sei eine Folge von zufälligen Matrizen, die in Wahrscheinlichkeit gegen die Matrix A konvergiert (d.h.: es gilt p lim $b_{ij} = a_{ij}$ für alle i,j). Wenn die quadratische Form $d_n'Ad_n$ in Verteilung konvergiert, dann konvergiert $d_n'B_nd_n$ in Verteilung gegen dieselbe Verteilung wie die erste Form.

Beweis: Es gilt .
$d_n'B_nd_n = \Sigma b_{ij}d_id_j = \Sigma a_{ij}d_id_j + \Sigma(b_{ij}-a_{ij})d_id_j$.
Die zweite Summe konvergiert in Wahrscheinlichkeit gegen 0, daher haben $\Sigma b_{ij}d_id_j$ und $\Sigma a_{ij}d_id_j$ die gleiche Grenzverteilung.

Die oben konstruierten Tests beruhen auf folgendem Prinzip: man wähle die "Parameter" S(x),S(y),S(x+y),resp.S(x),S(px). Außerdem ist unter der Nullhypothese ein funktionaler Zusammenhang bekannt, der diese "Parametervektoren" auf 0 abbildet:

$$S(x)S(y) - S(x+y) = 0$$

respektive

$$S^p(x) - S(px) = 0.$$

Danach betrachtet man konsistente Schätzungen für diese "Parameter" und ersetzt in ebendieser Funktionsbeziehung die Parameter durch deren Schätzungen. Die daraus resultierende Statistik darf unter der Nullhypothese nicht "zu groß" werden - genauer: man berechnet die Grenzverteilungen dieser Statistiken und konstruiert daraus eine Teststatistik. Der so gewonnene Test ist dann

konsistent gegen jede Alternative, für die die Funktionsbeziehung zwischen den gewählten Parametern nicht erfüllt ist.
Um die Menge der Alternativen, gegen die der so zu konstruierende Test konsistent ist, zu erhöhen, kann man - insbesondere bei großen Stichproben - die Zahl der zu schätzenden und zu vergleichenden Parameter erhöhen:
Exakte Vorgangsweise: man wähle $n \in N$ und unterteile das Intervall $(0,C)$ in n (gleiche) Teile. Für $i = 1,...n$ sei $S_i := S(iC/n)$; unter H_0 gilt die Beziehung $S_{i+j} = S_i S_j$ (für alle i,j mit $i+j \leq n$). Die Idee besteht nun darin, alle S_i zu schätzen und alle derartigen Beziehungen auszunützen.

Lemma: Es ist $S_{i+j} = S_i S_j$ für alle i,j mit $i+j \leq n$ genau dann erfüllt, wenn $S_{1+j} = S_1 S_j$ gilt für alle $j \leq n-1$. D.h.: die ersten $n(n-1)/4$ Bedingungen sind in den letzten n-1 Bedingungen enthalten.
Beweis: Es muß gezeigt werden, daß aus $S_1 S_j = S_{1+j}$ für alle $j \leq n-1$ folgt, daß $S_i S_j = S_{i+j}$ gilt für alle i,j mit $i+j \leq n$. Dies geschieht durch Induktion nach i: für i=1 ist die Behauptung trivial; angenommen, für i ist die Behauptung gezeigt: es gilt $S_i S_j = S_{i+j}$ für $j \leq n-i$. Sei j derart, daß $i+1+j \leq n$; dann ist $S_{i+1} S_j = S_i S_1 S_j = S_i S_{1+j}$ nach Voraussetzung; es ist $i+1+j \leq n$, daher $j \leq n-1$, daher gilt nach Induktionsannahme $S_i S_{1+j} = S_{i+1+j}$. Daraus folgt dann die Behauptung.
Für $i = 1,...n$ sei $\hat{S}_i = \hat{S}(iC/n)$; setzt man $X_{ij} := I\{T_j > iC/n\}$ (für i=n muß man wieder die zensierten Beobachtungen als "eigentlich" ein infinitesimales Zeitintervall nach C eintreffend interpretieren). Dann gilt

$$EX_{ij} = S_i \text{ und } \hat{S}_i = N^{-1} \Sigma_j X_{ij}.$$

Nach der schon zitierten Fassung des zentralen Grenzwertsatzes strebt die Statistik:

$$\sqrt{N}(\hat{S}_1-S_1,\hat{S}_2-S_2,\ldots\hat{S}_n-S_n)$$

mit wachsendem N in Verteilung gegen eine multivariate Normalverteilung mit Mittelwert 0 und Varianz-Kovarianz-Matrix T, wobei T gegeben ist durch

$$T = \mathrm{Cov}(X_{1j},\ldots X_{nj})'.$$

Zunächst muß also $\mathrm{Cov}(X_i,X_k)$ bestimmt werden.(Wir lassen dabei den Index j weg.) Ist $i\leq k$, dann ist $X_iX_k = I\{T_j>iC/n\}\cdot I\{T_j>kC/n\} = I\{T_j>kC/n\} = X_k$. Daher ist in diesem Fall die Kovarianz gegeben durch $\mathrm{Cov}(X_i,X_k) = E(X_iX_k) - E(X_i)E(X_k) = E(X_k) - E(X_i)E(X_k) = S_k(1-S_i) = S_kF_i$, wobei $F_i = 1-S_i$ die Verteilungsfunktion bezeichnet. Die Matrix T ist damit gegeben durch

$$T = (S_{\max(i,k)}F_{\min(i,k)}).$$

Lemma: Die Matrix T ist genau dann regulär, wenn $1>S_i>S_{i+1}>0$ gilt für alle i,d.h. wenn die Überlebensfunktion in keinem der Intervalle $[iC/n,(i+1)C/n]$ konstant ist.

Beweis: Ist die Voraussetzung erfüllt, dann müssen wir zeigen, daß die Spalten von T linear unabhängig sind. Dazu dividieren wir die i-te Spalte durch F_i und zeigen durch Induktion, daß dann jeweils die i+1-te Spalte von den ersten i Spalten linear unabhängig ist. Die i-te Spalte hat nach erfolgter Division die Gestalt

$$t_i = (S_iF_1/F_i,\ldots S_iF_{i-1}/F_i,S_i,S_{i+1},\ldots S_n)'.$$

Für i=0 oder 1 ist die Behauptung trivialerweise richtig; angenommen, die Behauptung ist für i richtig und die i+1-te Spalte ist von den ersten i Spalten linear abhängig. Dann gäbe es eine

nichttriviale Darstellung

$$t_{i+1} = \sum_{k=1}^{i} r_k t_k.$$

Die i+1-te Eintragung in t_{i+1} wäre $S_{i+1} = (\Sigma r_k)S_{i+1}$, damit wäre $\Sigma r_k = 1$. Für die i-te Eintragung in t_{i+1} würde gelten,daß $\Sigma r_k S_i = S_{i+1}F_i/F_{i+1}$. Da aber $\Sigma r_k = 1$, wäre damit $S_i = S_{i+1}F_i/F_{i+1}$ oder $S_i(1-S_{i+1}) = S_{i+1}(1-S_i)$ woraus sich $S_i = S_{i+1}$ ergibt, ein Widerspruch.
Die Umkehrung ist trivial, da in diesem Fall (mindestens) zwei Zeilen und zwei Spalten identisch wären.

Es gibt nun mehrere Varianten mittels Vergleich der geschätzten S_i Teststatistiken zu konstruieren:
1.Vergleiche alle $\hat{S}_1\hat{S}_j$ mit $\hat{S}_{j+1}$.Der Kürze halber wollen wir folgende Schreibweise vereinbaren:

$$S := (S_1,\ldots S_n)' \text{ und } \hat{S} := (\hat{S}_1,\ldots \hat{S}_n)'.$$

Sei nun der Prüfvektor d definiert durch:
$d := (\hat{S}_1\hat{S}_1-\hat{S}_2,\ldots \hat{S}_1\hat{S}_i-\hat{S}_{i+1},\ldots \hat{S}_1\hat{S}_{n-1}-\hat{S}_n)'$.
Weiters sei $g:R^n \dashrightarrow R^{n-1}$ definiert durch
$g(u_1,\ldots u_n) = (u_1u_1-u_2,\ldots u_1u_i-u_{i+1},\ldots u_1u_{n-1}-u_n)$.
Dann ist $g(S) = 0$ und $g(\hat{S}) = d$. Nach dem oben zitierten Satz aus der Theorie der Grenzverteilungen gilt daher, daß sie Statistik

$$\sqrt{N}(g(\hat{S})-g(S)) = \sqrt{N}d$$

asymptotisch multivariat normalverteilt ist mit Mittelwert 0 und Varianz-Kovarianz-Matrix V, die gegeben ist durch:

$$V = D'TD$$

wobei D gegeben ist durch $D = Dg(S)$. Explizit erhalten wir daher

für die Matrix D:

$$\begin{aligned} D_{11} &= 2S_1 \\ D_{1i} &= S_i \qquad \text{für } 2\le i\le n-1 \\ D_{ii} &= S_1 \qquad \text{für } 2\le i\le n-1 \\ D_{i+1,1} &= -1 \qquad \text{für } 2\le i\le n-1 \\ D_{ij} &= 0 \qquad \text{sonst} \end{aligned}$$

Unter der Voraussetzung, daß H_0 zutrifft, wollen wir uns V explizit ausrechnen:

$A_{ij} := (TD)_{ij} = \Sigma_k T_{ik} D_{kj} = T_{i1}D_{1j} + T_{ij}D_{jj} + T_{i,j+1}D_{j+1,j} =$

$= S_iF_1S_j + S_{max(i,j)}F_{min(i,j)}S_1 - S_{max(i,j+1)}F_{min(i,j+1)}$

(dies gilt auch für j=1).

Wir betrachten nun die verschiedenen Fälle:

$i\le j$: $A_{ij} = S_iS_j(1-S_1) + S_jS_1(1-S_i) - S_{j+1}(1-S_i) = S_iS_j - S_iS_jS_1 +$

$+ S_jS_1 - S_iS_jS_1 - S_{j+1} + S_{j+1}S_i$

Unter der Nullhypothese gilt $S_iS_jS_1 = S_{j+1}S_i$ und $S_jS_1 = S_{j+1}$; daher gilt für diesen Fall:

$$A_{ij} = S_iS_j - S_iS_jS_1 = S_iS_j(1-S_1)$$

$i>j$: $A_{ij} = S_iS_j(1-S_1) + S_iS_1(1-S_j) - S_i(1-S_{j+1}) = S_iS_j - S_iS_jS_1 +$

$+ S_iS_1 - S_iS_jS_1 - S_i + S_iS_{j+1}$

Unter der Nullhypothese gilt $S_iS_jS_1 = S_iS_{j+1}$, daher gilt für diesen Fall:

$$A_{ij} = S_iS_j(1-S_1) - S_i(1-S_1) = S_i(1-S_1)(S_j-1)$$

Wir erhalten somit:

$$\begin{aligned} A_{ij} &= S_iS_j(1-S_1) \qquad \text{für } i\le j \\ A_{ij} &= -S_i(1-S_j)(1-S_i) \qquad \text{für } i>j \end{aligned}$$

Weiters ist

$V_{ij} = (D'A)_{ij} = D'_{i1}A_{1j} + D'_{ii}A_{ij} + D'_{i,i+1}A_{i+1,j} =$

$= S_iA_{1j} + S_1A_{ij} - A_{i+1,j}.$

Wir unterscheiden wieder die möglichen Fälle:

$i<j$: $V_{ij} = S_iS_1S_j(1-S_1) + S_1S_iS_j(1-S_1) - S_{i+1}S_j(1-S_1) =$

$= S_iS_jS_1 - S_iS_jS_1^2 + S_iS_jS_1 - S_iS_jS_1^2 - S_{i+1}S_j + S_{i+1}S_jS_1.$

Unter der Nullhypothese gilt $S_iS_jS_1 = S_{i+1}S_j$ und $S_iS_jS_1^2 = S_{i+1}S_jS_1$. Daher erhalten wir für diesen Fall:

$$V_{ij} = S_iS_jS_1(1-S_1).$$

i=j: $V_{ii} = S_iS_1S_i(1-S_1) + S_1S_iS_i(1-S_1) + S_{i+1}(1-S_i)(1-S_1) = 2S_i^2S_1$
$- 2S_i^2S_1^2 + S_{i+1} - S_{i+1}S_i - S_{i+1}S_1 + S_{i+1}S_iS_1.$

Unter der Nullhypothese gilt $S_i^2S_1 = S_iS_{i+1}$ und $S_i^2S_1^2 = S_{i+1}S_iS_1$. Daher erhalten wir für diesen Fall:

$V_{ii} = S_i^2S_1 - S_i^2S_1^2 + S_{i+1} - S_{i+1}S_1 = S_i^2S_1(1-S_1) + S_{i+1}(1-S_1)$
$= S_{i+1}(1-S_1)(1+S_i) = S_1(1-S_1)S_i(1+S_i).$

i>j: $V_{ij} = S_iS_1S_j(1-S_1) - S_1S_i(1-S_j)(1-S_1) + S_{i+1}(1-S_j)(1-S_1) =$
$= S_iS_jS_1 - S_iS_jS_1^2 - S_iS_1 + S_iS_jS_1 + S_iS_1^2 - S_iS_jS_1^2 +$
$+ S_{i+1} - S_{i+1}S_j - S_{i+1}S_1 + S_{i+1}S_jS_1.$

Unter der Nullhypothese gilt $S_iS_1^2 = S_{i+1}S_1$, $S_iS_jS_1^2 = S_{i+1}S_jS_1$, $S_{i+1}S_j = S_iS_jS_1$ und $S_iS_1 = S_{i+1}$. Daher ist wieder

$$V_{ij} = S_iS_jS_1(1-S_1)$$

Es muß dasselbe Ergebnis wie für i<j herauskommen, da V eine symmetrische Matrix ist. Zusammenfassend erhalten wir für die Matrix V:

$$V = S_1(1-S_1)U = S_1(1-S_1)(u_{ij})$$

wobei

$$u_{ij} = S_iS_j \quad \text{für } i \neq j$$
$$u_{ii} = S_i^2 + S_i$$

Die Matrix V ist nicht singulär, was man direkt beweisen kann (ähnlich wie gezeigt wurde, daß T nicht singulär ist), oder daraus folgern kann, daß V = D'TD, T nicht singulär (und symmetrisch) ist und rg D = n-1 ist.

Nach dem oben bewiesenen Lemma haben die Statistiken

$$Nd'V^{-1}d \text{ und } Nd'\hat{V}^{-1}d$$

dieselbe Grenzverteilung. Wir erhalten daher als Teststatistik:

$$H = \frac{N}{\frac{h(1)(N-h(1))}{N^2}} * d'\hat{U}^{-1}d = \frac{N^3}{h(1)(N-h(1))} * d'\hat{U}^{-1}d$$

welche asymptotisch chi^2-verteilt ist mit n-1 Freiheitsgraden, wobei Nd der folgende Vektor sein soll

$$Nd = \begin{bmatrix} h(1)h(1)/N - h(2) \\ h(1)h(2)/N - h(3) \\ * \\ * \\ h(1)h(n-1)/N - h(n) \end{bmatrix}$$

und $\hat{U} = (\hat{u}_{ij})$ die Matrix mit den folgenden Eintragungen bezeichnen soll:

$\hat{u}_{ij} = h(i)h(j)/N^2$ für $i \neq j$

$\hat{u}_{ii} = h^2(i)/N^2 + h(i)/n$

wobei h(i) = N - Anzahl der Beobachtungen bis zum Zeitpunkt iC/n. Dann gilt: der Test ist konsistent gegen jede Alternative, für die für mindestens ein Paar (i,j) mit $i+j \leq n$ $S_iS_j \neq S_{i+j}$ erfüllt ist. Wenn die Nullhypothese nicht gilt, dann kann die von uns explizit berechnete Matrix V nicht mehr als (asymptotische) Varianz-Kovarianz-Matrix des zufälligen Vektors d interpretiert werden. Nichtsdestoweniger konvergiert aber $\hat{V}$ in Wahrscheinlichkeit gegen die (endliche) Matrix V, insbesondere ist $\hat{V}$ in Wahrscheinlichkeit beschränkt, (d.h. es existiert ein $K>0$, sodaß zu jedem $\delta>0$ ein $N_0 \in N$ existiert, sodaß $P(|V|>K)<\delta$, wenn nur $N>N_0$). Damit ist sichergestellt, daß für den Fall einer derartigen Alternative gilt: p lim H = ∞.

Zwei weitere Varianten dieses Tests beruhen darauf, daß der Prüfvektor d modifiziert wird. Im bisher betrachteten Fall kamen in

vor, daher waren drei Eintragungen in jeder Spalte von D ungleich 0. Die Matrix D kann nun insoferne vereinfacht werden,als für den Prüfvektor d folgende Form gewählt wird:

$d = (\hat{S}_1^2-\hat{S}_2,\ldots\hat{S}_1^n-\hat{S}_n)'$.

Wir wollen wieder die Schreibweise von vorher verwenden:

$\hat{S} = (\hat{S}_1,\ldots\hat{S}_n)'$ und $S = (S_1,\ldots S_n)'$. Wie vorher ist die Statistik

$$\sqrt{N}(\hat{S} - S)$$

asymptotisch multivariat normalverteilt mit Mittelwert 0 und Varianz-Kovarianz-Matrix T. Setzt man

$g(u_1,\ldots u_n) = (u_1^2-u_2,u_1^3-u_3,\ldots u_1^n-u_n)$,

dann gilt $g(S) = 0$ und $g(\hat{S}) = d$. Nach dem bereits zitierten Satz aus der asymptotischen Verteilungstheorie ist die Statistik

$$\sqrt{N}(g(\hat{S})-g(S)) = \sqrt{N}g(\hat{S}) = \sqrt{N}d$$

asymptotisch normalverteilt mit Mittelwert 0 und Varianz-Kovarianz-Matrix $V = D'TD$, wobei $D = Dg(S)$ bedeutet. Wir müssen also wieder D und V berechnen. Bildet man die partiellen Ableitungen von g, dann erhält man:

$$D_{1i} = (i+1)S_1^i \quad \text{für } 1\leq i\leq n-1$$

$$D_{i+1,i} = -1 \quad \text{für } 1\leq i\leq n-1$$

$$D_{ij} = 0 \quad \text{sonst}$$

Wir wollen wieder $V = D'TD$ berechnen:

$$A_{ij} = (TD)_{ij} = T_{i1}(j+1)S_1^j - T_{i,j+1} =$$

$$= (j+1)S_iS_1^j(1-S_1) - S_{\max(i,j+1)}(1-S_{\min(i,j+1)}).$$

Daher ist

$$V_{ij} = (D'A)_{ij} = D'_{i1}A_{1j} + D'_{i,i+1}A_{i+1,j} = (i+1)S_1^iA_{1j} - A_{i+1,j} =$$

$$= (i+1)S_1^i[(j+1)S_1S_1^j(1-S_1) - S_{j+1}(1-S_1)] -$$

$$- (j+1)S_{i+1}S_1^j(1-S_1) + S_{\max(i+1,j+1)}(1-S_{\min(i+1,j+1)}) =$$

$$= (i+1)(j+1)S_1^{i+1}S_1^j(1-S_1) - (i+1)S_1^iS_{j+1}(1-S_1) -$$

$$- (j+1)S_{i+1}S_1^j(1-S_1) + S_{\max(i,j)+1}(1-S_{\min(i,j)+1}) =$$

$$= (ij+i+j+1)S_1^{i+1}S_1^j - (ij+i+j+1)S_1^{i+1}S_1^{j+1} - (i+1)S_1^iS_{j+1} +$$
$$+ (i+1)S_1^{i+1}S_{j+1} - (j+1)S_{i+1}S_1^j + (j+1)S_{i+1}S_1^{j+1} +$$
$$+ S_{\max(i,j)+1}(1-S_{\min(i,j)+1}).$$

Unter der Nullhypothese gilt: $S_1^{i+1}S_1^j = S_1^iS_{j+1} = S_{i+1}S_1^j$ und $S_1^{i+1}S_1^{j+1} = S_1^{i+1}S_{j+1} = S_{i+1}S_1^{j+1}$. Damit läßt sich unter H_0 die Matrix V angeben durch:

$$V_{ij} = (-ij-i-j-1+i+1+j+1)S_{i+1}S_{j+1} +$$
$$+ (ij+i+j+1-i-1-j-1)S_iS_jS_1 + S_{\max(i,j)+1}(1-S_{\min(i,j)+1}) =$$
$$= (ij-1)S_iS_jS_1(1-S_1) + S_{\max(i,j)+1}(1-S_{\min(i,j)+1})).$$

Man sieht also, daß in diesem Fall die Varianz-Kovarianz-Matrix sogar komplizierter wird als in der vorhergehenden Variante. Ersetzt man die S_i durch $\hat{S}_i$, dann erhält man eine konsistente Schätzung $\hat{V}$ für V, und wieder ist die Statistik

$$H := Nd'\hat{V}^{-1}d$$

unter H_0 asymptotisch chi^2-verteilt mit n-1 Freiheitsgraden. Unter H_1: $S_iS_j \neq S_{i+j}$ für mindestens ein Paar (i,j) mit $i+j \leq n$ gilt wieder p lim H = ∞. (Daß die Menge der Alternativen, gegen die dieser Test konsistent ist, mit der des vorigen Tests übereinstimmt, ergibt sich daraus, daß die n-1 Bedingungen $\{S_1^i=S_i \mid i=2,\ldots n\}$ und $\{S_1S_i=S_{i+1} \mid i=1,\ldots n-1\}$ äquvalent sind.) Neben dem, daß bei diesem Test die Varianz-Kovarianz-Matrix eher komplizierte Gestalt hat, hat auch der Prüfvektor d einen Nachteil: in jeder Eintragung von d kommt $\hat{S}_1$ vor, d.h. es werden alle anderen $\hat{S}_i$ mit dem *einem* $\hat{S}_1$ verglichen, sodaß den Beobachtungen, die im ersten Intervall gemacht werden, eine größere Bedeutung zukommt. Dies wird durch folgende Modifizierung des Prüfvektors d vermieden. Sei:

$$d = (\hat{S}_1^2-\hat{S}_2, \hat{S}_2^3-\hat{S}_3^2, \ldots \hat{S}_{n-1}^n-\hat{S}_n^{n-1})'.$$

Setzt man ferner:

$g(u_1,\ldots u_n) = (u_1^2-u_2, u_2^3-u_3^2, \ldots u_{n-1}^n-u_n^{n-1})$,

dann ist $g(S) = 0$ und $g(\hat{S}) = d$, daher ist wieder

$$\sqrt{N}(g(\hat{S}) - g(S)) = \sqrt{N}g(\hat{S}) = \sqrt{N}d$$

asymptotisch multivariat normalverteilt mit Mittelwert 0 und Varianz-Kovarianz-Matrix $V = D'TD$, wobei $D = Dg(S)$. Bildet man die partiellen Ableitungen von g und wertet sie an der Stelle $S = (S_1,\ldots S_n)$ aus, dann erhält man:

$$D_{ii} = (i+1)S_i^i \qquad \text{für } 1\leq i\leq n-1$$

$$D_{i+1,i} = -iS_{i+1}^{i-1} \qquad \text{für } 1\leq i\leq n-1$$

$$D_{ij} = 0 \qquad \text{sonst}$$

Wir wollen nun wieder D'TD explizit berechnen:

$A_{ij} = (TD)_{ij} = T_{ij}D_{jj} + T_{i,j+1}D_{j+1,j}$.

Für $i\leq j$ erhalten wir:

$A_{ij} = S_j(1-S_i)(j+1)S_j^j - S_{j+1}(1-S_i)jS_{j+1}^{j-1}$.

Für $i>j$ erhalten wir:

$A_{ij} = S_i(1-S_j)(j+1)S_j^j - S_i(1-S_{j+1})jS_{j+1}^{j-1}$.

Unter der Nullhypothese gilt nun zunächst $S_{j+1} = S_jS_1$, $S_1^j = S_j$ und $S_j^{j+1} = S_{j+1}^j$. Für $i\leq j$ ergibt sich daher:

$A_{ij} = (j+1)(1-S_i)S_j^{j+1} - j(1-S_i)S_{j+1}^j = S_j^{j+1}(1-S_i)(j+1-j) =$

$= S_j^{j+1}(1-S_i)$.

Weiters gilt unter der Nullhypothese $S_{j+1}^{j-1}S_1 = S_j^{j-1}S_1^{j-1}S_1 = S_j^{j-1}S_j = S_j^j$. Ist nun $i>j$, dann ist $i>1$ und wir können unter H_0 schreiben:

$A_{ij} = S_j^j[(j+1)S_i(1-S_j) - jS_{i-1}(1-S_{j+1})] =$

$= S_j^j[(j+1)S_i - (j+1)S_iS_j - jS_{i-1} + jS_{i-1}S_{j+1}] =$

$= S_j^j[(j+1)S_i - jS_{i-1} - S_iS_j]$.

Da $i>j$, kann S_j herausgehoben werden, und es ergibt sich:

$$A_{ij} = S_j^{j+1}(1-S_i) \qquad \text{für } i\leq j$$

$$A_{ij} = S_j^{j+1}[(j+1)S_{i-j} - jS_{i-1-j} - S_i] \text{ für } i>j$$

wobei $S_0 := 1$ gesetzt wird. Damit erhalten wir für die Matrix V:

$V_{ij} = (D'A)_{ij} = D'_{ii}A_{ij} + D'_{i,i+1}A_{i+1,j}$, wobei $D'_{ii} = (i+1)S_i^i$ und $D'_{i,i+1} = -iS_{i+1}^{i-1}$.

Wir unterscheiden wieder die verschiedenen Fälle:

Ist $i<j$, dann ist

$$V_{ij} = (i+1)S_i^iS_j^{j+1}(1-S_i) - iS_{i+1}^{i-1}S_j^{j+1}(1-S_{i+1}) =$$
$$= (i+1)S_i^iS_j^{j+1} - (i+1)S_i^{i+1}S_j^{j+1} - iS_{i+1}^{i-1}S_j^{j+1} +$$
$$+ iS_{i+1}^iS_j^{j+1}.$$

Unter der Nullhypothese gilt $S_{i+1}^i = S_i^{i+1}$ und $S_{i+1}^{i-1}S_j^{j+1} = S_i^iS_j^{j+1}/S_1$. Daher gilt für diesen Fall:

$$V_{ij} = S_i^iS_j^{j+1}(i + 1 - i/S_1) - S_i^{i+1}S_j^{j+1} =$$
$$= S_i^iS_j^{j+1}(i + 1 - i/S_1 - S_i).$$

Ist $i=j$, dann erhalten wir:

$$V_{ii} = (i+1)S_i^iS_i^{i+1}(1-S_i) - iS_{i+1}^{i-1}S_i^{i+1}[(i+1)S_1 - i - S_{i+1}] =$$
$$= (i+1)S_i^iS_i^{i+1} - (i+1)S_i^{2(i+1)} - (i^2 + i)S_{i+1}^{i-1}S_i^{i+1}S_1 +$$
$$+ i^2S_{i+1}^{i-1}S_i^{i+1} + iS_{i+1}^iS_i^{i+1} =$$
$$= (i + 1 - i^2 - i)S_i^iS_i^{i+1} - (i+1)S_i^{2(i+1)} + i^2S_i^iS_i^{i+1}/S_1 +$$
$$+ iS_{i+1}^iS_i^{i+1} = S_i^iS_i^{i+1}[i^2/S_1 - i^2 + 1 - (i+1)S_i + iS_i] =$$
$$= S_i^iS_i^{i+1}[i^2/S_1 - i^2 + 1 - S_i].$$

Ist $i>j$, dann erhalten wir:

$$V_{ij} = (i+1)S_i^iS_j^{j+1}[(j+1)S_{i-j} - jS_{i-1-j} - S_i] -$$
$$- iS_{i+1}^{i-1}S_j^{j+1}[(j+1)S_{i+1-j} - jS_{i-j} - S_{i+1}].$$

Unter der Nullhypothese gilt nun: hebt man aus der Klammer des zweiten Ausdrucks S_1 heraus, dann ist das in der Klammer verbleibende genau die Klammer des ersten Ausdrucks. Multipliziert man $S_{i+1}^{i-1}S_j^{j+1}$ mit S_1, dann er hält man gerade $S_i^iS_j^{j+1}$, d.h. der erste Ausdruck ist genau das $(i+1)/i$-fache des zweiten Ausdrucks. Damit erhält man:

$$V_{ij} = S_i^iS_j^{j+1}[(j+1)S_{i-j} - jS_{i-1-j} - S_i].$$

$i>j$, dann kann man S_{i-j} aus der Klammer herausheben und wegen $S_i^i S_j^{j+1} S_{i-j} = S_j^j S_i^{i+1}$ erhält man gerade V_{ji}, wie es für den ersten Fall berechnet wurde. Zusammenfassend läßt sich die Varianz-Kovarianz-Matrix also folgendermaßen schreiben:

$V_{ij} = S_i^i S_j^{j+1}(i + 1 - i/S_1 - S_i)$ für $i<j$

$V_{ij} = V_{ji} = S_i^i S_j^{j+1}[(j+1)S_{i-j} - jS_{i-1-j} - S_i]$ für $j<i$

$V_{ii} = S_i^i S_i^{i+1}[1 - i^2 + i^2/S_1 - S_i]$

Ersetzt man die in V auftretenden S_i durch die Schätzungen $\hat{S}_i$, dann ist unter der Nullhypothese

$$H := Nd'\hat{V}^{-1}d$$

asymptotisch chi^2-verteilt mit n-1 Freiheitsgraden, und der so gewonnene Test ist konsistent gegen jede Alternative, für die gilt: $S_iS_j \neq S_{i+j}$ für ein Paar (i,j) mit $i+j\leq n$.(Mittels vollständiger Induktion kann man zeigen, daß die Bedingungen $\{S_i^{i+1} = S_{i+1}^i | i=1,...n-1\}$ äquivalent sind zu den Bedingungen $\{S_1^i = S_i | i=2,...n\}$ und damit äquivalent zu den Bedingungen $\{S_iS_j = S_{i+j} | 1\leq i,j\leq n-1, i+j\leq n\}$.)

Man hat also die folgende Situation: je größer das gewählte n, desto größer die Menge der Alternativen, gegen die der Test konsistent ist; genauer: wenn n_2 von n_1 geteilt wird, dann ist die Menge der Alternativen, gegen die $H(n_1)$ konsistent ist, in der Menge der Alternativen, gegen die $H(n_2)$ konsistent ist, enthalten. Andrerseits wird, bei gegebener Stichprobengröße der Test mit wachsendem n schlechter, da aus den gegebenen Daten eine größere Anzahl von Parametern geschätzt werden muß, was sich auf die Power des Tests und auf die "Entfernung" der Nullverteilung der Teststatistik von ihrer Grenzverteilung auswirken wird.

Weitere Bemerkungen:

i) Es stört nicht, wenn bei Gelten einer Alternativ-Hypothese V singulär ist, und damit $\hat{V}$ gegen eine singuläre Matrix konvergiert, wenn man vereinbart, daß $H = \infty$, wenn $\hat{V}$ singulär ist. Hinreichend dafür, daß H unter H_1 gegen ∞ konvergiert ist ja, daß eine Eintragung in d gegen einen Wert ungleich Null konvergiert und daß $\hat{V}$ in Wahrscheinlichkeit beschränkt ist, d.h. daß ein $K>0$ existiert, sodaß für jedes $\delta>0$ ein N_0 existiert, sodaß $P(|\hat{V}|\geq K)<\delta$, wenn nur $N\geq N_0$.

ii) Notwendig und hinreichend dafür, daß die Matrix $\hat{V}$ nicht singulär ist, ist, daß die $\hat{S}_i$ paarweise verschieden sind, d.h. daß in jedem Intervall $(iC/n,(i+1)C/n]$ mindestens eine Beobachtung vorkommt.

Betrachtet man $\hat{D}'\hat{T}\hat{D}$, dann ist dies unmittelbar einsichtig, da dann $\hat{T}$ regulär und symmetrisch ist und $\hat{D}$ Rang n-1 hat. Betrachtet man die explizite Form von $\hat{V}$, wie sie vorher ausgerechnet wurde und nur unter der Nullhypothese eine konsistente Schätzung von D'TD darstellt, so kann man zeigen, daß sich $\hat{V}$ für die erste Variante durch zulässige Matrixoperationen auf die Gestalt $E+\Gamma$ reduzieren läßt, wobei E nur aus Einsen besteht und Γ eine Diagonalmatrix mit vollem Rang (=n-1) ist, wenn in jedem derartigen Intervall mindestens eine Beobachtung vorkommt. Für die anderen Varianten gilt eine ähnliche, etwas kompliziertere Überlegung.

iii) Zu jeder möglichen Alternative H_1 gibt es ein $n\in N$, sodaß der mit diesem n konstruierte Test konsistent gegen H_1 ist: Angenommen, eine Überlebensfunktion S(t) erfüllt $S(x+y) = S(x)S(y)$ für alle $x,y \in \{iC/n|i,n\in N\}$ mit $x+y\leq C$. Ist $S(C) = 0$, dann ist S auf $\{iC/n|i,n\in N\}$ konstant 0 und wegen der Monotonie auf $[0,C]$ identisch 0, was wir ausschließen können. Sei $\alpha := -\log S(C)/C$.

Dann ist $S(C) = \exp(-\alpha C)$; weiters ist $S(C/n) = S^{1/n}(C) =$
$= \exp(-\alpha C/n)$ und $S(iC/n) = S^{i/n}(C) = \exp(-\alpha iC/n)$ für $i,n \in N$.
Damit stimmt $S(t)$ auf der dichten Teilmenge $\{iC/n\}$ mit $\exp(-\alpha t)$ überein. $S(t)$ ist monoton und auf $\{iC/n\}$ stetig, hat also keine Sprungstellen; daher ist $S(t)$ auf $[0,C]$ stetig, die Fortsetzung von $\{iC/n\}$ auf $[0,C]$ ist eindeutig, daher ist $S(t) = \exp(-\alpha t)$ für alle $t \in [0,C]$.

1.2.Tests mittels der Maximumlikelihood-Schätzung des Parameters α

Wir wollen als Nullhypothese wieder annehmen, daß die beobachteten Daten von einem Prozeß erzeugt worden sind, der im Beobachtungsintervall [0,C] eine konstante Hazardfunktion $r(t) \equiv \alpha$ besitzt; für die Überlebensfunktion S(t) heißt dies, daß im Intervall [0,C] gilt: $S(t) \equiv \exp(-\alpha t)$.

Der Maximum-Likelihood-Schätzer für α ist gegeben durch (siehe z.B.Diekmann-Mitter(1984)):

$$\hat{\alpha}^{-1} = K^{-1} \sum_{i=1}^{K} T_i + (N-K)C/K$$

wobei K die Anzahl der tatsächlich eingetretenen Todesfälle und T_i die Zeiten ebendieser Todesfälle bezeichnen (K ist also selbst eine zufällige Variable). Man kann diese Formel etwas anders schreiben und erhält:

$$\hat{\alpha}^{-1} = (N/K) \cdot N^{-1} \sum_{i=1}^{N} \min(T_i,C) \qquad (*)$$

K ist die Anzahl der Beobachtungen im Intervall (0,C); daher gilt $p \lim K/N = P(T_i \leq C) = 1 - S(C)$. Damit ist aus der Formel (*) sofort ersichtlich, daß

$$p \lim \hat{\alpha}^{-1} = (1-S(C))^{-1} \cdot E(\min(T_i,C))$$

gilt.

Wir wollen zunächst $E(\min(T_i,C))$ bestimmen. Da $dF(t) = -dS(t)$, erhalten wir:

$$E(\min(T_i,C)) = -\int_0^\infty \min(t,C)dS(t) = -\int_0^C t dS(t) - C\int_C^\infty dS(t) = -tS(t)\Big|_0^C + \int_0^C S(t)dt - C(S(\infty)-S/C)) = -C \cdot S(C) + \int_0^C S(t)dt - C \cdot 0 + C \cdot S(C) = \int_0^C S(t)dt.$$

Wir wollen dabei voraussetzen, daß alle Regularitätsbedingungen für S(t) erfüllt sind, die notwendig für die Gültigkeit der

durchgeführten Manipulationen sind - wenn eine Dichtefunktion existiert, d.h. wenn die Überlebensfunktion absolutstetig ist, ist sicher alles erlaubt. Allgemein gilt also:

$$p \lim \hat{\alpha}^{-1} = (1-S(C))^{-1}\int_0^C S(t)dt.$$

Dies gilt nicht nur unter der Nullhypothese, sondern für jede Verteilung, für die die oben durchgeführten Operationen erlaubt sind. Unter der Nullhypothese, wenn also $S(t) = \exp(-\alpha t)$, erhalten wir folgendes:

$$\int_0^C \exp(-\alpha t)dt = -\alpha^{-1}\exp(-\alpha t)\Big|_0^C = -\alpha^{-1}(\exp(-\alpha C) - 1) = \alpha^{-1}(1-S(C)).$$

Unter H_0 gilt also

$$p \lim \hat{\alpha}^{-1} = \alpha^{-1}$$

Und dies würde auch gelten, wenn wir den ML-Schätzer bezüglich einer anderen Zensierungszeit $C'<C$ berechnet hätten. Unter der Nullhypothese gilt also:

$$H(z) := p \lim \hat{\alpha}^{-1}(z) = (1-S(z))^{-1}\int_0^z S(t)dt \equiv \alpha^{-1}.$$

Sei umgekehrt eine Überlebensfunktion $S(t)$ gegeben, für die im Intervall $(0,C)$ die oben definierte Funktion $H(z)$ konstant ist. Dann gilt

$$\int_0^z S(t)dt \equiv \alpha^{-1}(1 - S(z)) \quad \text{auf } (0,C).$$

Setzen wir voraus, daß $S(t)$ im ganzen Intervall $(0,C)$ differenzierbar ist, dann erhalten wir:

$$S(z) \equiv -\alpha^{-1}S'(z)$$

oder

$$-\alpha \equiv S'(z)/S(z) \equiv d(\log S(z))/dz.$$

Daraus ergibt sich:

$$S(t) = \exp(-\alpha t) \text{ für alle } t \in (0,C).$$

Daraus ergibt sich, daß die Exponentialverteilung dadurch charakterisiert ist, daß der in (*) definierte Ausdruck *unabhängig vom Wert der Zensierungszeit* C in Wahrscheinlichkeit gegen denselben Wert α^{-1} konvergiert.

Damit ist wieder die Möglichkeit gegeben, einen Hausmantest zu konstruieren: dazu unterteile man das Intervall [0,C] in n Teile mit den Teilungspunkten $0 < C_1 < C_2 < \ldots < C_n = C$. Faßt man die Punkte C_i als n verschiedene Zensierungszeiten auf, und bildet für jede dieser Zeiten den ML-Schätzer $\hat{\alpha}^{-1}(C_i)$, dann sollten unter der Nullhypothese die Differenzen zwischen diesen alternativen Schätzungen für α^{-1} nicht "zu groß" werden.

Ausführung: Sei U_i der zufällige (n x 1) -Vektor, dessen j-te Eintragung gegeben ist durch:

$$U_{ij} := (1-S(C_j))^{-1}.\min(T_i,C_j)$$

und A der (n x 1) -Vektor mit $A_j = \alpha^{-1}$ für alle j.

Nach der im vorigen Abschnitt zitierten Fassung des zentralen Grenzwertsatzes ist die Statistik

$$N^{-\frac{1}{2}} \sum_{i=1}^{N} (U_i - A)$$

asymptotisch normalverteilt mit Mittelwert 0 und VarianzKovarianz-Matrix $T = \mathrm{Cov}(U_i)$. Wir wollen daher zuerst die VarianzKovarianz-Matrix T berechnen.

Sei $j \leq k$, dann ist

$$t_{jk} = (1-S(C_j))^{-1}.(1-S(C_k))^{-1} \cdot \mathrm{Cov}(\min(T_i,C_j),\min(T_i,C_k)).$$

Wir wollen diesen Wert speziell für $S(t) = \exp(-\alpha t)$ bestimmen. Es ist (wir lassen den Index der Variablen T_i weg):

$$E(\min(T,C_j)\min(T,C_k)) = \int_0^\infty \min(t,C_j)\min(t,C_k)\,dF(t) =$$

$$= \int_0^{C_j} t^2\alpha\exp(-\alpha t)dt + \int_{C_j}^{C_k} C_j t\alpha\exp(-\alpha t)dt + C_jC_k\int_{C_k}^{\infty}\alpha\exp(-\alpha t)dt.$$

Für den ersten Ausdruck ergibt sich:

$$\int_0^{C_j} t^2\alpha\exp(-\alpha t)dt = -\alpha^{-1}t^2\alpha\exp(-\alpha t)\Big|_0^{C_j} + 2\int_0^{C_j} t\exp(-\alpha t)dt =$$

$$= -C_j^2\exp(-\alpha C_j) - 2t\alpha^{-1}\exp(-\alpha t)\Big|_0^{C_j} + 2\int_0^{C_j}\alpha^{-1}\exp(-\alpha t)dt =$$

$$= -C_j^2\exp(-\alpha C_j) - 2C_j\alpha^{-1}\exp(-\alpha C_j) - 2\alpha^{-2}\exp(-\alpha t)\Big|_0^{C_j} =$$

$$= -C_j^2\exp(-\alpha C_j) - 2C_j\alpha^{-1}\exp(-\alpha C_j) - 2\alpha^{-2}\exp(-\alpha C_j) + 2\alpha^{-2}.$$

Für den zweiten Ausdruck ergibt sich:

$$\int_{C_j}^{C_k} t\alpha\exp(-\alpha t)dt = -t\exp(-\alpha t)\Big|_{C_j}^{C_k} + \int_{C_j}^{C_k}\exp(-\alpha t)dt =$$

$$= C_j\exp(-\alpha C_j) - C_k\exp(-\alpha C_k) - \alpha^{-1}\exp(-\alpha t)\Big|_{C_j}^{C_k} =$$

$$= C_j\exp(-\alpha C_j) - C_k\exp(-\alpha C_k) + \alpha^{-1}\exp(-\alpha C_j) - \alpha^{-1}\exp(-\alpha C_k).$$

Für den dritten Ausdruck erhalten wir:

$$\int_{C_k}^{\infty}\alpha\exp(-\alpha t)dt = -\exp(-\alpha t)\Big|_{C_k}^{\infty} = \exp(-\alpha C_k).$$

Damit erhalten wir:

$$E(\min(T,C_j)\min(T,C_k)) = -C_j^2S(C_j) - 2C_j\alpha^{-1}S(C_j) - 2\alpha^{-2}S(C_j) + 2\alpha^{-2}$$

$$+ C_j^2S(C_j) - C_jC_kS(C_k) + C_j\alpha^{-1}S(C_j) - C_j\alpha^{-1}S(C_k) + C_jC_kS(C_k) =$$

$$= 2\alpha^{-2}(1-S(C_j)) - C_j\alpha^{-1}(S(C_j)+S(C_k)).$$

Vorher haben wir schon berechnet, daß

$$E(\min(T,C_j)) = \alpha^{-1}(1-S(C_j)) = \alpha^{-1}F(C_j).$$

Insgesamt erhalten wir daher für die Kovarianz:

$$Cov(U_{ij},U_{ik}) = 2\alpha^{-2}F(C_k)^{-1} - \alpha^{-2} - C_j\alpha^{-1}[S(C_j)+S(C_k)]/[F(C_j)F(C_k)]$$

(wobei wir $C_j \leq C_k$ vorausgesetzt hatten). Man sieht also, daß die Varianz-Kovarianz-Matrix hier eher kompliziert ist.

Jenachdem, welche der verschiedenen Schätzungen für α^{-1} man nun vergleicht, erhält man verschiedene Varianten eines Hausmantests:

i) Sei $g(u_1,\ldots u_n) = (u_1-u_n,\ldots u_{n-1}-u_n)$. Dann ist

$$N^{-\frac{1}{2}}\Sigma g(U_i)$$

asymptotisch normalverteilt mit Mittelwert 0 und Varianz-Kovarianz-Matrix $V = D'TD$, wobei $T = (Cov(U_{ij},U_{ik}))$ wie eben bestimmt wurde und D gegeben ist durch:

$$D_{ii} = 1 \quad \text{für } 1\leq i\leq n-1$$

$$D_{ni} = -1 \quad \text{für } 1\leq i\leq n-1$$

$$D_{ij} = 0 \quad \text{sonst.}$$

ii) Sei $g(u_1,\ldots u_n) = (u_1-u_2,\ldots u_i-u_{i+1},\ldots u_{n-1}-u_n)$. Dann ist

$$N^{-\frac{1}{2}}\Sigma g(U_i)$$

asymptotisch normalverteilt mit Mittelswert 0 und Varianz-Kovarianz-Matrix $V = D'TD$, wobei T oben bestimmt wurde und D gegeben ist durch:

$$D_{ii} = 1 \quad \text{für } 1\leq i\leq n-1$$

$$D_{i+1,i} = -1 \quad \text{für } 1\leq i\leq n-1$$

$$D_{ij} = 0 \quad \text{sonst.}$$

Ersetzt man nun alle in den Eintragungen von T vorkommenden unbekannten Ausdrücke durch konsistente Schätzungen:
für α^{-1} wählt man am besten den ML-Schätzer bezüglich der größten Zensierungszeit C_n und für $F(C_j)$ setzt man:

$$\hat{F}(C_j) := K_j/N$$

wobei K_j die Anzahl der Beobachtungen $<C_j$ bezeichnet, dann erhält man für beide Varianten eine Teststatistik der Form

$$H = Nd'\hat{V}^{-1}d$$

die unter der Nullhypothese asymptotisch chi^2-verteilt ist mit n-1 Freiheitsgraden, wobei $\hat{V} = D'\hat{T}D$ und $\hat{T}$ die oben erwähnte Schätzung für T ist.

Weiters ist für die erste Variante $d = (d_1,\ldots d_{n-1})'$ gegeben durch

$$d_i = K_i^{-1} \sum_{j=1}^{N} \min(T_j,C_i) - K_n^{-1} \sum_{j=1}^{N} \min(T_j,C_n)$$

Für die zweite Variante ist $d =(d_1,\ldots d_{n-1})'$ gegeben durch

$$d_i = K_i^{-1} \sum_{j=1}^{N} \min(T_j,C_i) - K_{i+1}^{-1} \sum_{j=1}^{N} \min(T_j,C_{i+1})$$

Bei der ersten Variante wird die ML-Schätzung bezüglich jeder "künstlichen" Zensierungszeit C_i ($i<n$) mit der ML-Schätzung bezüglich der eigentlichen Zensierungszeit C_n verglichen. Bei der zweiten Variante werden jeweils die ML-Schätzungen bezüglich zweier aufeinanderfolgender "Zensierungszeiten" verglichen.
Beide Varianten sind konsistent gegen jede Alternative H_1, für die gilt: $H(C_i) \neq H(C_{i+1})$ für ein i, wobei H(z) defniert ist durch:

$$H(z) = (1-S(z))^{-1} \int_0^z S(t)dt = E(\min(T_j,z)).$$

Wählt man speziell $C_i = iC/n$, dann gilt wieder: zu jeder möglichen Alternative H_1 gibt es ein $n\in N$, sodaß $H(iC/n) \neq H((i+1)C/n)$ für ein $i<n$. Sonst wäre H(z) auf einer dichten Teilmenge von (0,C) konstant und daher selbst konstant. Wählt man dann als künstliche Zensierungszeiten $C_i = iC/n$, dann ist der so konstruierte Test konsistent gegen die gegebene Alternative.
Ein offensichtlicher Nachteil der so konstruierten Tests ist die relative Kompliziertheit der asymptotischen Varianz-Kovarianz-Matrix, damit die Kompliziertheit der Teststatistik insgesamt; außerdem müssen sehr viele Eintragungen durch Schätzungen ersetzt werden (In jeder Eintragung der Varianz-Kovarianz-Matrix sind sowohl Schätzungen für den Parameter α als auch Schätzungen für die Verteilungsfunktion $F(C_j)$ einzusetzen.)

2. Typ II zensierte Daten

liegen vor, wenn nicht die Beobachtungsdauer fix gewählt ist, sondern eine feste Anzahl r, sodaß nach der r-ten Beobachtung der Beobachtungsprozeß abgebrochen wird. Unterstellt man wieder einen wachsenden Stichprobenumfang, wie er zur Konstruktion von Hausman-Tests erforderlich ist, bleibt natürlich r nicht fix, sondern wächst mit dem Stichprobenumfang, sodaß $r/N \equiv p < 1$ oder wenigstens $r/N \dashrightarrow p < 1$ für $N \dashrightarrow \infty$ gilt. Sei also $o < p < 1$; wir wollen annehmen, daß die Beobachtung eingestellt wird, wenn die ersten [pN] Todesfälle eingetreten sind. Eine Möglichkeit, ähnlich wie im vorigen Abschnitt einen Hausman-Test zu konstruieren, besteht darin, verschiedene Quantile mittels Order Statistiken zu schätzen und diese Schätzungen miteinander zu vergleichen.
Definition: Sei $0 < q < 1$; dann ist das q-te Qantil x_q einer gegebenen Verteilung definiert durch die Gleichung:

$$q = F(x_q).$$

Das q-te Qantil ist also jener Punkt x_q auf der Zahlengerade, sodaß die Wahrscheinlichkeit, daß eine Beobachtung $\leq x_q$ auftritt gerade q ist. (Wir wollen voraussetzen, daß das q-te Qantil eindeutig definiert ist.)

2.1 Tests mittels durch Order Statistiken geschätzte Quantile

Ist $(T_1,\ldots T_{[pN]})$ der Vektor der Beobachtungen und ist $(T_{(1)},\ldots T_{([pN])})$ der Vektor, den man aus ersterem durch Umordnen der Größe nach erhält, dann heißt $T_{(i)}$ die i-te Order Statistik von $(T_1,\ldots T_{[pN]})$.

Seien nun $0 < q_1 < q_2 < \ldots < q_n \leq p$. Unter gewissen Regularitätsbedingungen (z.B.: F ist in einer Umgebung eines jeden q_i-ten Quantils differenzierbar und es gilt $f(x_r) = F'(x_r) \neq 0$ für $r = q_i$, $1 \leq i \leq n$, - näheres siehe z.B. Lawless (1982), Sarhan & Greenberg (1962), oder Mosteller (1946)) ist die Statistik

$$\sqrt{N}\begin{bmatrix} T([Nq_1]) - x_{q_1} \\ T([Nq_2]) - x_{q_2} \\ \vdots \\ T([Nq_n]) - x_{q_n} \end{bmatrix}$$

asymptotisch multivariat normalverteilt mit Mittelwert 0 und Varianz-Kovarianz-Matrix $T = (t_{ij})$ wobei t_{ij} gegeben ist durch

$$t_{ij} = \frac{q_i(1-q_j)}{f(x_{q_i})f(x_{q_j})}$$

für $q_i \leq q_j$.

Was bedeutet dies, wenn die Ankunftszeiten T_i exponentialverteilt sind? Wir wollen zuerst die Quantile berechnen:

$$q = F(x_q) = 1 - \exp(-\alpha x_q).$$

Daher ist $\exp(-\alpha x_q) = 1-q$ und $x_q = -\log(1-q)/\alpha$. Insbesondere ist $x_{q_i}/x_{q_j} = \log(1-q_i)/\log(1-q_j)$. Definiert man wieder einen Prüfvektor $d = (d_1,\ldots d_{n-1})'$ durch:

$$d_i := T([Nq_i])/T([Nq_{i+1}]) - \log(1-q_i)/\log(1-q_{i+1})$$

dann ist unter H_0 $\sqrt{N}d$ asymptotisch multivariat normalverteilt mit Mittelwert 0 und Varianz-Kovarianz-Matrix V und daher die Statistik

$$H = Nd'\hat{V}^{-1}d$$

asymptotisch chi^2-verteilt mit n-1 Freiheitsgraden, wobei V gegeben gegeben ist durch $V = D'TD$, T wie oben angegeben und $D = Dg(x_{q_1},\ldots x_{q_n})$ mit folgender Abbildung g: $(R^+)^n \dashrightarrow R^{n-1}$: $g(x_1,\ldots x_n) := (x_1/x_2, x_2/x_3, \ldots x_{n-1}/x_n)$.

und $\hat{V}$ eine konsistente Schätzung von V ist. Es gelten wieder ähnliche Aussagen wie im vorhergehenden Abschnitt: z.B.: der Test ist konsistent gegen jede Alternative, für die für ein Paar (i,j) gilt $x_{q_i}/x_{q_j} \neq \log(1-q_i)/\log(1-q_j)$.

Obwohl diese Statistik insgesamt sehr kompliziert anmutet, hat sie den Vorteil, daß unter der Nullhypothese die Matrix V nicht geschätzt werden muß, sondern exakt angegeben werden kann. Wir wollen zunächst T bestimmen: es ist $f(t) = \alpha\exp(-\alpha t)$ und $x_{qi} = -\log(1-q_i)/\alpha$, daher ist $f(x_{q_i}) = \alpha(1-q_i)$, daher ist T gegeben durch

$$t_{ij} = \frac{1}{\alpha^2}\frac{q_i}{1-q_i}$$

wobei $i \leq j$ angenommen wurde.

Weiters ist D gegeben durch:

$$D_{ii} = -\alpha/\log(1-q_i) \qquad 1 \leq i \leq n-1$$

$$D_{i+1,1} = \alpha\log(1-q_i)/\log^2(1-q_{i+1}) \qquad 1 \leq i \leq n-1$$

$$D_{ij} = 0 \quad \text{sonst.}$$

Es kann also α aus D und D', und α^{-2} aus T herausgehoben werden; in D'TD kommen dann keine unbekannten Größen mehr vor.

Man erhält daher eine Teststatistik

$$H = Nd'U^{-1}d$$

die unter der Nullhypothese asymptotisch chi^2-verteilt ist mit n-1 Freiheitsgraden; die Matrix U ist gegeben durch

$$U = D'TD$$

wobei

$$T = (t_{ij}) = (q_{\min(i,j)}/(1-q_{\min(i,j)}))$$

und

$$D_{ii} = -1/\log(1-q_i) \qquad 1 \le i \le n-1$$

$$D_{i+1,i} = \log(1-q_i)/\log^2(1-q_{i+1}) \qquad 1 \le i \le n-1$$

$$D_{ij} = 0 \quad \text{sonst}$$

T ist regulär (siehe 3.1), damit ist auch U=D'TD regulär. Insbesondere ist die Nullverteilung von H unabhängig von α, da sich der Parameter α auch im Vektor d herauskürzt. Die explizite Gestalt von V = D'TD soll hier nicht mehr ausgerechnet werden - die Elemente dieser Matrix sind sehr komplizierte Ausdrücke in den Variablen q_i und $\log(1-q_i)$.

Wählt man speziell $q_n = p$, $n = 2$ und $q_1 = p/2$, dann ist

$$d = \frac{T([Np/2])}{T([Np])} - \frac{\log(1-p/2)}{\log(1-p)}$$

Weiters ist

$$T = 2p\begin{bmatrix}(2-p)^{-1} & (2-p)^{-1} \\ (2-p)^{-1} & (2-2p)^{-1}\end{bmatrix}$$

und

$$D = \begin{bmatrix}-1/\log(1-p/2) \\ \log(1-p/2)/\log^2(1-p)\end{bmatrix}$$

womit sich als asymptotische Varianz von $\sqrt{N}d$ ergibt:

$$\sigma^2 = \frac{p}{(1-p/2)\log^2(1-p/2)} + \frac{p\log^2(1-p/2)}{(1-p)\log^4(1-p)} - \frac{2p}{(1-p/2)\log^2(1-p)}$$

Für die so bestimmte Varianz σ^2 und das oben angegebene d gilt dann:

i) die Statistik

$$\sqrt{N}d/\sigma$$

ist unter der Nullhypothese asymptotisch normalverteilt mit Mittelwert 0 und Varianz 1.

ii) die Statistik

$$Nd^2/\sigma^2$$

ist asymptotisch chi^2-verteilt mit einem Freiheitsgrad.

2.2. Eine Bemerkung über exakte Tests bei Typ-II zensierten Daten

Es sei darauf hingewiesen, daß für den Fall der Typ-II zensierten Daten viel schönere - nämlich exakte Spezifikationstests existieren. Im folgenden sollen kurz einige Resultate von Epstein (1960) referiert werden:

Seien $T_{(1)} \leq T_{(2)} \leq \ldots \leq T_{(r)}$ die ersten r Überlebenszeiten aus einem Sample der Größe N. Dann heißt die Statistik

$$t(i) := T_{(1)} + T_{(2)} + \ldots + T_{(i-1)} + (N-i+1)T_{(i)}$$

die bis zum Zeitpunkt $T_{(i)}$ totale beobachtete Lebenszeit. Für $i<j$ heißt

$$t(i,j) := t(j) - t(i)$$

die im Intervall $(T_{(i)},T_{(j)})$ totale beobachtete Lebenszeit.

Es gilt nun, daß $2\alpha t(i,j)$ unter der Nullhypothese (exakt) chi^2-verteilt ist mit $2j-2i$ Freiheitsgraden; weiters sind die Statistiken $t(i,j)$ und $t(k,l)$ stochastisch unabhängig, wenn sich die dazugehörigen Zeitintervalle nicht überlappen. Diese Tatsache erlaubt die Konstruktion verschiedener Teststatistiken aus Brüchen derartiger totaler Überlebenszeiten, die als Nullverteilung (exakt) eine F- oder eine maximum-F-ratio-Verteilung besitzen.

3. Zufällig zensierte Daten

Das Problem der zufällig zensierten Daten läßt sich folgendermaßen formulieren: Seien (T_i,C_i) $i = 1,...n$, n unabhängige, identisch verteilte Paare von Zufallsvariablen, deren Verteilungsfunktionen mit $F_1(t) = P(T_i \leq t)$ und $F_2(t) = P(C_i \leq t)$ bezeichnet werden. Das Problem besteht darin, diese Verteilungsfunktionen zu schätzen (zumindest die Verteilungsfunktion der Variablen T_i - das Problem ist aber symmetrisch in T_i und C_i), obwohl man nicht Realisierungen (t_i,c_i) die Paare (T_i,C_i) beobachten kann, sondern nur Realisierungen der Paare (X_i,k_i), wobei diese Variablen definiert sind durch:

$$X_i := \min(T_i,C_i)$$

und

$$k_i := I\{T_i \leq C_i\}$$

Obwohl das Problem symmetrisch in T_i und C_i ist, bezieht man sich (vom Ausgangspunkt her) auf T_i als die interessierende Variable und auf C_i als die Zensierungsvariable. Die Indikatorvariable k_i zeigt an, ob es sich bei der gemachten Beobachtung um eine tatsächliche oder eine zensierte handelt.

3.1. Der Produkt-Limit-Schätzer der Überlebensfunktion und sein asymptotisches Verhalten

Der Produkt-Limit- oder Kaplan-Meier-Schätzer (PLS) für die Überlebensfunktion $S(t) = 1 - F_1(t)$, dem Komplement der Verteilungsfunktion, ist folgendermaßen definiert: In Abhängigkeit

vom Sample $\{(X_1,k_1),\ldots(X_N,k_N)\}$ sei

$$\hat{S}(t) := \begin{cases} \prod_{i:X_{(i)}<t} \left[\frac{N-i}{N-i+1}\right]^{k_i} & \text{für } t \leq X_{(N)} \\ 0 & \text{für } t > X_{(N)} \text{ und } k_{(N)}=1 \\ \text{undefiniert} & \text{für } t > X_{(N)} \text{ und } k_{(N)}=0 \end{cases}$$

Dabei bedeutet $X_{(i)}$ das i-te Element von $\{X_1,\ldots X_N\}$, nachdem das Sample $\{(X_i,1-k_i)\}$ lexikographisch geordnet wurde. Anschaulich heißt dies: die X_i werden der Größe nach geordnet; kommen gleiche vor, dann jedes so oft, wie es vorkommt, und zwar werden die unzensierten vor den zensierten gereiht (die man immer als zumindest ein winziges Zeitintervall später eintreffend interpretiert).

Eine andere (äquivalente) Definition des PLS ist folgende: Seien $t_{(1)} < t_{(2)} < \ldots < t_{(k)}$ die k verschiedenen "tatsächlich beobachteten" Überlebenszeiten, d.h. jene X_i, für die $k_i = 1$ gilt (und jede auftretende Zeit nur einmal genommen). Weiters sei n_i $(1 \leq i \leq k)$ die Mächtigkeit der Risikomenge zum Zeitpunkt $t_{(i)}$, d.h. die Anzahl der Individuen, die unmittelbar vor dem Zeitpunkt $t_{(i)}$ am Leben und unzensiert sind, oder in der obigen Terminologie:

$$n_i := |\{j|X_j \geq t_{(i)}\}|$$

Weiters sei d_i die Anzahl der Individuen, die zum Zeitpunkt $t_{(i)}$ sterben, oder formalisiert ausgedrückt:

$$d_i := |\{j|X_j = t_{(i)}, k_i = 1\}|.$$

Der Produkt-Limit-Schätzer ist dann definiert durch:

$$\hat{S}(t) := \prod_{i:t_{(i)}<t} \frac{n_i-d_i}{n_i} \qquad \text{für alle } t$$

Außerdem wird $\hat{S}(t)$ als nicht definiert gesetzt, wenn $t > X_{(N)}$ und

$k_{(N)}=0$.

Der formal einfachste Beweis für die Äquivalenz der beiden Definitionen ist wohl mittels vollständiger Induktion nach der Sample-Größe zu führen.
Anschaulich ist die Äquivalenz jedoch klar: die Faktoren, die in der ersten Definition einer zensierten Beobachtung entsprechen, fallen weg, da $k_{(i)} = 0$. Sei nun ein tatsächlich beobachtetes $t_{(j)}$ gegeben und $X_{(i)}$ das erste in der obigen Ordnung auftretende X_l mit $X_{(i)} = t_{(j)}$; dann ist klarerweise $n_j = N-i+1$. Wenn nun d_j Individuen zum Zeitpunkt $t_{(j)}$ sterben, dann gilt
$t_{(j)} = X_{(i)} = X_{(i+1)} = \ldots = X_{(i+d_j-1)}$
und $k_{(l)}=1$ für $i \leq l \leq i+d_j-1$ und entweder $X_{(i)} < X_{(i+d_j)}$ oder $k_{i+d_j}=0$ (jenachdem ob zum Zeitpunkt $t_{(j)}$ auch noch eine zensierte Beobachtung vorkommt oder nicht).

Dem Faktor $\frac{n_j-d_j}{n_j}$ im zweiten Produkt entspricht im ersten Produkt der folgende Ausdruck:

$$\frac{N-i}{N-i+1} * \frac{N-(i+1)}{N-i} \cdots \frac{N-(i+d_j-1)}{N-(i+d_j-2)} = \frac{(N-i+1)-d_j}{N-i+1}$$

Die anschauliche Bedeutung des Produkt-Limit-Schätzers ist ebenfalls klar: für $t_{(j)}<t\leq t_{(j+1)}$ ist der Ausdruck

$$\frac{n_j-d_j}{n_j}$$

eine Schätzung für die Wahrscheinlichkeit, das Intervall $[t_{(j)},t)$ zu überleben, unter der Bedingung bis zum Zeitpunkt $t_{(j)}$ überlebt

zu haben:

$$\frac{n_j-d_j}{n_j} = \hat{P}(T_i>t|T_i>t_{(j)}) = \frac{\hat{P}(T_i>t \text{ und } T_i>t_{(j)})}{\hat{P}(T_i>t_{(j)})} = \frac{\hat{P}(T_i>t)}{\hat{P}(T_i>t_{(j)})} =$$

$$= \frac{\hat{S}(t)}{\hat{S}(t_{(j)})}$$

Bildet man das Produkt über diese Ausdrücke, dann erhält man eben eine Schätzung für S(t). Kommen nur unzensierte Daten vor, dann stimmt der PLS mit der empirischen Überlebensfunktion, wie sie im ersten Abschnitt für das Intervall [0,C), in dem nur unzensierte Daten auftraten, betrachtet wurde, überein.

Asymptotische Eigenschaften des PLS (bei diversen Annahmen über den Zensierungsprozeß) wurden vor allem von Meier (1975) und von Breslow & Crowley (1974) studiert.

Um Hausman-Tests mittels des PLS konstruieren zu können, brauchen wir derartige asymptotische Eigenschaften des PLS. Wir verwenden im folgenden einige Resultate von Breslow & Crowley (1974). Wir wollen zunächst einige Definitionen und Annahmen anführen:

S(t) sei die Überlebensfunktion der interessierenden Variablen T_i; G(t) sei die Überlebensfunktion der Zensierungsvariablen C_i; die T_i als auch die C_i seien i.i.d., außerdem seien interessierende und zensierende Variable voneinander unabhängig; beide Überlebensfunktionen S(t) und G(t) seien stetig.

Weiters sei

$$F(t) := P(X_i \leq t, k_i=1)$$

die Sub-Verteilungsfunktion der nicht zensierten Beobachtungen.

Außerdem existiere ein K>0 derart, daß

$$R(K) := S(K)G(K) = P(X_i>K) > 0$$

d.h.: die Wahrscheinlichkeit, daß nach dem Zeitpunkt K noch (zensierte und/oder unzensierte) Beobachtungen vorkommen sei

positiv. Unter diesen Voraussetzungen haben Breslow und Crowley folgendes gezeigt:

Satz: Unter den eben gemachten Voraussetzungen konvergiert für $0 < t < K$ der Prozeß

$$\sqrt{N}(\hat{S}(t) - S(t))$$

schwach gegen einen Gauß'schen Prozeß $Z(t)$ mit Mittelwert 0 und Kovarianzfunktion

$$\mathrm{Cov}(Z(v),Z(t)) = S(v)S(t)\int_0^v \frac{-dS(u)}{S(u)R(u)} = S(v)S(t)\int_0^v \frac{dF(u)}{R^2(u)} \qquad (v \le t)$$

Es ist zu vermuten, daß dieser Satz oder ein ähnlicher auch unter schwächeren Bedingungen bewiesen werden kann; z.B. daß die Stetigkeitsforderung für $S(t)$ und $G(t)$ weggelassen wird und nurmehr verlangt wird, daß $S(t)$ und $G(t)$ keine gemeinsamen Sprungstellen besitzen - eventuell auch, daß die Sprungstellen der beiden Überlebensfunktionen keinen Häufungspunkt besitzen; dann müßte aber die Formel für die Kovarianz-Funktion modifiziert - nämlich um einen diskreten Anteil "ergänzt" werden (vgl. Peterson (1977)). Nicht weggelassen werden kann $S(K)G(K)>0$: ist nämlich $S(K) = 0$ und $t^* = \sup\{t|S(t)>0\}$, dann ist $\hat{S}(t)$ für $t \ge t^*$ entweder = 0 oder undefiniert; es wäre daher sinnlos, das Verhalten von $\sqrt{N}(\hat{S}(t) - S(t))$ zu studieren. Dasselbe gilt, wenn $G(K) = 0$. Außerdem wäre dann das Integral in der Kovarianzfunktion nicht definiert. Diese Bedingung stellt ein Analogon zu der Forderung aus 1.1. dar, daß nur Aussagen über die Überlebensfunktion im Beobachtungsintervall $[0,C]$ gemacht werden können, und besagt, daß ab einem Zeitpunkt t^*, ab dem mit

Wahrscheinlichkeit 1 keine (zensierten oder unzensierten) Beobachtungen mehr gemacht werden können, keine Aussagen über die Verteilung der Variablen T_i getroffen werden können.
Daraus folgt nun nach einem bekannten Satz aus der Theorie der stochastischen Prozesse (siehe z.B. Billingsley (1968) oder Hall & Heyde (1980),Appendix), daß die endlichdimensionalen Verteilungen dieses Prozesses asymptotisch normalverteilt sind:

Satz: Es seien die Voraussetzungen des vorhergehenden Satzes erfüllt und $K := \inf\{t|S(t)G(t)=0\}$. Seien $0 < t_1 < \ldots < t_n < K$, dann ist die Statistik

$$\sqrt{N}(\hat{S}(t_1) - S(t_1),\ldots\hat{S}(t_n) - S(t_n))$$

asymptotisch multivariat normalverteilt mit Mittelwert 0 und Varianz-Kovarianz-Matrix $T = (t_{ij})$ mit

$$t_{ij} = \mathrm{Cov}(Z(t_i),Z(t_j)).$$

Wir wollen zunächst untersuchen, unter welchen Bedingungen die Matrix T regulär ist. Die Eintragungen t_{ij} dieser Matrix lassen sich schreiben als:

$$t_{ij} = p_i p_j l_{\min(i,j)}$$

wobei

$$p_i = S(t_i), p_j = S(t_j)$$

und

$$l_i = \int_0^{t_i} \frac{-dS(u)}{S(u)R(u)}$$

Erstens muß $p_i>0$ für alle $1\leq i\leq n$ gelten. Dies ist aber schon in den Bedingungen des obigen Satzes enthalten.
Zweitens darf kein $l_i = 0$ sein, da sonst (mindestens) eine Zeile und eine Spalte in der Matrix verschwinden würden. Dies ist genau

dann erfüllt, wenn $S(t_i) < S(0) = 1$ ist für alle $1 \le i \le n$. Für das Stieltjes-Integral gilt nämlich:

$$\int_0^{t_i} \frac{-dS(u)}{S(u)R(u)} \geq \int_0^{t_i} -dS(u) = S(0) - S(t_i) > 0$$

da $1 \le (S(u)R(u))^{-1} \le (S(t_i)R(t_i))^{-1}$ für alle $0 \le u \le t_i$ gilt.

Wäre umgekehrt $S(t_i) = S(0) = 1$, dann wäre

$$0 \le \int_0^{t_i} \frac{-dS(u)}{S(u)R(u)} \le S(t_i)R(t_i) \int_0^{t_i} -dS(u) = 0$$

Sind alle $p_i \neq 0$, dann kann man sie aus den Zeilen und Spalten der Matrix T eliminieren, und wir erhalten

$$\text{rg } T = \text{rg}(l_{\min(i,j)}).$$

Da $-S(u)$ monoton wachsend und der Integrand $((S(u)R(u))^{-1}$ positiv ist, gilt $l_1 \le l_2 \le \ldots \le l_n$. Wäre nun $l_i = l_{i+1}$ für ein i, dann würden die i-te und die i+1-te Zeile in $(l_{\min(i,j)})$ übereinstimmen und daher rg T < n gelten. Als weitere notwendige Bedingung dafür, daß T regulär ist, ergibt sich damit: $0 < l_1 < l_2 < \ldots < l_n$, oder anders: es muß gelten

$$l_i - l_{i+1} = \int_{t_i}^{t_{i+1}} \frac{-dS(u)}{S(u)R(u)} > 0$$

für alle $1 \le i \le n-1$.

Dies ist nach der selben Argumentation wie vorher genau dann erfüllt, wenn $S(t_i) > S(t_{i+1})$ gilt für $1 \le i \le n-1$:

Satz: Die Kovarianzmatrix T ist genau dann regulär, wenn $0 < S(t_n) < S(t_{n-1}) < \ldots < S(t_1)$ gilt.

Beweis: Seien $k_1,\ldots k_n$ paarweise verschiedene Zahlen, alle von 0 verschieden; durch Induktion nach n zeigen wir, daß $rg\ (k_{min(i,j)}) = n$.
Für n=1 ist die Behauptung trivial; angenommen, die Behauptung stimmt für n-1. Wir subtrahieren in der Matrix die erste Spalte von allen anderen und erhalten eine Matrix für die gilt:
i) in der ersten Zeile sind alle Eintragungen gleich 0 außer in der ersten Spalte;
ii) streicht man die erste Zeile und erste Spalte, dann erhält man eine (n-1)x(n-1), auf die die Induktionsvoraussetzung zutrifft;
aus i) und ii) ergibt sich, daß der Rang der ursprünglichen Matrix gleich n ist.

Daraus folgt übrigens auch daß die Varianz-Kovarianz-Matrix aus 2.1. regulär ist.

Die wahren Werte der Eintragungen von T sind nicht bekannt; in den zu konstruierenden Teststatistiken sind sie daher durch konsistente Schätzungen zu ersetzen.
Es ist also eine Schätzung für den Ausdruck

$$t_{ij} = S(t_i)S(t_j)\int_0^{t_j} \frac{-dS(u)}{S(u)R(u)}$$

zu finden.
Als Schätzung für S(u) nimmt man klarerweise den Produkt-Limit-Schätzer $\hat{S}(u)$. Weiters ist $R(u) = S(u)G(u) = P(X_i>u)$ die Wahrscheinlichkeit, daß eine beobachtete (zensierte oder

unzensierte) Zeit größer als u ist. Eine natürliche Schätzung für R(u) ist daher gegeben durch

$$\hat{R}(u) := n_u/N,$$

wobei

$$n_u = |\{j|X_j \geq u\}|$$

die Anzahl der Individuen, die zum Zeitpunkt u noch unter Beobachtung stehen (die Risikomenge zum Zeitpunkt u), bedeutet. (Genau genommen wird damit R(u-0) geschätzt; aufgrund der Stetigkeit von R(t) macht dies nichts aus; außerdem werden wir die Schätzung mit allen derartigen Möglichkeiten durchspielen.)

Ersetzt man nun in $\int_0^t \frac{-dS(u)}{S(u)R(u)}$ S durch $\hat{S}$ und R durch $\hat{R}$, dann

dann sind Integrand und Integrator Treppenfunktionen, die beide zu den Zeitpunkten $t_{(j)}$ Sprungstellen besitzen. Das Riemann-Stieltjes-Integral ist dann nicht definiert. (Jede konkrete Realisierung von) $-d\hat{S}(u)$ kann aber als diskretes Wahrscheinlichkeitsmaß auf R^+ aufgefaßt werden, wenn man für den Fall, daß $\hat{S}(t)$ nicht definiert ist, $\hat{S}(t)$ gleich 0 setzt. Dann ist ein derartiger Ausdruck sehr wohl sinnvoll, der Wert des Integrals hängt aber wesentlich davon ab, welchen Wert der Integrand in den Massepunkten $t_{(j)}$ des Wahrscheinlichkeitsmaßes $-d\hat{S}(u)$ hat, d.h. ob man dem Integranden in diesen Punkten den rechtsseitigen oder den linksseitigen Limes zuordnet. Zunächst gilt:

$-d\hat{S}(u) = 0$ für $u \neq t_{(j)}$

$$-d\hat{S}(t_{(j)}) = \hat{S}(t_{(j)}-0) - \hat{S}(t_{(j)}+0) =$$
$$= \hat{S}(t_{(j)}-0) - \hat{S}(t_{(j+1)}) =$$

$$= \prod_{l:t_{(l)}<t_{(j)}} \frac{n_l-d_l}{n_l} - \prod_{l:t_{(l)}<t_{(j+1)}} \frac{n_l-d_l}{n_l} =$$

$$= \hat{S}(t_{(j+1)})\left(\frac{n_j}{n_j-d_j} - 1\right) = \hat{S}(t_{(j+1)})\frac{d_j}{n_j-d_j} =$$

$$= \hat{S}(t_{(j)})\left(1 - \frac{n_j-d_j}{n_j}\right) = \hat{S}(t_{(j)})\frac{d_j}{n_j}$$

Wir wollen nun die verschiedenen Möglichkeiten für den Integranden durchspielen:

1.Wählt man in den Sprungstellen $t_{(j)}$ für den Integranden jeweils den rechtsseitigen Limes, dann erhält man:

$$\hat{S}(t_{(j)}+0) = \hat{S}(t_{(j+1)})$$

und

$$\hat{R}(t_{(j)}+0) = (u_j+0)/N = (n_j-d_j-c_j)/N$$

wobei c_j die Anzahl der zensierten Beobachtungen zum Zeitpunkt $t_{(j)}$ bedeutet. Als erste Schätzung für t_{ij} erhalten wir daher:

$$\hat{t}_{ij}(1) = \hat{S}(t_i)\hat{S}(t_j)N \sum_{l:t_{(l)}<t_i} \frac{d_l}{(n_l-d_l)(n_l-d_l-c_l)}$$

wobei $t_i \leq t_j$ angenommen wurde.

2.Wählt man im Integranden jeweils den linksseitigen Limes, dann erhält man

$$\hat{S}(t_{(j)}-0) = \hat{S}(t_{(j)})$$

und

$$\hat{R}(t(j)-0) = \hat{R}(t_{(j)}) = n_j/N$$

Als zweite Schätzung für t_{ij} erhalten wir somit

$$\hat{t}_{ij}(2) = \hat{S}(t_i)\hat{S}(t_j)N \sum_{l:t_{(l)}<t_i} \frac{d_l}{n_l^2}$$

wieder für $t_i \leq t_j$.

3.Lawless schlägt einen "Mittelweg" vor, nämlich in den Sprungstellen im Integranden für $\hat{S}$ den rechtsseitgen Limes, für $\hat{R}$ den linksseitigen Limes zu verwenden:

$$\hat{S}(t_{(j)}+0) = \hat{S}(t_{(j+1)})$$

und

$$\hat{R}(t_{(j)}-0) = \hat{R}(t_{(j)}) = n_j/N$$

Als dritte Schätzung für t_{ij} erhält man dann:

$$\hat{t}_{ij}(3) = \hat{S}(t_i)\hat{S}(t_j)N \sum_{l:t_{(l)}<t_i} \frac{d_l}{n_l(n_l-d_l)}$$

4. Wählt man umgekehrt für $\hat{S}$ den linksseitigen, für $\hat{R}$ den rechtseitigen Limes, erhält man als vierte Schätzung für t_{ij}:

$$\hat{t}_{ij}(4) = \hat{S}(t_i)\hat{S}(t_j)N \sum_{l:t_{(l)}<t_i} \frac{d_l}{n_l(n_l-d_l-c_l)}$$

Offensichtlich gilt $\hat{t}_{ij}(2) \leq \hat{t}_{ij}(3) \leq \hat{t}_{ij}(4) \leq \hat{t}_{ij}(1)$.

5.Andererseits kann das Integral in der Kovarianzfunktion auch noch dargestellt werden als

$$t_{ij} = S(t_i)S(t_j)\int_0^{t_i} \frac{dF(u)}{R^2(u)}$$

Eine natürliche Schätzung für F(u) ist die empirische Verteilungsfunktion der tatsächlich eingetretenen Todesfälle:

$$\hat{F}(u) := N^{-1} \sum_{l:x_{(l)}<u} k_{(l)} = N^{-1} \sum_{l:t_{(l)}<u} d_l$$

Dann ist

$$d\hat{F}(t_{(k)}) = \hat{F}(t_{(k)}+0) - \hat{F}(t_{(k)}-0) =$$
$$= \hat{F}(t_{(k+1)}) - \hat{F}(t_{(k)}) = d_k/N$$

Wählt man für den Integranden $\hat{R}^{-2}(u)$ in den Punkten $t_{(k)}$ den

linkseitigen Limes:

$$\hat{R}(t_{(k)}-0) = \hat{R}(t_{(k)}) = n_k/N,$$

dann erhält man gerade die vorher erwähnte Schätzung $\hat{t}_{ij}(2)$.
Wählt man für den Integranden $\hat{R}^{-2}(u)$ in den Punkten $t_{(k)}$ den rechtseitigen Limes:

$$\hat{R}(t_{(k)}+0) = (n_k-d_k-c_k)/N,$$

dann erhält man eine weitere Schätzung für die Kovarianz:

$$\hat{t}_{ij}(5) = \hat{S}(t_i)\hat{S}(t_j)N \sum_{l:t_{(l)}<t_i} \frac{d_l}{(n_l-d_l-c_l)^2}$$

Zunächst gilt für diese Schätzungen:

$\hat{t}_{ij}(2) < \hat{t}_{ij}(3) \leq \hat{t}_{ij}(4) < \hat{t}_{ij}(1) \leq \hat{t}_{ij}(5)$.

In unserem Fall gilt aber:

Die Verteilungsfunktionen S(t) und G(t) sind stetig, daher gilt mit Wahrscheinlichkeit 1 für alle l:

i) $d_l = 1$

ii) $c_l = 0$

Das erste folgt daraus, daß die Ankunftszeiten T_i unabhängig sind und stetige Verteilung besitzen, das zweite daraus, daß die Zensierungsvariablen von den Ankunftszeiten unabhängig sind und ebenfalls stetige Verteilung besitzen. Daher gilt fast sicher:

$\hat{t}_{ij}(2) = \hat{S}(t_i)\hat{S}(t_j)N\Sigma 1/n_k^2 < \hat{t}_{ij}(3) = \hat{t}_{ij}(4) =$

$= \hat{S}(t_i)\hat{S}(t_j)N\Sigma 1/n_k(n_k-1) < \hat{t}_{ij}(1) = \hat{t}_{ij}(5) =$

$= \hat{S}(t_i)\hat{S}(t_j)N\Sigma 1/(n_k-1)^2$, wobei über alle Todeszeiten $t_{(k)}<t_i$ summiert wird.

Damit gilt fast sicher:

$0 \leq \hat{t}_{ij}(5) - \hat{t}_{ij}(2) = \hat{S}(t_i)\hat{S}(t_j)N\Sigma[1/(n_k-1)^2 - 1/n_k^2] = (*)$

wobei wieder über alle $t_{(k)}<t_i$ summiert wird. Für den einzelnen

Summanden gilt:

$$\frac{1}{(n_k-1)^2} - \frac{1}{n_k^2} = \frac{n_k^2-n_k^2+2n_k-1}{n_k^2(n_k-1)^2} = \frac{2-1/n_k}{n_k(n_k-1)^2} \leq \frac{2}{(n_k-1)^3}$$

Sei nun r-1 die Anzahl der Beobachtungen $X_j < t_i$,

d.h.: $r = \Sigma I\{X_j<t_i\} + 1$;

dann läßt sich (*) folgendermaßen abschätzen:

$$(*) \leq \hat{S}(t_i)\hat{S}(t_j)\, N \sum_{i=0}^{r} \frac{2}{(N-i)^3} \leq \hat{S}(t_i)\hat{S}(t_j)N(r+1)\frac{2}{(N-r)^3} =$$

$$\hat{S}(t_i)\hat{S}(t_j)N(r+1)\frac{2}{N^3(1-r/N)^3} = \hat{S}(t_i)\hat{S}(t_j)\frac{2(r+1)/N}{N(1-r/N)^3} = O(1)/N$$

Dieser letzte Ausdruck strebt in Wahrscheinlichkeit gegen 0, da (r+1)/N und r/N in Wahrscheinlichkeit gegen einen Wert,der kleiner als 1 ist, strebt, nämlich gegen $1 - R(t_i)$.

Für unseren Fall streben also alle fünf Schätzungen gegen den gleichen Wahrscheinlichkeitslimes. Es bleibt noch zu zeigen, daß sie tatsächlich gegen das in der Formel für die Kovarianzfunktion vorkommende Integral streben.

$\hat{F}$ - oder genauer: jede konkrete Realisierung $\hat{F}(\omega)$, aufgefaßt als Funktion von t (wir wollen diese Richtigstellung fürderhin stillschweigend voraussetzten) - definiert ein diskretes Wahrscheinlichkeitsmaß auf R^+. $1/\hat{R}^2(t)$ ist monoton wachsend, daher ist:

$$\int_0^u \frac{d\hat{F}(t)}{\hat{R}^2(t-0)} \leq \int_0^u \frac{d\hat{F}(t)}{\hat{R}^2(t)} \leq \int_0^u \frac{d\hat{F}(t)}{\hat{R}^2(t+0)}$$

Es wurde schon gezeigt, daß der erste und der dritte Ausdruck in Wahrscheinlichkeit gegen den gleichen Limes konvergieren. Es ist noch zu zeigen, daß tatsächlich eine konsistente Schätzung des in der Formel für die Kovarianzfunktion vorkommenden Integrals vorliegt. Es muß also gezeigt werden, daß

$$p\lim_{N\to\infty} \int_0^u \frac{d\hat{F}(t)}{\hat{R}^2(t)} = \int_0^u \frac{dF(t)}{R^2(t)} \qquad \text{gilt}$$

Seien $\epsilon_1,\epsilon_2>0$ beliebig. Sei $u<v$ derart, daß $0<R(v)<R(u)$. Die Abbildung $x \to 1/x^2$ ist gleichmäßig stetig auf dem Intervall $[R(v),1]$, daher gibt es zu dem gegebenen $\epsilon_1>0$ ein $\delta>0$, o.B.d.A.: $\delta<R(u)-R(v)$, sodaß aus $|x-y|<\delta$ für $x,y \in [R(v),1]$ stets folgt $|1/x^2-1/y^2|<\epsilon_1$. Nach dem Satz von Glivenko-Cantelli (siehe z.B. A.Rényi (1977)) existiert ein $N_1 \in N$, sodaß für $N\geq N_1$ gilt:

$$P(\sup_{t\in R^+}|\hat{R}(t)-R(t)|<\delta) \geq 1-\epsilon_2$$

Ist $t \in [0,u]$, dann nimmt $R(t)$, und $\hat{R}(t)$ mit Wahrscheinlichkeit $\geq 1-\epsilon_2$ nur Werte im Intervall $[R(u)-\delta,1] \subseteq [R(v),1]$ an. Daraus folgt aus der Wahl von δ, daß

$$P(\sup_{t\in[0,u]}|1/\hat{R}^2(t) - 1/R^2(t)| \leq \epsilon_1) \geq 1-\epsilon_2$$

wenn $N\geq N_1$.

Nun muß die Differenz zwischem exaktem und geschätztem Wert des Integrals abgeschätzt werden:

$$\left|\int_0^u \frac{dF(t)}{R^2(t)} - \int_0^u \frac{d\hat{F}(t)}{\hat{R}^2(t)}\right| \leq$$

$$\leq \left|\int_0^u \frac{dF(t)}{R^2(t)} - \int_0^u \frac{d\hat{F}(t)}{R^2(t)}\right| + \left|\int_0^u (1/R^2(t)-1/\hat{R}^2(t))d\hat{F}(t)\right|$$

Für den zweiten Ausdruck =: B gilt:

$$B \leq \int_0^u |1/R^2(t)-1/\hat{R}^2(t)|d\hat{F}(t) \leq \sup_{t\in[0,u]}|1/R^2(t)-1/\hat{R}^2(t)|$$

Damit ist für $N\geq N_1$: $P(B\leq\epsilon_1) \geq 1-\epsilon_2$.

Um zu zeigen, daß auch der erste Ausdruck =: A mit

Wahrscheinlichkeit 1 gegen 0 konvergiert, bedarf es einiger zusätzlicher Überlegungen. Zunächst besagt der Satz von Glivenko-Cantelli, daß

$$P(\sup_{t\in R^+} |\hat{F}(t) - F(t)| \underset{N\to\infty}{\longrightarrow} 0) = 1$$

Bezeichnet man nun die von F(t) resp. $\hat{F}(t)$ definierten Wahrscheinlichkeitsmaße mit dF resp. $d\hat{F}$, dann folgt aus diesem Satz, daß für alle halboffenen Intervalle $(a,b] \subseteq R^+$ gilt, daß

$$P(d\hat{F}(a,b] \underset{N\to\infty}{\longrightarrow} dF(a,b]) = 1$$

Die Klasse der halboffenen Intervalle (inklusive der leeren Menge) auf R^+ ist (i) abgeschlossen unter Bildung endlicher Durchschnitte; (ii) jede offene Menge in R^+ ist eine endliche oder abzählbare Vereinigung von derartigen halboffenen Intervallen. Damit folgt aus einem Theorem aus der Theorie der Wahrscheinlichkeitsmaße auf metrischen Räumen (siehe Billingsley (1968),Theorem 2.2.,p.14), daß $d\hat{F}$ (fast sicher) schwach gegen dF konvergiert: es gilt

$$P(\lim_{N\to\infty} \int_{R^+} f d\hat{F} = \int_{R^+} f dF) = 1$$

für jede beschränkte, gleichmäßig stetige Funktion f auf R^+.
Da F stetig ist, gibt es ein $\delta_1>0$, sodaß
$dF(u,u+\delta_1] = F(u+\delta_1) - F(u) < \epsilon_1$; weiters gibt es ein $N_2 \in N$, sodaß $P(\hat{F}(u+\delta_1)-\hat{F}(u)<\epsilon_1) \geq 1-\epsilon_2$ für $N\geq N_2$.
Sei nun f(t) definiert durch:

$$f(t) := \begin{cases} 1/R^2(t) & \text{für } t\leq u \\ 0 & \text{für } t>u+\delta_1 \\ \text{linear dazwischen} & \end{cases}$$

f ist dann sicherlich beschränkt und gleichmäßig stetig auf R^+.

Daher gilt mit Wahrscheinlichkeit 1:

$$\lim_{N\to\infty} \int_{R^+} f d\hat{F} = \int_{R^+} f dF$$

Nun ist

$$\int_0^u \frac{dF(t)}{R^2(t)} - \int_0^u \frac{d\hat{F}(t)}{R^2(t)} =$$

$$\int_{R^+} f(t)dF(t) - \int_{R^+} f(t)d\hat{F}(t) - \int_u^{u+\delta_1} f(t)dF(t) + \int_u^{u+\delta_1} f(t)d\hat{F}(t)$$

Daher ist:

$$\left| \int_0^u \frac{dF(t)}{R^2(t)} - \int_0^u \frac{d\hat{F}(t)}{R^2(t)} \right| \leq$$

$$\leq \left| \int_{R^+} f(t)dF(t) - \int_{R^+} f(t)d\hat{F}(t) \right| + \frac{1}{R^2(u)} \int_u^{u+\delta_1} dF(t) + \frac{1}{R^2(u)} \int_u^{u+\delta_1} d\hat{F}(t)$$

Der erste Ausdruck konvergiert mit Wahrscheinlichkeit 1 gegen 0, der zweite ist nach der Wahl von δ_1 kleiner als $\epsilon_1 \cdot R^{-2}(u)$, und vom dritten wissen wir, daß er mit Wahrscheinlichkeit $\geq 1-\epsilon_2$ kleiner als $\epsilon_1 \cdot R^{-2}(u)$ wird, wenn nur N hinreichend groß wird; es ist also

$$P\left(\lim_{N\to\infty} \sup \int_u^{u+\delta_1} d\hat{F}(t) \leq \epsilon_1 \cdot R^{-2}(u)\right) \geq 1-\epsilon_2$$

Insgesamt haben wir damit gezeigt, daß

$$P\left(\lim_{N\to\infty} \sup(A+B) \leq K_1 \cdot \epsilon_1\right) \geq 1-K_2 . \epsilon_2$$

Da $\epsilon_1, \epsilon_2 > 0$ beliebig waren, haben wir gezeigt, daß für $N \to \infty$

$\int_0^u \frac{d\hat{F}(t)}{\hat{R}^2(t)}$ mit Wahrscheinlichkeit 1 gegen $\int_0^u \frac{dF(t)}{R^2(t)}$ konvergiert

Die Schätzungen $\hat{t}_{ij}(1)$ bis $\hat{t}_{ij}(5)$ sind für den Fall, daß die Überlebensfunktionen S(t) und G(t) stetig sind, konsistente Schätzungen der Kovarianzen t_{ij}. Wir wollen in Hinkunft die von

Lawless vorgeschlagene Schätzung $\hat{t}_{ij}(3)$ verwenden, weil sie "in der Mitte liegt"; jede andere wäre aber auch "erlaubt".

3.2. Tests mittels Schätzung der Überlebensfunktion durch den PLS

Damit ergibt sich wieder die Möglichkeit mit Hilfe des Hausman-Prinzips Tests für die Nullhypothese

$$H_0: T_i \text{ ist exponentialverteilt}$$

zu konstruieren. Wie in 1. benützt man wieder die Tatsache, daß die T_i genau dann exponentialverteilt sind, wenn die Überlebensfunktion der Funktionalgleichung

$$S(x+y) = S(x)S(y)$$

genügt. Die Differenz $|\hat{S}(x+y) - \hat{S}(x)\hat{S}(y)|$ darf dann nicht zu groß werden.

Vorgangsweise: man wähle $x,y,x<y$ mit $S(x+y)G(x+y) > 0$. Bildet man für die Werte $x,y,x+y$ die Produkt-Limit-Schätzungen, dann ist

$$\sqrt{N}(\hat{S}(x)-S(x),\hat{S}(y)-S(y),\hat{S}(x+y)-S(x+y))$$

asymptotisch normalverteilt mit Mittelwert 0 und Varianz-Kovarianz-Matrix T wie vorher beschrieben.

Nach dem im ersten Abschnitt zitierten Satz aus der Theorie der Grenzverteilungen folgt dann:

$$\sqrt{N}(\hat{S}(x+y)-S(x+y)-\hat{S}(x)\hat{S}(y)+S(x)S(y)) =$$
$$= \sqrt{N}(\hat{S}(x+y)-\hat{S}(x)\hat{S}(y))$$

ist ebenfalls asymptotisch normalverteilt mit Varianz D'TD, wobei $D' = (-S(y),-S(x),1)$. Setzt man

$$Q(z) = \int_0^z \frac{-dS(u)}{S(u)R(u)}$$

dann ist T gegeben durch

$$T = \begin{bmatrix} S^2(x)Q(x) & S(x)S(y)Q(x) & S(x)S(x+y)Q(x) \\ S(x)S(y)Q(x) & S^2(y)Q(y) & S(y)S(x+y)Q(y) \\ S(x)S(x+y)Q(x) & S(y)S(x+y)Q(y) & S^2(x+y)Q(x+y) \end{bmatrix}$$

Daraus erhält man:

$$TD = \begin{bmatrix} (-2S^2(x)S(y) + S(x)S(x+y))Q(x) \\ -S(x)S^2(y)Q(x) - S^2(y)S(x)Q(y) + S(y)S(x+y)Q(y) \\ -S(x)S(y)S(x+y)Q(x) - S(x)S(y)S(x+y)Q(y) + S^2(x+y)Q(x+y) \end{bmatrix}$$

Für die asymptotische Varianz ergibt sich damit wiederum:

$$\sigma^2 = D'TD = (2S^2(x)S^2(y) - S(x)S(y)S(x+y))Q(x) +$$
$$+ S^2(x)S^2(y)Q(x) + S^2(y)S^2(x)Q(y) - S(x)S(y)S(x+y)Q(y) -$$
$$- S(x)S(y)S(x+y)Q(x) - S(x)S(y)S(x+y)Q(y) + S^2(x+y)Q(x+y)$$

Unter H_0 gilt $S(x)S(y) = S(x+y)$. Damit vereinfacht sich die Formel für die asymptotische Varianz folgendermaßen:

$$\sigma^2 = S^2(x)S^2(y)Q(x) - S(x)S(y)S(x+y)Q(y) + S^2(x+y)Q(x+y) =$$
$$= S(x)S(y)S(x+y)(Q(x+y)-Q(y)+Q(x))$$

Wenn wir als Schätzung für Q(z) die von Lawless vorgeschlagene wählen:

$$\hat{Q}(z) = N \sum_{l:t_{(l)}<z} \frac{d_l}{n_l(n_l-d_l)}$$

dann ist

$$\hat{Q}(x+y)-\hat{Q}(y)+\hat{Q}(x) = N\Sigma\frac{d_l}{n_l(n_l-d_l)}$$

wobei in der letzten Summe über alle Zeitpunkte $t_{(l)}$ summiert wird, für die entweder $t_{(l)}<x$ oder $y\leq t_{(l)}<x+y$ gilt.

Zusammenfassend ergibt sich damit als Teststatistik:

$$H := \frac{\hat{S}(x+y) - \hat{S}(x)\hat{S}(y)}{\left[\hat{S}(x)\hat{S}(y)\hat{S}(x+y) \sum \frac{d_1}{n_1(n_1-d_1)}\right]^{\frac{1}{2}}}$$

wobei sich der Summationsindex wieder über alle Zeitpunkte $t_{(1)}$ mit $t_{(1)}<x$ oder $y\leq t_{(1)}<x+y$ erstreckt. Die Statistik H ist unter der Nullhypothese wieder asymptotisch N(0,1)-verteilt. Der Zähler in H kann folgendermaßen dargestellt werden:

$$\hat{S}(x+y) - \hat{S}(x)\hat{S}(y) = \prod_{t_{(1)}<y} \frac{n_1-d_1}{n_1}\left[\prod_{y\leq t_{(1)}<x+y} \frac{n_1-d_1}{n_1} - \prod_{t_{(1)}<x} \frac{n_1-d_1}{n_1}\right]$$

Wenn wir in der Formel für die Varianz für den Ausdruck $S(x)S(y)S(x+y)$ einen äquivalenten Ausdruck wählen, erhalten wir folgende, unter der Nullhypothese asymptotisch äquivalente Teststatistiken:

Für $S(x)S(y)S(x+y) = S^2(x)S^2(y)$ ergibt sich:

$$H = \frac{\sqrt{N}}{\sqrt{(\hat{Q}(x+y)-\hat{Q}(y)+\hat{Q}(x))}}\left[\frac{\hat{S}(x+y)}{\hat{S}(x)\hat{S}(y)} - 1\right]$$

Für $S(x)S(y)S(x+y) = S^2(x+y)$ ergibt sich

$$H = \frac{\sqrt{N}}{\sqrt{(\hat{Q}(x+y)-\hat{Q}(y)+\hat{Q}(x))}}\left[1 - \frac{\hat{S}(x)\hat{S}(y)}{\hat{S}(x+y)}\right]$$

Explizit lassen sich diese Statistiken folgendermaßen formulieren:

Für die erste Variante:

$$H = \frac{1}{\left[\sum_{t_{(1)}<x,\, y\le t_{(1)}<x+y} \frac{d_1}{n_1(n_1-d_1)}\right]^{\frac{1}{2}}} \left[\frac{\prod_{y\le t_{(1)}<x+y} \frac{n_1-d_1}{n_1}}{\prod_{t_{(1)}<x} \frac{n_1-d_1}{n_1}} - 1\right]$$

Für die zweite Variante:

$$H = \frac{1}{\left[\sum_{t_{(1)}<x,\, y\le t_{(1)}<x+y} \frac{d_1}{n_1(n_1-d_1)}\right]^{\frac{1}{2}}} \left[1 - \frac{\prod_{t_{(1)}<x} \frac{n_1-d_1}{n_1}}{\prod_{y\le t_{(1)}<x+y} \frac{n_1-d_1}{n_1}}\right]$$

Beide Statistiken sind unter der Nullhypothese wieder asymptotisch N(0,1)-verteilt.

Wie in Abschnitt 1. besteht eine weitere Variante dieses Verfahrens darin, die Beziehung $S^p(x) = S(px)$, die unter der Nullhypothese gilt, auszunützen: wähle ein x mit $S(x)G(x)>0$ und ein p mit $0<p<1$. Dann ist die Statistik

$$\sqrt{N}(\hat{S}(px) - S(px), \hat{S}(x) - S(x))$$

asymptotisch bivariat normalverteilt mit Mittelwert 0 und Varianz-Kovarianz-Matrix T, gegeben durch

$$T = \begin{bmatrix} S^2(px)Q(px) & S(px)S(x)Q(px) \\ S(px)S(x)Q(px) & S^2(x)Q(x) \end{bmatrix}$$

Wegen des schon oft zitierten Satzes gilt damit auch, daß

$$\sqrt{N}(\hat{S}(x) - \hat{S}^{1/p}(px))$$

asymptotisch normalverteilt ist mit Mittelwert 0 und Varianz σ^2 =

D'TD, wobei D'= $(-1/pS^{(1-p)/p}(px),1)$ ist. Für TD ergibt sich damit

$$TD = \begin{bmatrix} -1/pS^{(1+p)/p}(px)Q(px) + S(px)S(x)Q(px) \\ -1/pS^{1/p}(px)S(x)Q(px) + S^2(x)Q(x) \end{bmatrix}$$

Daher ist

$$\sigma^2 = 1/p^2S^{2/p}(px)Q(px) - 1/pS^{1/p}(px)S(x)Q(px) -$$
$$- 1/pS^{1/p}(px)S(x)Q(px) + S^2(x)Q(x)$$

Unter der Nullhypothese gilt $S^2(x) = S^{1/p}(px)S(x) = S^{2/p}(px)$, daher ist für diesen Fall:

$$\sigma^2 = S^{1/p}(px)S(x)(1/p(1/p-2)Q(px)+Q(x)).$$

Unter H_0 ist daher die Teststatistik

$$H := \frac{\sqrt{N}(\hat{S}(x) - \hat{S}^{1/p}(px))}{[\hat{S}^{1/p}(px)\hat{S}(x)(1/p(1/p-2)\hat{Q}(px)+\hat{Q}(x))]^{\frac{1}{2}}}$$

asymptotisch N(0,1)-verteilt.

Wählen wir speziell $p = \frac{1}{2}$, dann ergibt sich für die Varianz $\sigma^2 = S^2(x/2)S(x)Q(x) = S^2(x)Q(x)$ $(= S^4(x/2)Q(x))$, woraus wir die folgende, einfache Teststatistik erhalten:

$$H := \frac{1}{\left[\sum_{t_{(1)}<x} \frac{d_1}{n_1(n_1-d_1)}\right]^{\frac{1}{2}}} \left[1 - \prod_{x/2 \le t_{(1)} < x} \frac{n_1-d_1}{n_1}\right]$$

die unter der Nullhypothese wieder asymptotisch N(0,1)-verteilt ist. Für die Menge der Alternativen, gegen die diese Tests konsistent sind, gilt im wesentlichen das schon in 1. Erwähnte. Die *beidseitigen* Tests sind konsistent gegen alle Alternativen,

für die $S(x)S(y) \neq S(x+y)$, resp. $S^p(x) \neq S(px)$ gilt. Weiters eignen sich die *einseitigen* Tests gegen Alternativen, deren Ratenfunktion monoton ist. Klarerweise sprechen diese Tests nicht auf solche Alternativen an, für die für die gewählten x,y,p $S(x+y) = S(x)S(y)$ resp. $S^p(x) = S(px)$ gilt. Wie im Fall einer festen Zensierungszeit kann man dem wieder dadurch begegnen, daß für mehrere Zeitpunkte t_i die Überlebensfunktion geschätzt wird, und die Schätzwerte "simultan" verglichen werden: Man wähle ein festes x mit $R(x)>0$ und ein $n \in N$. Für $i = 1,...n$ sei $x_i = ix/n$. Für jedes dieser x_i bildet man die Produkt-Limit-Schätzung $\hat{S}(x_i)$. Wir verwenden folgende Abkürzungen: $S_i = S(x_i)$, $\hat{S}_i = \hat{S}(x_i)$, $Q_i = Q(x_i)$ und $\hat{Q}_i = \hat{Q}(x_i)$.

Die Statistik

$$\sqrt{N}(\hat{S}_1 - S_1,...\hat{S}_n - S_n)$$

ist wieder asymptotisch multivariat normalverteilt mit Mittelwert 0 und Varianz-Kovarianz-Matrix $T = (t_{ij})$, wobei t_{ij} für $i \leq j$ gegeben ist durch

$$t_{ij} = S(x_i)S(x_j)\int_0^{x_i} \frac{-dS(u)}{S(u)R(u)} = S_iS_jQ_i$$

Es gibt wieder mehrere Varianten, die Schätzungen $\hat{S}_i$ zu vergleichen:

1. Wähle als Prüfvektor

$$d = (\hat{S}_1\hat{S}_1 - \hat{S}_2,...\hat{S}_1\hat{S}_{n-1} - \hat{S}_n)'.$$

Unter der Nullhypothese ist

$\sqrt{N}d$ asymptotisch multivariat normalverteilt mit Mittelwert 0 und Varianz-Kovarianz-Matrix V, und daher

$$H := Nd'\hat{V}^{-1}d$$

asymptotisch chi^2-verteilt mit n-1 Freiheitsgraden; V ist gegeben durch V = D'TD mit folgender n x (n-1) Matrix D:

$$\begin{aligned} D_{11} &= 2S_1 \\ D_{1i} &= S_i \text{ für } 2 \le i \le n-1 \\ D_{ii} &= S_1 \text{ für } 2 \le i \le n-1 \\ D_{i+1,i} &= -1 \text{ für } 1 \le i \le n-1 \\ D_{ij} &= 0 \text{ sonst} \end{aligned}$$

Wir wollen nun wieder die explizite Gestalt von V unter H_0 ausrechnen:

$A_{ij} := (TD)_{ij} = T_{i1}S_j + T_{ij}S_1 - T_{i,j+1}$ (da die anderen Eintragungen in der j-ten Spalte von D gleich 0 sind). Daher ist

$(TD)_{ij} = S_iS_1Q_1S_j + S_iS_jQ_{min(i,j)}S_1 - S_iS_{j+1}Q_{min(i,j+1)}$

Unter H_0 ist $S_jS_1 = S_{j+1}$, daher gilt in diesem Fall:

$A_{ij} = S_iS_{j+1}(Q_1 + Q_{min(i,j)} - Q_{min(i,j+1)})$

$(D'TD)_{ij} = S_iA_{1j} + S_1A_{ij} - A_{i+1,j}$ (da die übrigen Eintragungen in der i-ten Zeile von D' gleich 0 sind). Daher ist

$$\begin{aligned} V_{ij} = (D'TD)_{ij} &= S_iS_1S_{j+1}Q_1 + \\ &+ S_1S_iS_{j+1}(Q_1 + Q_{min(i,j)} - Q_{min(i,j+1)}) - \\ &- S_{i+1}S_{j+1}(Q_1 + Q_{min(i+1,j)} - Q_{min(i+1,j+1)}) \end{aligned}$$

Unter der Nullhypothese gilt $S_iS_1 = S_{i+1}$, daher erhalten wir für $i \le j$:

$V_{ij} = S_{i+1}S_{j+1}(Q_1 + Q_i - Q_i - Q_{min(i+1,j)} + Q_{i+1})$.

Insgesamt erhalten wir für die Varianz-Kovarianz-Matrix:

$$V_{ij} = \begin{cases} S_{i+1}S_{j+1}Q_1 & \text{wenn } i<j \\ S_{i+1}S_{j+1}(Q_1 + Q(i,i+1)) & \text{wenn } i=j \end{cases}$$

wobei

$$Q(i,i+1) = \int_{x_i}^{x_{i+1}} \frac{-dS(u)}{S(u)R(u)}$$

Für $j<i$ erhält man natürlich das symmetrische Ergebnis. In den Teststatistiken ist wieder S durch $\hat{S}$ und Q durch $\hat{Q}$ zu

ersetzen; dabei ist

$$\hat{Q}(i,i+1) = \sum_{x_i \le t_{(l)} < x_{i+1}} \frac{d_l}{n_l(n_l-d_l)}$$

2. Als weitere Variante vergleichen wir $S_1{}^i$ mit S_i: Wir wählen als Prüfvektor

$d := (\hat{S}_1{}^2 - \hat{S}_2, \ldots \hat{S}_1{}^i - \hat{S}_i, \ldots \hat{S}_1{}^n - \hat{S}_n)'$

Unter der Nullhypothese ist $\sqrt{N}d$ asymptotisch multivariat normalverteilt mit Mittelwert 0 und Varianz-Kovarianz-Matrix V, daher ist

$$H = Nd'\hat{V}^{-1}d$$

asymptotisch chi^2-verteilt mit n-1 Freiheitsgraden, wobei $\hat{V}$ eine konsistente Schätzung von V ist, und V selbst gegeben ist durch $V = D'TD$ mit der folgenden n x (n-1) Matrix D:

$$D_{1i} = (i+1)S_1{}^i$$
$$D_{i+1,i} = -1$$
$$D_{ij} = 0 \text{ sonst}$$

Wir wollen V unter der Nullhypothese explizit ausrechnen. Zunächst ist

$$A_{ij} = (TD)_{ij} = T_{i1}D_{1j} + T_{i,j+1}D_{j+1,j} = S_iS_1Q_1(j+1)S_1{}^j - S_iS_{j+1}Q_{\min(i,j+1)}$$

Weiters ist $D'_{i1} = (i+1)S_1{}^i$, $D'_{i,i+1} = -1$ und $D'_{ij} = 0$ sonst. Daher ist

$$V_{ij} = D'_{i1}A_{1j} - A_{i+1,j} =$$
$$= (i+1)S_1{}^i[S_1S_1Q_1(j+1)S_1{}^j - S_1S_{j+1}Q_1] -$$
$$- S_{i+1}S_1Q_1(j+1)S_1{}^j + S_{i+1}S_{j+1}Q_{\min(i+1,j+1)}.$$

Unter der Nullhypothese gilt

$$V_{ij} = (i+1)S_1{}^ijS_1S_{j+1}Q_1 - S_{i+1}S_1Q_1(j+1)S_1{}^j + S_{i+1}S_{j+1}Q_{\min(i+1,j+1)}$$
$$= S_iS_jS_1{}^2Q_1(ij+j-j-1) + S_{i+1}S_{j+1}Q_{\min(i+1,j+1)} =$$
$$= S_{i+1}S_{j+1}[(ij-1)Q_1 + Q_{\min(i+1,j+1)}]$$

In der Teststatistik müssen die S und Q wieder durch $\hat{S}$ und $\hat{Q}$ ersetzt werden.

Der Nachteil dieser zweiten Variante besteht wieder darin, daß im Vektor d *eine* Schätzung: $\hat{S}_1$ mit allen anderen verglichen wird, und daher dieser einen Schätzung, resp. dem Zeitintervall $[0,x_1]$ ein "größeres Gewicht" zukommt.

3. Um letzeres zu vermeiden, können wir folgenden Prüfvektor wählen:

$d := (\hat{S}_1^2 - \hat{S}_2, \ldots \hat{S}_i^{i+1} - \hat{S}_{i+1}^{i}, \ldots \hat{S}_{n-1}^{n} - \hat{S}_n^{n-1})$

Wieder ist unter der Nullhypothese $\sqrt{N}d$ asymptotisch multivariat normalverteilt mit Mittelwert 0 und Varianz-Kovarianz-Matrix V und daher die Teststatistik

$$H := Nd'\hat{V}^{-1}d$$

asymptotisch chi^2-verteilt mit n-1 Freiheitsgraden, wobei $\hat{V}$ eine konsistente Schätzung von V ist und V = D'TD mit folgender n x n-1 Matrix D:

$$D_{ii} = (i+1)S_i^{\,i}$$
$$D_{i+1,i} = -iS_{i+1}^{\,i-1}$$

Wir wollen wieder V unter H_0 explizit berechnen:

$$A_{ij} := (TD)_{ij} = T_{ij}D_{jj} + T_{i,j+1}D_{j+1,j} =$$
$$= S_iS_jQ_{\min(i.j)(j+1)}S_j^{\,j} - jS_iS_{j+1}Q_{\min(i,j+1)}S_{j+1}^{\,j-1}$$

Unter H_0 ist $S_{j+1}^{\,j} = S_j^{\,j+1}$, daher ist

$$A_{ij} = S_iS_j^{\,j+1}[(j+1)Q_{\min(i,j)} - jQ_{\min(i,j+1)}].$$

Weiters ist

$$(D'TD)_{ij} = (D'A)_{ij} = D'_{ii}A_{ij} + D_{i,i+1}A_{i+1,j} =$$
$$= (i+1)Si^iS_iS_j^{\,j+1}[(j+1)Q_{\min(i,j)} - jQ_{\min(i,j+1)}] -$$
$$- iS_{i+1}^{\,i-1}S_{i+1}S_j^{\,j+1}[(j+1)Q_{\min(i+1,j)} - jQ_{\min(i+1,j+1)}].$$

Unter H_0 ist $S_{i+1}^{\,i-1}S_{i+1} = S_{i+1}^{\,i} = S_i^{\,i+1}$. Daher ist für diesen Fall:

$$V_{ij} = S_i^{\,i+1}S_j^{\,j+1}[(ij+i+j+1)Q_{\min(i,j)} - (ij+j)Q_{\min(i,j+1)} -$$

$$- (ij+i)Q_{\min(i+1,j)} + ijQ_{\min(i+1,j+1)}]$$

Für $i<j$ ist daher

$$V_{ij} = S_i^{i+1}S_j^{j+1}[Q_i(ij+i+j+1-ij-j) - Q_{i+1}(ij+i-ij)] =$$

$$= S_i^{i+1}S_j^{j+1}[(i+1)Q_i - iQ_{i+1}] =$$

$$= \underline{S_i^{i+1}S_j^{j+1}(Q_i - iQ(i,i+1))}$$

Für $i=j$ erhält man

$$V_{ii} = S_i^{2(i+1)}[(1-i^2)Q_i + i^2Q_{i+1}] = \underline{S_i^{2(i+1)}[Q_i + i^2Q(i,i+1)]}$$

Für $i>j$ erhält man

$$V_{ij} = S_i^{i+1}S_j^{j+1}[(ij+i+j+1)Q_j - (ij+j)Q_{j+1} - (ij+i)Q_j + ijQ_{j+1}] =$$

$$= S_i^{i+1}S_j^{j+1}[(j+1)Q_j - jQ_{j+1}] =$$

$$= S_i^{i+1}S_j^{j+1}(Q_j - jQ(j,j+1))$$

was wir eigentlich nicht mehr berechnen hätten müssen, da die Matrix V ja symmetrisch ist.

In der Teststatistik müssen dann die unbekannten Größen durch konsistente Schätzungen ersetzt werden.

In allen drei Fällen gilt: der Test ist konsistent gegen alle Alternativen, für deren Überlebensfunktion für mindestens ein Paar (i,j) mit $i+j \leq n$ gilt: $S(x_i)S(x_j) \neq S(x_{i+j})$.

Dies gilt, obwohl die abgeleitete explizite Gestalt von V unter H_1 nicht mehr mit D'TD übereinstimmt. Wählt man nämlich diese Darstellung, dann gilt unter

$$H_1: S^k(x_1) \neq S(x_k) \text{ für ein } k \leq n$$

daß $p \lim \sup |\hat{V}| < \infty$ und daher $p \lim Nd'\hat{V}^{-1}d = \infty$.

Wie in Abschnitt 1. macht es auch nichts aus, wenn für eine Alternative H_1 die Matrix V singulär ist, und damit $\hat{V}$ in Wahrscheinlichkeit gegen eine singuläre Matzrix konvergiert. Setzt man nämlich den Wert der Teststatistik $H = \infty$, wenn $\hat{V}$ singulär ist, dann ist der Test gegen diese Alternative sicherlich konsistent. Ähnlich wie im Fall der Typ-I-Zensierung kann man sich auch überlegen, daß in jedem Intervall $(x_i, x_{i+1}]$ mindestens eine

Beobachtung liegen muß, damit $\hat{V}$ nicht singulär ist.
Wie schon am Ende von 1.1. ausgeführt gilt auch hier: mit wachsendem n wächst die Menge der Alternativen, gegen die der Test konsistent ist, gleichzeitig nimmt bei gegebener Stichprobengröße N die Güte des Tests mit wachsendem n ab.

3.3. Tests mittels Schätzung der kummulierten Hazardfunktion durch die empirische kummulierte Hazardfunktion

Eine weitere Möglichkeit derartige Tests zu konstruieren, besteht darin, die sogenannte kummulierte Hazardfunktion:

$$H(t) := -\log S(t)$$

zu betrachten. Es gilt: T_i ist exponentialverteilt genau dann, wenn

$$H(t) \equiv \alpha t \text{ für ein } \alpha > 0.$$

Für H(t) bieten sich folgende Schätzungen an:

$$\hat{H}(t) := -\log \hat{S}(t)$$

wobei $\hat{S}$ die Produkt-Limit-Schätzung ist. Oder:

$$\tilde{H}(t) := \sum_{j:t_{(j)}<t} \frac{d_j}{n_j}$$

$\tilde{H}(t)$ heißt die empirische kummulative Hazardfunktion. Es gilt

$$\hat{H}(t) = -\sum_{t_{(j)}<t} \log\left(1 - \frac{d_j}{n_j}\right) = \sum_{t_{(j)}<t} \left(\frac{d_j}{n_j} + \frac{d_j^2}{2n_j^2} + \ldots\right.$$

$\tilde{H}(t)$ ist damit eine Approximation erster Ordnung von $\hat{H}(t)$, die recht brauchbar ist, wenn t nicht allzugroß ist. Asymptotisch sind die beiden Schätzungen wieder äquivalent: es gilt
$p\lim (\hat{H}(t) - \tilde{H}(t)) = 0$ (ein Resultat, das Breslow und Crowley unter der Voraussetzung der Stetigkeit von S(t) und G(t) bewiesen haben).

Man kann nun dieselben Tests konstruieren wie vorher unter Verwendung der Funktionalgleichung für die Überlebensfunktion einer exponentialverteilten Variablen:
$S(x+y) = S(x)S(y)$ resp. $S(px) = S^p(x)$. Da die kummulierte Hazardfunktion der negative Logarithmus der Überlebensfunktion

ist, ist eine exponentialverteilte Variable dadurch charakterisiert, daß ihre kummulierte Hazardfunktion H(t) eine lineare Funktion mit positiver Steigung ist. Dadurch werden die Teststatistiken einfacher, allerdings sollen die Zeitpunkte, zu denen die kummulierte Hazardfunktion geschätzt wird, nicht zu groß sein. Über das asymptotische Verhalten von $\tilde{H}(t)$ haben nun Breslow & Crowley unter den Annahmen, die in 3.1 gemacht wurden, folgendes gezeigt:

Satz: Sei T>0, sodaß R(T)>0; für 0<t<T konvergiert der Prozeß $\sqrt{N}(\tilde{H}(t) - H(t))$ schwach gegen einen Gauß'schen Prozeß Z(t) mit Mittelwert 0 und Kovarianzfunktion

$$\mathrm{Cov}(Z(s),Z(t)) = \int_0^s \frac{-dS(u)}{S(u)R(u)} = \int_0^s \frac{dF(u)}{R^2(u)} \qquad \text{für } s \le t$$

Bemerkung: Breslow & Crowley haben diesen Satz für eine etwas anderes definierte empirische kummulierte Hazardfunktion bewiesen:

$$\tilde{H}(t) := \sum_{X_{(i)}<t} \frac{k_{(i)}}{N+1-i}$$

Wegen der Stetigkeit der zugrundeliegenden Verteilungsfunktionen S(t) und G(t) sind die beiden Prozesse fast sicher identisch. Wie vorher läßt sich damit folgendes über die endlichdimensionalen Verteilungen feststellen:

Satz: Seien $0 < t_1 < t_2 < \ldots < t_n < T$, dann ist die Statistik

$$\sqrt{N}(\tilde{H}(t_1) - H(t_1), \ldots \tilde{H}(t_n) - H(t_n))'$$

asymptotisch multivariat normalverteilt mit Mittelwert 0 und Varianz-Kovarianz-Matrix $T = (t_{ij})$, wobei

$$t_{ij} = \int_0^{t_i} \frac{-dS(u)}{R^2(u)} = \int_0^{t_i} \frac{-dS(u)}{S(u)R(u)} \quad (\text{für } t_i \leq t_j)$$

Die einfachste Möglichkeit einen derartigen Test zu konstruieren ist wieder die folgende: Wähle x,y mit $0<x<y$ und $R(x+y)>0$; dann ist die Statistik

$$\sqrt{N}(\tilde{H}(x)-H(x), \tilde{H}(y)-H(y), \tilde{H}(x+y)-H(x+y))$$

unter der Nullhypothese asymptotisch normalverteilt mit Mittelwert 0 und Varianz-Kovarianz-Matrix T gegeben durch:

$$T = \begin{bmatrix} Q(x) & Q(x) & Q(x) \\ Q(x) & Q(y) & Q(y) \\ Q(x) & Q(y) & Q(x+y) \end{bmatrix}$$

Damit ist aber auch

$$\sqrt{N}(\tilde{H}(x+y) - \tilde{H}(x) - \tilde{H}(y))$$

asymptotisch normalverteilt mit Mittelwert 0 und Varianz $\sigma^2 = D'TD$, wobei $D' = (-1,-1,1)$. Dann ist

$$TD = \begin{bmatrix} -Q(x) \\ -Q(x) \\ -Q(x)-Q(y)+Q(x+y) \end{bmatrix} \quad \text{und daher ergibt sich für die Varianz:}$$

$$\sigma^2 = D'TD = Q(x) + Q(x) - Q(x) - Q(y) + Q(x+y) =$$

$$= Q(x+y) - Q(x,y) = Q(x) + Q(y,x+y)$$

wobei

$$Q(w,v) = \int_w^v \frac{dF(u)}{R^2(u)} = \int_w^v \frac{-dS(u)}{S(u)R(u)}$$

Ersetzt man dieses Integral wieder durch die schon erwähnte

konsistente Schätzung, dann erhält man folgende einfache Teststatistik:

$$H := \frac{\sum_{y\le t_{(k)}<x+y} \frac{d_k}{n_k} - \sum_{t_{(k)}<x} \frac{d_k}{n_k}}{\left[\sum_{t_{(k)}<x} \frac{d_k}{n_k(n_k-d_k)} + \sum_{y\le t_{(k)}<x+y} \frac{d_k}{n_k(n_k-d_k)}\right]^{\frac{1}{2}}}$$

welche unter der Nullhypothese asymptotisch N(0,1)-verteilt ist. Ist $H(x+y) > H(x)+H(y)$, dann gilt $p\lim H = \infty$; ist $H(x+y) < H(x)+H(y)$, dann gilt $p\lim H = -\infty$. Insbesondere eignet sich dieser Test gegen Alternativen mit monotoner Hazardrate: ist r(t) auf (0,z) monoton wachsend (fallend), dann gilt $H(z) > H(x)+H(y)$ ($H(z) < H(x)+H(y)$) für beliebige $x,y\ge 0$ mit $x+y=z$.

Eine weitere Möglichkeit benützt die Beziehung $H(rx) \equiv rH(x)$, welche für die kummulierte Hazardfunktion einer exponentialverteilten Variablen charakteristisch ist:

Man wähle ein $x>0$ mit $R(x)>0$ und ein p mit $0<p<1$; dann ist

$$\sqrt{N}(\tilde{H}(px) - H(px), \tilde{H}(x) - H(x))$$

asymptotisch bivariat normalverteilt mit Mittelwert 0 und Varianz-Kovarianz-Matrix T gegeben durch

$$T = \begin{bmatrix} Q(px) & Q(px) \\ Q(px) & Q(x) \end{bmatrix}$$

Damit ist die Statistik

$$\sqrt{N}(p\tilde{H}(x) - \tilde{H}(px))$$

asymptotisch $N(0,\sigma^2)$-verteilt, wobei die Varianz gegeben ist durch $\sigma^2 = D'TD$ mit $D' = (-1,p)$. Damit ergibt sich für die Varianz:

$\sigma^2 = Q(px) - pQ(px) - pQ(px) + p^2Q(x) = (1-2p)Q(px) + p^2Q(x)$.

Als Teststatistik erhalten wir somit:

$$H := \frac{p\sum_{t_{(k)}<x} \frac{d_k}{n_k} - \sum_{t_{(k)}<px} \frac{d_k}{n_k}}{\left[(1-2p)\sum_{t_{(k)}<px} \frac{d_k}{n_k(n_k-d_k)} + p^2\sum_{t_{(k)}<x} \frac{d_k}{n_k(n_k-d_k)}\right]^{\frac{1}{2}}}$$

Wählt man $p = \frac{1}{2}$, dann erhält man folgende einfachere Teststatistik:

$$H := \frac{\sum_{t_{(k)}<x} \frac{d_k}{n_k} - 2\sum_{t_{(k)}<x/2} \frac{d_k}{n_k}}{\left[\sum_{t_{(k)}<x} \frac{d_k}{n_k(n_k-d_k)}\right]^{\frac{1}{2}}} = \frac{\sum_{x/2\le t_{(k)}<x} \frac{d_k}{n_k} - \sum_{t_{(k)}<x/2} \frac{d_k}{n_k}}{\left[\sum_{t_{(k)}<x} \frac{d_k}{n_k(n_k-d_k)}\right]^{\frac{1}{2}}}$$

Diese Statistiken sind unter der Nullhypothese wieder asymptotisch N(0,1)-verteilt. Außerdem gilt $p\lim H = \infty$, wenn $pH(x) > H(px)$, und $p\lim H = -\infty$, wenn $pH(x) < H(px)$.

Der Wert $p = \frac{1}{2}$ ist aber i.a. nicht der Wert von p, sodaß der Test bei gegebenem x und bei einer gegebenen Alternative die größte Power hat. Die wichtigste Bestimmungsgröße für die Power bei einer bestimmten Alternative H_1 ist wohl die Größe

$$f(p) := |H(px) - pH(x)|$$

Auf der Schätzung dieser Differenz beruht ja dieser Test. Bildet man die Ableitung f'(p), dann erhält man:

$$f'(p) = \pm(H'(px)x - H(x)) = \pm(r(px)x - \int_0^x r(t)dt)$$

Setzt man f'(p) = 0, dann erkennt man, daß die beste Teilung des Intervalls durch einen Punkt gegeben wäre, in dem r(t) den mittleren Wert annimmt, d.h. der die Gleichung

$$r(px) = x^{-1}\int_0^x r(t)dt$$

erfüllt. Ist r(t) streng monoton und stetig, dann ist dieser Punkt

immer eindeutig bestimmt.

Im folgenden sollen noch drei Varianten von Teststatistiken konstruiert werden, die darauf abzielen, die Menge der Alternativen, gegen die der Test konsistent ist zu vergrößern. Dabei werden wieder an mehreren Stellen in einem Intervall [0,x] die Werte für H(t) geschätzt und miteinander verglichen.

Vorgangsweise: Man Wähle ein festes x>0 mit R(x)>0 und ein n∈N. Für i = 1,...n sei $x_i = ix/n$, weiters seien $\tilde{H}_i = \tilde{H}(x_i)$, $H_i = H(x_i)$ und Q_i, $\hat{Q}_i$ etc. wie im vorigen Abschnitt. Die Statistik

$$\sqrt{N}(\tilde{H}_1 - H_1,\ldots\tilde{H}_n - H_n)'$$

ist asymptotisch multivariat normalverteilt mit Mittelwert 0 und Varianz-Kovarianz-Matrix $T = (t_{ij}) = (Q_{\min(i,j)})$.

1.Variante: Wähle als Prüfvektor:

$d := (2\tilde{H}_1 - \tilde{H}_2,..\tilde{H}_1 + \tilde{H}_i - \tilde{H}_{i+1},..\tilde{H}_1 + \tilde{H}_{n-1} - \tilde{H}_n)'$

Dann ist $\sqrt{N}d$ asymptotisch multivariat normalverteilt mit Mittelwert 0 und Varianz-Kovarianz-Matrix V = D'TD und daher die Teststatistik

$$H := Nd'\hat{V}^{-1}d$$

asymptotisch chi^2-verteilt mit n-1 Freiheitsgraden, wobei $\hat{V}$ eine konsistente Schätzung von V ist und die n x (n-1) -Matrix D gegeben ist durch

$$\begin{aligned} D_{11} &= 2 \\ D_{1i} &= 1 \quad 2\le i\le n-1 \\ D_{ii} &= 1 \quad 2\le i\le n-1 \\ D_{i+1,i} &= -1 \quad 1\le i\le n-1 \\ D_{ij} &= 0 \quad \text{sonst} \end{aligned}$$

Daher ergibt sich für D'TD:

$$A_{ij} := (TD)_{ij} = T_{i1} + T_{ij} - T_{i,j+1} =$$

$$= Q_1 + Q_{\min(i,j)} - Q_{\min(i,j+1)}$$

$$V_{ij} = (D'TD)_{ij} = A_{1j} + A_{ij} - A_{i+1,j} =$$

$$= Q1 + Q_1 - Q_1 +$$

$$+ Q_1 + Q_{\min(i,j)} - Q_{\min(i,j+1)} +$$

$$- Q_1 - Q_{\min(i+1,j)} + Q_{\min(i+1,j+1)}$$

Wir erhalten damit für die Matrix V:

$$V_{ij} = \begin{cases} Q_1 & \text{wenn } i \neq j \\ Q_1 + Q(i+1,i) & \text{wenn } i = j \end{cases}$$

Für die Teststatistik sind die Q_i wieder durch die Schätzungen $\hat{Q}_i$ etc. zu ersetzen.

2. Variante: hier betrachtet man den Prüfvektor

$$d := (2\tilde{H}_1 - \tilde{H}_2, \ldots n\tilde{H}_1 - \tilde{H}_n)'$$

Unter der Nullhypothese ist $\sqrt{N}d$ asymptotisch normalverteilt mit Mittelwert 0 und Varianz-Kovarianz-Matrix V und daher ist die Teststatistik

$$H := Nd'\hat{V}^{-1}d$$

asymptotisch chi^2-verteilt mit n-1 Freiheitsgraden, wobei $\hat{V}$ eine konsistente Schätzung von V ist und V gegeben ist durch V = D'TD mit folgender n x (n-1) -Matrix D:

$$\begin{aligned} D_{1i} &= i+1 \\ D_{i+1,i} &= -1 \\ D_{ij} &= 0 \text{ sonst} \end{aligned}$$

Wir wollen wieder die explizite Gestalt von V unter H_0 bestimmen:

$$A_{ij} = (TD)_{ij} = (j+1)T_{i1} - T_{i,j+1} =$$

$$= (j+1)Q_1 - Q_{\min(i,j+1)}$$

Damit ist weiters

$$V_{ij} = (D'TD)_{ij} = (i+1)A_{1j} - A_{i+1,j} =$$

$$= (i+1)[(j+1)Q_1 - Q_1] - (j+1)Q_1 + Q_{\min(i+1,j+1)} =$$

$$= (i+1)jQ_1 - (j+1)Q_1 + Q_{\min(i+1,j+1)} =$$

$$= Q_1(ij+j-j-1) + Q_{\min(i+1,j+1)}$$

Für die Matrix V erhalten wir daher

$$V_{ij} = (ij-1)Q_1 + Q_{\min(i+1,j+1)}$$

Wieder sind die exakten Werte durch konsistente Schätzungen zu ersetzen. Diese Statistik hat den Nachteil, daß das kleinste

Intervall $[0,x_1]$ mit allen anderen verglichen wird und insofern den dort liegenden Beobachtungen größere Bedeutung zukommt.

3. Variante: dies wird vermieden, wenn man als Prüfvektor

$d := (2\tilde{H}_1 - \tilde{H}_2, \ldots (i+1)\tilde{H}_i - i\tilde{H}_{i+1}, \ldots n\tilde{H}_{n-1} - (n-1)\tilde{H}_n)'$

wählt. Wieder ist unter H_0 $\sqrt{N}d$ asymptotisch multivariat normalverteilt mit Mittelwert 0 und Varianz-Kovarianz-Matrix V und daher

$$H := Nd'\hat{V}^{-1}d$$

asymptotisch chi^2-verteilt mit n-1 Freiheitsgraden, wobei $\hat{V}$ eine konsistente Schätzung von V und V gegeben ist durch $V = D'TD$ mit folgender n x (n-1) -Matrix D:

$$\begin{aligned} D_{ii} &= i+1 \\ D_{i+1,i} &= -i \\ D_{ij} &= 0 \text{ sonst} \end{aligned}$$

Unter der Nullhypothese läßt sich V wieder explizit angeben:

$$A_{ij} := (TD)_{ij} = (j+1)T_{ij} - jT_{i,j+1} =$$

$$= (j+1)Q_{\min(i,j)} - jQ_{\min(i,j+1)}$$

Weiters ist

$$V_{ij} = (D'A)_{ij} = (i+1)A_{ij} - iA_{i+1,j} =$$

$$= (i+1)(j+1)Q_{\min(i,j)} - (i+1)jQ_{\min(i,j+1)} -$$

$$- i(j+1)Q_{\min(i+1,j)} + ijQ_{\min(i+1,j+1)}$$

Für i=j ergibt sich:

$$V_{ii} = Q_i(ij+i+j+1-ij-j-ij-i) + ijQ_{i+1} =$$

$$= (1-i^2)Q_i + i^2Q_{i+1} =$$

$$= \underline{i^2Q(i,i+1) + Q_i}$$

Für $i < j$ ergibt sich

$$V_{ij} = Q_i(ij+i+j+1-ij-j) - Q_{i+1}(ij-ij+i) = (i+1)Q_i - iQ_{i+1} =$$

$$= \underline{Q_i - iQ(i,i+1)}$$

In den Teststatistiken müssen wieder die exakten Werte durch konsistente Schätzungen ersetzt werden. Alle Bemerkungen, die in 1.1. und 3.2. über diese Tests gemacht wurden, gelten auch für die letzten Tests und sollen hier nicht mehr wiederholt werden.

III. Zusammenfassung

Für die drei verschiedenen Zensierungsmuster

1. single-type-I-censoring
2. type-II-censoring
3. random censoring

wurden auf verschiedene Arten Spezifikationstests vom Hausmantyp für die Nullhypothese

$$H_0: r(t) \text{ ist konstant}$$

entwickelt. (Dabei ist r(t) die Hazardfunktion des betrachteten Prozesses.)

Die erste Grundidee für einen derartigen Test besteht darin, eine Funktionalgleichung für die Überlebensfunktion S exponentialverteilter Daten zu verwenden: diese erfüllt nämlich

$$S(x+y) = S(x)S(y), \text{ resp. } S^p(x) = S(px).$$

Bildet man nun für fix gewählte Werte x,y, resp. x,p Schätzungen $\hat{S}(x+y)$, $\hat{S}(x)$, $\hat{S}(y)$, $\hat{S}(px)$, dann sollten unter der Nullhypothese $|\hat{S}(x+y)-\hat{S}(x)\hat{S}(y)|$ resp. $|\hat{S}(px)-\hat{S}^p(x)|$ nicht zu groß werden.

Ausgeführt heißt dies: bei single-type-I-censoring wählt man als Schätzung $\hat{S}$ die empirische Überlebensfunktion, bei random-censoring den sogenannten Produkt-Limit-Schätzer. In beiden Fällen sind dann unter der Nullhypothese die Statistiken

$$\sqrt{N}(\hat{S}(x+y)-\hat{S}(x)\hat{S}(y)) \text{ resp. } \sqrt{N}(\hat{S}(px)-\hat{S}^p(x))$$

asymptotisch normalverteilt. In beiden Fällen wird eine konsistente Schätzung der asymptotischen Varianz angegeben, sodaß daraus Teststatistiken konstruiert werden können, die unter H_0 asymptotisch N(0,1)-verteilt oder chi²-verteilt mit einem

Freiheitsgrad sind.

Bei random-censoring wird dieses Verfahren noch dahingehend modifiziert, daß die sogenannte kummulierte Hazardfunktion

$$H(t) = -\log S(t)$$

betrachtet wird: diese ist unter der Nullhypothese eine lineare Funktion mit positivem Anstieg. Die charakteristischen Funktionalgleichungen lauten nun:

$$H(x+y) = H(x) + H(y) \text{ und } H(px) = pH(x).$$

Analog zum vorher skizzierten Verfahren für den Produkt-Limit-Schätzer $\hat{S}(t)$ wird nun mit der empirischen kummulierten Hazardfunktion $\tilde{H}(t)$ verfahren. Die analogen Teststatistiken werden hier einfacher, allerdings ist nun - insbesondere für große t - die Güte der Schätzung schlechter als die des Produkt-Limit-Schätzers, was sich auf die Geschwindigkeit der Konvergenz der Teststatistik gegen die Grenzverteilung auswirken kann.

In allen Fällen gilt nun: die so konstruierten Tests sind konsistent gegen alle Alternativen, für die für die fix gewählten Werte x,y die angegebene Funktionalgleichung nicht erfüllt ist. Weiters wird gezeigt, daß die Tests konsistent sind gegen die Klasse der Alternativen mit monotoner Hazardfunktion - unabhängig von der konkreten Wahl der x,y, für die die Schätzungen gemacht und verglichen werden.

Alle diese Verfahren können nun folgendermaßen verallgemeinert werden: man wähle ein $x > 0$ und unterteile das Intervall $[0,x]$ in n Teile, z.B. durch die Punkte ix/n. Anstatt die gegebene Funktionalgleichung der Überlebensfunktion (der kummulierten Hazardfunktion) bei zwei Werten zu "überprüfen", wird dies nun für alle Teilungspunkte ix/n des Intervall $[0,x]$ durchgeführt. Daraus konstruiert man eine Teststatistik, die unter H_0 asymptotisch chi^2-verteilt ist mit n-1 Freiheitsgraden. Dieser Test ist

konsistent gegen alle Alternativen, für die

$$S(ix/n)S(jx/n) \neq S((i+j)x/n)$$

gilt für mindestens ein Paar (i,j) mit $i+j \leq n$. Außerdem gilt: zu *jeder* Alternative H_1 gibt es ein n sodaß der dazugehörige Test gegen H_1 konsistent ist. Erkauft wird die Vergrößerung dieser Alternativenmenge durch eine Erhöhung der Freiheitsgrade (es werden ja n verschiedene Werte geschätzt), was sich dahingehend auswirkt, daß die Teststatistik langsamer gegen die Grenzverteilung konvergiert.
Für den Fall von type-II-censierten Daten wurde folgendes Verfahren gewählt: mittels Orderstatistiken werden bestimmte Quantile geschätzt, die unter H_0 in einem bestimmten Verhältnis zueinander stehen. Auch hier kann das Verfahren verallgemeinert werden auf den "simultanen" Vergleich der Schätzungen für n verschieden Quantile. Bezüglich der Menge der konsistenten Alternativen etc. gelten wieder ähnliche Aussagen wie vorher.

Für single-type-I-censierte Daten wurde schließlich noch eine weitere Idee für einen derartigen Test ausgeführt: man bildet den Maximum-Likelihood-Schätzer für den Parameter α bezüglich der eigentlichen Zensierungszeit C und bezüglich einer (oder mehrerer) "künstlicher" Zensierungszeit(en) $C' < C$. Unter der Nullhypothese dürfen diese Schätzungen nicht zu sehr voneinander abweichen. Daraus resultiert wieder eine Teststatistik, die unter H_0 asymptotisch chi^2-verteilt ist mit k Freiheitsgraden (k ist die Anzahl der "künstlichen" Zensierungszeiten).
Ausdrücklich sei darauf hingewiesen, daß es sich bei allen Tests um asymptotische Tests handelt, die daher nur für große Stichproben geeignet sind. Die exakte Nullverteilung für kleine Stichproben dürfte außerordentlich schwierig zu bestimmen sein.

Wollte man mittels Monte-Carlo-Simulation Signifikanzpunkte für kleine Stichproben ermitteln, muß man beachten, daß diese

1. bei single-Type-I-censoring von der Zensierungszeit C
2. bei type-II-censoring von p, dem Anteil der exakten Beobachtungen
3. bei random-censoring von der Verteilung der Zensierungsvariablen C_i abhängen.

Für viele Anwendungen in den Sozialwissenschaften ist die Modellannahme einer homogenen Population nicht realistisch. Das in derartigen Anwendungen betrachtete "Grundmodell" lautet:

$$r(t) = r \cdot \exp(x_1\beta_1 + \ldots + x_k\beta_k)$$

Die Hazardfunktion ändert sich nicht während der Zeit. Die Hazardrate ist aber nicht für alle Individuen der Grundgesamtheit gleich, vielmehr abhängig von sogenannten Kovariaten x_i, die bestimmte Merkmale des Individuums ausdrücken. Die oben erwähnten Tests lassen sich also nicht ohne weiteres auf diesen Fall anwenden. Am ehesten scheint sich der Test mittels des Maximum-Likelihood-Schätzers auf diesen Fall verallgemeinern zu lassen. Es müßte nachgewiesen werden,daß die Maximum-Likelihood-Schätzungen für (r,β) im obigen Modell bezüglich verschiedener Zensierungszeiten asymptotisch eine gemeinsame (multivariate) Normalverteilung besitzen. Sollte dies zutreffen, müßte eine konsistente Schätzung der asymptotischen Varianz-Kovarianz-Matrix gefunden werden. Am schwierigsten dürfte dies sein, wenn kontinuierliche Kovariaten x_i auftreten: dann läßt sich nämlich der Maximum-Likelihood-Schätzer nicht mehr explizit als Funktion der Ankunftszeiten darstellen.

LITERATUR

Anderson,T.W., An Introduction to Multivariate Statistical Analysis, Wiley, New York, 1958

Billingsley,P., Convergence of Probability Measures, Wiley, New York, 1968

Breslow,N. and J. Crowley, A large sample study of the life table and the product limit estimates under random censorship, Ann.Stat.,2,3,437-453,1974

Diekmann,A. and P. Mitter, Methoden zur Analyse von Zeitverläufen, Teubner, Stuttgart,1984

Epstein,B., Tests for the validity of the asumption that the underlying distribution of life is exponential I, Technometrics,2,1,83-101,1960

Hall,P. and C.C. Heyde, Martingale Limit Theory and Its Applications, Academic Press, New York, 1980

Hausman,J.A., Specification Tests in Econometrics, Econometrica,46,1978,1251-1271.

Kalbfleisch,J.A. and R.L. Prentice, The Statistical Analysis of Failure Time Data, Wiley, New York, 1980

Lawless,J.F.,Statistical Models and Methods for Lifetime Data, New York, 1982

Meier,P., Estimation of a distribution function from incomplete observations, in: Perspectives in Probability and Statistics, Edited by J. Gani, Academic Press, London, 1975

Mosteller,F.C., On some useful "inefficient" statistics, Ann.Math.Stat.,17,377-408

Peterson,A.V.Jr., Expressing the Kaplan-Meier Estimator as a

function of empirical subsurvival functions, JASA, 72,360,854-848,1977

Rao,C.R., Linear Statistical Inference and Its Applications, Wiley, New York, 1965

Rényi,A., Wahrscheinlichkeitsrechnung, Veb Deutscher Verlag der Wissenschaften, Berlin, 1977

Sarhan,A.E. and B.G. Greenberg, Contributions to Order Statistics, Wiley, New York, 1982