Beschleunigte Lebensdauertests

Vorlesung Zuverlässigkeit & Qualitätssicherung (Master)

B.Eng. René Schwarz
mail@rene-schwarz.com — http://www.rene-schwarz.com

Hochschule Merseburg (FH)
Fachbereich Informatik- und Kommunikationssysteme

in Vertretung für Herrn Prof. Dr. rer. nat. habil. Eckhard Liebscher

07. Juli 2010

Belegthemen

Prüfungshinweis

Am Ende des Semesters ist ein ausgearbeiteter Beleg abzugeben (vgl. Modul „Computergestützte Datenanalyse"), der die folgenden Themen umfasst:

1. **Lebensdauerverteilungen**
   Berechnung von Wahrscheinlichkeiten, Berechnung der Parameter der Verteilung, Ausfallrate, Restnutzungsdauer
2. **Systeme**
   Verfügbarkeit, Lebensdauerverteilung, Wichtigkeit von Elementen
3. **Statistik von Lebensdauerverteilungen**
   Weibull-Verteilung, zensierte Stichproben
4. **Instandhaltung** (optimale Strategien)
5. **Statistische Qualitätssicherung in der Prozessüberwachung** (Kontrollkarten)
6. **Stichprobenverfahren zur Attributprüfung**
7. **Beschleunigte Lebensdauertests**

Einführung

Hintergrund

Beschleunigte Lebensdauertests werden für Systeme oder Bauteile eingesetzt, die langlebig sind. In der Regel sind Lebensdauerexperimente für diese Systeme unpraktikabel, da sie die Entwicklung erheblich verzögern würden. Dennoch möchte man die Lebensdauer dieser Systeme bestimmen.

Ansatz

Die zu untersuchenden Systeme werden außerhalb der Grenzen ihrer Spezifikation belastet. Der Betrieb außerhalb der Spezifikationsgrenzen verursacht eine schnellere Alterung und damit ein frühzeitiges Versagen. Durch die gewonnenen Lebensdauerdaten unter Überlastbedingungen lassen sich Rückschlüsse auf die Lebensdauerverteilung der Systeme unter normalen Bedingungen ziehen.

Probleme

- Beschleunigte Lebensdauertests müssen sorgfältig geplant werden, um jedwede andere, unerwünschte Störgröße während der Experimente nicht mitzuerfassen. Durch eine verunreinigte Datenbasis wird das Endergebnis unbrauchbar.
- Die Schätzung der Lebensdauerverteilung unter Normalbedingungen ist nicht trivial. Alle Schätzungen weichen naturgemäß von der Realität ab; es gilt, den Fehler zu minimieren.
- Die Lebensdauer hängt i.d.R. von vielen Faktoren (Kovariablen) ab, die es zu bedenken gilt.

Beispiel: Voyager 1

Voyager 1 3D-Modell

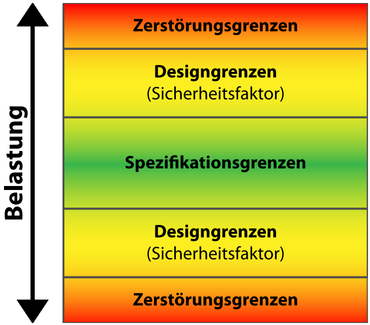
NASA-Sonde zur Erforschung des äußeren Sonnensystems

Start:
05. September 1977
voraussichtliches Missionsende:
2025
voraussichtliche Lebensdauer:
48 Jahre

Bei solch langen Einsatzdauern sind unbeschleunigte Lebensdauertests für Lebensdaueranalysen undenkbar.

Credit:
National Aeronautics and Space Administration (NASA), USA.
Re-meshed and converted to Universal 3D Format (u3d) by B.Eng. René Schwarz (rene-schwarz.com). See license information under http://www.nasa.gov/audience/formedia/features/MP_Photo_Guidelines.html.
Source: http://www.nasa.gov/multimedia/3d_resources/assets/voyager_c.html

Design von Lebensdauerexperimenten



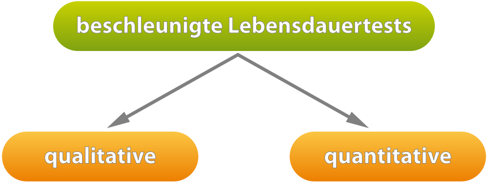
(nach [Weibull.com 2010])

Spezifikationsgrenze
Belastungen im normalen Betrieb; normale Lebensdauerexperimente – direkte Erfassung der Ausfallrate (Lebensdauerverteilung)

Designgrenze
mögliche Belastung des Systems gemäß Design (mit Sicherheitsfaktor); beschleunigte Lebensdauerexperimente

Zerstörungsgrenze
sofortige Zerstörung des Systems; keine Rückschlüsse auf Lebensdauer möglich

Arten von beschleunigten Lebensdauerexperimenten (I)



Qualitative, beschleunigte Lebensdauerexperimente werden generell für eine kleine Anzahl von System- oder Bauteilmustern durchgeführt. Sie werden mit einer einzigen Belastung oder mehreren Belastungsarten durchgeführt (z.B. zyklische Belastungen, Warm-Kalt-Experimente etc.). Wenn das System diesen Test überlebt, hat es den Test bestanden. Andernfalls werden geeignete Maßnahmen getroffen, um die Ursachen des Versagens zu beseitigen.
Qualitative Tests werden vorrangig durchgeführt, um mögliche Fehlerquellen zu entdecken. Generell quantifizieren diese Tests aber die Lebensdauerverteilung (oder die Zuverlässigkeit) des Systems nicht. Diese Testmethode kann aber wertvolle Informationen zu Einflussgrößen auf die Lebensdauer liefern, die in quantitativen Tests verwendet werden können. Qualititative Lebensdauerexperimente werden auch *Highly Accelerated Life Tests (HALT)* genannt.

Beispiel: Falltests bei Handys (zyklische Beschleunigungsbelastung)

Arten von beschleunigten Lebensdauerexperimenten (II)

Quantitative, beschleunigte Lebensdauertests bestehen – im Gegensatz zu der qualitativen Methode – aus Versuchen, die zur Quantifizierung der Lebensdauercharakteristik eines Produktes, einer Komponente oder eines Systems unter normalen Einsatzbedingungen dienen; sie stellen also Kennzahlen zur Zuverlässigkeit bereit.
Zuverlässigkeitskennzahlen können die Ausfallwahrscheinlichkeit des Produktes unter normalen Einsatzbedingungen, die durchschnittliche Lebensdauer (MTTF), die Anzahl der Rückläufer und die Garantiekosten beinhalten. Sie können ebenfalls als Grundlage einer Kosten-Risikoeinschätzung einer Produkteinführung, für Vergleiche von verschiedenen Produktentwürfen etc. pp. verwendet werden.
Quantitative, beschleunigte Lebensdauertests werden üblicherweise als *Quantitative Accelerated Life Tests (QALT)* bezeichnet.

Beispiel: erhöhte Stromstärke zur schnelleren Alterung von LEDs

(nach [Weibull.com 2010])

Maximum-Likelihood-Methode

- engl. *Methode der maximalen Wahrscheinlichkeit*
- parametrisches Schätzverfahren in der Statistik
- schätzt Parameter einer bekannten Dichtefunktion $f(x)$ einer Wahrscheinlichkeitsverteilung (z.B. Erwartungswert μ und Varianz σ^2 einer Normalverteilung) auf Grundlage einer Stichprobe (Daten)
- zu schätzende Parameter: Parametermenge ϑ

Voraussetzungen

- n Beobachtungen (Realisierungen) der Stichprobe X mit $x_1, x_2, \ldots, x_n$ statistisch unabhängig
- Wahrscheinlichkeitsverteilung und deren Dichtefunktion $f(x)$ a priori bekannt

Ansatz

- gemeinsame Wahrscheinlichkeitsfunktion $f(x_1, x_2, \ldots, x_n)$ aller Realisationen der Stichprobe ist die multiplikative Verknüpfung der Einzelwahrscheinlichkeiten (bei unzensierten Stichproben):

$$f(x_1, x_2, \ldots, x_n) = f(x_1) \cdot f(x_2) \cdot \ldots \cdot f(x_n)$$

- Stichprobe wird als gegeben betrachtet; Parametermenge ϑ gesucht

$$\mathcal{L}(\vec{x}\,|\vartheta) = \prod_{i=1}^{n} f(x_i|\vartheta)$$

- Schätzung eines Parameters der Parametermenge ϑ: Maximierungsaufgabe! („Welcher Parameterwert passt mit maximaler Wahrscheinlichkeit zu der Stichprobe?“)
- Log-Likelihood-Funktion bilden (Vereinfachung der Gleichung)

$$\frac{\partial \ln(\mathcal{L})}{\partial \vartheta} \stackrel{!}{=} 0$$

- Log-Likelihood-Funktion nach gesuchten Parameter partiell ableiten, Nullsetzen und auflösen

Herleitung: Likelihood-Funktion für Normalverteilung (I)

Ausgangssituation

Zufallsgröße X, Realisierungen $x_1, x_2, \ldots, x_n$, normalverteilt mit Erwartungswert μ und Varianz σ^2: $X \sim \mathcal{N}(\mu, \sigma^2)$, keine Zensierung der Daten

- Verteilung $f(x)$ der Wahrscheinlichkeitsdichte einer Normalverteilung (bekannt):

$$f(x, \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)$$

- Ansatz für Likelihood-Funktion $\mathcal{L}$ bezüglich Dichteverteilung $f(x)$:

$$\mathcal{L}(\vec{x} \,|\vartheta) = \prod_{i=1}^{n} \left(f(x_i, \mu, \sigma^2)\right)$$

- Parametermenge ϑ enthält die beiden zu bestimmenden Parameter der Normalverteilung: $\vartheta = \{\mu, \sigma^2\}$

Herleitung: Likelihood-Funktion für Normalverteilung (II)

- einsetzen, umformen:

$$\begin{aligned}\mathcal{L}(\vec{x}\,|\vartheta) &= \prod_{i=1}^{n}\left(\frac{1}{\sigma\sqrt{2\pi}}\cdot\exp\left(\frac{-(x_i-\mu)^2}{2\sigma^2}\right)\right)\\ &= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n\cdot\prod_{i=1}^{n}\exp\left(\frac{-(x_i-\mu)^2}{2\sigma^2}\right)\end{aligned}$$

Potenzgesetz

$$\mathrm{e}^{\frac{x_1}{z}}\cdot\mathrm{e}^{\frac{x_2}{z}}\equiv\mathrm{e}^{\frac{x_1+x_2}{z}}\equiv\mathrm{e}^{\frac{\sum x_i}{z}}$$

$$\begin{aligned}\mathcal{L}(\vec{x}\,|\vartheta) &= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n\cdot\exp\left(\frac{1}{2\sigma^2}\sum_{i=1}^{n}(-(x_i-\mu)^2)\right)\\ &= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n\cdot\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)\end{aligned}$$

Herleitung: Likelihood-Funktion für Normalverteilung (III)

- Gleichung durch Bildung der Log-Likelihood-Funktion $\ln(\mathcal{L})$ vereinfachen

$$
\begin{aligned}
\ln(\mathcal{L}) &= \ln\left[\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \cdot \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)\right] \\
&\overset{\ln(xy)=}{\underset{\ln(x)+\ln y}{=\!=\!=}} \ln\left[\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n\right] + \ln\left[\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)\right] \\
&\overset{\ln(a^r)=}{\underset{r\cdot\ln(a)}{=\!=\!=}} n\cdot\ln\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 \\
&\overset{\ln\left(\frac{x}{y}\right)=}{\underset{\ln(x)-\ln(y)}{=\!=\!=}} \overbrace{n\ln(1)}^{=0} - n\ln(\sigma\sqrt{2\pi}) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 \\
&\overset{\ln(xy)=}{\underset{\ln(x)+\ln y}{=\!=\!=}} -n\ln(\sigma) - n\ln(\sqrt{2\pi}) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2
\end{aligned}
$$

Herleitung: Likelihood-Funktion für Normalverteilung (IV)

- Schätzer $\hat{\mu}$ für Erwartungswert μ ermitteln:
  Log-Likelihood-Funktion nach μ maximieren: partielle Ableitung nach μ bilden und Nullsetzen; nach μ umstellen

$$\frac{\partial \ln(\mathcal{L})}{\partial \mu} = \frac{2}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu) \xlongequal{\mu \to \text{max!}} 0 = \sum_{i=1}^{n} (x_i - \mu)$$

$$n\mu = \sum_{i=1}^{n} x_i$$

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x}$$

- Et voilà! Wir haben mit $\hat{\mu} = \bar{x}$ den erwartungstreuen Schätzer für μ hergeleitet. Dies entspricht unserem Vorwissen der Normalverteilung.
- Analog kann der Schätzer für σ^2 hergeleitet werden (der im Übrigen nur asymptotisch erwartungstreu ist).

Proportional Hazards Model (Cox-Modell, Cox-Regression)

- entwickelt und veröffentlicht 1972 von Sir David Roxbee Cox (britischer Statistiker)
- weit verbreitetes Regressionsmodell für die elegante Analyse von Überlebenszeiten
- basiert auf dem Konzept der Ausfallrate $h(t) = \frac{f(t)}{1-F(t)}$
- **Eigenschaften:** semi-parametrisch, proportional
- Stärke des Modells: dient zur Schätzung des **relativen** Risikos, nicht des absoluten (Vergleich unterschiedlicher Ausprägungen der Einflussgrößen auf die Lebensdauer untereinander möglich)

Grundidee

Das Cox-Modell geht davon aus, dass die Lebensdauerverteilung von einem Grundrisiko, dem sog. Baseline Hazard $h_0(t)$, abhängt. Andere Einflüsse wirken sich lediglich multiplikativ auf dieses Grundrisiko aus, die Funktion der Lebensdauerverteilung wird lediglich gestreckt oder gestaucht – ihre Form ändert sich aber hingegen nicht.

Motivation

Beispiel

Jeder Mensch hat ein Grundrisiko, an Lungenkrebs zu erkranken. Faktoren wie das Alter oder Rauchverhalten wirken sich jedoch auf das Risiko aus. Das Cox-Modell gestattet es nun, eine Aussage darüber zu treffen, ob ein Raucher in Gegenüberstellung eines Nichtrauchers oder ein älterer Mensch gegenüber einem jüngeren Menschen ein größeres Risiko hat, an Lungenkrebs zu erkranken (relative Schätzung des Hazard Ratio). Das Modell macht hingegen aber keine Aussage über das eigentliche Grundrisiko, an Lungenkrebs zu erkranken (absolute Schätzung).

Eine absolute Schätzung kann zwar vorgenommen werden, ist aber in den wenigsten Fällen sinnvoll – es wäre eine Schätzung von der Schätzung. Es ist äußerst fraglich, ob tatsächlich alle Einflussfaktoren (Kovariablen) auf das Grundrisiko bekannt und in dem Modell enthalten sind. Beispielsweise könnten das Geschlecht, die Ernährungsweise u.a. Kovariablen wesentliche Einflussgrößen sein, die eine absolute Schätzung vollkommen verändern würden.

Die relative Schätzung durch Bildung eines Verhältnisses zweier Individuen mit unterschiedlichen Ausprägungen der Einflussfaktoren liefert jedoch eine wertvolle Abschätzung über die Veränderung bezüglich des Grundrisikos. Diese Schätzung kann mit dem Cox-Modell hinreichend genau vorgenommen werden.

Ansatz (I)

- Lebensdaueranalyse untersucht Beziehung zwischen Lebensdauerverteilung und Kovariablen
- häufig folgt aus einer Lebensdaueruntersuchung ein annähernd lineares Modell für logarithmierte Ausfallrate $\ln(h(t))$
- **Beispiel:** parametrisches Modell basierend auf der Exponentialverteilung für $h_i(t)$:

  $$\ln(h_i(t)) = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik}$$

  oder umgeformt:

  $$h_i(t) = \exp(\alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik})$$

  $\Rightarrow$ lineares Modell für $\ln(h_i(t))$ (i - Beobachtung, x - Kovariablen)
- Konstante α impliziert eine Art von „Grundrisiko“, da $\alpha = \ln(h_i(t))$ (oder $h_i(t) = \mathrm{e}^{\alpha}$), wenn alle Kovariablen Null sind
- Cox-Modell lässt nun $\mathrm{e}^{\alpha(t)} = h_0(t)$ undefiniert:

  $$\ln(h_i(t)) = \alpha(t) + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik}$$

  oder wieder umgeformt:

  $$h_i(t) = h_0(t) \cdot \exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik})$$

Ansatz (II)

Cox-Modell

$$h_i(t) = h_0(t) \cdot \mathrm{e}^{\underline{X}\vec{\beta}}$$

$$= h_0(t) \cdot \exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik}) = h_0(t) \cdot \exp\left(\sum_k \beta_k x_{ik}\right)$$

$$\vec{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} \qquad \underline{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix} = x_{(ij)}$$

- **semi-parametrisch:** $h_0(t)$ kann jede beliebige Form haben; die Kovariablen gehen unabhängig davon linear in das Modell ein
- **proportional:** zwei Individuen i und j mit den linearen Schätzern

$$\eta_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} \quad \text{und} \quad \eta_j = \beta_1 x_{j1} + \beta_2 x_{j2} + \ldots + \beta_k x_{jk}$$

Risiko-Verhältnis (engl. Hazard Ratio) für beide Individuen:

$$\frac{h_i(t)}{h_j(t)} = \frac{h_0(t)\mathrm{e}^{\eta_i}}{h_0(t)\mathrm{e}^{\eta_j}} = \frac{\mathrm{e}^{\eta_i}}{\mathrm{e}^{\eta_j}}$$

  - Lebensdauerverteilung behält ihre Form bei, wird lediglich gestaucht oder gestreckt
  - Kovariablen gehen unabhängig von der Zeit in das Modell ein: sie müssen über den gesamten Betrachtungszeitraum konstant sein (**Nachteil** des Modells!)

Ansatz (III)

- auch wenn Grundrisiko $h_0(t)$ nicht bekannt ist, können die Modellparameter $(\vec{\beta}, h_0(t))$ geschätzt werden
- **Schätzung der Parameter ist nicht trivial:** Maximum-Likelihood-Methode versagt, da in den Maximierungsfunktionen Abhängigkeiten zum jeweils anderen Parameter auftreten $\Rightarrow$ mathematisch nicht lösbar (eine Gleichung mit zwei unbekannten Variablen)
- David Cox lieferte eine neue Methode zur Schätzung der Modellparameter: Partial Likelihood-Methode
- Resultate sind nicht so effizient wie bei der Maximimum-Likelihood-Methode, dafür muss aber auch keine (möglicherweise falsche) Annahme über die Form des Baseline Hazard $h_0(t)$ getroffen werden (großer Vorteil des Cox-Modells)
- nach Bestimmung der Modellparameter können Individuenvergleiche erstellt werden – das Hazard Ratio für unterschiedliche Ausprägungen der einzelnen Kovariablen kann bestimmt werden

Schätzung der Modellparameter (I)

- Herleitung der Likelihood-Funktion für das Cox-Modell

Indikatorfunktion δ_i

$$\delta_i = \begin{cases} 1 & \text{wenn } i\text{-te Beobachtung nicht zensiert} \\ 0 & \text{wenn } i\text{-te Beobachtung zensiert} \end{cases}$$

- Besonderheit: rechts zensierte Daten $\Rightarrow$ Anpassung der Likelihood-Funktion:

$$\begin{aligned}
\mathcal{L}(\vec{t}, \vec{\delta}\,|\vartheta) &= \prod_{i:\delta_i=1} P(T = t_i|\vartheta) \prod_{i:\delta_i=0} P(T > t_i|\vartheta) \\
&= \prod_{i:\delta_i=1} f_i(t_i) \prod_{i:\delta_i=0} S_i(t_i) = \prod_{i=1}^{n} (f_i(t_i))^{\delta_i} \prod_{i=1}^{n} (S_i(t_i))^{1-\delta_i} \\
&= \prod_{i=1}^{n} (f_i(t_i))^{\delta_i} (S_i(t_i))^{1-\delta_i} = \prod_{i=1}^{n} \left(\frac{f_i(t_i)}{S_i(t_i)}\right)^{\delta_i} S_i(t_i) \\
&= \prod_{i=1}^{n} h_i^{\delta_i}(t_i) S_i(t_i)
\end{aligned}$$

Schätzung der Modellparameter (II)

- $h_i(t)$ ist bekannt: $h_i(t) = h_0(t) \cdot \exp\left(\sum_k \beta_k x_{ik}\right)$
- Überlebensfunktion $S_i(t)$ (engl. Survival Function) muss aus $h_i(t)$ hergeleitet werden:

$$
\begin{aligned}
h_i(t) &= \frac{f_i(t)}{1 - F_i(t)} = \frac{f_i(t)}{S_i(t)} = \frac{\frac{\mathrm{d}}{\mathrm{d}t} F_i(t)}{S_i(t)} = \frac{\frac{\mathrm{d}}{\mathrm{d}t}(1 - S_i(t))}{S_i(t)} = -\frac{\dot{S}_i(t)}{S_i(t)} \\
&= -\frac{\mathrm{d}}{\mathrm{d}t} \int \frac{\dot{S}_i(t)}{S_i(t)} \mathrm{d}t = -\frac{\mathrm{d}}{\mathrm{d}t} \ln(S_i(t))
\end{aligned}
$$

$$
\begin{array}{lrcll}
\Rightarrow & h_i(t) & = & -\frac{\mathrm{d}}{\mathrm{d}t} \ln(S_i(t)) & \mid \int \mid \cdot (-1) \\
 & -\int\limits_0^t h_i(\tau)\mathrm{d}\tau & = & \ln(S_i(t)) & \mid \cdot \mathrm{e}^{\cdots} \\
 & S_i(t) & = & \exp\left(-\int\limits_0^t h_i(\tau)\mathrm{d}\tau\right) &
\end{array}
$$

Schätzung der Modellparameter (III)

- Einsetzen von $h_i(t)$ und $S_i(t)$ in $\mathcal{L}(\vec{t}, \vec{\delta}\,|\vartheta)$, wobei $\vartheta = \{\vec{\beta}, h_0(t)\}$:

$$\begin{aligned}\mathcal{L}(\vec{t}, \vec{\delta}\,|\vartheta) &= \prod_{i=1}^{n} h_i^{\delta_i}(t_i) S_i(t_i) \\ &= \prod_{i=1}^{n} \left[\left(h_0(t_i) \cdot \exp\left(\sum_k \beta_k x_{ik} \right) \right)^{\delta_i} \right. \\ &\qquad \left. \cdot \exp\left(-\int_0^{t_i} h_0(\tau) \cdot \exp\left(\sum_k \beta_k x_{ik} \right) \mathrm{d}\tau \right) \right]\end{aligned}$$

- Produkt auflösen: Faktorisierung; kleinere Umformungsarbeiten

$$\begin{aligned}\mathcal{L}(\vec{t}, \vec{\delta}\,|\vartheta) = \prod_{i=1}^{n} h_0^{\delta_i}(t_i) \cdot \exp\left(\delta_i \sum_k \beta_k x_{ik} \right) \\ \cdot \prod_{i=1}^{n} \exp\left(-\int_0^{t_i} h_0(\tau) \mathrm{d}\tau \right)^{\exp\left(\sum_k \beta_k x_{ik} \right)}\end{aligned}$$

Schätzung der Modellparameter (IV)

- Erweiterung des ersten Faktors (verändert die Gleichung nicht):

$$\mathcal{L}(\vec{t},\vec{\delta}\,|\vartheta) = \prod_{i=1}^{n} h_0^{\delta_i}(t_i) \cdot \exp\left(\delta_i \sum_k \beta_k x_{ik}\right) \cdot \frac{\sum\limits_{j:R_{t_i}(j)>0} \exp\left(\sum\limits_k \beta_k x_{jk}\right)}{\sum\limits_{j:R_{t_i}(j)>0} \exp\left(\sum\limits_k \beta_k x_{jk}\right)}$$

$$\cdot \prod_{i=1}^{n} \exp\left(-\int_0^{t_i} h_0(\tau)\mathrm{d}\tau\right)^{\exp\left(\sum\limits_k \beta_k x_{ik}\right)}$$

- $h_0^{\delta_i}(t)$ kann in den Zähler gezogen werden; die Potenz/der Faktor δ_i wird eliminiert, da die Bedingung dem Produkt hinzugefügt wird

$$\mathcal{L}(\vec{t},\vec{\delta}\,|\vartheta) = \prod_{\substack{i=1,\ldots,n\\ i:\delta_i=1}}^{n} \exp\left(\sum_k \beta_k x_{ik}\right) \cdot \frac{\sum\limits_{j:R_{t_i}(j)>0} h_0(t_i) \cdot \exp\left(\sum\limits_k \beta_k x_{jk}\right)}{\sum\limits_{j:R_{t_i}(j)>0} \exp\left(\sum\limits_k \beta_k x_{jk}\right)}$$

$$\cdot \prod_{i=1}^{n} \exp\left(-\int_0^{t_i} h_0(\tau)\mathrm{d}\tau\right)^{\exp\left(\sum\limits_k \beta_k x_{ik}\right)}$$

Schätzprobleme mit der Maximum-Likelihood-Methode

Likelihood-Funktion für das Cox-Modell

$$\mathcal{L}(\vec{t},\vec{\delta}\,|\vartheta) = \prod_{\substack{i=1,\ldots,n\\ i:\delta_i=1}}^{n} \frac{\exp\left(\sum\limits_k \beta_k x_{ik}\right)}{\sum\limits_{j:R_{t_i}(j)>0} \exp\left(\sum\limits_k \beta_k x_{jk}\right)} \cdot \sum_{j:R_{t_i}(j)>0} h_0(t_i)\cdot\exp\left(\sum_k \beta_k x_{jk}\right)$$

$$\cdot \prod_{i=1}^{n} \exp\left(-\int_0^{t_i} h_0(\tau)\mathrm{d}\tau\right)^{\exp\left(\sum\limits_k \beta_k x_{ik}\right)}$$

- nach Vorgabe der Maximum-Likelihood-Methode (vgl. vorheriger Abschnitt) müsste nun die Log-Likelihood-Funktion gebildet und diese partiell nach den Elementen der Parametermenge ϑ abgeleitet werden
- bei diesem Vorgehen würde man feststellen können, dass sowohl in der partiellen Ableitung nach $\vec{\beta}$ als auch nach $h_0(t)$ Abhängigkeiten zum jeweils anderen Parameter bestehen bleiben $\Rightarrow$ Gleichung nicht lösbar (eine Gleichung, zwei Unbekannte); auf den Beweis der Herleitung wird an dieser Stelle verzichtet

Problemlösung: Partial Likelihood

- David Cox schlug vor, nur die von $h_0(t)$ unabhängigen Teile der Likelihood-Funktion zu verwenden, um der Schätzung und der komplexen Formel Herr zu werden:

$$\mathcal{L}_{\text{part}}(\vec{t}, \vec{\delta} \,|\vartheta) = \prod_{\substack{i=1,\ldots,n \\ i:\delta_i=1}}^{n} \frac{\exp\left(\sum_k \beta_k x_{ik}\right)}{\sum_{j:R_{t_i}(j)>0} \exp\left(\sum_k \beta_k x_{jk}\right)}$$

- Schätzung der Regressionskoeffizienten β_i erfolgt über die Nullsetzung der partiellen Ableitung von $\ln(\mathcal{L}_{\text{part}})$ nach $\vec{\beta}$: $\frac{\partial \ln(\mathcal{L}_{\text{part}})}{\partial \vec{\beta}}$
- danach erfolgt die Schätzung von $h_0(t_i)$ über die vollständige Likelihood-Gleichung des Cox-Modells mit eingesetzem $\hat{\beta}$
- man bedient sich bei der Schätzung numerischer Lösungsverfahren, da eine algebraische Lösung i.d.R. äußerst komplex oder unmöglich ist ($\Rightarrow$ Lösung durch Software wie *Mathematica*, *Statistica* etc.)

Hörsaalexperiment

Animation Büroklammer

Lebensdauer einer Büroklammer

- Ziel: Lebensdauer unter normaler Belastung (20° Biegung) ermitteln
- Ansatz: beschleunigter Lebensdauertest: Belastung innerhalb der Designgrenzen: 90°, 135° und 180° Biegung

Wichtig:

- Biegung unter konstanter Geschwindigkeit (s. Animation links)
- identische, unbenutzte Büroklammern
- Biegung lediglich entlang der dargestellten Achse

Bitte zählen Sie die Anzahl an Biegungen bei dem Ihnen vorgegebenen Winkel, die Sie durchführen können bis die Büroklammer bricht!

Bildung eines Modells

Cox-Modell für das Hörsaalexperiment

$$h_i(t) = h_0(t) \cdot \exp(\beta x_i)$$

β	Regressionskoeffizient
x	Kovariable: Biegewinkel
i	Laufindex der Beobachtungen

Schätzung der Parametermenge ϑ

$\hat{\beta}$, $\hat{h}_0(t)$ werden mit Hilfe von Software geschätzt

Schätzung der Lebensdauer bei $20°$-Biegungen

Hazard Ratio:

$$\psi_{90/20} = \frac{h_i(t)}{h_j(t)} = \frac{h_0(t) \cdot \exp(\beta \cdot 90)}{h_0(t) \cdot \exp(\beta \cdot 20)} = \mathrm{e}^{\beta(90-20)} = \mathrm{e}^{\beta \cdot 70}$$

Die Probe auf's Exempel

- Überprüfen wir nun die Übereinstimmung unseres Modells mit der Realität!
- $20°$-Biegungen sind schwierig im Hörsaalexperiment abzumessen $\Rightarrow$ wir weichen daher auf $45°$-Biegungen aus

Schätzung der Lebensdauer bei $45°$-Biegungen

Hazard Ratio:

$$\psi_{90/45} = \frac{h_i(t)}{h_j(t)} = \frac{h_0(t) \cdot \exp(\beta \cdot 90)}{h_0(t) \cdot \exp(\beta \cdot 45)} = \mathrm{e}^{\beta(90-45)} = \mathrm{e}^{\beta \cdot 45}$$

Biegen Sie nun bitte Ihre Büroklammer um $45°$ mit der gleichen, konstanten Geschwindigkeit wie im vorherigen Experiment und zählen Sie die Anzahl der Biegungen, die Sie durchführen müssen, bis es zum Versagen der Büroklammer kommt.

Wie gut ist unser Modell? :o)

Literatur- und Quellenverzeichnis

Weibull.com
Accelerated Life Testing, this Issue's Hot Topic
Website, 2010.
http://www.weibull.com/hotwire/issue18/hottopics18.htm

Fox, John
Cox Proportional-Hazards Regression for Survival Data
in: *Appendix to An R and S-PLUS Companion to Applied Regression*
R Project, PDF-Datei, 2002.
http://cran.r-project.org/doc/contrib/Fox-Companion/appendix-cox-regression.pdf

Stocker, Herbert
Einführung in die angewandte Ökonometrie, Kapitel 14: Die Maximum-Likelihood-Methode
Vorlesungsskript a.d. Universität Innsbruck, PDF-Datei, 2010.
http://cran.r-project.org/doc/contrib/Fox-Companion/appendix-cox-regression.pdf

Held, Leonhard; Wolbers, Marcel
Biostatistische Methoden, Abschnitt 7: Regressionsmodelle für Überlebenszeiten
Vorlesungspräsentation a.d. Universität Zürich, PDF-Datei, 2010.
http://www.biostat.uzh.ch/teaching/master/methods2008-1/slides7.pdf

Noack, Marcel
Cox-Regression
Vortrag a.d. Universität Duisburg, PDF-Datei, 2008.
http://www.uni-due.de/imperia/md/content/soziologie/stein/beamer-ereignis2nopause.pdf