

# Bayes'sche Untersuchung von zensierten Daten

Martin Hellmich

19. März 2001

## Zusammenfassung

- Einführung
- Typen von Ereigniszeiten
- Semi-parametrische Modelle
- A-priori-Prozesse
- Bayes'sche Analyse univariater Überlebenszeiten
- WinBUGS-Beispiele
- Literatur

# Einführung (1)

- Im Zentrum medizinischer Untersuchungen stehen oft die Beziehungen zwischen der Zeit bis zu einem bestimmten Ereignis (z.B. Tod, Erkrankung) und verschiedenen erklärenden Variablen (z.B. Behandlungen, Design-Variablen).
- Nicht- und semi-parametrische Modelle sind in der Überlebenszeitanalyse beliebt, da sie trotz weitgehender Allgemeinheit und Anwendbarkeit genügend Struktur für Interpretation und Verständnis einer konkreten Anwendung enthalten.
- Entsprechend einer Arbeit von Cox (1972) sollen im folgenden Überlebenszeit-Modelle betrachtet werden, die auf *Hazard*- bzw. *Intensity*-Funktionen basieren.

(Sinha und Dey, 1997, 1998)

## Einführung (2)

- “The field of survival analysis has been dominated by the frequentist philosophy. The Bayesian approach has important potential in survival analysis, especially for handling complex models and data structures, such as dependent censoring, noncompliance, interval censoring, and multivariate failure times. Considerable progress has already been made in this area [. . .]. One expects to see further developments and refinements of the Bayesian methods as well as their application in clinical trials over the coming decades.”

(Fleming und Lin, 2000)

# Typen von Ereigniszeiten (1)

- **Univariate Ereigniszeiten** – den einzelnen Beobachtungseinheiten einer Studie kann das Ereignis höchstens einmal widerfahren.
  - *Finkelstein und Wolfe (1985), Finkelstein (1986)*
    - \* Vergleich von primärer Radiotherapie plus adjuvanter Chemotherapie ( $n_{RC} = 48$ ) und primärer Radiotherapie allein ( $n_R = 46$ ) bei frühem Brustkrebs
    - \* Nachuntersuchungen in regelmäßigen Abständen (alle 3-4 Wochen)
    - \* Die Ereigniszeit ist die Zeit ab Behandlungsbeginn bis zur Verschlechterung des kosmetischen Zustands der Brust.
    - \* Die Patientinnen werden, gegeben ihre zufällig zugeteilte Therapie, als voneinander unabhängig angenommen.
    - \* *Intervall-zensierte* Ereigniszeit, z.B. (6, 18]
    - \* *Rechts-zensierte* Ereigniszeit, z.B. 23+

## Typen von Ereigniszeiten (2)

- **Multiple Ereigniszeiten** – von Interesse ist die Intensität des wiederholten Auftretens eines wichtigen, nicht notwendig fatalen Ereignisses bei einzelnen Beobachtungseinheiten; diese Intensität kann als Maß für die Morbidität oder den physischen Status einer Beobachtungseinheit dienen.
  - *Gail et al. (1980)*
    - \* Ereigniszeiten sind die Zeiten bis zum Auftreten von Tumoren in 48 Ratten, denen ein Karzinogen injiziert wurde und die danach in zwei Behandlungsgruppen randomisiert wurden.
    - \* Regelmäßige Nachuntersuchungen (nur zwei pro Woche)
    - \* Pro Ratte wird die Anzahl der jeweils neuen Tumoren erfasst.
    - \* Viele Bindungen von Ereigniszeiten innerhalb und zwischen einzelnen Ratten
    - \* Alle Beobachtungen sind rechts-zensiert nach dem 128-sten Tag.

## Typen von Ereigniszeiten (3)

- **Multivariate Ereigniszeiten** – die Ereigniszeiten verschiedener Beobachtungseinheiten einer Studie sind voneinander abhängig (z.B. bei *Matched-pair-Studien*).
  - *Freireich et al. (1963)*
    - \* 21 matched pairs (*clusters*) von Leukämie-Patienten
    - \* Ereigniszeit ist die Zeit in Remission.
    - \* Die Remissionszeiten innerhalb desselben Clusters sind wahrscheinlich voneinander abhängig.
    - \* Kovariablen wie Behandlung (6MP, Placebo) und Remissionsstatus (partiell, vollständig) können die Ereigniszeit beeinflussen.

# Semi-parametrische Modelle (1)

- Univariate Überlebenszeiten

- *Proportional-hazards-Modell (Cox, 1972)*

$$h(t|\mathbf{x}) = \lim_{\Delta t \downarrow 0} P(t \leq T < t + \Delta t | \mathbf{x}) / \Delta t$$

$$= h_0(t) \exp(\boldsymbol{\beta}^T \mathbf{x})$$

$$H_0(t) = \int_0^t h_0(u) du$$

$$S(t|\mathbf{x}) = P(T > t | \mathbf{x})$$

$$= \exp\{-H_0(t) \exp(\boldsymbol{\beta}^T \mathbf{x})\}$$

- \* Überlebenszeit  $T$  absolut-stetig
- \*  $\mathbf{x}$  Kovariablen
- \*  $\boldsymbol{\beta}$  unbekannte Regressionskoeffizienten
- \*  $h(t|\mathbf{x})$  Hazardfunktion
- \*  $h_0(\cdot)$  *baseline* Hazardfunktion
- \*  $H_0(\cdot)$  kumulative *baseline* Hazardfunktion
- \*  $S(t|\mathbf{x})$  Überlebensfunktion

## Semi-parametrische Modelle (2)

– Diskretes Cox-Modell (Cox, 1972)

$$S(a_j | \boldsymbol{\beta}, \mathbf{x}) = \prod_{k=1}^j (1 - \alpha_k)^{\exp(\boldsymbol{\beta}^T \mathbf{x})}$$

für  $j = 1, 2, \dots, g$

\* Überlebenszeit  $T$  nicht absolut-stetig

\* Partition der Zeitachse  $[0, \infty)$ :  $I_1, I_2, \dots, I_{g+1}$  mit  $I_j = [a_{j-1}, a_j)$ ,  $j = 1, 2, \dots, g + 1$ ,  $a_0 = 0$  und  $a_{g+1} = \infty$

\* Diskrete baseline Hazardrate  $\alpha_k = P(a_{k-1} < T \leq a_k | T > a_{k-1})$

– Grenzen frequentistischer Methodik

\* Methodik für rechts-zensierte Daten erlaubt zeitabhängige Kovariablen, aber keine zeitabhängigen Kovariablen-Effekte (Cox und Oakes, 1984).

\* Keine Methodik zur Überprüfung von Modellannahmen bei intervall-zensierten Daten (proportional hazards, konstante Kovariablen-Effekte)



# Semi-parametrische Modelle (3)

- Multiple Ereigniszeiten

- *Proportional-intensity-Modell (Sinha, 1993)* – Ereignisse werden durch einen bedingten Poisson-Prozess beschrieben mit bedingter Intensitätsfunktion

$$b_i(t|\mathbf{x}_i, w_i) = b_0(t) \exp(\boldsymbol{\beta}^T \mathbf{x}_i) w_i$$

- \* Unbeobachtete *frailties*  $w_i \sim_{\text{iid}} \text{Gamma}$
- \* Deren unbekannte Varianz  $\kappa$  bestimmt die Abhängigkeiten zwischen Ereignissen innerhalb von sowie die Heterogenität der Intensitätsprozesse zwischen Beobachtungseinheiten.
- Das Modell ist nützlich, falls
  - \* der Einfluss von Kovariablen bzw. die die individuelle Ereignisrate beeinflussende Heterogenität zwischen Beobachtungseinheiten untersucht werden soll,
  - \* jeder Beobachtungseinheit mehrere Ereignisse widerfahren können,
  - \* nur mäßig viele Beobachtungseinheiten in einer relativ kurzen Zeitspanne beobachtet werden können.

## Semi-parametrische Modelle (4)

- Grenzen anderer frequentistischer Modelle (Oakes, 1992)
  - \* Individuelles Ereignis-Risiko bleibt durch vorangehende Ereignisse unbeeinflusst.
  - \* Heterogenität zwischen Beobachtungseinheiten kann nicht erfasst werden.
  - \* Bindungen können nicht berücksichtigt werden.

### • Multivariate Überlebenszeiten

- *Frailty-Modell* (Andersen et al., 1993) – bedingte Hazardfunktion der  $i$ -ten Überlebenszeit  $T_{li}$  im  $l$ -ten Cluster

$$h_{li}(t|x_{li}, w_l) = h_0(t) \exp(\beta x_{li}) w_l$$

- \*  $x_{li}$  beobachtete Kovariable
- \*  $\beta$  unbekannter Regressionskoeffizient
- \*  $h_0(\cdot)$  unbekannte baseline Hazardfunktion
- \* Frailties  $w_l \sim_{\text{iid}}$  Gamma beschreiben das verbleibende zufällige Risiko des  $l$ -ten Clusters nach Kovariablenadjustierung.

# A-priori-Prozesse (1)

- Nicht- und semi-parametrische Bayes-Ansätze erfordern A-priori-Zufallsprozesse, die Verteilungs- oder Hazardfunktionen (z.B.  $h_0$ ) generieren.
  - Dirichlet-a-priori-Prozess (Ferguson, 1973)
    - \* war erstes zufälliges Wahrscheinlichkeitsmaß in nicht-parametrischem Bayes-Kontext;
    - \* generiert als A-priori-Verteilung eine Verteilungsfunktion  $F(\cdot)$ , die wiederum die Daten generiert;
    - \* ist bei natürlicher Ordnung des Stichprobenraums wenig geeignet;
    - \* hat keine einfache Darstellung bzw. Interpretation über die Hazardfunktion,  $h(t) = F'(t)/(1 - F(t))$ .
  - Gamma-a-priori-Prozess (Kalbfleisch, 1978; Burridge, 1981)
    - \* Verteilungsfunktion von  $T$ ,  $F(t) = 1 - \exp\{-H(t)\}$ , wobei  $H(\cdot)$  Realisation eines nicht-fallenden Prozesses mit positiven unabhängigen Zuwächsen in disjunkten Intervallen (Lévy-Prozess)

## A-priori-Prozesse (2)

- \* Häufigst benutzt:  $H(t_2) - H(t_1)$  Gamma-verteilt mit Erwartungswert  $\Lambda^*(t_2) - \Lambda^*(t_1)$  und Varianz  $\{\Lambda^*(t_2) - \Lambda^*(t_1)\}/c$  für  $t_2 > t_1$  und  $c > 0$ ; nicht-fallende positive Funktion  $\Lambda^*$  ist Erwartung des A-priori-Prozesses; Hyperparameter  $c$  interpretierbar als Grad initialer Übereinstimmung von  $\Lambda^*$  und  $H(\cdot)$
- \* Drei Nachteile: 1. Modelltyp kann, falls angewendet auf stetige Überlebenszeiten, zu falschen Schlussfolgerungen führen; 2. Zuwächse in angrenzenden Intervallen sind unabhängig; 3. diskrete Realisierung des Stichprobenpfads fast-sicher, also haben Bindungen unter der Randverteilung positive Wahrscheinlichkeit
- Weitere A-priori-Prozesse
  - \* Korrelierter-a-priori-Prozess (Leonard, 1978)
  - \* Beta-a-priori-Prozess (Hjort, 1990)
  - \* *Random-finite-mixture*-Prozess (Gelfand und Mallick, 1995)

# Bayes'sche Analyse univariater Überlebenszeiten (1)

- Cox-Modell:  $S(t|\mathbf{x}) = \exp\{-e^{\beta^T \mathbf{x}} H_0(t)\}$ 
  - $\beta$  hat A-priori-Verteilung  $\tau(\cdot)$
  - $H_0(t)$  (oder  $h_0(t)$ ) hat geeigneten A-priori-Prozess
  - Partition der Zeitachse  $[0, \infty)$ :  $I_1, I_2, \dots, I_{g+1}$  mit  $I_j = [a_{j-1}, a_j)$ ,  $j = 1, 2, \dots, g+1$ ,  $a_0 = 0$  und  $a_{g+1} = \infty$
  - $H_j = H_0(a_j) - H_0(a_{j-1})$  für  $j = 1, 2, \dots, g$
  - $T_1, T_2, \dots, T_N$  Überlebenszeiten,

$$P(T_k \in I_j | \mathbf{H}, \mathbf{x}_k) = \exp(-e^{\beta^T \mathbf{x}_k} \sum_{m=1}^{j-1} H_m) \times [1 - \exp(-e^{\beta^T \mathbf{x}_k} H_j)]$$

für  $k = 1, 2, \dots, N$

# Bayes'sche Analyse univariater Überlebenszeiten (2)

– Vollständige Likelihood

$$L(\boldsymbol{\beta}, \mathbf{H} | \text{gruppierte Daten}) \propto \prod_{j=1}^g G_j \quad \text{wobei}$$

$$G_j = \exp(-H_j \sum_{k \in R_j} e^{\boldsymbol{\beta}^T \mathbf{x}_k}) \times \prod_{l \in D_j} [1 - \exp(-e^{\boldsymbol{\beta}^T \mathbf{x}_l} H_j)]$$

- \*  $R_j$  Patienten unter Risiko im  $j$ -ten Zeitintervall
- \*  $D_j$  Patienten mit Ereignis im  $j$ -ten Zeitintervall

– Im Falle des diskreten Cox-Modells ist

$$G_j = (1 - \alpha_j)^{\sum_{k \in R_j} \exp(\boldsymbol{\beta}^T \mathbf{x}_k)} \times \alpha_j^{\sum_{l \in D_j} \exp(\boldsymbol{\beta}^T \mathbf{x}_l)}$$

# Bayes'sche Analyse univariater Überlebenszeiten (3)

- Gemeinsame A-posteriori-Verteilung

$$p(\boldsymbol{\beta}, \mathbf{H} | \text{gruppierte Daten}) \propto$$

$$L(\boldsymbol{\beta}, \mathbf{H} | \text{gruppierte Daten}) \times \phi(\mathbf{H}) \times \tau(\boldsymbol{\beta})$$

- \*  $\phi(\mathbf{H})$  gemeinsame Verteilung der  $H_j$
  - \*  $\tau(\boldsymbol{\beta})$  A-priori-Verteilung für  $\boldsymbol{\beta}$
- Spezifikation von A-priori-Verteilungen (Geisser, 1993; Ibrahim und Sinha, 1998)
    - Verwendung historischer Daten zur Konstruktion von informativen A-priori-Verteilungen für das Cox-Modell
    - *Elicitation* basiert auf A-priori-Vorhersage von Beobachtungen  $y_0$  und deren Unsicherheit  $a_0$
    - $y_0$  und  $a_0$  führen *semi-automatisch* zu A-priori-Verteilungen für  $\boldsymbol{\beta}$
    - Eine Hauptanwendung besteht in der Modellselektion.

# WinBUGS-*Mice* - Beispiel (1)

## Weibull-Regression

- *Grieve (1987), Dellaportas und Smith (1993)*
  - Untersuchungen zur Fotokarzinogenität an vier Gruppen à 20 Mäusen
  - Überlebenszeit (\*Zensierung)
  - Für zensierte Beobachtungen wird eine gestutzte Weibull-Verteilung angenommen.

$$f(t_i, \mathbf{z}_i) = r e^{\beta^T \mathbf{z}_i} t_i^{r-1} \exp(-e^{\beta^T \mathbf{z}_i} t_i^r)$$

$$\lambda_0(t_i) = r t_i^{r-1} \quad \text{baseline Hazardfunktion}$$

$$t_i \sim \text{Weibull}(r, \mu_i) \quad \text{mit} \quad \mu_i = e^{\beta^T \mathbf{z}_i}$$

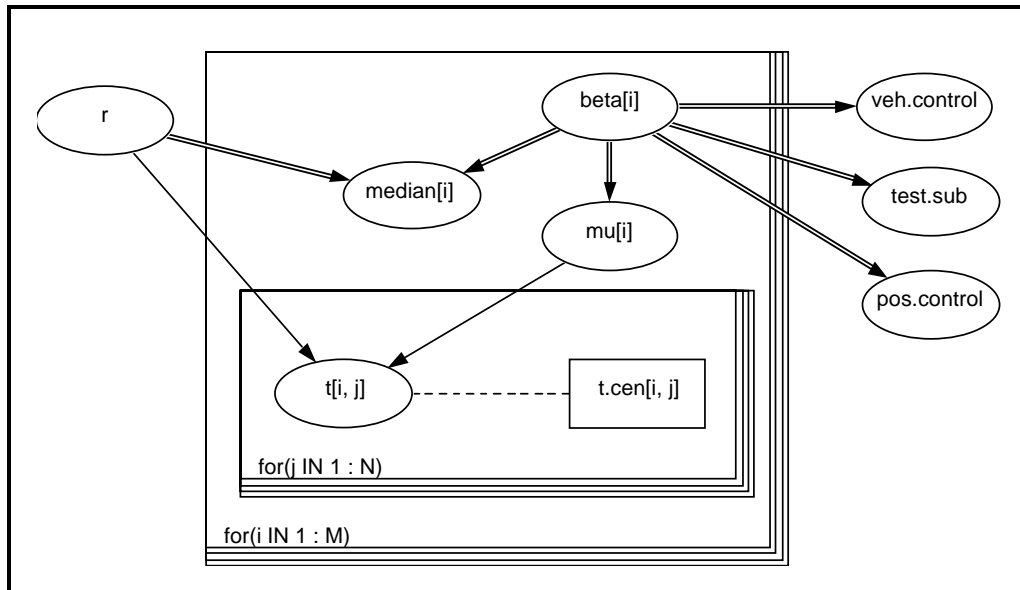
$$t_{\text{median}} = \sqrt[r]{\log(2e^{-\beta^T \mathbf{z}_i})}$$

Mouse	Irradiated control	Vehicle control	Test substance	Positive control
1	12	32	22	27
...				
18	*40	30	24	12
19	31	37	37	17
20	36	27	29	26



# WinBUGS-*Mice* - Beispiel (2)

## Weibull-Regression



```

model
{
  for(i in 1:M){ # M=4
    for(j in 1:N){ # N=20
      t[i, j]~dweib(r, mu[i])I(t.cen[i, j],)
    }
    mu[i]<-exp(beta[i])
    beta[i]~dnorm(0.0, 0.001)
    median[i]<-pow(log(2)*exp(-beta[i]), 1/r)
  }
  r~dexp(0.001)
  veh.control<-beta[2]-beta[1]
  test.sub<-beta[3]-beta[1]
  pos.control<-beta[4]-beta[1]
}

```

# WinBUGS-*Kidney*-Beispiel (1)

## Weibull-Regression mit zufälligen Effekten

- *McGilchrist und Aisbett (1991)*
  - Erst- und Zweitinfektion von 38 Dialyse-Patienten
  - McGilchrist und Aisbett benutzen Cox-Modell mit individueller multiplikativer Frailty
  - Hier: Weibull-Modell mit individuellem additivem zufälligem Effekt  $b_i$  im Exponenten des Hazard-Modells
    - \* Risikovariablen: Alter, Geschlecht und zugrundeliegende Erkrankung
    - \* Für zensierte Beobachtungen wird wieder eine gestutzte Weibull-Verteilung angenommen.

Patient Number	Recurrence time $t$	Event <sup>†</sup>	Age at time $t$	Sex*	Disease**
1	8,16	1,1	28,28	0	0
2	23,13	1,2	48,48	1	1
3	22,28	1,1	32,32	0	0
4	447,318	1,1	31,32	1	0
...					
35	119,8	1,1	22,22	1	1
36	54,16	2,2	42,42	1	1
37	6,78	2,1	52,52	1	3
38	63,8	1,2	60,60	0	3

<sup>†</sup>2=censored; \*1=female; \*\*0=other, 1=GN, 2=AN, 3=PKD

# WinBUGS-*Kidney*-Beispiel (2)

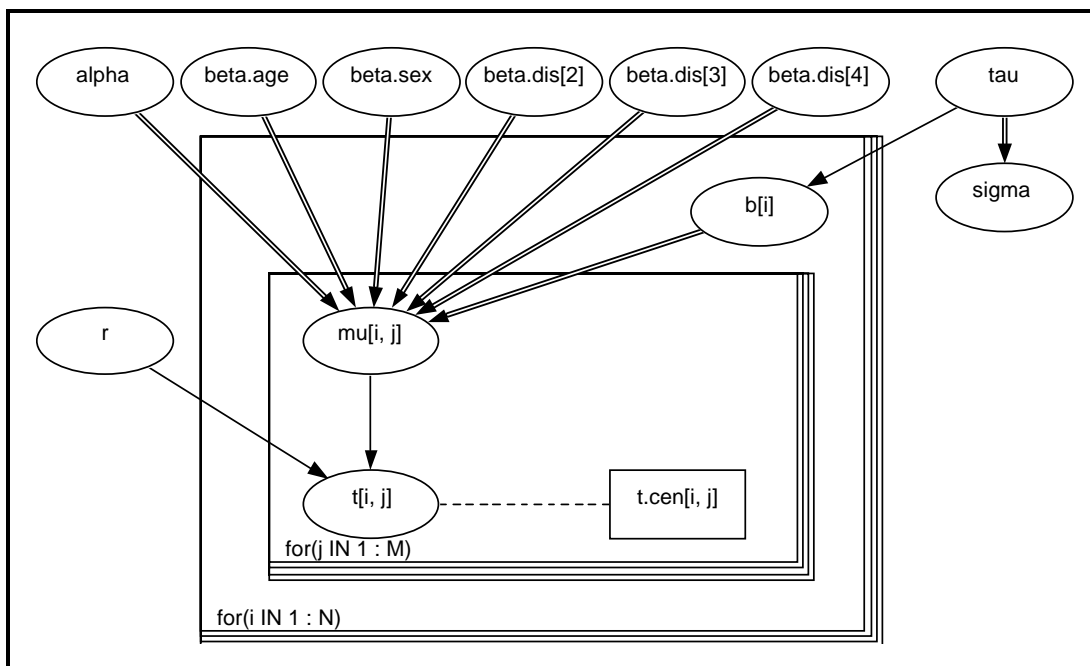
## Weibull-Regression mit zufälligen Effekten

\* Für  $i = 1, 2, \dots, 38$  und  $j = 1, 2$

$$t_{ij} \sim \text{Weibull}(r, \mu_{ij})$$

$$\log \mu_{ij} = \alpha + \beta_{\text{age}} \text{age}_{ij} + \beta_{\text{sex}} \text{sex}_i + \beta_{\text{dis1}} \text{disease}_{i1} + \beta_{\text{dis2}} \text{disease}_{i2} + \beta_{\text{dis3}} \text{disease}_{i3} + b_i$$

$$b_i \sim \text{Normal}(0, \tau)$$



# WinBUGS-*Kidney*-Beispiel (3)

## Weibull-Regression mit zufälligen Effekten

```
model
{
  for(i in 1:N){
    for(j in 1:M){
# Survival times bounded below by censoring times:
      t[i,j]~dweib(r,mu[i,j])I(t.cen[i,j],)
      log(mu[i,j])<-alpha+beta.age*age[i,j]
                          + beta.sex*sex[i]
                          + beta.dis[disease[i]]+b[i]
    }
# Random effects:
      b[i]~dnorm(0.0,tau)
  }
# Priors:
      alpha~dnorm(0.0,0.0001)
      beta.age~dnorm(0.0,0.0001)
      beta.sex~dnorm(0.0,0.0001)
# beta.dis[1]<-0; # corner-point constraint
      for(k in 2:4){
        beta.dis[k]~dnorm(0.0,0.0001)
      }
      tau~dgamma(1.0E-3,1.0E-3)
      r~dgamma(1.0,1.0E-3)
      sigma<-1/sqrt(tau) # s.d. of random effects
}
```

# WinBUGS-*Leuk*-Beispiel (1)

## Cox-Regression

- *Freireich et al. (1963)*
  - Vergleich zweier Leukämie-Therapien
  - Zeit in Remission (21 *matched pairs*)
  - Cox-Modell in Zählprozess-Notation (Andersen und Gill, 1982; Clayton, 1991)
    - \*  $N_i(t)$  Anzahl der Ereignisse bis zur Zeit  $t$  für Individuum  $i = 1, 2, \dots, n$
    - \* Der zugehörige Intensitätsprozess  $I_i(t)$  ist gegeben durch

$$I_i(t) = E(dN_i(t) | \mathcal{F}_{t-})$$

wobei  $dN_i(t)$  den Zuwachs von  $N_i$  im Zeitintervall  $[t, t + dt)$ ,  $\mathcal{F}_{t-}$  die Daten bis zur Zeit  $t$  und  $E(dN_i(t) | \mathcal{F}_{t-})$  das Risiko von Individuum  $i$  im Zeitintervall  $[t, t + dt)$  beschreibt.

Treatment	Survival time in weeks						
Placebo	1	1	2	2	3	4	4
	5	5	8	8	8	8	11
	11	12	12	15	17	22	23
6-MP	*6	6	6	6	7	9	*10
	10	*11	13	16	*17	*19	*20
	22	23	*25	*32	*32	*34	*35

\*indicates censoring

## WinBUGS-*Leuk*-Beispiel (2)

### Cox-Regression

- \* Für  $dt \rightarrow 0$  ist der *instantaneous hazard*

$$I_i(t) = Y_i(t) \lambda_0(t) \exp(\beta^T \mathbf{z}_i)$$

wobei  $Y_i(t)$  die Werte 1 bzw. 0 annimmt, falls Individuum  $i$  zur Zeit  $t$  unter Risiko steht oder nicht.

- \* Mit bekanntem  $D = \{N_i(t), Y_i(t), \mathbf{z}_i; i = 1, 2, \dots, n\}$ , sowie unbekanntem  $\beta$  (parametrisch) und  $\Lambda_0(t) = \int_0^t \lambda_0(u) du$  (nicht-parametrisch) ist die gemeinsame A-posteriori-Verteilung gegeben durch

$$P(\beta, \Lambda_0(\cdot) | D) \propto P(D | \beta, \Lambda_0(\cdot)) P(\beta) P(\Lambda_0(\cdot))$$

- \* Unter der Annahme nicht-informativer Zensierungen ist die Likelihood

$$P(D | \beta, \Lambda_0(\cdot)) \propto \prod_{i=1}^n \left( \prod_{t \geq 0} I_i(t)^{dN_i(t)} \right) \exp\left(- \int_{t \geq 0} I_i(t) dt\right)$$

# WinBUGS-*Leuk*-Beispiel (3)

## Cox-Regression

- \* Im Wesentlichen gilt für  $dN_i(t)$  in  $[t, t + dt)$

$$dN_i(t) \sim \text{Poisson}(I_i(t)dt)$$

- \* Bezeichnet  $d\Lambda_0(t) = \lambda_0(t)dt$  den Zuwachs der kumulativen baseline Hazardfunktion im Zeitintervall  $[t, t + dt)$ , so ist

$$I_i(t)dt = Y_i(t) \exp(\beta^T z_i) d\Lambda_0(t)$$

- \* Kalbfleisch (1978)

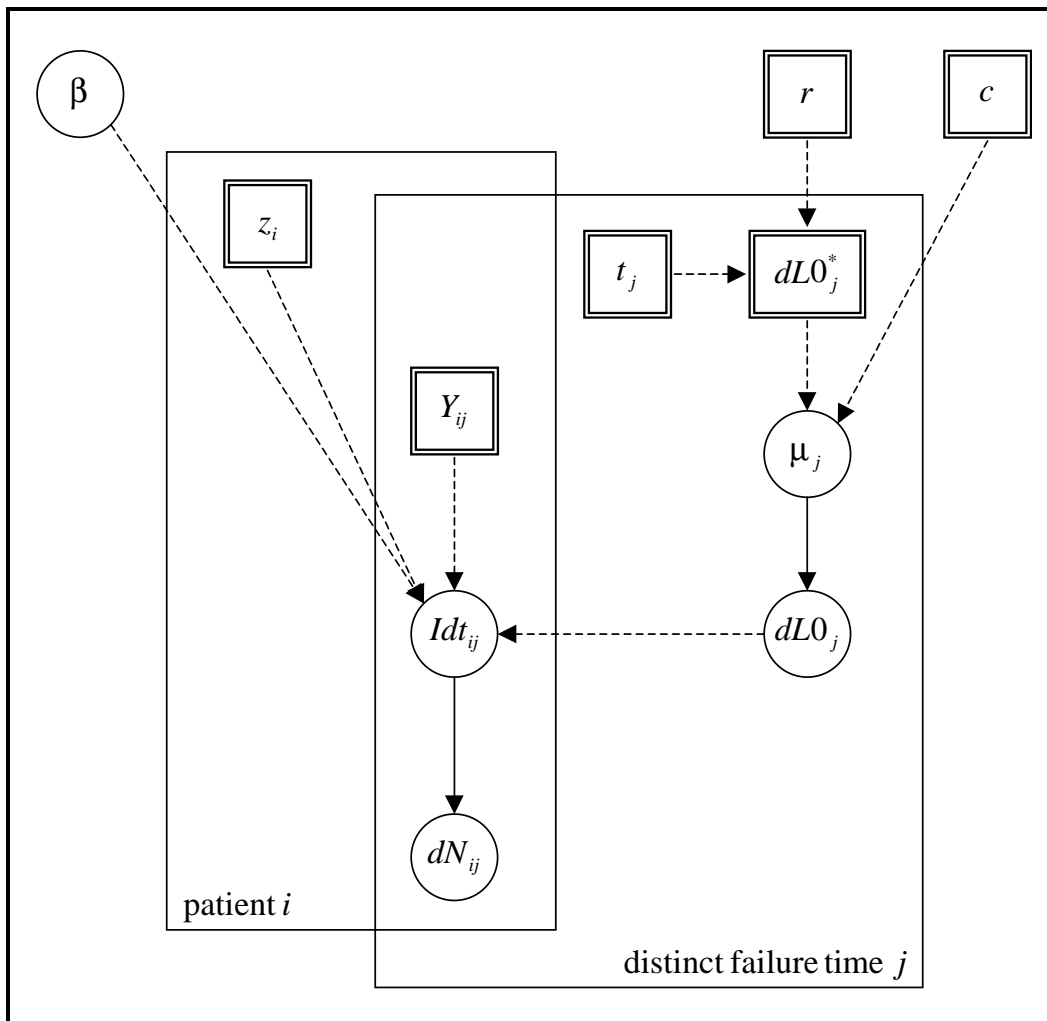
$$d\Lambda_0(t) \sim \text{Gamma}(c \cdot d\Lambda_0^*(t), c)$$

da die Gamma-Verteilung die konjugierte A-priori-Verteilung im Zusammenhang mit dem Poisson-Parameter ist.  $d\Lambda_0^*(t)$  kann als initiale Schätzung mit Konfidenzgrad  $c$  (kleine  $c$ -Werte entsprechen geringem A-priori-Wissen) der unbekanntem Hazardfunktion interpretiert werden. Im Beispiel ist  $d\Lambda_0^*(t) = rdt$ , dabei ist  $r$  die Schätzung der Ereignisrate pro Zeiteinheit.

# WinBUGS-*Leuk* - Beispiel (4)

## Cox-Regression

– Modell ohne Berücksichtigung des *Matching*





# WinBUGS-*Leuk* - Beispiel (5)

## Cox-Regression

```

model
{
# Set up data
  for(i in 1:N){ # N=42
    for(j in 1:T){ # T=17 unique failure times
# risk set=1 if obs.t>=t
      Y[i,j]<-step(obs.t[i]-t[j]+eps)
# counting process jump=1 if obs.t in [t[j],t[j+1])
#           i.e. if t[j]<=obs.t<t[j+1]
      dN[i,j]<-Y[i,j]*step(t[j+1]-obs.t[i]-eps)*fail[i]
    }
  }
# Model
  for(j in 1:T){
#   beta0[j]~dnorm(0,0.001) # include for Poisson trick
    for(i in 1:N){
      dN[i,j]~dpois(Idt[i,j]) # Likelihood
      Idt[i,j]<-Y[i,j]*exp(beta*Z[i])*dL0[j] # Intensity
# Poisson trick - independent log-normal hazard increments
#           - enables dL0, c, r, mu to be dropped
# Idt[i,j]<-Y[i,j]*exp(beta0[j]+beta*Z[i]) # Intensity
    }
    dL0[j]~dgamma(mu[j],c)
    mu[j]<-dL0.star[j]*c # prior mean hazard
# Survivor function=exp(-Integral{l0(u)du})^exp(beta*z)
    S.treat[j]<-pow(exp(-sum(dL0[1:j])),exp(beta*-0.5))
    S.placebo[j]<-pow(exp(-sum(dL0[1:j])),exp(beta*0.5))
  }
  c<-0.001; r<-0.1
  for(j in 1:T){dL0.star[j]<-r*(t[j+1]-t[j])}
  beta~dnorm(0.0,0.000001)
}

```

# WinBUGS-*LeukFr*-Beispiel (1)

## Cox-Regression mit zufälligen Effekten

- *Freireich et al. (1963)*
  - 21 Patienten-Paare gematched nach Remissionsstatus (partiell oder vollständig)
  - Modellierung der Clusterung von Ereigniszeiten innerhalb von Paaren durch zufälligen Effekt (*frailty term*) im *Proportional-hazards*-Modell

$$I_i(t)dt = Y_i(t) \exp(\beta^T \mathbf{z}_i + b_{\text{pair}_i}) d\Lambda_0(t)$$

$$b_{\text{pair}_i} = \text{Normal}(0, \tau)$$

für  $i = 1, 2, \dots, 42$  und  $\text{pair}_i = 1, 2, \dots, 21$ .

# WinBUGS-*LeukFr*-Beispiel (2)

## Cox-Regression mit zufälligen Effekten

```
[...]  
# Model  
  for(j in 1:T){  
    for(i in 1:N){  
      # Likelihood  
      dN[i,j]~dpois(Idt[i,j])  
      # Intensity  
      Idt[i,j]<-Y[i,j]*exp(beta*Z[i]+b[pair[i]])*dL0[j]  
    }  
  }  
[...]  
  for(k in 1:Npairs){  
    b[k]~dnorm(0.0,tau)  
  }  
  tau~dgamma(0.001,0.001)  
[...]
```

# Literatur

- [1] O O Aalen. Frailty models. In B S Everitt and G Dunn, editors, *Statistical Analysis of Medical Data. New Developments*, pages 59–74. Arnold, London, 1998.
- [2] K R Abrams. Regression models for survival data. In B S Everitt and G Dunn, editors, *Statistical Analysis of Medical Data. New Developments*, pages 13–58. Arnold, London, 1998.
- [3] P K Andersen, O Borgan, R D Gill, and N Keiding. *Statistical Models Based on Counting Processes*. Springer, New York, 1993.
- [4] P K Andersen and R D Gill. Cox's regression model for counting process: a large sample study. *Annals of Statistics*, 10:1100–1120, 1982.
- [5] J Burridge. Empirical Bayes analysis of survival time data. *Journal of the Royal Statistical Society, Series A*, 43:65–75, 1981.
- [6] D G Clayton. A Monte Carlo method for Bayesian inference in frailty models. *Biometrics*, 47:467–485, 1991.
- [7] D R Cox. Regression models and life tables. *Journal of the Royal Statistical Society, Series B*, 34:187–220, 1972.
- [8] D R Cox and D Oakes. *Analysis of Survival Data*. Chapman & Hall, London, 1984.
- [9] P Dellaportas and A F M Smith. Bayesian inference for generalized linear and proportional hazards model via gibbs sampling. *Applied Statistics*, 42:443–460, 1993.
- [10] T S Ferguson. A Bayesian analysis of some nonparametric problems. *The Annals of Statistics*, 1:209–230, 1973.
- [11] D M Finkelstein. A proportional hazards model for interval-censored failure time data. *Biometrics*, 42:845–854, 1986.
- [12] D M Finkelstein and R A Wolfe. A semiparametric model for regression analysis of interval-censored failure time data. *Biometrics*, 41:933–945, 1985.
- [13] T R Fleming and D Y Lin. Survival analysis in clinical trials: past developments and future directions. *Biometrics*, 56:971–983, 2000.
- [14] E J Freireich, E Gehan, and *others*. The effect of 6-mercaptopurine on the duration of steroid-induced remissions in acute leukaemia: a model for evaluation of other potentially useful therapy. *Blood*, 21:699–716, 1963.
- [15] M H Gail, T J Santner, and C C Brown. An analysis of comparative carcinogenesis experiments with multiple times to tumor. *Biometrics*, 36:255–266, 1980.

- [16] S Geisser. *Predictive Inference*. Chapman & Hall, New York, 1993.
- [17] A E Gelfand and B Mallick. Bayesian analysis of proportional hazards model built from monotone functions. *Biometrics*, 51:843–852, 1995.
- [18] B K Ghosh and *others*. Bayesian analysis of censored data. *Statistics & Probability Letters*, 41:255–265, 1999.
- [19] A P Grieve. Applications of Bayesian software: two examples. *Statistician*, 36:283–288, 1987.
- [20] N L Hjort. Nonparametric Bayes estimators based on beta processes in models for life history data. *The Annals of Statistics*, 18:1259–1294, 1990.
- [21] J G Ibrahim and D Sinha. Prior elicitation for semiparametric Bayesian survival analysis. In D Dey, P Müller, and D Sinha, editors, *Practical Nonparametric and Semiparametric Bayesian Statistics*, pages 273–292. Springer, New York, 1998.
- [22] J D Kalbfleisch. Nonparametric Bayesian analysis of survival time data. *Journal of the Royal Statistical Society, Series B*, 40:214–221, 1978.
- [23] T Leonard. Density estimation, stochastic processes and prior information. *Journal of the Royal Statistical Society, Series B*, 40:113–146, 1978.
- [24] C McGilchrist and C Aisbett. Regression with frailty in survival analysis. *Biometrics*, 47:461–466, 1991.
- [25] D Oakes. Frailty models for multiple event-times. In J P Klein and P K Goel, editors, *Survival Analysis: State of the Art*, pages 371–379. Kluwer, Dordrecht, 1992.
- [26] D Sinha. Semiparametric Bayesian analysis of multiple event time data. *Journal of the American Statistical Association*, 88:979–983, 1993.
- [27] D Sinha and D K Dey. Semiparametric Bayesian analysis of survival data. *Journal of the American Statistical Association*, 92:1195–1212, 1997.
- [28] D Sinha and D K Dey. Survival analysis using semiparametric Bayesian methods. In D Dey, P Müller, and D Sinha, editors, *Practical Nonparametric and Semiparametric Bayesian Statistics*, pages 195–211. Springer, New York, 1998.