

## 6. Wielokrotne zdarzenia

Struktura danych

„Odporny” estymator wariancji

Modelowanie intensywności zdarzeń

Modele brzegowe/warunkowe

Modele „podatności”

Kopuły

# „Odporne” szacowanie wariancji dla modelu PH

- ◆ Niech  $\mathbf{J}$  będzie macierzą z  $i$ -tym rzędem równym  $(\hat{\beta}_{(i)} - \hat{\beta})^T$
- ◆ Oszacowanie wariancji metodą „syczorykową”:

$$\text{Var}(\hat{\beta}) \approx \frac{n-p}{n} (\mathbf{J} - \bar{\mathbf{J}})^T (\mathbf{J} - \bar{\mathbf{J}}) \equiv V_J$$

( $p$  – liczba parametrów;  $\bar{\mathbf{J}}$  to macierz zawierająca średnie kolumn)

# Reszty „score” i „odporne” oszacowanie wariancji

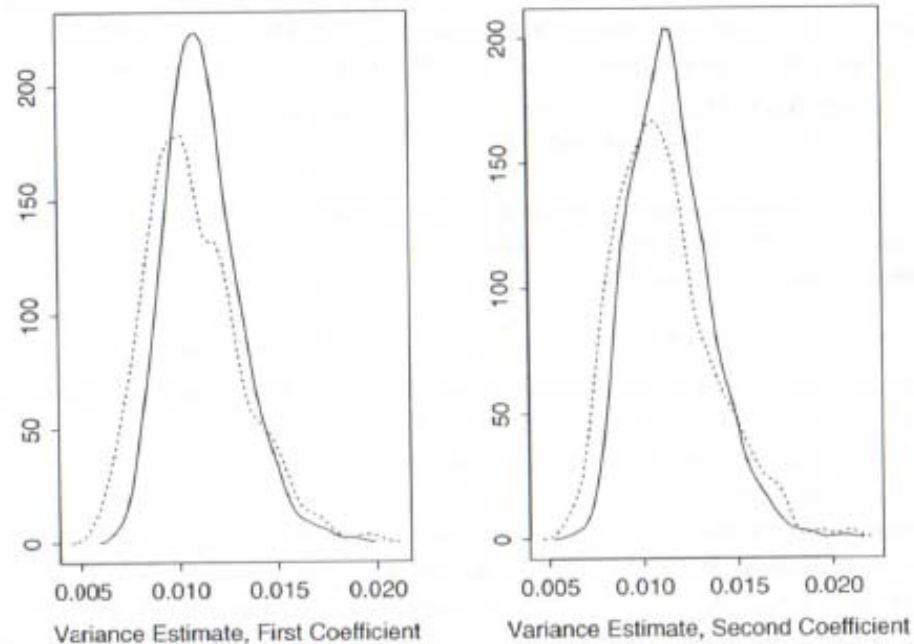
- ◆ *Score process dla obserwacji  $i$*   $U_i(\beta, t) = \int_0^t \{Z_i(u) - \bar{Z}(\beta, u)\}^T dM_i(u)$ 
  - wektor !
- ◆ *Reszta „score” dla obserwacji  $i$  i zmiennej  $j$ :*  $\hat{r}_{Uij} = U_{ij}(\hat{\beta}, \infty)$ 
  - Macierz reszt,  $\mathbf{U}$ , rozmiaru  $n \times p$
  - $\sum_i U_{ij} = 0$
- ◆ *Skalowana reszta „score”:*  $\hat{r}_{Ui}^* = \hat{r}_{Ui} I(\hat{\beta})^{-1} = \hat{r}_{Ui} \text{var}(\hat{\beta}) \approx \hat{\beta}_{(i)} - \hat{\beta}$
- ◆ *Czyli:*  $V_J \approx (U^*)^T U^*$

# „Odporne” oszacowanie wariancji

- ◆ Mamy:  $V_J \approx (U^*)^T U^* = \text{var}(\hat{\beta})(U^T U) \text{var}(\hat{\beta}) = I^{-1}(U^T U)I^{-1}$
- ◆ Estymator „kanapkowy” („sandwich estimator”)
  - $U^T U$  „poprawia” oszacowanie, uwzględniając korelację czasów
- ◆ Dla regresji linowej:  $V_J = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}(\sigma^2\mathbf{X}'\mathbf{R}\mathbf{X}\sigma^2)\sigma^2(\mathbf{X}'\mathbf{X})^{-1}$ 
  - $\mathbf{R}$  – diagonalna macierz z kwadratami reszt na przekątnej
  - Jeśli model jest poprawny,  $\mathbf{R} = \sigma^2 \mathbf{I} \rightarrow V_J = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$
  - Jeśli nie,  $(\sigma^2\mathbf{X}'\mathbf{R}\mathbf{X}\sigma^2)$  „poprawia” oszacowanie wariancji
    - Np. jeśli wariancja resztowa nie jest stała

# Zmienność „odpornego” oszacowania wariancji

- ◆ 1000 symulacji, rozkład wykładniczy, dwie zmienne niezależne  $\sim N(0,1)$  z  $\beta_1=0$  i  $\beta_2=0.5$ , 200 obserwacji, cenzurowanie: 50%



- ◆ Większa zmienność „odpornego” oszacowania wariancji

# Wielokrotne zdarzenia

- ◆ Czasy  $T_1^*, T_2^*, \dots$  dla zdarzeń
  - od rozpoczęcia obserwacji
- ◆ Możliwe cenzurowanie (rozważamy prawostronne)
- ◆ Różne scenariusze:
  - „Nawracające” zdarzenia („recurrent events”):  $T_1^*, T_2^*, \dots$  to czasy *kolejnych* zdarzeń tego samego typu
  - „Skupione” („clustered”) czasy zdarzeń:  $T_1^*, T_2^*, \dots$  to czasy zdarzeń tego samego typu, ale bez kolejności
  - „Wielowymiarowe” („multivariate”) czasy zdarzeń:  $T_1^*, T_2^*, \dots$  to czasy zdarzeń *różnych* typów, bez ustalonej kolejności

# Modele dla wielokrotnych zdarzeń

- ◆ Różne podejścia do modelowania:
  - modele dla intensywności/średniej liczby zdarzeń
  - czas do pierwszego zdarzenia
    - strata informacji!
  - modele brzegowe/warunkowe
  - modele z efektami losowymi („podatności”, frailty)
  - kopuły
  - ...
- ◆ Wybór modelu w zależności od tego, czy rozróżniamy typy zdarzeń, ich kolejność itd.

# Wielokrotne zdarzenia: charakterystyki rozkładu

- ◆ Łączna funkcja przeżycia:  $S(t_1, \dots, t_m) = P(T_1^* \geq t_1, \dots, T_m^* \geq t_m)$
- ◆ Brzegowe funkcje hazardu:  $\lambda_j(t) = \lim_{h_j \rightarrow 0^+} \frac{P(t_j \leq T_j^* < t_j + h_j \mid T_j^* \geq t_j)}{h_j}$



# Model Coxa dla brzegowych funkcji hazardu

- ♦ Model Coxa dla brzegowych funkcji hazardu:

$$\lambda_j(t) = \lambda_{0j}(t) \exp\{\mathbf{Z}_j'(t)\boldsymbol{\beta}\}$$

- ♦ Zakładamy (roboczo) niezależność zdarzeń:

$$L(\boldsymbol{\beta}) = \prod_{j=1}^m \prod_{i=1}^n \left\{ \frac{e^{\boldsymbol{\beta}' \mathbf{Z}_{ji}(t_{ji})}}{\sum_{l=1}^n Y_{jl}(t_{ji}) e^{\boldsymbol{\beta}' \mathbf{Z}_{jl}(t_{ji})}} \right\}^{\delta_{ji}}$$

- ♦ Nie jest częściową funkcją wiarygodności...
- ♦ ... ale rozwiązanie  $d\ln L/d\boldsymbol{\beta} = 0$  jest zgodne i  $\hat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, \mathbf{V})$ 
  - $\mathbf{V}$  można szacować przy użyciu „odpornego” estymatora  $\mathbf{V}_J$

# „Odporne” szacowanie wariancji dla skorelowanych czasów zdarzeń

- ◆ Jednostki z wielokrotnymi zdarzeniami
- ◆ Podobnie jak dla niezależnych obserwacji, używamy metody „scyzorykowej”
- ◆ Ale z wyłączeniem całych jednostek (a nie pojedynczych czasów)
  - dopasowujemy model
  - sumujemy skalowane reszty „score” dla jednostki
  - używamy macierzy  $(\mathbf{U}^*)^T \mathbf{U}^*$  jako przybliżenia  $\mathbf{V}_j$

# Skorelowane czasy zdarzeń: retinopatia cukrzycowa (1)

- ◆ Retinopatia cukrzycowa: uszkodzenie naczyń krwionośnych odżywiających siatkówkę u chorych na cukrzycę. Prowadzi do ślepoty.
- ◆ Próba kliniczna (1972-1975)
  - leczenie losowo wybranego oka fotokoagulacją
  - czas od randomizacji do poważnej utraty wzroku dla obu oczu
    - nie rozróżniamy kolejności zdarzeń ani typu (prawe/lewe oko)
  - 1742 chorych
  - analiza 197 chorych „wysokiego ryzyka”

# Skorelowane czasy zdarzeń: retinopatia cukrzycowa (2)

## ◆ Dane dla 197 chorych „wysokiego ryzyka”:

- subject id
- laser type: 1=xenon, 2=argon
- treated eye: 1=right 2=left
- age at diagnosis of diabetes
- type of diabetes: 1= juvenile (age at dx < 20), 2=adult
- risk group: 6-12 (for each eye)
- status: 0=censored, 1=blindness (for each eye)
- follow-up time (for each eye)

```
> head(diab)
```

	id	laser	trt_eye	age_dx	adult	time	status	treat	risk
1	5	2	2	28	2	46.23	0	1	9
2	5	2	2	28	2	46.23	0	0	9
3	14	2	1	12	1	42.50	0	1	8
4	14	2	1	12	1	31.30	1	0	6
5	16	1	1	9	1	42.27	0	1	11
6	16	1	1	9	1	42.27	0	0	11

# Skorelowane czasy zdarzeń: retinopatia cukrzycowa (3)

```
> diab.fit <- coxph(Surv(time,status) ~ adult + treat + cluster(id), data=diab)
```

```
> summary(diab.fit)
```

```
n= 394, number of events= 155
```

	coef	exp(coef)	se(coef)	robust se	z	Pr(> z )
adult	0.05388	1.05536	0.16211	0.17864	0.302	0.763
treat	-0.77893	0.45890	0.16893	0.14851	-5.245	1.56e-07 ***

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

	exp(coef)	exp(-coef)	lower .95	upper .95
adult	1.0554	0.9475	0.7436	1.4978
treat	0.4589	2.1791	0.3430	0.6139

```
Concordance= 0.589 (se = 0.023 )
```

```
Rsquare= 0.055 (max possible= 0.988 )
```

```
Likelihood ratio test= 22.48 on 2 df, p=1.312e-05
```

```
Wald test = 27.85 on 2 df, p=8.942e-07
```

```
Score (logrank) test = 22.36 on 2 df, p=1.395e-05,
```

```
Robust = 26.36 p=1.891e-06
```

(Note: the likelihood ratio and score tests assume independence of observations within a cluster, the Wald and robust score tests do not).

# „Nawracające” zdarzenia (1)

- ◆ Niech  $N^*(t)$  będzie liczbą zdarzeń w  $(0, t]$ .
  - Dopuszczamy  $N^*(t) > 1$
- ◆ Niech  $N(t)$  będzie *obserwowaną* liczbą zdarzeń w  $(0, t]$ .
  - $N(t) \leq N^*(t)$  z racji cenzurowania
- ◆ Rozważmy  $d\Lambda(t) = E\{dN^*(t) | N^*(u), 0 \leq u < t, \mathbf{Z}(t)\}$ 
  - *oczekiwany przyrost* liczby (intensywność) zdarzeń
- ◆ Niezależne cenzurowanie, gdy
$$E\{dN(t) | N(u), Y(u), 0 \leq u < t, \mathbf{Z}(t)\} = Y(t) d\Lambda(t)$$
  - dla jednostek pozostających pod obserwacją, oczekiwany przyrost obserwowanych zdarzeń odpowiada  $d\Lambda(t)$

# „Nawracające” zdarzenia (2)

- ◆  $d\Lambda(t) = E\{dN^*(t) \mid N^*(u), 0 \leq u < t, \mathbf{Z}(t)\}$  warunkuje ze względu na całą historię  $N^*(u)$

- ◆ W praktyce, rozważa się *brzegowe* intensywności, np.

$$d\Lambda_b(t) = E\{dN^*(t) \mid \mathbf{Z}(t)\}$$

- Po wyciąkowaniu względem rozkładu  $\{N^*(u) \mid 0 \leq u < t, \mathbf{Z}(t)\}$
  - Cenzurowanie nie może zależeć historii (liczby, czasów) uprzednich zdarzeń dla danej jednostki
- ◆ Dla „zewnątrznych” zmiennych zależnych od czasu mamy  $E\{dN^*(u) \mid \mathbf{Z}(u)\} = E\{dN^*(u) \mid \mathbf{Z}(t)\}$  dla  $u \leq t \rightarrow \Lambda_b(t) = E\{N^*(t) \mid \mathbf{Z}(t)\}$ 
    - oczekiwana liczba zdarzeń w  $(0, t]$

# „Nawracające” zdarzenia (3)

- ◆ Inna *brzegowa* funkcja intensywności, dla warunkowania ze względu na  $q=N^*(t^-)$ :

$$d\Lambda_q(t)=E\{dN^*(t) \mid N^*(t^-), \mathbf{Z}(t)\}$$

- ◆ Niezależne cenzurowanie, gdy

$$E[dN(t) \mid N(t^-), \{Y(u), 0 \leq u < t\}, \mathbf{Z}(t)] = Y(t) d\Lambda_q(t)$$

- cenzurowanie może zależeć od liczby uprzednich zdarzeń dla danej jednostki, ale nie od ich historii (czasów)



# „Nawracające” zdarzenia: modelowanie funkcji hazardu

- ◆ Jeśli  $dN^*(t) \leq 1$  (czasy absolutnie ciągłe), to  $\Lambda(t) = \int_0^t \lambda(u)du$  i  $d\Lambda(t) = \lambda(t)dt = P\{dN^*(t)=1 \mid N^*(u), 0 \leq u < t, \mathbf{Z}(t)\}$
- ◆ Otrzymujemy modele dla funkcji hazardu, np.
  - $\lambda(t) = \lambda_0(t) \exp\{\mathbf{Z}'(t)\boldsymbol{\beta}\}$ 
    - model Andersona i Gilla (1982)
  - $\lambda(t) = \lambda_0(t) \exp[ I\{N(t^-)=1\}\beta_1 + I\{N(t^-)=2\}\beta_2 + \dots ]$ 
    - uwzględnia liczbę uprzednich zdarzeń
  - $\lambda(t) = \lambda_{0q}(t) \exp\{\mathbf{Z}'(t)\boldsymbol{\beta}_q\}$  dla  $q = N(t^-)$ 
    - warstwowy ze względu na liczbę uprzednich zdarzeń
    - Prentice, Williams, Peterson (1981) (PWP)
  - $\lambda(t) = \lambda_{0q}(v) \exp\{\mathbf{Z}'(t)\boldsymbol{\beta}_q\}$  dla  $q = N(t^-)$  i  $v = t - T_q$ 
    - uwzględniający czas od chwili zaobserwowania  $q$  zdarzeń (PWP)

# „Nawracające” zdarzenia: modelowanie średniej liczby zdarzeń

- ◆ Model dla *brzegowej* intensywności

$$d\Lambda_b(t) = E\{dN^*(t) \mid \mathbf{Z}(t)\} = d\Lambda_0(t)e^{\mathbf{Z}'(t)\boldsymbol{\beta}}$$

- ◆ Nie może uwzględniać zmiennych typu  $I\{N(t^-)=q\}$  itp.
  - Bo zależą od *historii*  $\{N^*(u) \mid 0 \leq u < t, \mathbf{Z}(t)\}$

- ◆ Interpretacja w terminach oczekiwanej liczby zdarzeń

$$\Lambda_b(t) = \int_0^t e^{\mathbf{Z}'(t)\boldsymbol{\beta}} d\Lambda_0(t)$$

- Dla  $\mathbf{Z}(t) \equiv \mathbf{Z} \rightarrow \Lambda_b(t) = \Lambda_0(t)e^{\mathbf{Z}'\boldsymbol{\beta}} \rightarrow$  model *proporcjonalnych średnich*

# Wielowymiarowe czasy zdarzeń: modele brzegowe i warunkowe

## ◆ Ogólna idea:

- wybieramy model (warstwy itd.)
- używamy modelu Coxa pomijając zależność czasów
- obliczamy “poprawione” wariancje oszacowanych współczynników
  - oparte o skalowane reszty „score”

## ◆ Najczęściej używane modele

- Andersen & Gill (1982) („niezależnych przyrostów”)
- Wei, Lin & Weissfeld (1989) (brzegowy)
- Prentice, Williams & Petersen (1981) (warunkowy)

## ◆ Wszystkie są „brzegowe” w tym sensie, że pomijają zależności czasów przy estymacji współczynników

# Andersen & Gill (1982)

- ◆ Funkcja hazardu (intensywności zdarzeń) dla „czasu od rozpoczęcia obserwacji”:

$$\lambda_{ij}(t) = Y_i(t) \lambda_0(t) e^{\mathbf{Z}'(t) \cdot \boldsymbol{\beta}}$$

$Y_i(t)$  pozostaje równe 1 nawet po zdarzeniu

- W „zwykłym” modelu Coxa, po zdarzeniu  $Y_i(t)=0$
- ◆ Model dla niezależnych czasów wewnątrz jednostki
  - „niezależne przyrosty” liczby zdarzeń → proces Poissonowski
- ◆ Możliwe również sformułowanie dla „odstępów” (“gap times”) między zdarzeniami:  $(0, t_1]$ ,  $(0, t_2 - t_1]$ , ...
  - „odstępy” z procesu odnowienia

# Wei, Lin & Weissfeld (1989)

- ◆ Funkcja hazardu (intensywności zdarzeń) dla „czasu od rozpoczęcia obserwacji”:

$$\lambda_{ij}(t) = Y_{ij}(t) \lambda_{0j}(t) \exp\{\mathbf{Z}_i'(t) \cdot \boldsymbol{\beta}_j\}$$

- ◆  $Y_{ij}(t) = 0$  gdy ma miejsce  $j$ -te zdarzenie (lub cenzurowanie)
- ◆ Dane traktowane jako „wielowymiarowe”

# Prentice, Williams & Peterson (1981)

- ◆ Funkcja hazardu (intensywności zdarzeń) dla „czasu od rozpoczęcia obserwacji”:

$$\lambda_{ij}(t) = Y_{ij}(t) \lambda_{0j}(t) \exp\{\mathbf{Z}_i'(t) \cdot \boldsymbol{\beta}\}$$

$Y_{ij}(t) = 0$  przed  $(j-1)$ -ym zdarzeniem i po  $j$ -ym zdarzeniu

- Narażenie na  $j$ -te zdarzenie dopiero po  $(j-1)$ -ym

- ◆ Możliwe również sformułowanie dla „odstępów” między zdarzeniami
- ◆ Dane traktowane jako czasy kolejnych zdarzeń

# Postać danych: czas od rozpoczęcia obserwacji

- ♦ Jednostka z trzema zdarzeniami dla  $t=10, 30$  i  $42$  (oraz dodatkowym czasem obserwacji do  $t=81$ )
  - „intervals” (AG, PWP) or „total time”

Model	Interval	Stratum	Event
AG	(0,10]	1	1
	(10,30]	1	1
	(30,42]	1	1
	(42,81]	1	0
WLW	(0,10]	1	1
	(0,30]	2	1
	(0,42]	3	1
	(0,81]	4	0
PWP	(0,10]	1	1
	(10,30]	2	1
	(30,42]	3	1
	(42,81]	4	0

# Postać danych: odstęp między zdarzeniami

- ◆ Jednostka z trzema zdarzeniami dla  $t=10, 30$  i  $42$  (i dodatkowym czasem obserwacji do  $t=81$ )
  - “gap times”

Model	Interval	Stratum	Event
AG	(0,10]	1	1
	(0,20]	1	1
	(0,12]	1	1
	(0,39]	1	0
PWP	(0,10]	1	1
	(0,20]	2	1
	(0,12]	3	1
	(0,39]	4	0



# Zbiory ryzyka: czas od rozpoczęcia obserwacji

- ◆ Jednostka z drugim zdarzeniem w chwili 32
- ◆ Jak są zdefiniowane zbiory ryzyka?
  - AG: wszyscy pozostający pod obserwacją dla  $t = 32$
  - WLW: wszyscy pozostający pod obserwacją dla  $t = 32$  i jeszcze bez drugiego zdarzenia
  - PWP: wszyscy pozostający pod obserwacją dla  $t = 32$  i jeszcze bez drugiego zdarzenia, ale już po pierwszym zdarzeniu

# Dane wielowymiarowe

- ◆ Najczęściej używany model: Wei, Lin & Weissfeld (1989)
  - brzegowy
- ◆ Generalna idea:
  - wybieramy postać modelu (warstwy itd.)
  - dopasowujemy model Coxa pomijając zależności czasów
  - obliczamy „poprawione” wariancje oszacowanych współczynników
    - przy użyciu skalowanych reszt „score”

# „Nawracające” zdarzenia

- ◆ AG & PWP traktują dane jako czasy kolejnych zdarzeń
- ◆ WLW może nie spełniać założenia PH (nawet jeśli prawdziwe czasy spełniają to założenie)

# Różne strategie analizy

data input	stratified by event number	
	no	yes
"intervals"	AG	PWP
"total time"	WLW (infrequent)	WLW

# Własności modeli: symulacje (1)

- ◆ Symulacje dla danych z rozkładu wykładniczego

$$\lambda = \lambda_0 \exp\{X_1 \cdot \beta_1 + X_2 \cdot \beta_2\}$$

$X_1$  – binarna (leczenie),  $X_2 \sim U(-2, 2)$

$$\beta_1 = -1, \quad \beta_2 = 1$$

- ◆ Niezależne kolejne zdarzenia; cenzurowanie po roku (średnia liczba zdarzeń 1.3); 2000 obserwacji

	Number of Events							
	0	1	2	3	4	5	6	7
Control	367	312	174	88	39	13	5	2
Treatment	680	250	50	14	6	0	0	0

# Własności modeli: symulacje (2)

- ◆ Pacjent nr 10: leczony,  $X_2=0.2$ , zdarzenia dla  $t=100, 200$ , cenzurowanie dla  $t=365$
- ◆ Układ danych dla modelu Andersena-Gilla/warunkowego :

Id	Start	Stop	Status	Enum	$x_1$	$x_2$
10	0	100	1	1	1	.2
10	100	200	1	2	1	.2
10	200	365	0	3	1	.2

- ◆ Układ danych dla modelu brzegowego:

Id	Time	Status	Enum	$x_1$	$x_2$
10	100	1	1	1	.2
10	200	1	2	1	.2
10	365	0	3	1	.2
10	365	0	4	1	.2
10	365	0	5	1	.2
10	365	0	6	1	.2
10	365	0	7	1	.2

# Własności modeli: symulacje (3)

	Fits with the Covariate		Fits without the Covariate
	$\beta_1$	$\beta_2$	$\beta_1$
Andersen–Gill			
Coefficient	−0.93	1.05	−0.92
Variance	.066, .066	.056, .056	.066, .084
WLW			
Coefficient	−1.60	1.82	−1.23
Variance	.069, .117	.063, .113	.066, .113
Conditional			
Coefficient	−0.91	1.03	−0.67
Variance	.073, .069	.065, .064	.070, .068

- ◆ AG odpowiada modelowi symulacyjnemu
- ◆ Dla modelu z  $X_1$  i  $X_2$ 
  - Oszacowania  $\beta_1$  i  $\beta_2$  bliskie prawdziwym wartościom
  - Wariancja „modelowa”=„odporna”
- ◆ Dla modelu z  $X_1$ 
  - Oszacowanie  $\beta_1$  bliskie prawdziwej wartości
  - Wariancja „modelowa” < „odporna” (pominięcie  $X_2$  indukuje zależność)

# Własności modeli: symulacje (4)

	Fits with the Covariate		Fits without the Covariate
	$\beta_1$	$\beta_2$	$\beta_1$
Andersen–Gill			
Coefficient	−0.93	1.05	−0.92
Variance	.066, .066	.056, .056	.066, .084
WLW			
Coefficient	−1.60	1.82	−1.23
Variance	.069, .117	.063, .113	.066, .113
Conditional			
Coefficient	−0.91	1.03	−0.67
Variance	.073, .069	.065, .064	.070, .068

- ◆ Dla warunkowego (PWP) modelu z  $X_1$  i  $X_2$ 
  - Oszacowania  $\beta_1$  i  $\beta_2$  bliskie prawdziwym wartościom
  - Wariancja „modelowa”  $\approx$  „odporna”
- ◆ Dla modelu z  $X_1$ 
  - Oszacowanie  $\beta_1$  obciążone
  - Wariancja „modelowa”  $\approx$  „odporna”



# Własności modeli: symulacje (5)

	Fits with the Covariate		Fits without the Covariate
	$\beta_1$	$\beta_2$	$\beta_1$
Andersen–Gill			
Coefficient	−0.93	1.05	−0.92
Variance	.066, .066	.056, .056	.066, .084
WLW			
Coefficient	−1.60	1.82	−1.23
Variance	.069, .117	.063, .113	.066, .113
Conditional			
Coefficient	−0.91	1.03	−0.67
Variance	.073, .069	.065, .064	.070, .068

- ◆ Dla brzegowego (WLW) modelu z  $X_1$  i  $X_2$ 
  - Oszacowania  $\beta_1$  i  $\beta_2$  są obciążone
  - Wariancja „modelowa” << „odporna”
- ◆ Dla modelu z  $X_1$ 
  - Oszacowanie  $\beta_1$  obciążone (mniej niż dla modelu z  $X_2$ !)
  - Wariancja „modelowa” << „odporna”

# Własności modeli: symulacje (6)

- ◆ Obciążenie  $\beta_1$  w modelu bez  $X_2$  wynika z niezrównoważenia rozkładu  $X_2$  dla  $X_1$ :
  - Warstwa 1: średnia  $X_2 = 0$  dla leczonych i kontrolnych
  - Warstwa 2: średnia 0.8 dla leczonych i 0.6 kontrolnych
    - leczenie jest skuteczne ( $\beta_1 = -1$ ), więc aby wystąpiło zdarzenie, chorzy leczeni muszą mieć średnio wyższe ryzyko zdarzenia
  - Warstwa 4: średnia większa o 40% dla leczonych

# Własności modeli: symulacje (7)

- ◆ Oszacowania  $\beta_1$  dla warstw (zdarzeń) w modelu bez  $X_2$  :

	trt1	trt2	trt3	trt4	trt5	trt6	trt7
Marginal	-0.99	-1.7	-2.1	-2.3	-4.2	-4.2	-4.2
Conditional	-0.99	-0.9	-0.6	-0.3	$-\infty$		

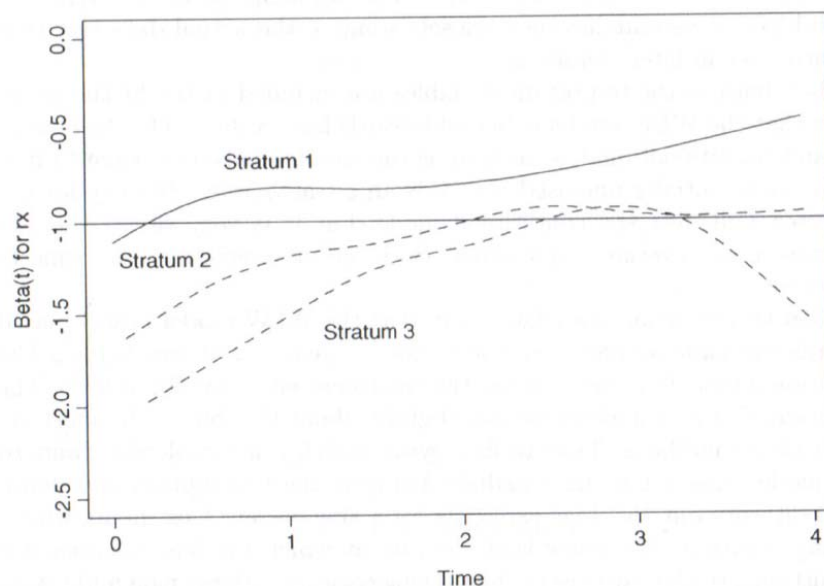
- ◆ Dla modelu warunkowego, oszacowania  $\beta_1$  maleją
  - warstwa 5: 0/6 zdarzeń dla chorych leczonych, 13/39 dla kontrolnych
  - warstwa 6 i 7: tylko kontrolni

# Własności modeli: symulacje (8)

- ◆ Oszacowania  $\beta_1$  dla warstw (zdarzeń) w modelu bez  $X_2$  :

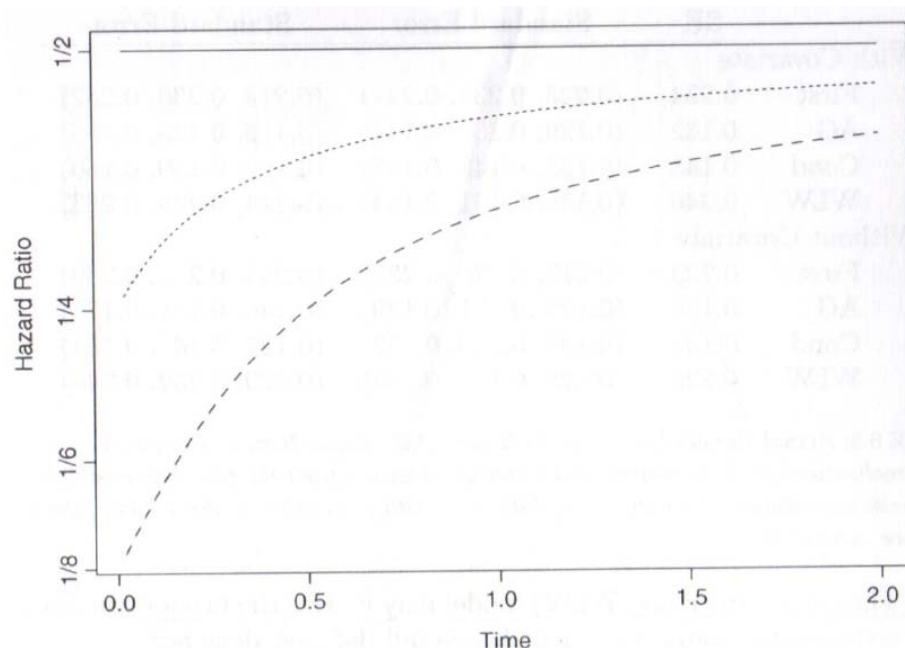
	trt1	trt2	trt3	trt4	trt5	trt6	trt7
Marginal	-0.99	-1.7	-2.1	-2.3	-4.2	-4.2	-4.2
Conditional	-0.99	-0.9	-0.6	-0.3	$-\infty$		

- ◆ Dla modelu brzegowego, oszacowania  $\beta_1$  rosną
- ◆ Problem z założeniem PH dla warstw  $> 1$



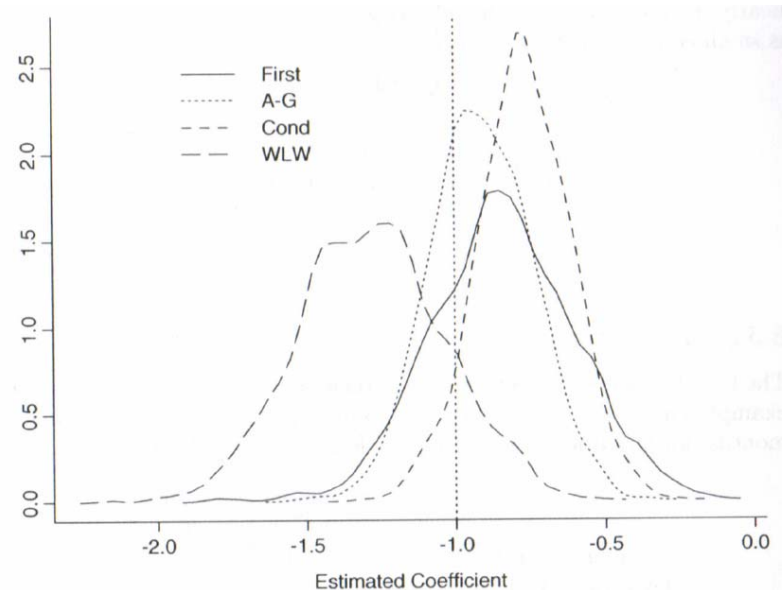
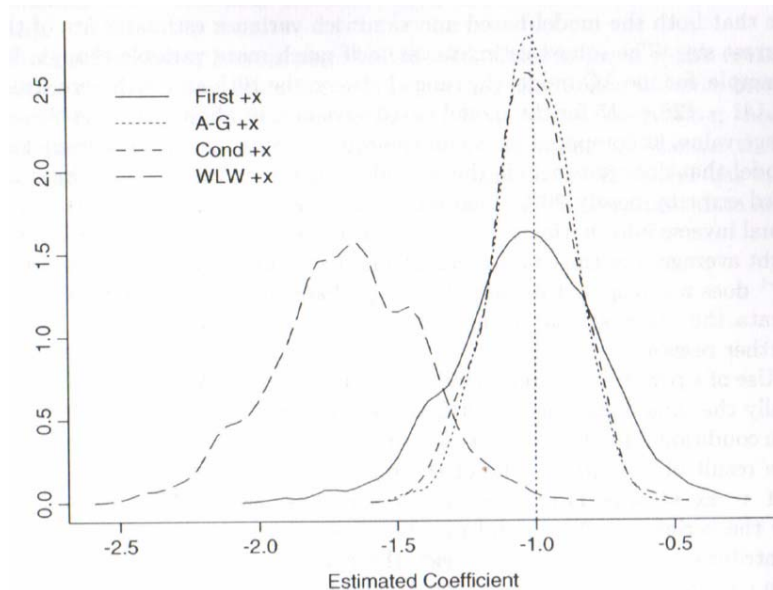
# Założenie PH dla modelu brzegowego

- ◆ Przyjmijmy model wykładniczy, tylko  $X_1$ , brak cenzurowania,  $HR = \lambda_1 / \lambda_0$
- ◆ Czas do  $k$ -tego zdarzenia ma rozkład gamma
  - $HR = (\lambda_1 / \lambda_0)^k$  dla  $t=0$  i dąży do  $\lambda_1 / \lambda_0$  z czasem
- ◆ Przykład dla  $\lambda_1 = 1$ ,  $\lambda_0 = 2$  :



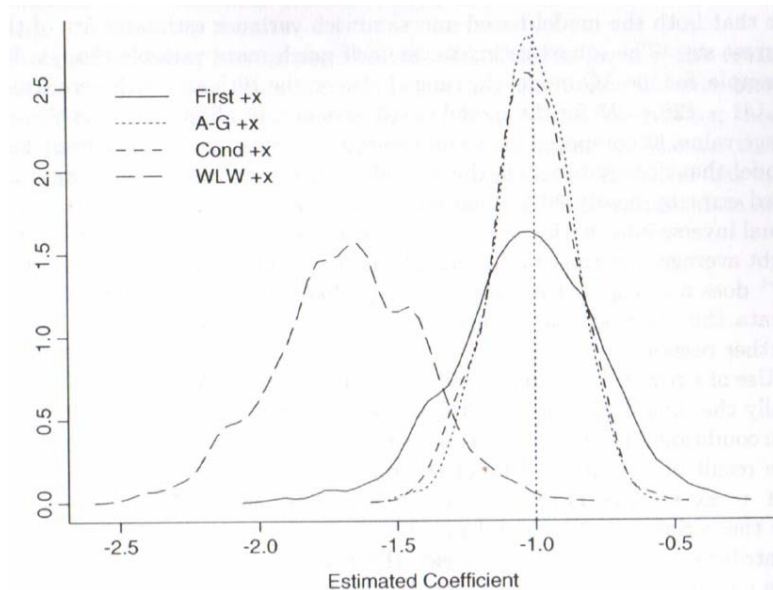
# Własności modeli: symulacje (9)

- ◆ 1000 symulacji dla 100 obserwacji, z  $X_2 \sim U(-1,1)$



# Własności modeli: symulacje (10)

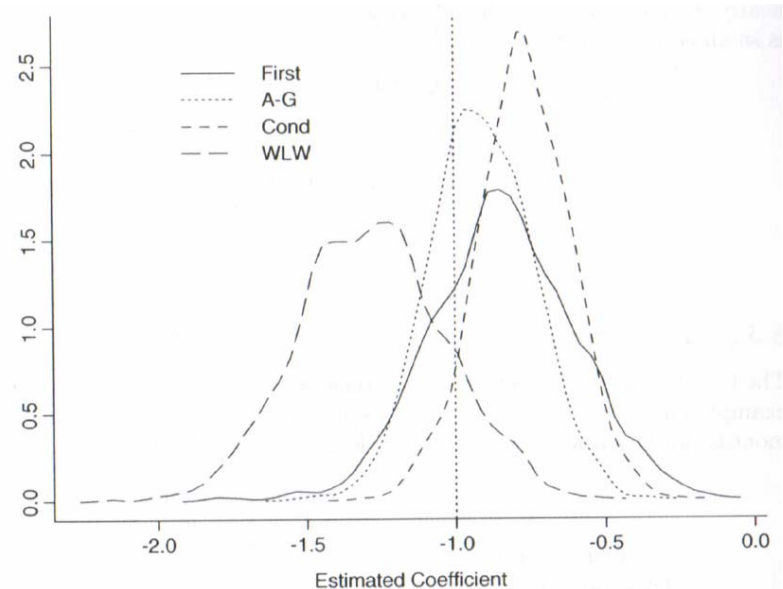
- ◆ 1000 symulacji dla 100 obserwacji, z  $X_2 \sim U(-1,1)$



- ◆ Dla modelu z  $X_1$  i  $X_2$ 
  - modele AG, warunkowy i „czasu do pierwszego zdarzenia” („zwykły” model PH) dają nieobciążone oszacowanie  $\beta_1$ 
    - AG i warunkowy bardziej efektywne niż „zwykły” model PH
  - oszacowanie dla modelu brzegowego jest obciążone

# Własności modeli: symulacje (11)

- ◆ 1000 symulacji dla 100 obserwacji, z  $X_2 \sim U(-1,1)$



- ◆ Dla modelu z  $X_1$ 
  - model AG daje najmniej (9%) obciążone oszacowanie  $\beta_1$



# Własności modeli: symulacje (12)

- ◆ 1000 symulacji dla 100 obserwacji, z  $X_2 \sim U(-1,1)$
- ◆ Oszacowania wariancji (1-szy, 5-ty, i 9-ty decyl):

	Actual SE	Model-Based Standard Error	Sandwich Standard Error
With Covariate			
First	0.234	(0.223, 0.234, 0.249)	(0.214, 0.230, 0.252)
AG	0.132	(0.126, 0.132, 0.141)	(0.116, 0.129, 0.146)
Cond	0.143	(0.135, 0.145, 0.157)	(0.122, 0.139, 0.160)
WLW	0.246	(0.132, 0.141, 0.153)	(0.214, 0.235, 0.263)
Without Covariate			
First	0.223	(0.217, 0.225, 0.235)	(0.214, 0.225, 0.239)
AG	0.159	(0.124, 0.131, 0.139)	(0.146, 0.161, 0.177)
Cond	0.143	(0.130, 0.139, 0.152)	(0.127, 0.142, 0.164)
WLW	0.229	(0.125, 0.133, 0.143)	(0.220, 0.232, 0.246)

- ◆ Model z  $X_1$  i  $X_2$ : oszacowania bliskie rzeczywistej wartości
  - dla AG, „odporne” ma większą zmienność niż „modelowe”
- ◆ Model z  $X_1$ : „odporne” bliskie rzeczywistej wartości
  - dla AG, „modelowe” obciążone

# Dane wielowymiarowe: marskość wątroby (1)

- ◆ Pierwotna marskość żółciowa wątroby: przewlekłe schorzenie o etiologii immunologicznej. Może prowadzić do niewydolności wątroby.
- ◆ Próba kliniczna (1988-1992)
  - 180 chorych, kwasem ursodeoksycholowy (UDCA) lub kontrola
  - kryteria oceny skuteczności leczenia:

	UDCA	Placebo
Death	6	10
Transplant	6	6
Drug toxicity	0	0
Voluntary withdrawal	11	18
Histologic progression	8	12
Development of varices	8	17
Development of ascites	1	5
Development of encephalopathy	3	1
Doubling of bilirubin	2	15
Worsening of symptoms	7	9

# Dane wielowymiarowe: marskość wątroby (2)

- ◆ Wykluczonych 10 chorych z niepełną obserwacją
- ◆ Analiza z wyłączeniem „voluntary withdrawal”
  - analiza czasu do pierwszego zdarzenia
  - analiza wszystkich zdarzeń (pomijamy kolejność i typ)
- ◆ Przynajmniej jedno zdarzenie: placebo - 45, UDCA - 27
- ◆ Wszystkich zdarzeń: placebo – 75, UDCA – 41

# Dane wielowymiarowe: marskość wątroby

## (3)

### ◆ Analiza czasu do pierwszego zdarzenia

	id	rx	entry	hi_stage	bili	risk	nevent	futime	status
1	1	1	21APR88	1	1.0	5.1	0	1896	0
2	2	0	27APR88	1	1.7	4.2	1	1456	1
3	3	1	25APR88	0	0.5	3.4	0	1892	0
4	4	1	27APR88	1	1.4	5.0	0	343	0
5	5	0	12MAY88	1	1.1	4.3	0	1875	0
6	6	1	18MAY88	1	1.4	5.9	2	768	1

Call:

```
coxph(formula=Surv(futime, status)~rx+log(bili)+hi_stage + cluster(id), data=udcal)
```

n= 170, number of events= 72

	coef	exp(coef)	se(coef)	robust se	z	Pr(> z )	
rx	-1.02728	0.35798	0.25520	0.26044	-3.944	8.00e-05	***
log(bili)	0.62144	1.86161	0.15124	0.15782	3.938	8.23e-05	***
hi_stage	-0.07357	0.92907	0.29099	0.30965	-0.238	0.812	

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(...)

Likelihood ratio test= 30.24 on 3 df, p=1.23e-06  
Wald test = 28.82 on 3 df, p=2.443e-06  
Score (logrank) test = 30.29 on 3 df, p=1.2e-06, Robust = 23.03 p=3.984e-05

# Dane wielowymiarowe: marskość wątroby (4)

- ♦ Analiza wszystkich zdarzeń (WLW,  $\beta_j \equiv \beta$ )

	id	rx	hi_stage	bili	risk	etype	futime	status	
1	1	1		1	1.0	5.1	1	1896	0
2	1	1		1	1.0	5.1	2	1896	0
3	1	1		1	1.0	5.1	3	1896	0
4	1	1		1	1.0	5.1	4	1896	0
5	1	1		1	1.0	5.1	5	1896	0
6	1	1		1	1.0	5.1	6	1896	0
7	1	1		1	1.0	5.1	7	1896	0
8	1	1		1	1.0	5.1	8	1896	0
9	2	0		1	1.7	4.2	1	1890	0

Call:

```
coxph(formula = Surv(futime, status) ~ rx + log(bili) + hi_stage +  
  cluster(id) + strata(etype), data = udca2)
```

```
n= 1360, number of events= 116
```

	coef	exp(coef)	se(coef)	robust se	z	Pr(> z )	
rx	-0.97386	0.37762	0.20306	0.27317	-3.565	0.000364	***
log(bili)	0.65999	1.93478	0.11975	0.17341	3.806	0.000141	***
hi_stage	0.03105	1.03153	0.24115	0.32556	0.095	0.924024	

(...)

Likelihood ratio test= 51.8 on 3 df, p=3.296e-11

Wald test = 29.69 on 3 df, p=1.604e-06

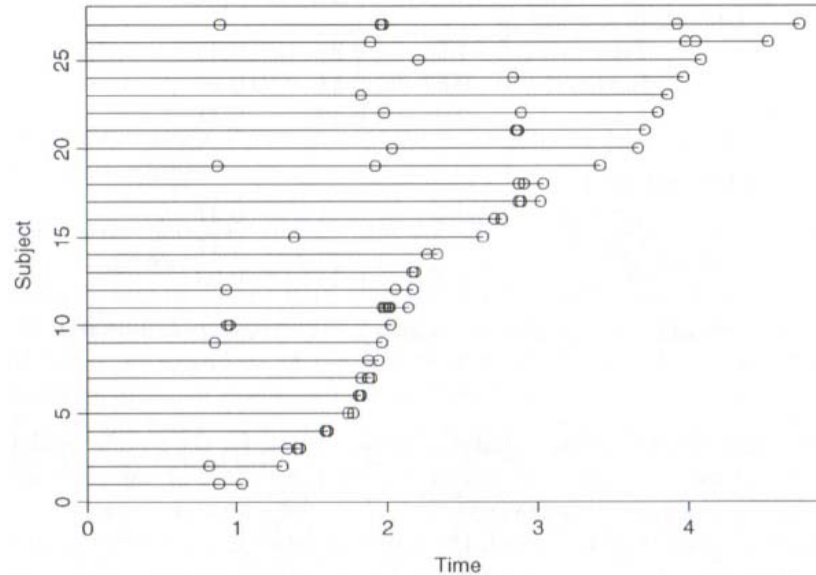
Score (logrank) test = 51.47 on 3 df, p=3.89e-11, Robust = 18.92 p=0.0002844

# Dane wielowymiarowe: marskość wątroby (5)

	$\beta$	SE( $\beta$ )	Robust SE
Time to first event			
Treatment	-1.03	0.26	0.26
Log(bili)	0.62	0.15	0.16
Stage	-0.07	0.29	0.31
Marginal model			
Treatment	-0.97	0.21	0.27
Bilirubin	0.66	0.12	0.17
Stage	0.03	0.24	0.33

- ◆ W analizie czasu do pierwszego zdarzenia oszacowania wariancji są podobne
- ◆ Dla wszystkich zdarzeń „odporne” jest większe
- ◆ „Odporne” dla wszystkich zdarzeń > pierwszego zdarzenia!
  - nie zyskał na efektywności !

# Dane wielowymiarowe: marskość wątroby (6)



- ◆ Wizyty kliniczne co roku
  - większość zdarzeń obserwowana dla tej samej wizyty
- ◆ Użycie różnych zdarzeń zwiększyło szanse wykrycia niewydolności wątroby
  - liczba zdarzeń nie miała znaczenia
  - analiza wszystkich zdarzeń nie zwiększyła mocy statystycznej

# „Nawracające” zdarzenia: rak pęcherza (1)

- ◆ Czasy wznów dla 86 chorych na raka pęcherza

	Number of Recurrences				
	0	1	2	3	4
Number of Subjects	39	18	7	8	14
Followup After Last Event	38	17	5	6	12

- ◆ Dane:

	id	rx	futime	number	size	recurrences
id	subject id, 1 to 85;					
futime	followup or recurrence time;					
status	1 = recurrence, 0 = censoring;					
number	initial number;					
size	initial size;					
rx	treatment code, 1 = placebo, 2 = thiotepa;					
enum	event number.					
	0	1	0	1	1	
	1	1	1	1	3	
	2	1	4	2	1	
	3	1	7	1	1	
	4	1	10	5	1	
	5	1	10	4	1	6
	6	1	14	1	1	
	7	1	18	1	1	
	8	1	18	1	3	5
	9	1	18	1	1	12, 16



# „Nawracające” zdarzenia: rak pęcherza (2)

- ◆ Układ danych dla modelu AG/warunkowego:

id	rx	futime	number	size	recurrences
0	1	0	1	1	
1	1	1	1	3	
2	1	4	2	1	
3	1	7	1	1	
4	1	10	5	1	
5	1	10	4	1	6
6	1	14	1	1	
7	1	18	1	1	
8	1	18	1	3	5
9	1	18	1	1	12, 16



id	time1	time2	status	rx	number	size	enum
1	0	1	0	1	1	3	1
2	0	4	0	1	2	1	1
3	0	7	0	1	1	1	1
4	0	10	0	1	5	1	1
5	0	6	1	1	4	1	1
5	6	10	0	1	4	1	2
6	0	14	0	1	1	1	1
7	0	18	0	1	1	1	1
8	0	5	1	1	1	3	1
8	5	18	0	1	1	3	2
9	0	12	1	1	1	1	1
9	12	16	1	1	1	1	2
9	16	18	0	1	1	1	3

# „Nawracające” zdarzenia: rak pęcherza (3)

- ◆ Układ danych dla modelu brzegowego:

id	rx	futime	number	size	recurrences
<del>0</del>	<del>1</del>	<del>0</del>	<del>1</del>	<del>1</del>	
1	1	1	1	3	
2	1	4	2	1	
3	1	7	1	1	
4	1	10	5	1	
5	1	10	4	1	6
6	1	14	1	1	
7	1	18	1	1	
8	1	18	1	3	5
9	1	18	1	1	12, 16



id	time	status	rx	number	size	enum
1	1	0	1	1	3	1
1	1	0	1	1	3	2
1	1	0	1	1	3	3
1	1	0	1	1	3	4
5	6	1	1	4	1	1
5	10	0	1	4	1	2
5	10	0	1	4	1	3
5	10	0	1	4	1	4
9	12	1	1	1	1	1
9	16	1	1	1	1	2
9	18	0	1	1	1	3
9	18	0	1	1	1	4

# „Nawracające” zdarzenia: rak pęcherza (4)

## ◆ Uwaga na dane!

Call:

```
coxph(formula=Surv(time1,time2,status) ~ rx+size+number+factor(enum),data=bladder1)
```

	coef	exp(coef)	se(coef)	z	Pr(> z )	
rx	-2.799e-01	7.558e-01	2.058e-01	-1.361	0.173671	
size	-3.751e-03	9.963e-01	7.032e-02	-0.053	0.957464	
number	1.403e-01	1.151e+00	5.142e-02	2.729	0.006347	**
factor(enum)2	5.893e-01	1.803e+00	2.568e-01	2.295	0.021745	*
factor(enum)3	1.680e+00	5.368e+00	3.024e-01	5.558	2.73e-08	***
factor(enum)4	1.338e+00	3.810e+00	3.510e-01	3.811	0.000139	***
factor(enum)5	-1.727e+01	3.173e-08	2.915e+03	-0.006	0.995275	

- ◆ Maksimum 4 zdarzenia w danych (w rzeczywistości, więcej)
- ◆ 12 chorych z obserwacjami poza 4-te zdarzenie
  - nieśmiertelni !
  - należy ich usunąć z danych

# „Nawracające” zdarzenia: rak pęcherza (5)

## ◆ Modele AG, brzegowy i warunkowy

```
> summary(AG.fit)
Call:coxph(formula = Surv(time1, time2, status) ~ rx + size +
number + cluster(id), data = bladder3)

      coef exp(coef) se(coef) robust se      z Pr(>|z|)
rx      -0.46469   0.62833  0.19973   0.26556 -1.750  0.08015 .
size     -0.04366   0.95728  0.06905   0.07762 -0.563  0.57376
number    0.17496   1.19120  0.04707   0.06304  2.775  0.00551 **
```

```
> summary(WLW.fit)
Call: coxph(formula = Surv(futime, status) ~ rx + size + number
+ cluster(id) + strata(enum), data = bladder2)

      coef exp(coef) se(coef) robust se      z Pr(>|z|)
rx      -0.58479   0.55722  0.20105   0.30795 -1.899  0.0576 .
size     -0.05162   0.94969  0.06973   0.09459 -0.546  0.5853
number    0.21029   1.23404  0.04675   0.06664  3.156  0.0016 **
```

```
> summary(PWP.fit)
Call: coxph(formula = Surv(time1, time2, status) ~ rx + size +
number + cluster(id) + strata(enum), data = bladder3)

      coef exp(coef) se(coef) robust se      z Pr(>|z|)
rx      -0.333489   0.716420  0.216168  0.204787 -1.628  0.1034
size     -0.008495   0.991541  0.072762  0.061635 -0.138  0.8904
number    0.119617   1.127065  0.053338  0.051387  2.328  0.0199 *
```

# „Nawracające” zdarzenia: rak pęcherza (6)

## ◆ Model dla czasu do pierwszego zdarzenia:

```
Call: coxph(formula = Surv(time2, status) ~ rx + number + cluster(id), data = bladder3, subset = (enum == 1))
```

```
n= 85, number of events= 47
```

	coef	exp(coef)	se(coef)	robust se	z	Pr(> z )
rx	-0.51218	0.59919	0.31299	0.31164	-1.644	0.10027
number	0.23079	1.25960	0.07542	0.07081	3.259	0.00112 **

## ◆ Model AG dla wszystkich zdarzeń:

```
Call:
```

```
coxph(formula = Surv(time1, time2, status) ~ rx + number + cluster(id), data = bladder3, init = TTFE.fit$coef, iter = 0)
```

```
n= 178, number of events= 112
```

	coef	exp(coef)	se(coef)	robust se	z	Pr(> z )
rx	-0.51218	0.59919	0.20096	0.27869	-1.838	0.066088 .
number	0.23079	1.25960	0.04314	0.06539	3.530	0.000416 ***

## ◆ Gdyby zdarzenia były niezależne, $(47/112)^{1/2}(0.313)=0.203$

## ◆ „Odporne” oszacowanie SE: $0.279 \approx (47/59.3)^{1/2}(0.313)$

- $112-47=65$ ,  $59.3-47=12.3$ ; nowe zdarzenie  $\approx 1/5$  nowego chorego

# Model PH z efektami losowymi (frailty) (1)

- ◆ Model „wspólnej podatności” („shared frailty”)

$$\lambda_{ij}(t) = \lambda_0(t) \cdot u_i \cdot e^{\mathbf{Z}' \cdot \boldsymbol{\beta}}$$

gdzie  $u_i$  jest zm. losową ze średnią 1 i wariancją  $\theta$

- z rozkładu gamma, log-normalnego, ...

- ◆  $u_i$  jest nazywana „podatnością” („frailty”)

- $u_i > 1 \rightarrow$  zwiększona funkcja hazardu (jednostka *podatna* na zdarzenia)
- $u_i < 1 \rightarrow$  zmniejszona funkcja hazardu

- ◆ Alternatywne sformułowanie (efekty losowe):

$$\lambda_{ij}(t) = \lambda_0(t) \cdot e^{w_i + \mathbf{Z}' \cdot \boldsymbol{\beta}}$$

gdzie  $w_i = \ln u_i$

# Model PH z efektami losowymi (frailty) (2)

- ◆ Implikuje zależność obserwacji
  - użyteczny dla „skupionych” czasów zdarzeń

- ◆ Może być rozszerzony do:

$$\lambda(t) = \lambda_0(t) \cdot e^{\mathbf{Z}' \cdot \boldsymbol{\beta} + \mathbf{X}' \cdot \mathbf{b}}$$

gdzie  $\mathbf{b}$  to wektor zmiennych losowych

- estymacja się komplikuje

# Estymacja modelu podatności

- ◆ Funkcja wiarygodności dla znanego  $\mathbf{w}=(w_1, \dots, w_N)^T$ :

$$L_{full}(\beta, \lambda_0; \theta) = f(\mathbf{w}; \theta) \prod_{i=1}^N \prod_{j=1}^{n_i} \left\{ \lambda_0(t_{ij}) e^{Z'_{ij} \beta + w_i} \right\}^{\delta_{ji}} \exp \left\{ - \Lambda_0(t_{ij}) e^{Z'_{ij} \beta + w_i} \right\}$$

- ◆ Dla obserwowanych danych, potrzebujemy

$$L_{obs}(\beta, \lambda_0; \theta) = \int L_{full}(\beta, \lambda_0; \theta) d\mathbf{w}$$

- ◆ Można pokazać, że

$$\ln L_{obs}(\beta, \lambda_0; \theta) = \sum_{i=1}^N \sum_{j=1}^{n_i} \left\{ \ln \Lambda_0(t_{ij}) + Z'_{ij} \beta \right\} + \sum_{i=1}^N \ln \left\{ (-1)^{d_i} \phi^{(d_i)}(A_i) \right\}$$

$\phi^{(k)}$  jest  $k$ -tą pochodną transformaty Laplace'a rozkładu  $u_i$ ,  $d_i$  jest sumą zdarzeń dla grupy  $i$ , a

$$A_i(\beta, \lambda_0) = \sum_{j=1}^{n_i} \left\{ \lambda_0(t_{ij}) e^{Z'_{ij} \beta} \right\}$$



# Estymacja przy użyciu algorytmu EM

- ♦ Wzór analityczny dla  $L_{obs}$  tylko w szczególnych przypadkach i przy parametrycznym  $\lambda_0(t)$ 
  - chcelibyśmy nie musieć zakładać formy  $\lambda_0(t)$

- ♦ Maksymalizacja  $\ln L_{obs}$  dla ustalonego  $\theta$  z pomocą EM:

- krok M(aximization): używając oszacowań  $\mathbf{w}$  ( $= \log \mathbf{u}$ ) jako *offsetu*, szacujemy  $\boldsymbol{\beta}$  i  $\lambda_0(t)$  jak w modelu PH
- krok E(xpectation): używając oszacowań  $\boldsymbol{\beta}$  i  $\lambda_0(t)$ , wyznaczamy oczekiwane wartości  $\mathbf{u}$  z

$$e^{w_i} = - \frac{\phi^{(d_i+1)} \{A_i(\hat{\boldsymbol{\beta}}, \hat{\lambda}_0)\}}{\phi^{(d_i)} \{A_i(\hat{\boldsymbol{\beta}}, \hat{\lambda}_0)\}}$$

- ♦ Szacujemy  $\theta$  z *profilowanej* funkcji wiarygodności

$$L_{obs}(\theta) = L_{obs} \{ \hat{\boldsymbol{\beta}}(\theta), \hat{\lambda}_0(\theta), \theta \}$$

# Podwójna pętla iteracyjna

- ◆ Użycie  $L_{obs}$  i algorytmu EM prowadzi do podwójnej pętli iteracyjnej:
  - Dla ustalonego  $\theta$ , szacujemy  $\beta$ ,  $\lambda_o(t)$ 
    - Uwaga: tutaj użycie EM wymaga wartości oczekiwanych  $w$
  - Dla ustalonych oszacowań  $\beta$  i  $w$ , szacujemy  $\theta$
- ◆ Podobny schemat pojawia się przy użyciu innych metod

# Estymacja przy użyciu penalizowanej częściowej funkcji wiarygodności (1)

- ◆ Gdybyśmy znali  $\theta$ , moglibyśmy rozważyć logarytm *penalizowanej częściowej funkcji wiarygodności*:

$$PPL = l(\boldsymbol{\beta}, \mathbf{w}) - g(\mathbf{w}, \theta)$$

gdzie

$$l(\boldsymbol{\beta}, \mathbf{w}) = \ln \prod_{i=1}^N \prod_{j=1}^{n_i} \frac{\left\{ e^{Z'_{ij}\boldsymbol{\beta} + w_i} \right\}^{\delta_{ji}}}{\sum_k \sum_l Y_{kl}(t_{ij}) e^{Z'_{kl}\boldsymbol{\beta} + w_k}}$$

a  $g(\mathbf{w}, \theta)$  jest logarytmem funkcji gęstości dla  $\mathbf{w}$ .

# Estymacja przy użyciu penalizowanej częściowej funkcji wiarygodności (2)

- ♦ Na podstawie  $PPL$  możemy szacować  $\mathbf{w}$ :

$$\frac{\partial PPL}{\partial w_i} = \frac{\partial l}{\partial w_i} - \frac{\partial g}{\partial w_i} = \left\{ d_i - e^{w_i} \sum_{j=1}^{n_i} \hat{\Lambda}_0(t_{ij}) e^{\mathbf{Z}_{ij}'\boldsymbol{\beta}} \right\} - \frac{\partial g}{\partial w_i} = 0$$

- ♦ Oszacowania dla  $\boldsymbol{\beta}$  otrzymujemy z  $\frac{\partial PPL}{\partial \boldsymbol{\beta}} = \frac{\partial l}{\partial \boldsymbol{\beta}} - \frac{\partial g}{\partial \boldsymbol{\beta}} = \frac{\partial l}{\partial \boldsymbol{\beta}} = 0$

- Czyli jak dla następującego modelu PH:

$$\lambda_{ij}(t) = \lambda_o(t) \exp\{\mathbf{Z}_{ij}'(t)\boldsymbol{\beta} + w_i\}$$

- $w_i$  traktowany jako *offset*

# Estymacja przy użyciu penalizowanej częściowej funkcji wiarygodności (3)

- ◆ Mając dane oszacowania  $\beta$  i  $w$ , możemy szacować  $\theta$
- ◆ W tym celu używamy *profilowanej PPL*:

$$PPL(\theta) = PPL\{\hat{\beta}(\theta), \hat{w}(\theta), \theta\}$$

- ◆ Czyli znowu mamy podwójną pętlę iteracyjną

# Profilowana $PPL$ dla modelu z podatnościami z rozkładu gamma

- ◆ Załóżmy, że  $u_i$  ma rozkład gamma z wariancją  $\theta=1/v$ .  
Wówczas

$$PPL(\beta, w, \theta) = l(\beta, w) - \sum_{i=1}^N (w_i - e^{w_i}) / \theta$$

- ◆ Można pokazać, że

$$\ln L_{obs}(\theta) = PPL\{\theta\} + \sum_{i=1}^N \left\{ \nu - (\nu + d_i) \ln(\nu + d_i) + \nu \ln \nu + \ln \frac{\Gamma(\nu + d_i)}{\Gamma(\nu)} \right\}$$

# Profilowana $PPL$ dla modelu ze podatnościami z rozkładu log-normalnego

- ♦ Załóżmy, że  $w_i$  ma rozkład normalny z wariancją  $\theta$ . Wówczas

$$PPL(\beta, w, \theta) = l(\beta, w) - \sum_{i=1}^N w_i^2 / 2\theta$$

- ♦ Można pokazać, że użycie  $PPL(\theta)$  prowadzi do  $\hat{\theta} = \frac{r + \sum_{i=1}^N \hat{w}_i^2}{N}$  gdzie

$r=0$  dla BLUP,  $\text{trace}\{(\mathbf{H}_{22})^{-1}\}$  dla ML, i  $\text{trace}\{(\mathbf{H}^{-1})_{22}\}$  dla REML  
oraz  $\mathbf{H} = -\partial^2 PPL / \partial(\beta, w)^T \partial(\beta, w)$

- ♦ W tym przypadku zewnętrzna pętla dla  $\theta$  nie wymaga iteracji

# Model podatności: nowotworzenie u szczurów (1)

- ◆ 50 miotów z 3 szczurami każdy
- ◆ Jeden ze szczurów poddany działaniu substancji rakotwórczej
  - zaobserwowano 40 przypadków raka
- ◆ Dane:

	<code>litter</code>	<code>treat</code>	<code>futime</code>	<code>status</code>
1	1	1	101	0
2	1	0	49	1
3	1	0	104	0
4	2	1	104	0
5	2	0	102	0
6	2	0	104	0
...				



# Model podatności: nowotworzenie u szczurów (2)

## ◆ Model brzegowy i podatności (log-gamma):

```
Call: coxph(formula = Surv(futime, status) ~ treat + cluster(litter), data = rats)
n= 150, number of events= 40
      coef exp(coef) se(coef) robust se      z Pr(>|z|)
treat  0.9047    2.4713   0.3175   0.3025 2.991  0.00278 **

---
Call: coxph(formula = Surv(futime, status) ~ treat + frailty(litter), data = rats)
n= 150
      coef se(coef) se2    Chisq DF    p
treat    0.914  0.323   0.319   8.01  1.0 0.0046
frailty(litter)      17.69 14.4 0.2400
```

## ◆ Podobne oszacowania współczynnika dla „treat”- przypadek !

- Model brzegowy szacuje *średni efekt w populacji*: stosunek ryzyka dla losowych próbek szczurów poddanych i nie poddanych działaniu substancji rakotwórczej
- Model podatności szacuje *efekt w danym miocie*

# Model podatności: nowotworzenie u szczurów (3)

```
Call: coxph(formula = Surv(futime, status) ~ treat + cluster(litter), data = rats)
n= 150, number of events= 40
      coef exp(coef) se(coef) robust se      z Pr(>|z|)
treat  0.9047    2.4713   0.3175    0.3025  2.991  0.00278 **
Likelihood ratio test= 7.98  on 1 df,   p=0.004741
```

```
> print(ratsm.fit$loglik)
[1] -185.6556 -181.6677
      β=0      β≠0
```

---

```
Call: coxph(formula = Surv(futime, status) ~ treat + frailty(litter), data = rats)
n= 150
      coef se(coef) se2   Chisq DF    p
treat    0.914 0.323   0.319   8.01  1.0 0.0046
frailty(litter)              17.69 14.4 0.2400
```

Iterations: 6 outer, 24 Newton-Raphson

Variance of random effect= 0.499 I-likelihood = -180.8

◆  $\tau = \theta/(\theta+2) = .2 \rightarrow$  korelacja wewnątrz-grupowa

# Model podatności: nowotworzenie u szczurów (4)

```
> print(ratsm.fit$loglik)
[1] -185.6556 -181.6677

Call: coxph(formula = Surv(futime, status) ~ treat + frailty(litter), data = rats)
n= 150

            coef se(coef) se2    Chisq DF    p
treat          0.914 0.323    0.319   8.01  1.0 0.0046
frailty(litter)          17.69 14.4 0.2400

Iterations: 6 outer, 24 Newton-Raphson
Variance of random effect= 0.499 I-likelihood = -180.8
Degrees of freedom for terms= 1.0 14.4
Likelihood ratio test= 37.6 on 15.4 df, p=0.00124

> print(ratsf.fit$loglik)
[1] -185.6556 -166.8325
```

- ◆ Test Walda dla podatności oparty na  $\{(\mathbf{H})^{-1}\}_{22}$  dla  $PPL$
- ◆ Test oparty na ilorazie wiarog.:  $2(181.7-180.8)=1.8$ ,  $p=0.18$
- ◆ Asymptotyczny rozkład testu dla  $(\boldsymbol{\beta}, \theta)$  to ważona suma  $\chi^2$

# Model podatności: nowotworzenie u szczurów (5)

## ◆ Modele z podatnościami log-gamma i normalnymi:

```
Call: coxph(formula = Surv(futime, status) ~ treat + frailty(litter), data = rats)

      coef se(coef) se2    Chisq DF    p
treat      0.914 0.323    0.319   8.01  1.0 0.0046
frailty(litter)              17.69 14.4 0.2400
```

Iterations: 6 outer, 24 Newton-Raphson

Variance of random effect= 0.499 I-likelihood = -180.8

```
Call: coxph(formula=Surv(futime, status) ~ treat + frailty(litter,dist="gaussian"),
data = rats)

      coef se(coef) se2    Chisq DF    p
treat      0.913 0.323    0.319   8.01  1.0 0.0046
frailty(litter, dist = "g")              15.57 11.9 0.2100
```

Iterations: 6 outer, 21 Newton-Raphson

Variance of random effect= 0.412

# Model podatności: retinopatia cukrzycowa

```
Call: coxph(formula = Surv(time, status) ~ adult + treat + cluster(id), data = diab)
      coef exp(coef) se(coef) robust se      z Pr(>|z|)
adult  0.05388   1.05536  0.16211   0.17864  0.302   0.763
treat -0.77893   0.45890  0.16893   0.14851 -5.245 1.56e-07 ***
Likelihood ratio test= 22.48 on 2 df,    p=1.312e-05
```

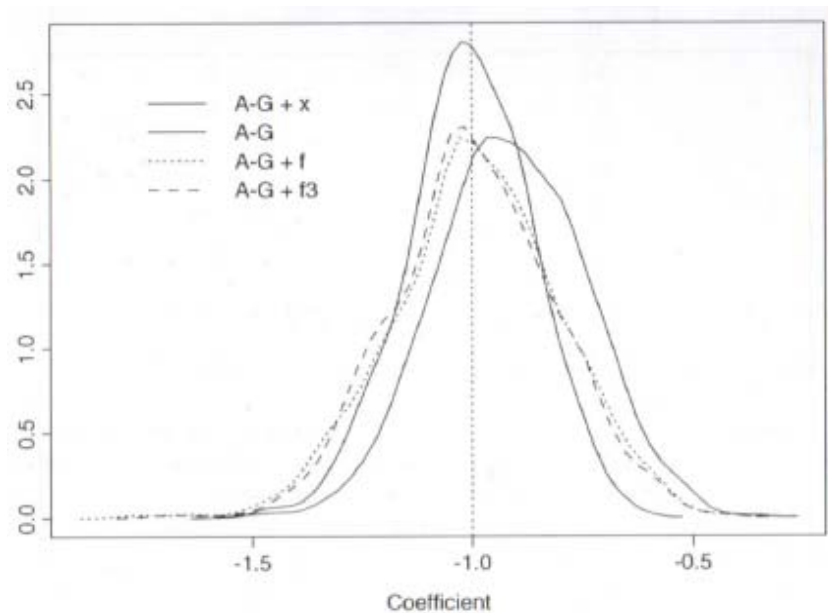
```
Call: coxph(formula = Surv(time, status) ~ adult + treat + strata(id), data = diab)
      coef exp(coef) se(coef)      z Pr(>|z|)
adult      NA      NA  0.0000      NA      NA
treat -0.9623   0.3820  0.2016 -4.773 1.82e-06 ***
Likelihood ratio test= 25.49 on 1 df,    p=4.455e-07
```

```
Call: coxph(formula = Surv(time, status) ~ adult + treat + frailty(id), data = diab)
      coef se(coef) se2 Chisq DF p
adult    0.041 0.221   0.166  0.03  1 8.5e-01
treat   -0.911 0.174   0.171 27.31  1 1.7e-07
frailty(id)                113.80 84 1.7e-02
Iterations: 6 outer, 30 Newton-Raphson
Variance of random effect= 0.851 I-likelihood = -850.8
Likelihood ratio test= 201 on 85.6 df,    p=2.77e-11
```

- ◆ Różnica oszacowań dla „treat” dla brzegowego i podatności
- ◆ Model warstwowy uwzględnia korelację, ale jest nieefektywny
  - „adult” to zmienna stała dla pacjenta, więc jej efekt jest nie-estymowalny

# Własności modeli: symulacje (13)

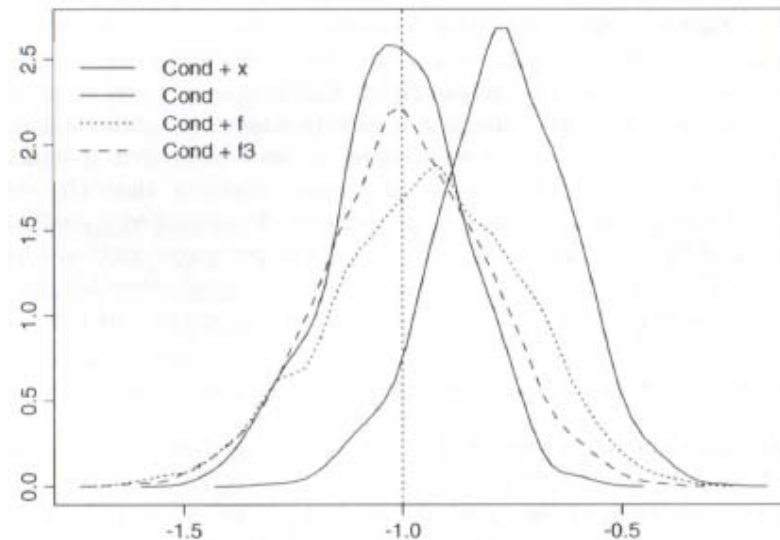
- ◆ 1000 symulacji dla 100 obserwacji, z  $X_2 \sim U(-1,1)$ 
  - Czyli model z efektem losowym o wariancji 1/3



- ◆ Model AG z  $X_2$  jest poprawny i nieobciążony
- ◆ Oszacowanie  $\beta_1$  dla AG bez  $X_2$  jest obciążone
- ◆ Włączenie efektu losowego usuwa obciążenie

# Własności modeli: symulacje (14)

- ◆ 1000 symulacji dla 100 obserwacji, z  $X_2 \sim U(-1,1)$ 
  - Czyli model z efektem losowym o wariancji 1/3



- ◆ Model warunkowy z  $X_2$  jest poprawny i nieobciążony
- ◆ Oszacowanie  $\beta_1$  dla modelu bez  $X_2$  jest obciążone
- ◆ Włączenie efektu losowego usuwa obciążenie jedynie jeśli użyjemy prawdziwej wartości  $\theta$

# Własności modeli: symulacje (15)

- ◆ 1000 symulacji dla 100 obserwacji, z  $X_2 \sim U(-1,1)$ 
  - Czyli model z efektem losowym o wariancji 1/3

	Andersen-Gill	Conditional
0-.1	9	921
.1-.2	106	14
.2-.25	125	5
.25-.3	190	7
.3-.35	197	10
.35-.4	152	5
.4-.5	160	8
.5-.6	42	11
.6+	9	18

- ◆ Model warunkowy z efektem losowym szacuje  $\theta=0$
- ◆ Dla AG z efektem losowym oszacowania skoncentrowane wokół  $\theta=1/3$



# Kopuły

- ◆ Dwuwymiarowa kopuła  $C(u,v)$  to funkcja spełniająca następujące warunki:
  - $C: [0,1] \times [0,1] \rightarrow [0,1]$
  - $C(u,0) = 0 = C(0,v)$  oraz  $C(u,1) = u$  i  $C(1,v) = v$  dla  $u,v \in [0,1]$
  - $C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \geq 0$  dla  $u_2 \geq u_1, v_2 \geq v_1$  z  $[0,1]$
- ◆ Przykłady specjalnych kopuł:
  - $M(u,v) = \min(u,v)$
  - $W(u,v) = \max(u+v-1, 0)$
  - $\Pi(u,v) = uv$
- ◆ Zachodzi:  $W(u,v) \leq C(u,v) \leq M(u,v)$

# Twierdzenie Sklara

- ◆ Niech  $H$  będzie dwuwymiarową dystrybuantą z brzegami  $F$  i  $G$ . Wówczas istnieje kopuła  $C$  taka, że dla  $x, y \in [-\infty, +\infty]$  mamy

$$H(x, y) = C\{F(x), G(y)\}.$$

I odwrotnie: jeśli  $C$  jest kopułą, a  $F$  i  $G$  dystrybuantami, to  $H$  zdefiniowane jak wyżej jest dwuwymiarową dystrybuantą z brzegami  $F$  i  $G$ .

- ◆ Dwie ciągłe zm. losowe są niezależne  $\Leftrightarrow C_{XY} = \Pi$

# Kopuły funkcji przeżycia

- ◆ Niech  $H^*$  będzie dwuwymiarową funkcją przeżycia z brzegowymi funkcjami przeżycia  $F^*=1-F$  i  $G^*=1-G$ . Mamy

$$\begin{aligned} H^*(x,y) &= 1-F(x)-G(y)+H(x,y) \\ &= F^*(x)+G^*(y)-1+C(F(x),G(y)) \\ &= F^*(x)+G^*(y)-1+C(1-F^*(x),1-G^*(y)) \end{aligned}$$

Czyli dla  $C^*(u,v)=u + v - 1 + C(1-u,1-v)$  dostajemy

$$H^*(x,y) = C^*(F^*(x),G^*(y))$$

- ◆  $C^*$  nazywane jest *kopułą funkcji przeżycia*

# Kopuły Archimedesowe

- ◆ Niech  $\varphi: [0,1] \rightarrow [0,\infty]$  będzie ciągłą, malejącą funkcją taką, że  $\varphi(1)=0$ . Wówczas
  - $C(u,v)=\varphi^{[-1]}(\varphi(u), \varphi(v))$  jest kopułą  $\Leftrightarrow \varphi$  jest wypukła.
  - $\varphi^{[-1]}(t)=\varphi^{-1}(t)$  dla  $t \in [0, \varphi(0)]$  i 0 w p.p.
- ◆  $\varphi$  jest nazywane *generatorem* kopuły
  - *Dokładnym* jeśli  $\varphi(0)=\infty$ . Wówczas  $\varphi^{[-1]}(t)=\varphi^{-1}(t)$  i generowana kopuła jest *dokładną* kopułą Archimedesową.
- ◆ Jeśli  $\varphi^{[-1]}$  jest transformatą Laplace'a dystrybuanty, to generowane przezeń kopuła funkcji przeżycia może być otrzymana z odpowiedniego modelu podatności

# Miary zależności: $\tau$ Kendalla

- ◆  $\tau_{X,Y} = P\{(X_1 - X_2)(Y_1 - Y_2) > 0\} - P\{(X_1 - X_2)(Y_1 - Y_2) < 0\}$
- ◆ Okazuje się, że  $\tau_{X,Y} = 4 \iint C(u,v) dC(u,v) - 1$
- ◆ Dla kopuł Archimedesowych,  $\tau_{X,Y} = 1 + 4 \int_0^1 \varphi(v)/\varphi'(t) dt$

# Miary zależności: $\rho$ Spearmana

- ◆  $\rho_{X,Y} = 3[P\{(X_1 - X_2)(Y_1 - Y_3) > 0\} - P\{(X_1 - X_2)(Y_1 - Y_3) < 0\}]$
- ◆ Okazuje się, że  $\rho_{X,Y} = 12 \iint C(u,v) du dv - 3$
- ◆ Alternatywnie,  $\rho_{X,Y} = 12 \iint \{C(u,v) - uv\} du dv$ 
  - Skalowana odległość rozkładu  $(X, Y)$  od niezależności
- ◆ Zachodzi:  $-1 \leq 3\tau_{X,Y} - 2\rho_{X,Y} \leq 1$

# Uzycie kopuł w analizie przeżycia

- ◆ Rozważmy przypadek dwóch czasów dla jednostki

- ◆ Łączna funkcja przeżycia w postaci

$$S(s,t) = C( S_S(s), S_T(t) )$$

- $S_S(s), S_T(t)$  to *brzegowe* funkcje przeżycia
- ◆  $S_S(s)$  i  $S_T(t)$  modelowane przy pomocy modelu PH lub parametrycznego
- ◆ Kopuła indukuje zależność czasów

# Przykład: kopuła Claytona

$$C^*(u,v) = \{ u^{1-\theta} + v^{1-\theta} \}^{1/(1-\theta)}, \text{ dla } \theta > 1$$

- ◆ Archimedesowa:  $\varphi(s) = (1+s)^{1/(1-\theta)}$ 
  - transformata Laplace'a rozkładu gamma

- ◆  $\tau_{X,Y} = (\theta-1)/(\theta+1)$

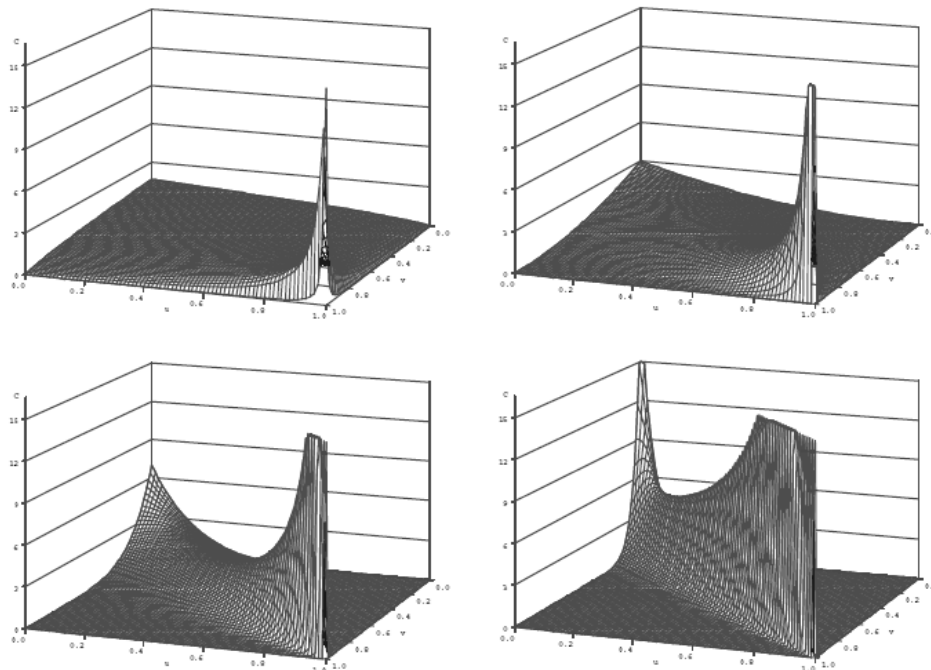


Figure 4.1: The density functions for the Clayton copulas with uniform margins for different values of Kendall's  $\tau$ .

NOTE: Clockwise from top-left:  $\tau = 0.25, 0.5, 0.75$  and  $0.9$ .



# Przykład: kopuła Gumbela-Hougaard

$$C^*(u,v) = \exp[- \{ (-\ln u)^{1/\theta} + (-\ln v)^{1/\theta} \}^\theta], \text{ dla } 0 < \theta < 1$$

- ◆ Archimedesowa:  $\varphi(s) = \exp(-s^\theta)$ 
  - transformata Laplace'a rozkładu dodatnio stabilnego

- ◆  $\tau_{X,Y} = 1 - \theta$

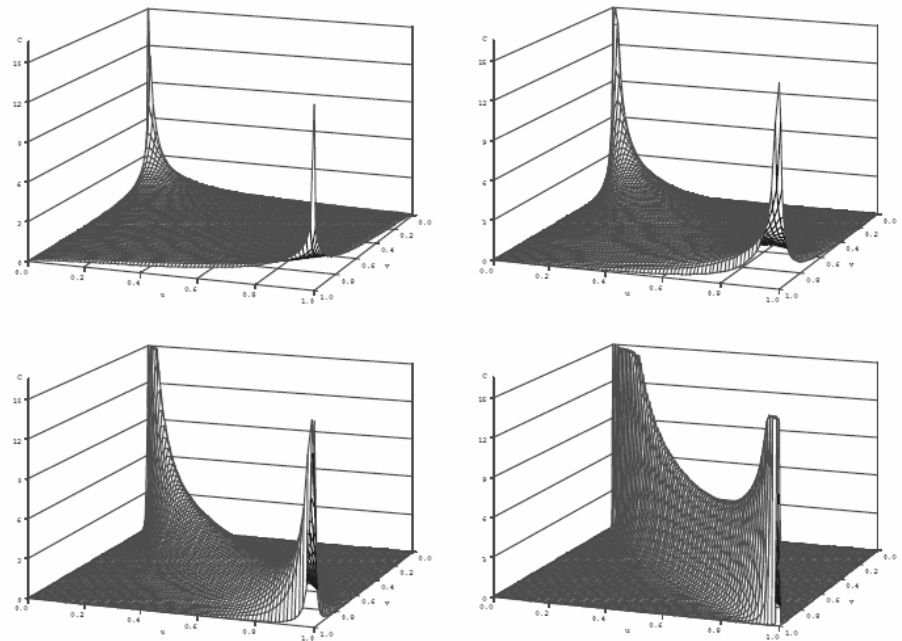


Figure 4.2: The density functions for the Hougaard copulas with uniform margins for different values of Kendall's  $\tau$ .

NOTE: Clockwise from top-left:  $\tau = 0.25, 0.5, 0.75$  and  $0.9$ .

# Przykład: kopuła Placketta

$$C^*(u,v) = \{1+(u+v)(\theta-1)-H_\theta(u,v)\}/2(\theta-1), \text{ dla } \theta \neq 1$$

$$H(u,v) = [ \{1+(u+v)(\theta-1)\}^2 + 4\theta(1-\theta)uv ]^{1/2}$$

$$C^*(u,v) = uv, \text{ dla } \theta=1$$

- ◆ Nie-Archimedesowa
- ◆  $C^*(u,v) = C(u,v)$
- ◆  $\rho_{X,Y} = (\theta^2 - 1 - 2\theta \ln \theta) / (\theta - 1)^2$

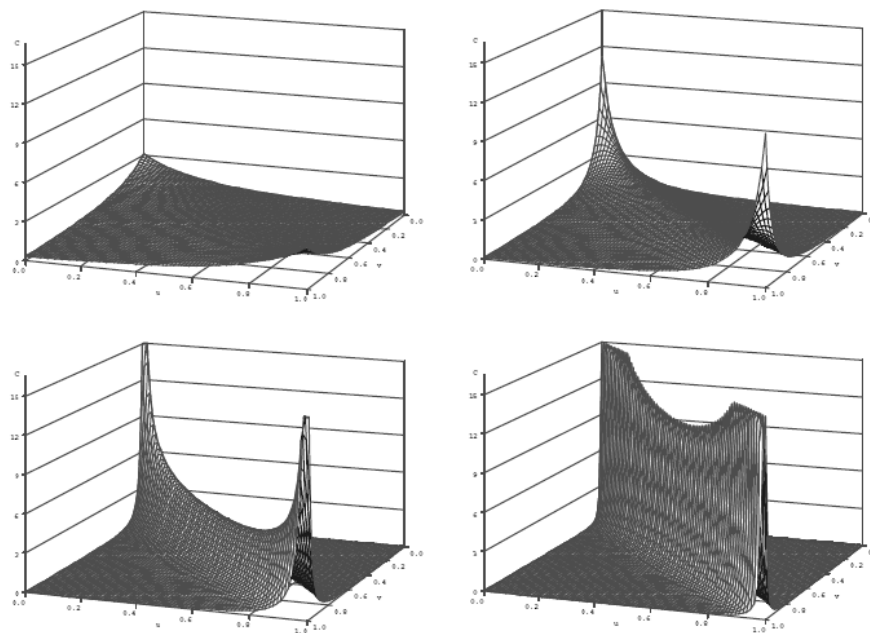


Figure 4.3: The density functions for the Plackett copulas with uniform margins for different values of Kendall's  $\tau$ .

NOTE: Clockwise from top-left:  $\tau = 0.25, 0.5, 0.75$  and  $0.9$ .

# Modele oparte na kopułach

- ◆ Możemy zapisać funkcję wiarygodności

$$L = \prod_{i=1}^N \{f(s_i, t_i)\}^{\delta_{iS}\delta_{iT}} \left\{ -\frac{\partial S(s_i, t_i)}{\partial s_i} \right\}^{\delta_{iS}(1-\delta_{iT})} \left\{ -\frac{\partial S(s_i, t_i)}{\partial t_i} \right\}^{(1-\delta_{iS})\delta_{iT}} \{S(s_i, t_i)\}^{(1-\delta_{iS})(1-\delta_{iT})}$$

- ◆ Dla parametrycznej formę rozkładów brzegowych, możemy maksymalizować  $L$  dla wszystkich parametrów
- ◆ Alternatywnie: oszacować parametry rozkładów brzegowych, użyć oszacowań w  $L$ , i maksymalizować tylko po  $\theta$ 
  - np. dla oszacowań Kaplana-Meiera brzegowych funkcji przeżycia albo modelu PH

# Modele oparte na kopułach: diagnostyka złamań u psów (1)

- ◆ Diagnostyka gojenia się złamań u psów
- ◆ Standard: radiografia (RX)
- ◆ Nowa technika: ultrasonografia (USG)
- ◆ 106 psów, ocena czasu do zagojenia się
- ◆ 7 psów cenzurowanych dla RX; 0 dla USG

Dogid	Time to diagnosis	Status	Method
1	63	1	RX
1	30	1	US
2	83	1	RX
2	83	1	US
...			
106	35	0	RX
106	35	1	US

# Modele oparte na kopułach: diagnostyka złamań u psów (2)

- ◆ Kopuła Claytona z  $\theta^* = \theta - 1$
  
- ◆ Brzegowe modele Weibulla:  $\lambda p t^{p-1}$  z  $(\lambda_R, p_R, \lambda_U, p_U)$ 
  - Oszacowania z jednoczesnej estymacji:  
 $\lambda_R=0.145, p_R=2.341, \lambda_U=0.233, p_U=2.212, \theta^*=1.066$
  - Oszacowania z dwustopniowej estymacji:  
 $\lambda_R=0.106, p_R=2.539, \lambda_U=0.219, p_U=2.323, \theta^*=0.89$
  
- ◆ Brzegowe modele Weibulla:  $\lambda p t^{p-1}$  i  $\lambda e^{\beta} p t^{p-1}$ 
  - Oszacowania z dwustopniowej estymacji:  
 $\lambda=0.119, \beta=0.522, p=2.42, \theta^*=0.87$  ( $\lambda_U=0.119e^{0.522} = 0.201$ )

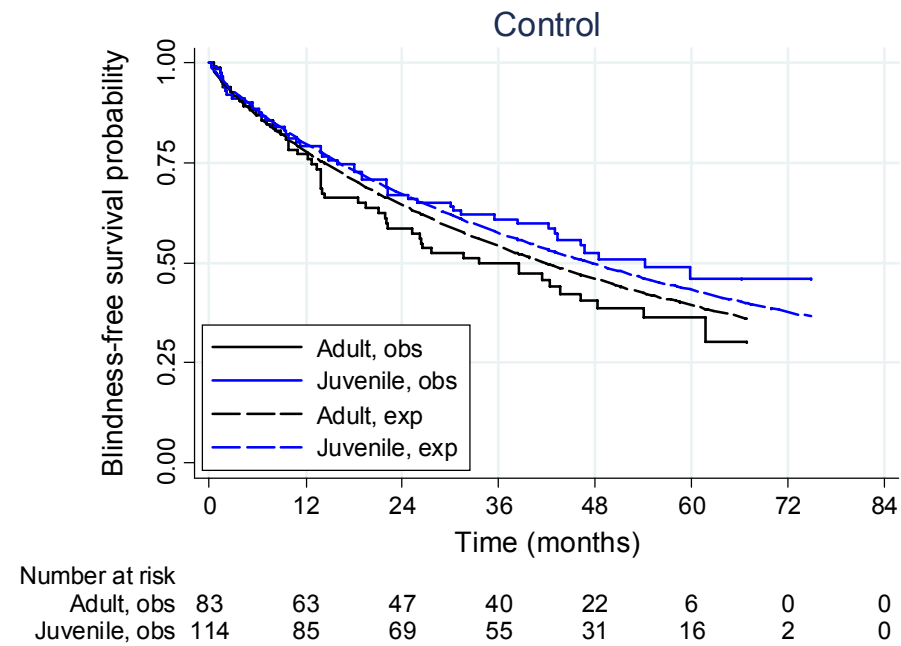
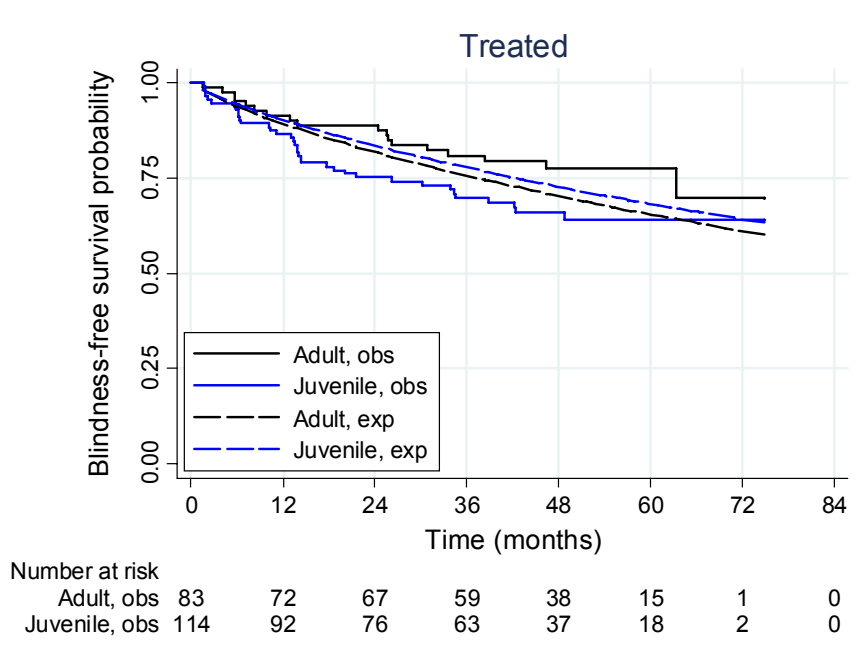
# Modele oparte na kopułach: retinopatia cukrzycowa (1)

- ◆ Brzegowe modele Weibulla:  $(\lambda e^\beta)^p p t^{p-1}$  ( $\lambda_C, \lambda_T, p, \beta$ )
  - Efekt typu cukrzycy (adult/juvenile):  $\beta$
- ◆ Oszacowania z jednoczesnej estymacji:

	Clayton	Hougaard	Plackett
$\lambda_C$	0.012 (0.083)	0.013 (0.079)	0.013 (0.083)
$\lambda_T$	0.004 (0.039)	0.005 (0.039)	0.005 (0.039)
$\beta$	0.125 (0.298)	0.117 (0.293)	0.127 (0.294)
$p$	0.812 (0.123)	0.793 (0.109)	0.816 (0.121)
$\ln(\lambda_T/\lambda_C)$	-0.973 (0.618)	-0.971 (0.619)	-0.966 (0.611)
$\tau$	0.32 (0.08)	0.21 (0.06)	0.25 (0.02)
$\ln L$	-829.4	-829.3	-828.6

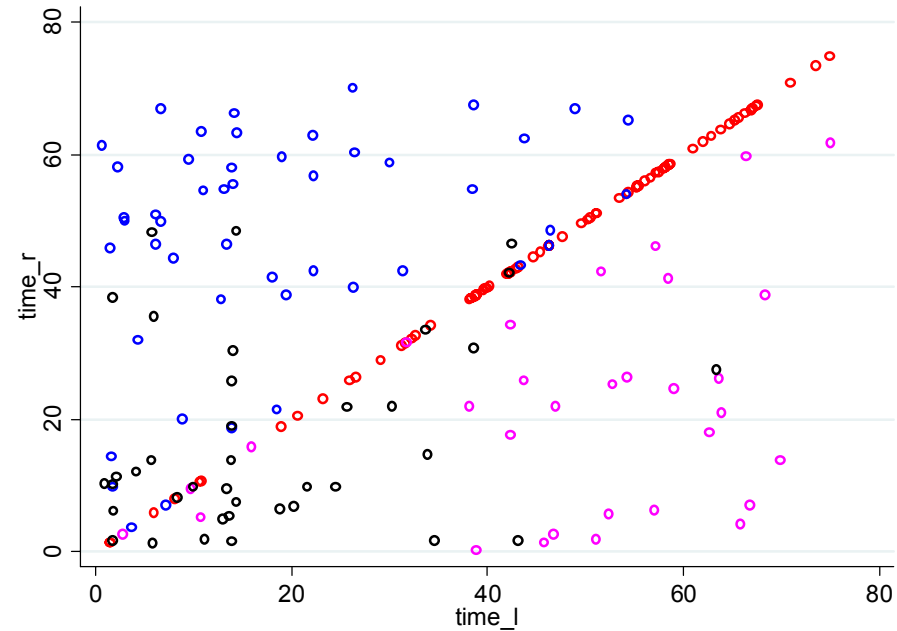
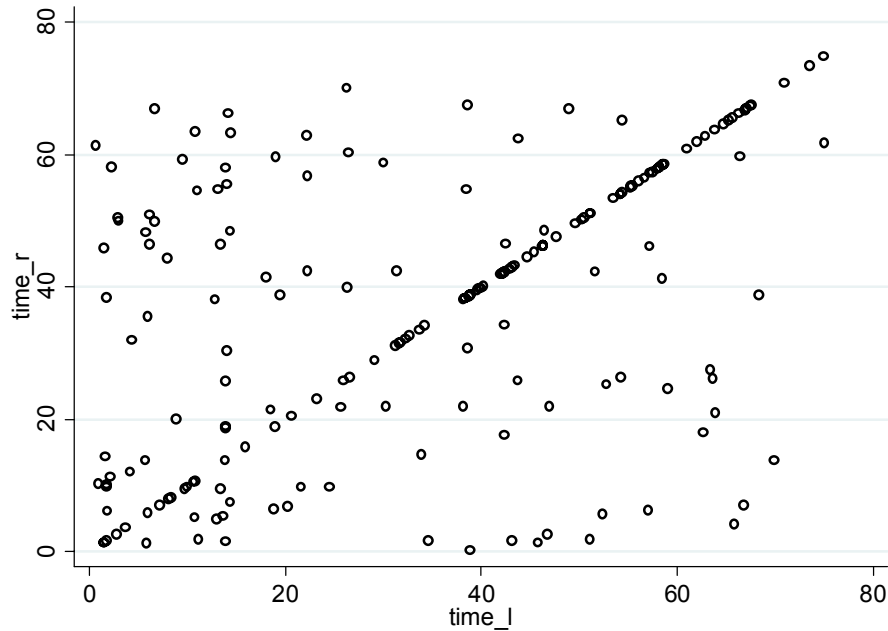
# Modele oparte na kopułach: retinopatia cukrzycowa (2)

- ◆ Brzegowe modele Weibulla:  $(\lambda e^\beta)^p p t^{p-1}$  ( $\lambda_C, \lambda_T, p, \beta$ )
  - Efekt typu cukrzycy (adult/juvenile):  $\beta$
- ◆ Obserwowane i oczekiwane (Plackett) funkcje przeżycia:



# Modele oparte na kopułach: retinopatia cukrzycowa (3)

- ◆ Czasy dla lewego i prawego oka:





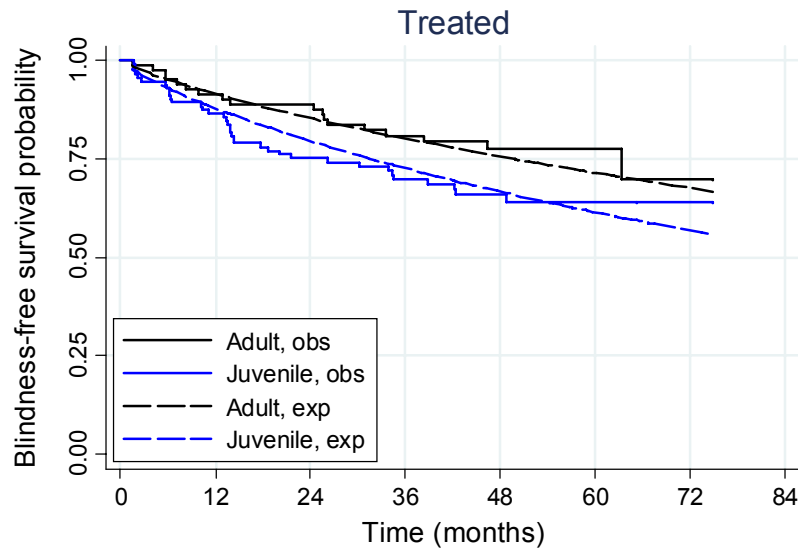
# Modele oparte na kopułach: retinopatia cukrzycowa (4)

- ◆ Brzegowe modele Weibulla:  $(\lambda e^\beta)^p p t^{p-1}$  ( $\lambda_C, \lambda_T, p, \beta_C, \beta_T$ )
- ◆ Oszacowania z jednoczesnej estymacji:

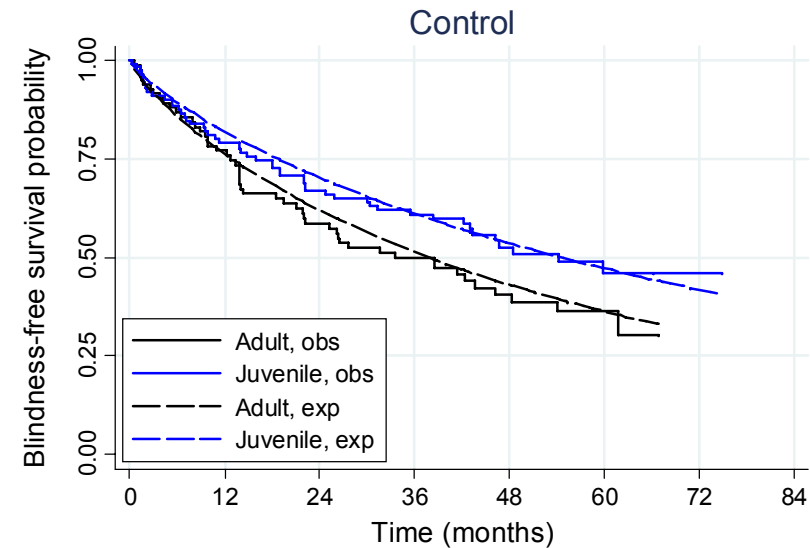
	Clayton	Hougaard	Plackett
$\lambda_C$	0.012	0.011	0.012
$\lambda_T$	0.007	0.007	0.007
$\beta_C$	0.370	0.364	0.367
$\beta_T$	-0.472	-0.429	-0.450
$p$	0.818	0.796	0.822
$\ln(\lambda_T/\lambda_C)$	-0.522	-0.539	-0.526
$\tau$	0.33	0.22	0.26
$\ln L$	-825.3	-825.5	-824.7

# Modele oparte na kopułach: retinopatia cukrzycowa (5)

- Brzegowe modele Weibulla:  $(\lambda e^\beta)^p p t^{p-1}$  ( $\lambda_C, \lambda_T, p, \beta_C, \beta_T$ )
- Obserwowane i oczekiwane (Plackett) funkcje przeżycia:



Number at risk	0	12	24	36	48	60	72	84
Adult, obs	83	72	67	59	38	15	1	0
Juvenile, obs	114	92	76	63	37	18	2	0



Number at risk	0	12	24	36	48	60	72	84
Adult, obs	83	63	47	40	22	6	0	0
Juvenile, obs	114	85	69	55	31	16	2	0

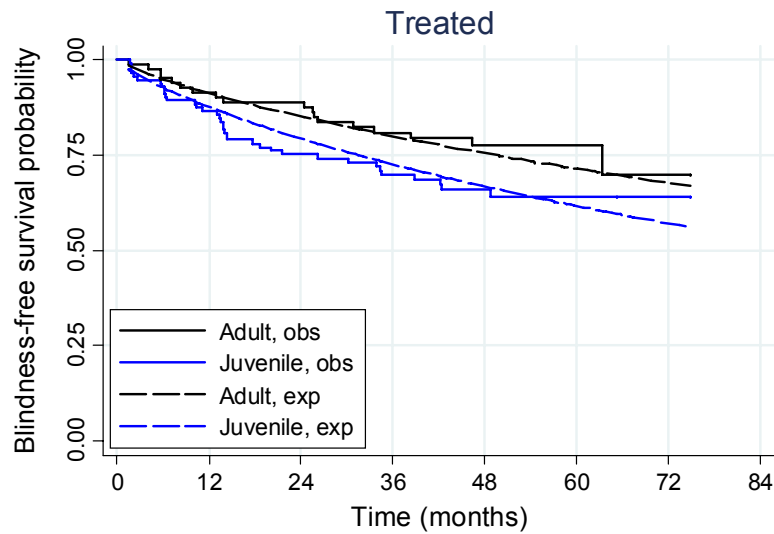
# Modele oparte na kopułach: retinopatia cukrzycowa (6)

- ◆ Brzegowe modele Weibulla:  $(\lambda e^\beta)^p p t^{p-1}$  ( $\lambda_C, \lambda_T, p_C, p_T, \beta_C, \beta_T$ )
- ◆ Oszacowania z jednoczesnej estymacji:

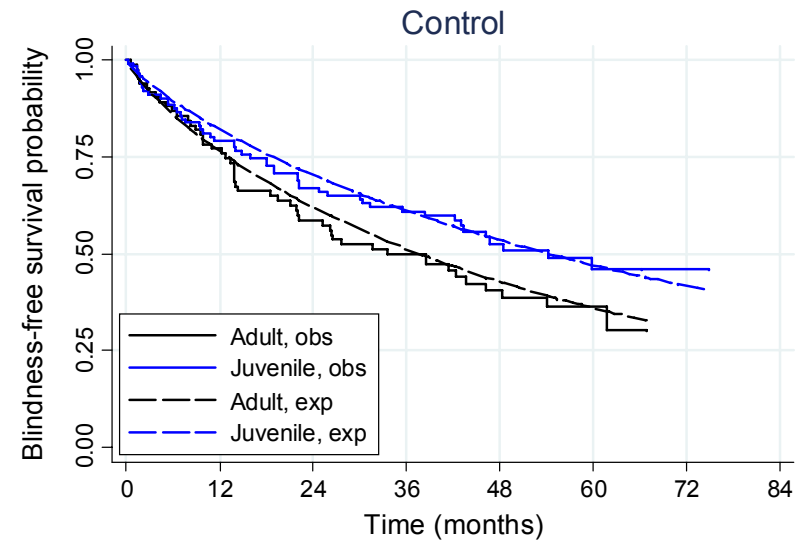
	Clayton	Hougaard	Plackett
$\lambda_C$	0.012	0.012	0.012
$\lambda_T$	0.007	0.006	0.007
$\beta_C$	0.371	0.363	0.368
$\beta_T$	-0.471	-0.428	-0.450
$p_C$	0.826	0.809	0.830
$p_T$	0.801	0.769	0.805
$\ln(\lambda_T/\lambda_C)$	*	*	*
$\tau$	0.33	0.22	0.26
$\ln L$	-825.2	-825.5	-824.7

# Modele oparte na kopułach: retinopatia cukrzycowa (7)

- Brzegowe modele Weibulla:  $(\lambda e^\beta)^p p t^{p-1}$  ( $\lambda_C, \lambda_T, p_C, p_T, \beta_C, \beta_T$ )
- Obserwowane i oczekiwane (Plackett) funkcje przeżycia:



Number at risk								
Adult, obs	83	72	67	59	38	15	1	0
Juvenile, obs	114	92	76	63	37	18	2	0



Number at risk								
Adult, obs	83	63	47	40	22	6	0	0
Juvenile, obs	114	85	69	55	31	16	2	0

# Kopuły wielowymiarowe

- ◆  $N$ -wymiarowa kopuła  $C$  to funkcja spełniająca następujące warunki:
  - $C: [0,1]^N \rightarrow [0,1]$
  - $C(\mathbf{u}) = 0$  jeśli co najmniej jedna ze współrzędnych  $\mathbf{u}$  jest równa 0
  - $C(\mathbf{u}) = u_k$  jeśli wszystkie ze współrzędnych  $\mathbf{u}$  (prócz  $k$ ) są równe 1
  - Dla  $\mathbf{v} \geq \mathbf{u}$ , objętość  $V_C([\mathbf{u}, \mathbf{v}]) \geq 0$
- ◆ Twierdzenie Sklara zachowuje się dla  $N$  wymiarów
- ◆  $X_1, \dots, X_N$  są niezależne  $\Leftrightarrow$  ich  $N$ -kopuła to  $\Pi^N(\mathbf{u}) = u_1 \cdots u_N$