

# 3. Model proporcjonalnych hazardów

Model Coxa

Model proporcjonalnych hazardów

Szacowanie współczynników

Wnioskowanie

Model warstwowy

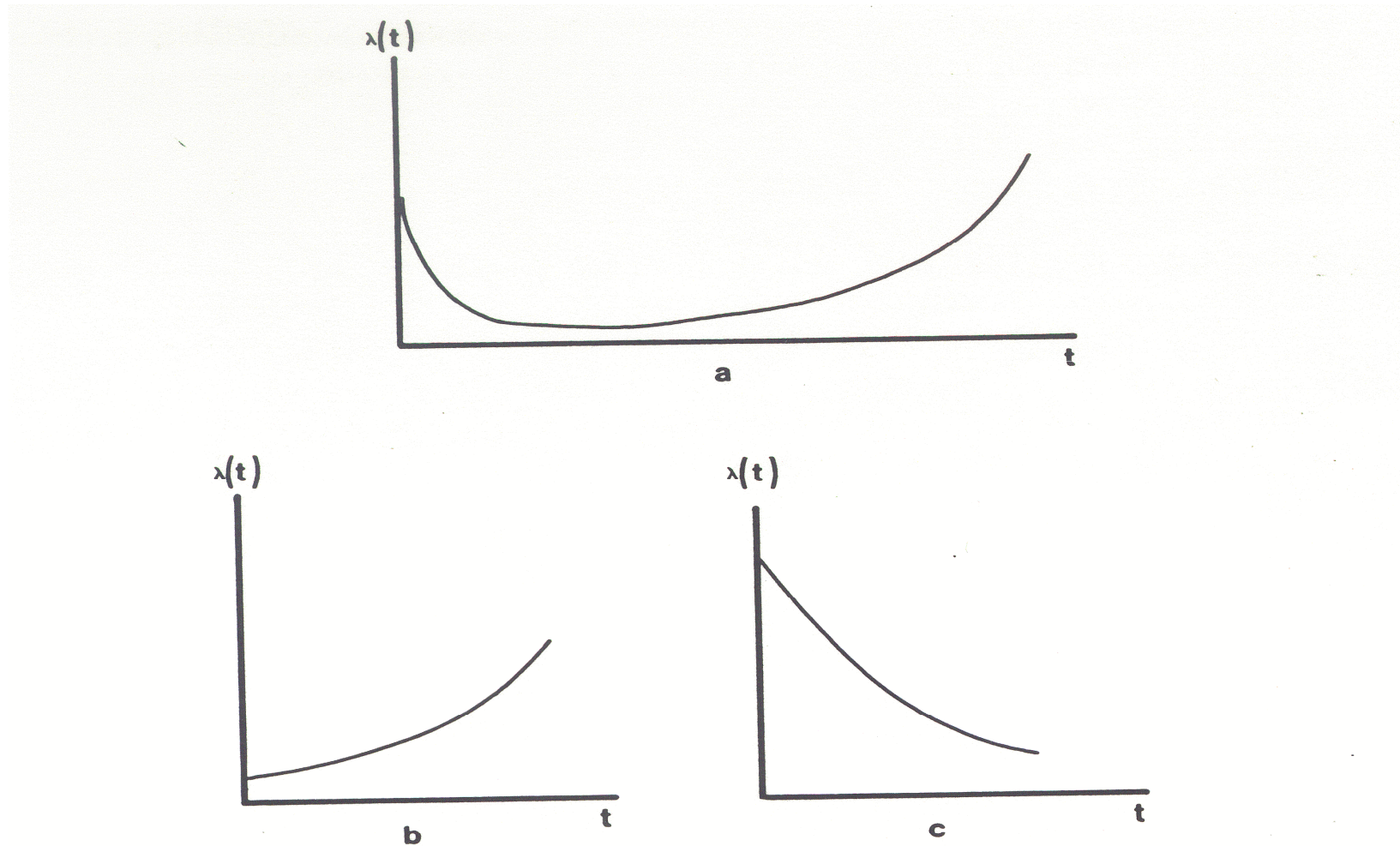
# Modele w analizie przeżycia

- ◆ Najczęściej używany *model proporcjonalnych hazardów*, często nazywany *modelem Coxa*.
- ◆ Alternatywna rodzina *modeli przyśpieszonego czasu do niepowodzenia* (*accelerated failure time model, AFT*).

# Hazard

- ◆ „Chwilowy współczynnik intensywności zdarzeń”, jako funkcja czasu
  - Oznaczany przez  $\lambda(t)$
- ◆ Jednostką są zdarzenia na jednostkę czasu
  - „Prędkość” występowania zdarzeń
  - $\lambda(t) \geq 0$
  - $\lambda(t)\Delta t \approx$  prawdopodobieństwo zdarzenia w  $(t, t + \Delta t]$ , pod warunkiem niewystąpienia zdarzenia przed  $t$

# Funkcje hazardu: przykłady



**Figure 1.1** Some types of hazard functions: (a) hazard for human mortality; (b) positive aging; (c) negative aging.

# Funkcja hazardu i przeżycia

- ◆ Funkcja przeżycia  $S(t)$  opisuje prawdopodobieństwo wystąpienia zdarzenia po chwili czasu  $t$
- ◆ Funkcja hazardu  $\lambda(t)$  opisuje możliwość wystąpienia zdarzenia tuż po chwili czasu  $t$  pod warunkiem, że nie wystąpiło do chwili  $t$
- ◆ Obie funkcje zawierają pełną informację o rozkładzie czasu do wystąpienia zdarzenia:

$$S(t) = e^{-\int_0^t \lambda(u) du}$$

# Funkcja skumulowanego hazardu

- ◆ Zdefiniowana jako

$$\Lambda(t) = \int_0^t \lambda(u) du$$

- ◆ Nieujemna i niemalejąca (gdyż  $\lambda(t) \geq 0$ )

- ◆ Zachodzi

$$-\ln S(t) = \int_0^t \lambda(u) du = \Lambda(t)$$

- ◆ Funkcja skumulowanego hazardu jest związana z  $\lambda(t)$  i  $S(t)$ , więc również zawiera pełną informację o rozkładzie czasu do wystąpienia zdarzenia

# Modelowanie funkcji hazardu

- ◆ Załóżmy, że mamy wektor zmiennych niezależnych

$$\mathbf{Z}(t)' = (Z_1(t), Z_2(t), \dots, Z_p(t))$$

- w szczególnym przypadku,  $\mathbf{Z}(t)$  może nie zależeć od czasu

- ◆ Model „względnego hazardu” (“względnego ryzyka”) otrzymujemy przez przyjęcie, że

$$\lambda(t; \mathbf{Z}(t)) = \lambda_0(t) \cdot r(\mathbf{Z}(t))$$

gdzie  $r()$  jest ustaloną funkcją, np.

- $r(\mathbf{Z}(t)) = 1 + \beta' \cdot \mathbf{Z}(t)$
- $r(\mathbf{Z}(t)) = \{1 + \beta' \cdot \mathbf{Z}(t)\}^{-1}$
- $r(\mathbf{Z}(t)) = e^{\beta' \cdot \mathbf{Z}(t)}$

# Model Coxa

- ♦ Zakłada, że

$$\lambda(t; \mathbf{Z}(t)) = \lambda_0(t) \cdot e^{\beta' \mathbf{Z}(t)}$$

- ♦ W szczególności, dla  $\mathbf{Z}(t) \equiv \mathbf{Z}$ , otrzymujemy *model proporcjonalnych hazardów*:

$$\lambda(t; \mathbf{Z}) = \lambda_0(t) \cdot e^{\beta' \mathbf{Z}}$$



# Model proporcjonalnych hazardów

- ◆ Zakłada, że wpływ zmiennej niezależnej  $Z$  na czas do wystąpienia niepowodzenia wyraża się poprzez przemnożenie funkcji hazardu przez stałą wartość liczbowa
- ◆ Przykład:
  - Funkcja hazardu dla kobiet:  $\lambda_F(t)$
  - Funkcja hazardu dla mężczyzn:  $\lambda_M(t) = \lambda_F(t) \cdot c$
- ◆ Interpretacja: hazard dla mężczyzn jest  $c$  razy mniejszy/wiekszy niż dla kobiet, *niezależnie od chwili czasu*

# Iloraz hazardów

- ◆ Równoważne sformułowanie modelu proporcjonalnych hazardów:

$$\frac{\lambda_M(t)}{\lambda_F(t)} = c$$

Iloraz hazardów  $HR(t)$

- ◆ Model proporcjonalnych hazardów (PH) zakłada, że iloraz hazardów dla dwóch wartości zmiennej niezależnej  $Z$  jest stały w czasie
  - często iloraz jest (niepoprawnie) nazywany *ryzykiem względnym* (*relative risk*)

# Iloraz hazardów w modelu PH (1)

$$\lambda(t) = \lambda_0(t) \cdot e^{\beta \cdot Z}$$

## ♦ Przykład:

- Przyjmijmy  $Z=0, 1, 2$ , dla TNM= I, II, III
- Funkcja hazardu dla TNM I:  $\lambda_I(t) = \lambda_0(t) \cdot e^{0 \cdot \beta} = \lambda_0(t)$
- Funkcja hazardu dla TNM II:  $\lambda_{II}(t) = \lambda_0(t) \cdot e^{1 \cdot \beta} = \lambda_0(t) \cdot e^{\beta}$
- Funkcja hazardu dla TNM III:  $\lambda_{III}(t) = \lambda_0(t) \cdot e^{2 \cdot \beta} = \lambda_0(t) \cdot e^{2\beta}$

# Iloraz hazardów w modelu PH (2)

## ◆ Przykład (cd.):

- Funkcja hazardu dla TNM I:  $\lambda_I(t) = \lambda_0(t) \cdot e^{0 \cdot \beta} = \lambda_0(t)$
- Funkcja hazardu dla TNM II:  $\lambda_{II}(t) = \lambda_0(t) \cdot e^{1 \cdot \beta} = \lambda_0(t) \cdot e^{\beta}$
- Funkcja hazardu dla TNM III:  $\lambda_{III}(t) = \lambda_0(t) \cdot e^{2 \cdot \beta} = \lambda_0(t) \cdot e^{2\beta}$

- Iloraz hazardów dla  $Z=1$  vs.  $Z=0$ :

$$\lambda_{II}(t)/\lambda_I(t) = \lambda_0(t) \cdot e^{\beta} / \lambda_0(t) = e^{\beta}$$

- Dla  $Z=2$  vs.  $Z=1$ :

$$\lambda_{III}(t)/\lambda_{II}(t) = \lambda_0(t) \cdot e^{2\beta} / \lambda_0(t) \cdot e^{\beta} = e^{\beta}$$

## ◆ Model zakłada, że wzrost $Z$ o 1 zwiększa hazard $e^{\beta}$ razy

- iloraz hazardu (HR) dla  $Z=z+1$  vs.  $Z=z$  wynosi  $e^{\beta}$

# Efekt zmiennej niezależnej w modelu PH

- ♦ Wniosek: w modelu PH z jedną zmienną niezależną:

$$\lambda_Z(t) = \lambda_0(t) \cdot e^{\beta \cdot Z}$$

$e^{\beta}$  wyraża proporcjonalną zmianę hazardu wynikającą ze wzrostu  $Z$  o 1 jednostkę

- ♦  $\beta > 0 \rightarrow$  hazard wzrasta;  $\beta < 0 \rightarrow$  maleje

- ♦ Przykład:

- Przyjmijmy  $Z=0, 1, 2$  dla TNM= I, II, III
- Dla  $\beta=1$ , otrzymujemy  $e^{\beta} = e^1 = 2.73$
- Wzrost TNM o jeden stopień zwiększa hazard (np. śmiertelność) 2.73 razy

# Poziom odniesienia zmiennej niezależnej (1)

## ♦ Przykład:

- $Z=0, 1, 2$ , dla TNM= I, II, III
- Funkcja hazardu dla TNM I:  $\lambda_I(t) = \lambda_0(t) \cdot e^{0\beta} = \lambda_0(t)$
- Funkcja hazardu dla TNM II:  $\lambda_{II}(t) = \lambda_0(t) \cdot e^{1\beta} = \lambda_I(t) \cdot e^\beta$
- Funkcja hazardu dla TNM III:  $\lambda_{III}(t) = \lambda_0(t) \cdot e^{2\beta} = \lambda_I(t) \cdot e^{2\beta}$

- ## ♦ Wniosek: funkcje hazardu dla TNM II i III są zdefiniowane *względem* funkcji hazardu dla TNM I

# Poziom odniesienia zmiennej niezależnej (2)

## ♦ Przykład (cd.):

- $Z=2, 1, 0$ , dla  $TNM= I, II, III$

- $TNM III: \lambda_{III}(t) = \lambda_0(t) \cdot e^{0\beta} = \lambda_0(t)$

- $TNM II: \lambda_{II}(t) = \lambda_0(t) \cdot e^{1\beta} = \lambda_{III}(t) \cdot e^{\beta}$

- $TNM I: \lambda_I(t) = \lambda_0(t) \cdot e^{2\beta} = \lambda_{III}(t) \cdot e^{2\beta}$

## ♦ Wniosek: funkcje hazardu dla $TNM I$ i $II$ są zdefiniowane względem funkcji hazardu dla $TNM III$

## ♦ Poziom odniesienia ( $Z=0$ ) uległ zmianie!

# Częściowa funkcja wiarygodności

- ◆ Przedstawmy dane  $\mathbf{Y}$  jako zbiór zmiennych  $A_1, B_1, \dots, A_m, B_m$
- ◆ Niech  $A^{(j)} = (A_1, \dots, A_j)$  and  $B^{(j)} = (B_1, \dots, B_j)$
- ◆ Załóżmy, że gęstość łącznego rozkładu  $(A^{(m)}, B^{(m)})$  można przedstawić jako

$$\prod_{j=1}^m f(b_j | b^{(j-1)}, a^{(j-1)}; \theta, \beta) \prod_{j=1}^m f(a_j | b^{(j)}, a^{(j-1)}; \beta)$$

częściowa funkcja  
wiarygodności

- ◆ Nie jest funkcją wiarygodności w „zwykłym” znaczeniu. Ale może być użyta do wnioskowania o  $\beta$ 
  - z pewną stratą na efektywności szacowania



# Częściowa funkcja wiarygodności dla modelu Coxa

$$\lambda(t; \mathbf{Z}(t)) = \lambda_0(t) e^{\beta' \mathbf{Z}(t)}$$

- ◆  $K$  obserwacji pełnych (zdarzeń) dla  $t_{(1)} < \dots < t_{(K)}$  (bez „zlepień”)
- ◆  $B_j$  = cenzurowanie w  $[t_{(k-1)}, t_{(k)})$  i zdarzenia w  $[t_{(k)}, t_{(k)} + h_k)$
- ◆  $A_j$  = zdarzenie dla jednostki  $j$  w  $[t_{(k)}, t_{(k)} + h_k)$
- ◆  $Y_i(t) = 1$  jeśli jednostka pod obserwacją dla  $t$ , 0 w p.p.
- ◆ Dla **niezależnego** cenzurowania

$$\begin{aligned} L(\beta) &= \prod_{k=1}^K f(a_k | b^{(k)}, a^{(k-1)}; \beta) = \prod_{k=1}^K \frac{\lambda(t_{(k)}; \mathbf{Z}_{(k)}(t_{(k)})) h_k}{\sum_{l=1}^n Y_l(t_{(k)}) \lambda(t_{(k)}; \mathbf{Z}_l(t_{(k)})) h_k} \\ &= \prod_{k=1}^K \frac{\lambda_0(t_{(k)}) h_k e^{\beta' \mathbf{Z}_{(k)}(t_{(k)})}}{\sum_{l=1}^n Y_l(t_{(k)}) \lambda_0(t_{(k)}) h_k e^{\beta' \mathbf{Z}_l(t_{(k)})}} = \prod_{k=1}^K \frac{e^{\beta' \mathbf{Z}_{(k)}(t_{(k)})}}{\sum_{l=1}^n Y_l(t_{(k)}) e^{\beta' \mathbf{Z}_l(t_{(k)})}} \end{aligned}$$

# Jednoczesne zdarzenia (1)

- ◆ Niech  $r_i(t) = \exp\{\beta' \mathbf{Z}_i(t)\}$
- ◆ Wkład pojedynczej obserwacji pełnej (zdarzenia) do  $L(\beta)$  to

$$\frac{r_i(t)}{\sum_{l=1}^n Y_l(t) r_l(t)}$$

- ◆ Rozważmy 3 jednoczesne zdarzenia w  $t$  wśród 5 obserwacji
- ◆ Gdybyśmy znali dokładne czasy dla dwóch pierwszych zdarzeń, wkład do  $L(\beta)$  wynosiłby

$$\frac{r_1}{r_1 + r_2 + r_3 + r_4 + r_5} \cdot \frac{r_2}{r_2 + r_3 + r_4 + r_5} \quad \text{lub} \quad \frac{r_2}{r_1 + r_2 + r_3 + r_4 + r_5} \cdot \frac{r_1}{r_1 + r_3 + r_4 + r_5}$$

# Jednoczesne zdarzenia (2)

- ◆ Metoda Breslawa: 
$$\frac{r_1 r_2 r_3}{(r_1 + r_2 + r_3 + r_4 + r_5)^3}$$
- ◆ Metoda Efrona: 
$$\frac{r_1 r_2 r_3}{(r_1 + r_2 + r_3 + r_4 + r_5) \left( \frac{2r_1}{3} + \frac{2r_2}{3} + \frac{2r_3}{3} + r_4 + r_5 \right) \left( \frac{r_1}{3} + \frac{r_2}{3} + \frac{r_3}{3} + r_4 + r_5 \right)}$$
- ◆ Metoda „dokładna”: rozważyć wszystkie 3! konfiguracje i użyć średniej odpowiednich wkładów do  $L(\beta)$ .
  - Problem: ogólnie,  $d! \cdot d$  mnożeń; np.  $8! \cdot 8 = 322560$
- ◆ Idea:
 
$$\frac{r_1}{r_1 + r_2 + r_3 + r_4 + r_5} \cdot \frac{r_2}{r_2 + r_3 + r_4 + r_5} + \frac{r_2}{r_1 + r_2 + r_3 + r_4 + r_5} \cdot \frac{r_1}{r_1 + r_3 + r_4 + r_5}$$

$$= \int_0^\infty \left\{ 1 - \exp\left(-\frac{r_1 t}{r_3 + r_4 + r_5}\right) \right\} \left\{ 1 - \exp\left(-\frac{r_2 t}{r_3 + r_4 + r_5}\right) \right\} e^{-t} dt$$
  - Całkujemy numerycznie; złożoność  $O(d^2)$

# Jednoczesne zdarzenia (3)

- ◆ Potraktujmy dane jako „dyskretne”. Wówczas możemy rozważyć

$$\frac{r_1 r_2 r_3}{\sum_{(i_1, i_2, i_3)} r_{i_1} r_{i_2} r_{i_3}}$$

- ◆ Problem numeryczny: dla 10 jednoczesnych zdarzeń wśród 1000 obserwacji, mianownik zawiera  $2.6 \cdot 10^{23}$  składników.
- ◆ Problem teoretyczny: odpowiedni model to

$$\frac{\lambda_i(t)}{1 - \lambda_i(t)} = \frac{\lambda_0(t)}{1 - \lambda_0(t)} e^{\beta' Z_i}$$

# Jednoczesne zdarzenia (4)

- ◆ Metoda Breslowa: najszybsza, lecz najmniej dokładna.
  - Bardzo często domyślna w różnych pakietach statystycznych!
- ◆ Metoda „dokładna”: złożona numerycznie.
- ◆ Metoda Efrona: szybka i daje dobre przybliżenie „dokładnej”.
  - O ile liczba jednoczesnych zdarzeń nie jest zbyt duża.
- ◆ Metoda dla danych dyskretnych: nie dla modelu PH.

# Pochodna logarytmu (częściowej) funkcji wiarygodności i macierz informacji

♦ Średnia ważona  $\mathbf{Z}(t)$ : 
$$\bar{Z}(\beta, t) = \sum_{i=1}^n Z_i(t) \frac{Y_i(t) e^{\beta' Z_i(t)}}{\sum_{j=1}^n Y_j(t) e^{\beta' Z_j(t)}} \equiv \sum_{i=1}^n w_i(t) Z_i(t)$$

♦ *Gradient („score”)* 
$$U(\beta) \equiv \frac{\partial \ln L(\beta)}{\partial \beta} = \sum_{k=1}^K \{Z_{(k)}(t_{(k)}) - \bar{Z}(\beta, t_{(k)})\}$$

♦ *(Obserwowana) Macierz informacji* 
$$I(\beta) = -\frac{\partial^2 \ln L}{\partial \beta \partial \beta^T} = \sum_{k=1}^K V(\beta, t_{(k)})$$

gdzie 
$$V(\beta, t) = \sum_{i=1}^n w_i(t) \{Z_i(t) - \bar{Z}(\beta, t)\} \{Z_i(t) - \bar{Z}(\beta, t)\}^T$$

- ważona wariancja  $\mathbf{Z}$  w chwili  $t$

# Szacowanie parametrów modelu Coxa

- ♦ Oszacowanie  $\beta$  uzyskujemy rozwiązując  $U(\hat{\beta}) = 0$
- ♦ Można pokazać, że asymptotycznie  $\hat{\beta} \sim N[\beta, \{E(I(\beta))\}^{-1}]$
- ♦ Estymator wariancji:  $I(\hat{\beta})^{-1}$

# Testowanie hipotez (1)

## ♦ Testowanie hipotez $H_0: \beta = \beta_0$ (np. $\beta = 0$ )

- Test ilorazu wiarygodności:  $2\{\ln L(\hat{\beta}) - \ln L(\beta_0)\} \equiv 2\{l(\hat{\beta}) - l(\beta_0)\}$
- Test Walda:  $(\hat{\beta} - \beta_0)^T I(\hat{\beta})(\hat{\beta} - \beta_0)$
- Test „score”:  $U^T(\beta_0)I(\beta_0)^{-1}U(\beta_0)$

## ♦ Rozkład $\chi^2$ z $p$ (liczba elementów $\beta$ ) stopniami swobody



# Testowanie hipotez (2)

- ◆ Test Walda, ilorazu wiarygodności i „score” równoważne asymptotycznie, ale nie dla „małych” próbek
  - Preferowany test ilorazu wiarygodności
- ◆ Test „score” dla jednej zmiennej kategorycznej równoważny testowi logrank.
- ◆ Uwaga na braki danych przy teście ilorazu wiarygodności! Modele powinny być szacowane na tym samym zbiorze danych.

# Równoważność asymptotyczna testów

- ◆ Test Walda, ilorazu wiarygodności i „score” równoważne asymptotycznie.

- Test Walda: przybliżenie  $l(\beta)$  funkcją kwadratową w  $\hat{\beta}$

$$l(\beta) \approx l(\hat{\beta}) + (\beta - \hat{\beta})^T U(\hat{\beta}) - (\beta - \hat{\beta})^T I(\hat{\beta})(\beta - \hat{\beta}) / 2$$

$$l(\hat{\beta}) - l(\beta_0) \approx (\hat{\beta} - \beta_0)^T I(\hat{\beta})(\hat{\beta} - \beta_0) / 2$$

- Test „score”: przybliżenie  $l(\beta)$  funkcją kwadratową w  $\beta_0$

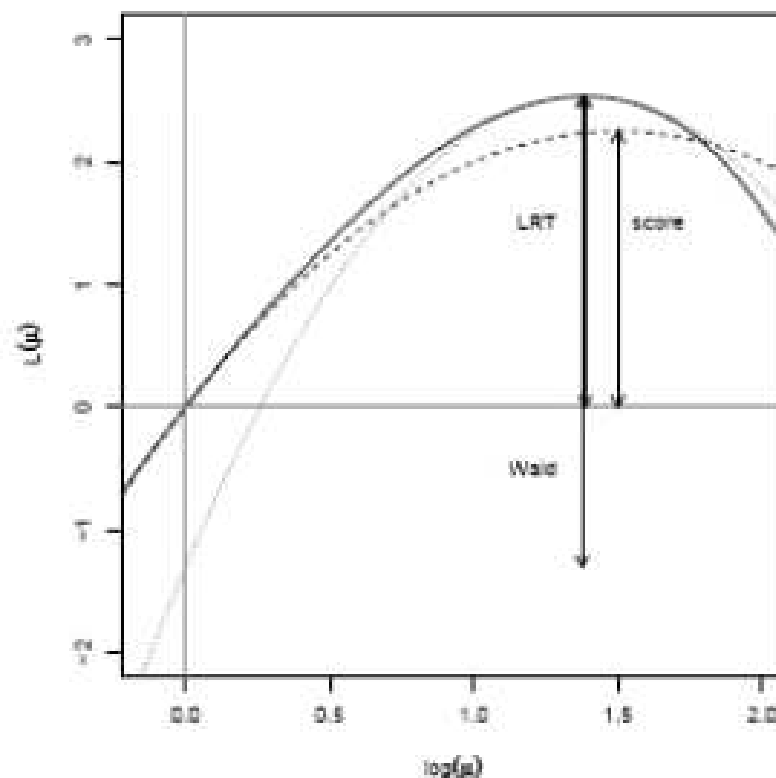
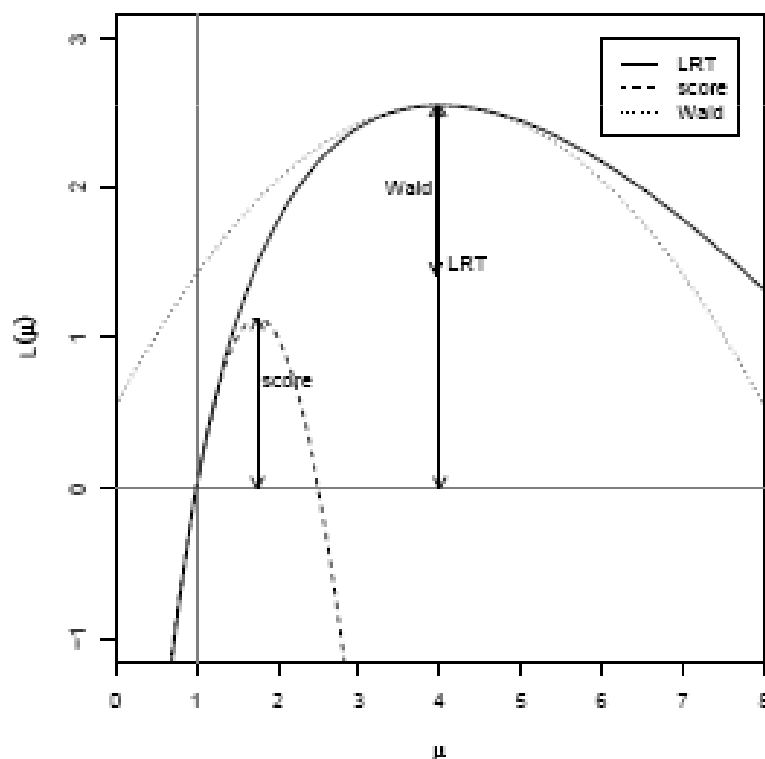
$$l(\beta) \approx l(\beta_0) + (\beta - \beta_0)^T U(\beta_0) - (\beta - \beta_0)^T I(\beta_0)(\beta - \beta_0) / 2$$

$$U(\hat{\beta}) \approx U(\beta_0) - I(\beta_0)(\hat{\beta} - \beta_0) \Rightarrow I(\beta_0)^{-1} U(\beta_0) \approx (\hat{\beta} - \beta_0)$$

$$l(\hat{\beta}) - l(\beta_0) \approx U(\beta_0)^T I(\beta_0)^{-1} U(\beta_0) / 2$$

# Własności testów dla „małych” próbek

- ◆ Test Walda, ilorazu wiarygodności i „score” równoważne asymptotycznie, ale nie dla „małych” próbek
  - Przykład: test  $H_0: \mu=1$  dla obserwacji  $Y=4$  z rozkładu Poissona



# Testowanie hipotez liniowych

## ♦ Testowanie liniowych hipotez $H_0: C\beta = \beta_0$

- Np.  $(0,0,1,-1) (\beta_1, \beta_2, \beta_3, \beta_4)^T = 0 \Leftrightarrow \beta_3 = \beta_4$

## ♦ Test Walda: $(C\hat{\beta} - \beta_0)^T \{CI(\hat{\beta})^{-1} C^T\}^{-1} (C\hat{\beta} - \beta_0) \sim \chi^2_{rank(C)}$

## ♦ Test ilorazu wiarygodności:

- Szacujemy model „alternatywny” ( $\beta$ )
- Szacujemy model odpowiadający  $H_0$  ( $\beta^*$ )

- Statystyka testowa:  $2\{l(\hat{\beta}^*) - l(\hat{\beta})\} \sim \chi^2_{rank(C)}$

# Przedziały ufności dla współczynników modelu

- ◆ Dopasowanie modelu do danych pozwala na uzyskanie oszacowania współczynników  $\beta$  i ich błędów standardowych ( $SE$ )
- ◆ Przy ich użyciu możemy skonstruować 95% przedział ufności dla  $\beta_i$ :  $[\hat{\beta}_i - 1.96 \cdot SE(\hat{\beta}_i), \hat{\beta}_i + 1.96 \cdot SE(\hat{\beta}_i)]$
- ◆ Przedział ufności dla ilorazu hazardów:

$$[e^{\hat{\beta}_i - 1.96 \cdot SE(\hat{\beta}_i)}, e^{\hat{\beta}_i + 1.96 \cdot SE(\hat{\beta}_i)}]$$

# Przedział ufności oparty na „profilowanej” funkcji wiarygodności

- ◆ Użyteczne jeśli problem z przybliżeniem rozkładem normalnym (Wald)
- ◆ Idea: dla różnych wartości  $\beta_i$  dopasowujemy model do danych i wykreślamy wartość  $l(\hat{\beta})$
- ◆ 95% przedział ufności dla  $\beta_i$  oparty na  $\arg \max \{l(\hat{\beta}) - 3.84/2\}$ 
  - wymaga „ręcznego” programowania

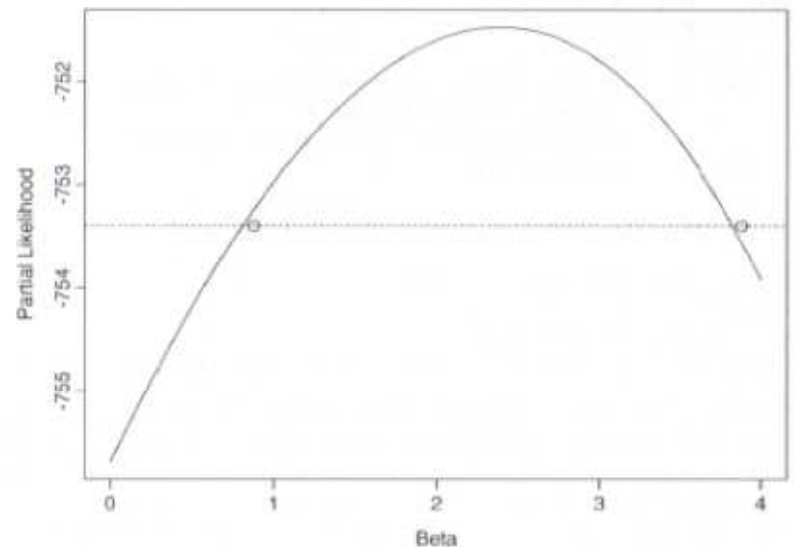


FIGURE 3.2: Profile likelihood curve and 95% confidence interval cutoff for prothrombin time

# „Nieskończone” współczynniki

- ◆ Gdy np. w jednej z grup nie było zdarzeń:

Grupa	Żyjący	Zgony
Leczona	40	0
Kontrolna	30	10

- ◆ Przykład (S+):

```
>fit <- coxph(Surv(futime, fustat) ~ rx + fustat, ovarian)
Warning messages:
  Loglik converged before variable 2 ; beta may be infinite.
```

```
>summary(fit)
n= 26

      coef exp(coef) se(coef)      z      p
rx    -0.557  5.73e-01   0.62 -0.898 0.37
fustat 10.582  3.94e+04  34.38  0.308 0.76
```

```
Likelihood ratio test= 30.8 on 2 df,  p=2.05e-07
Wald test               = 0.9 on 2 df,  p=0.637
Efficient score test = 29.1 on 2 df,  p=4.83e-07
```

# Testy dla „nieskończonych” współczynników

## ♦ Test ilorazu wiarygodności:

$$l(10.582) - l(0) \approx l(\infty) - l(0)$$

- wynik ma sens

## ♦ Test „score” : przybliżenie funkcją kwadratową w 0

- max dla 3.4
- wynik ma sens

## ♦ Test Walda : przybliżenie funkcją kwadratową w 10.6

- B. złe, wynik nie ma sensu

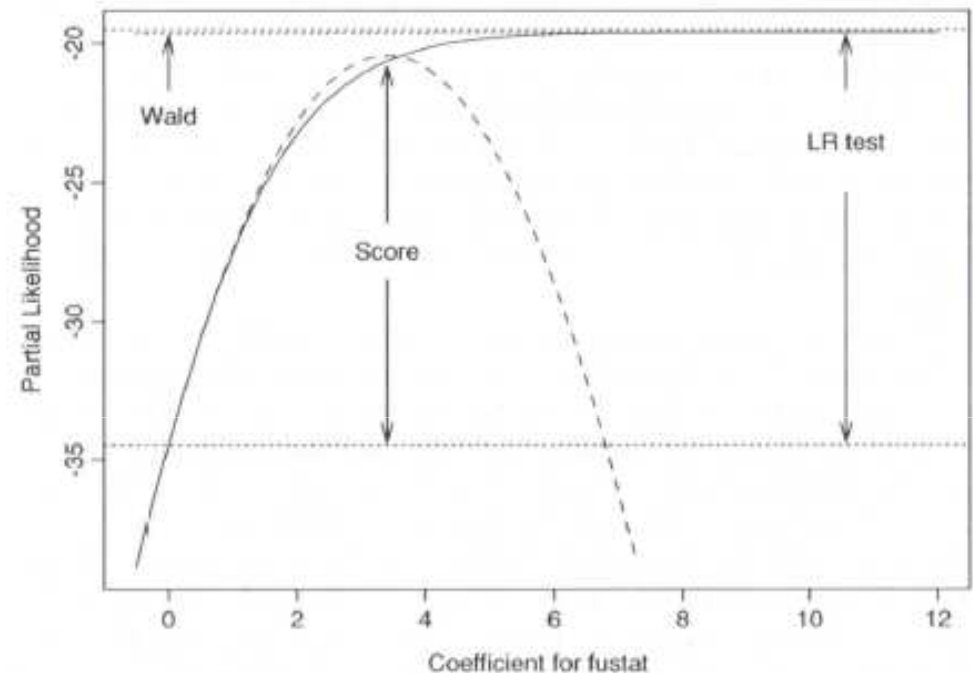
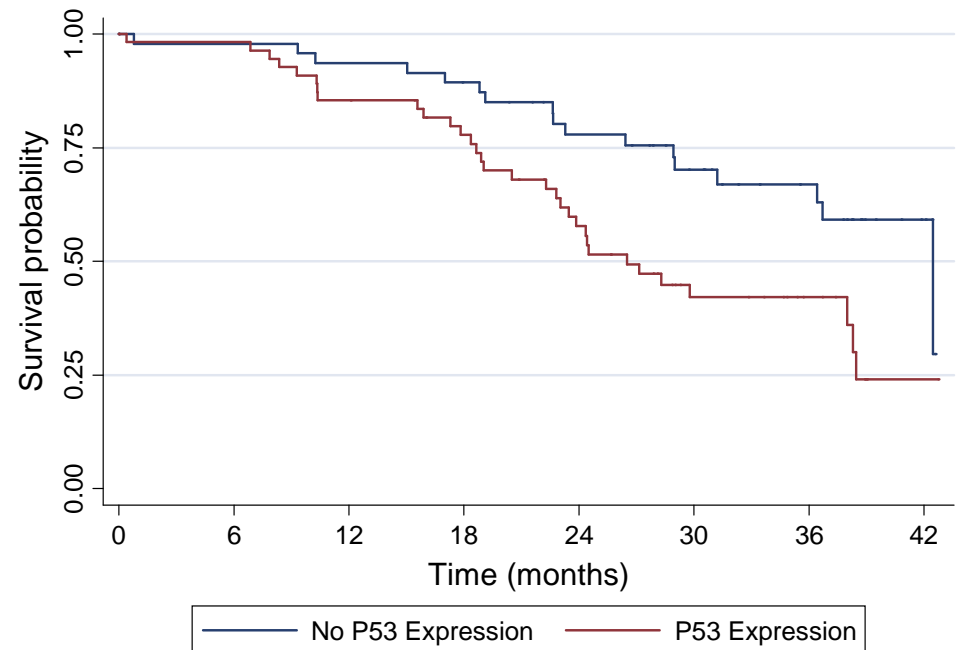


FIGURE 3.3: Likelihood ratio, score, and Wald tests for an infinite coefficient



# Model PH: niedrobnokomórkowy rak płuca (NSCLC) (1)

- ♦ Laudański *et al.*, *Eur Respir J* (2001)
- ♦ Wpływ ekspresji białka P53 na czas przeżycia 102 pacjentów
- ♦ Test logrank:  $p=0.008 \rightarrow$



## Model PH: NSCLC (2)

- ♦ Zmienna „P53e” = 0 jeśli brak ekspresji, 1 jeśli ekspresja

_t	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
P53e	.7859412	.3021812	2.601	0.009	.1936769	1.378205
	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
P53e	2.194471	.663128	2.601	0.009	1.213704	3.967775

- ♦ Wniosek: ekspresja P53 zwiększa hazard zgonu (śmiertelność) w przybliżeniu 2 razy
- ♦ Efekt jest istotnie statystyczny na poziomie istotności 5%
  - Uwaga: dla testu logrank,  $p=0.008$

# Alternatywna postać modelu PH (1)

- ♦ Zauważmy, że

$$\lambda(t) = \lambda_0(t) \cdot e^{\beta \cdot Z}$$

implikuje

$$\Lambda(t) = \Lambda_0(t) \cdot e^{\beta \cdot Z}$$

- ♦ W konsekwencji

$$e^{-\Lambda(t)} = e^{-\Lambda_0(t) \cdot \exp(\beta \cdot Z)}$$

oraz

$$S(t) = S_0(t)^{\exp(\beta \cdot Z)}$$

gdzie  $S_0(t)$  jest funkcją przeżycia odpowiadającą bazowej funkcji hazardu

## Alternatywna postać modelu PH (2)

$$S(t) = S_0(t)^{\exp(\beta \cdot Z)}$$

- ♦ Wzrost  $Z = z \rightarrow Z = z + 1$  powoduje:

$$S_z(t) = S_0(t)^{\exp(\beta \cdot z)} \rightarrow S_{z+1}(t) = S_0(t)^{\exp(\beta \cdot z + \beta)} = S_z(t)^{\exp(\beta)}$$

- ♦  $\beta > 0$  oznacza zmniejszenie wartości p-stwa przeżycia
  - wyższą wartość hazardu

# Model PH na skali logarytmicznej

- ◆ Ponieważ

$$\lambda(t) = \lambda_0(t) \cdot e^{\beta \cdot Z}$$

$$\Lambda(t) = \Lambda_0(t) \cdot e^{\beta \cdot Z}$$

to

$$\ln \lambda(t) = \ln \lambda_0(t) + \beta \cdot Z$$

$$\ln \Lambda(t) = \ln \Lambda_0(t) + \beta \cdot Z$$

- ◆ Na skali logarytmicznej otrzymujemy model liniowy

- interpretacja  $\beta$  jak w modelu prostej regresji liniowej

- ◆  $\Lambda(t) = -\ln S(t)$ , skąd

$$\ln \{ -\ln S(t) \} = \ln \{ -\ln S_0(t) \} + \beta \cdot Z$$

- Związek użyteczny dla graficznego sprawdzenia założenia PH

# Szacowanie bazowej funkcji hazardu

- ◆ Dla zaobserwowanych czasów zdarzeń  $t_{(1)} < \dots < t_{(K)}$  kładziemy

$$\hat{\lambda}_0(t_{(k)}) = 1 - \hat{\xi}_k$$

gdzie

$$\sum_{j: t_j = t_{(k)}} \frac{\delta_j e^{\hat{\beta}' Z_j(t_{(k)})}}{1 - (\hat{\xi}_k)^{e^{\beta' Z_j(t_{(k)})}}} = \sum_{i=1}^n Y_i(t_{(k)}) e^{\hat{\beta}' Z_i(t_{(k)})}$$

- ◆ Jeśli nie ma jednoczesnych czasów zdarzeń, to

$$\hat{\xi}_k = \left\{ 1 - \frac{e^{\hat{\beta}' Z_{(k)}(t_{(k)})}}{\sum_{i=1}^n Y_i(t_{(k)}) e^{\hat{\beta}' Z_i(t_{(k)})}} \right\} e^{-\hat{\beta}' Z_{(k)}(t_{(k)})}$$

# Szacowanie bazowej funkcji przeżycia

- ◆ Podobnie jak w metodzie Kaplana-Meiera, używamy

$$\hat{S}_0(t) = \prod_{t_{(k)} \leq t} \hat{\xi}_k$$

- Bez zmiennych niezależnych mamy

$$\sum_{j: t_j = t_{(k)}} \frac{\delta_j}{1 - \hat{\xi}_k} \equiv \frac{d_k}{1 - \hat{\xi}_k} = \sum_{i=1}^n Y_i(t_{(k)}) \equiv n_k \Rightarrow \hat{\xi}_k = \frac{n_k - d_k}{n_k}$$

czyli estymator Kaplana-Meiera

# Szacowanie bazowej funkcji skumulowanego hazardu

- ◆ Korzystając z oszacowania funkcji przeżycia, mamy

$$\hat{\Lambda}_0(t) = -\ln \hat{S}_0(t) = -\sum_{t_{(k)} \leq t} \ln \hat{\xi}_k$$

- ◆ Alternatywnie, kładąc  $(\xi_k)^{e^{-\hat{\beta}'Z_j(t_{(k)})}} \approx 1 + e^{-\hat{\beta}'Z_j(t_{(k)})} \ln \xi_k$

dostajemy

$$\sum_{j:t_j=t_{(k)}} \frac{\delta_j}{-\ln \tilde{\xi}_k} = \sum_{i=1}^n Y_i(t_{(k)}) e^{\hat{\beta}'Z_i(t_{(k)})} \Rightarrow \tilde{\xi}_k = \exp \left( -\frac{d_k}{\sum_{i=1}^n Y_i(t_{(k)}) e^{\hat{\beta}'Z_i(t_{(k)})}} \right)$$

i w konsekwencji estymator Breslowa:

$$\tilde{\Lambda}_0(t) = -\ln \tilde{S}_0(t) = \sum_{t_{(k)} \leq t} \frac{d_k}{\sum_{i=1}^n Y_i(t_{(k)}) e^{\hat{\beta}'Z_i(t_{(k)})}}$$



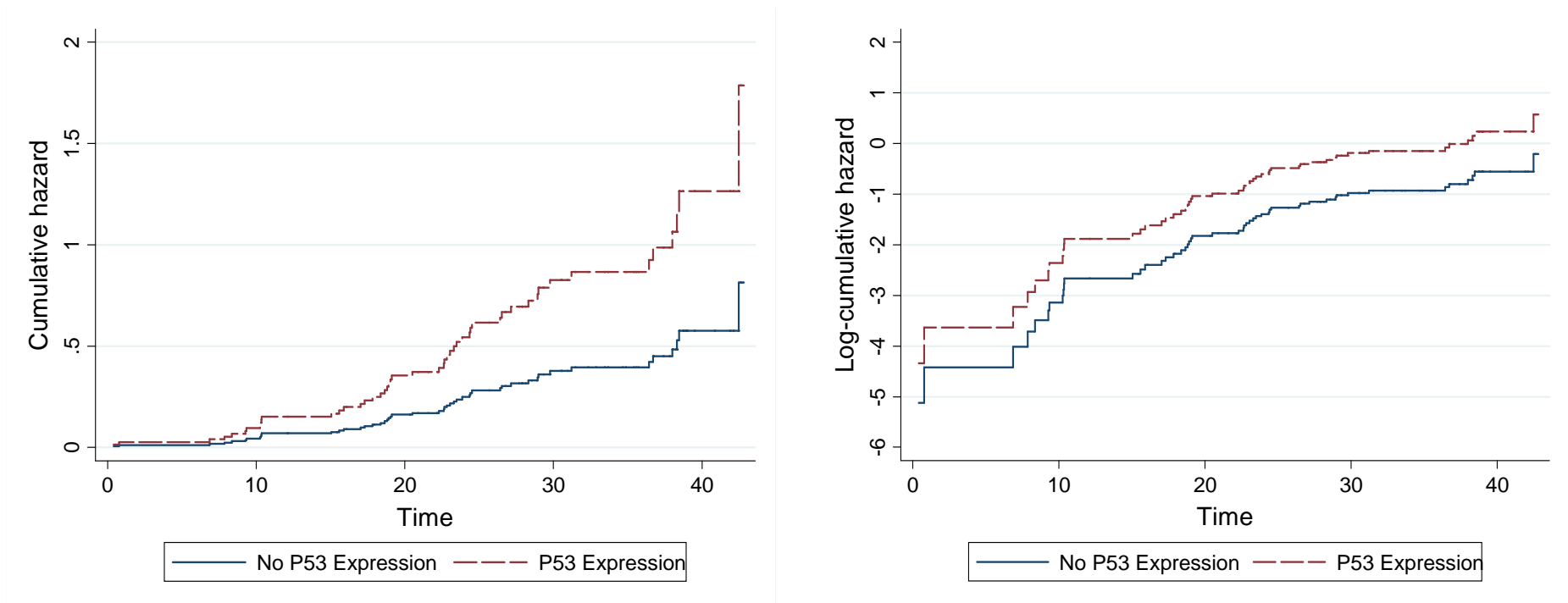
## Model PH: NSCLC (3)

- ♦ Zmienna „P53e” = 0 jeśli brak ekspresji, 1 jeśli ekspresja

_t	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
P53e	.7859412	.3021812	2.601	0.009	.1936769	1.378205
	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
P53e	2.194471	.663128	2.601	0.009	1.213704	3.967775

- ♦ Wniosek: ekspresja P53 zwiększa śmiertelność w przybliżeniu 2 razy w porównaniu do braku ekspresji
  - logarytm hazardu wzrasta o 0.78

# Model PH: NSCLC, ekspresja P53



- ◆ Oszacowanie (logarytmu) funkcji skumulowanego hazardu

# Model PH dla wielu zmiennych niezależnych

- ♦ Model dla trzech zmiennych  $Z_1$ ,  $Z_2$ , i  $Z_3$  przyjmuje postać:

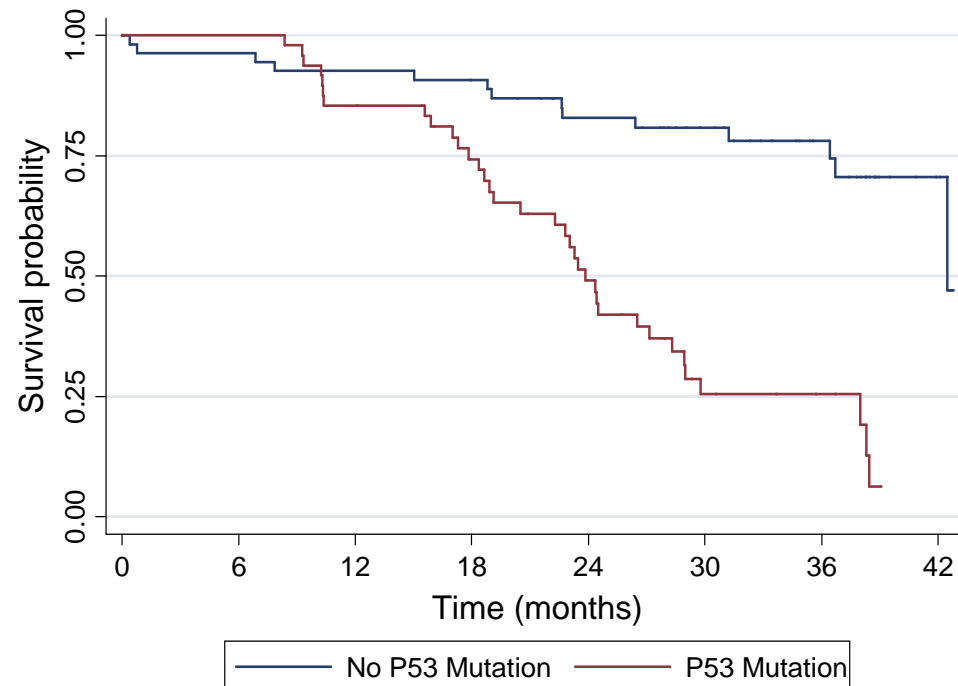
$$\lambda(t) = \lambda_0(t) e^{\beta_1 Z_1 + \beta_2 Z_2 + \beta_3 Z_3}$$

- ♦  $\exp(\beta_1)$  opisuje zmianę hazardu wynikającą ze wzrostu  $Z_1$  o 1 jednostkę dla ustalonych wartości  $Z_2$  i  $Z_3$
- ♦ Model umożliwia skorygowanie efektu  $Z_1$  ze względu na możliwy efekt  $Z_2$  i  $Z_3$

# Model PH: NSCLC (4)

- ♦ Laudański *et al.*, *Eur Respir J* (2001)
- ♦ Mutacja genu P53 (analiza DNA)
- ♦ Wpływ na czas przeżycia 102 chorych
- ♦ Test logrank:  $p < 0.001 \rightarrow$

mutacja	ekspresja +
-	27.8%
+	82.3%



# Model PH: NSCLC (5)

- ◆ Zmienna „P53m” = 0 jeśli brak mutacji, 1 jeśli mutacja

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
P53m	5.670694	2.308899	4.262	0.000	2.553043	12.59547
P53e	.8191471	.3038457	-0.538	0.591	.3959371	1.694718

- ◆ Wniosek: skorygowanie ze względu na efekt mutacji „usuwa” efekt ekspresji białka P53 (współliniowość?)
- ◆ Ocena ekspresji białka nie „dodaje informacji” jeśli wiemy czy występuje mutacja

# Model PH: NSCLC (6)

_t	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
P53m	1.735312	.4071634	4.26	0.000	.937286	2.533337
P53e	-.1994916	.3709294	-0.54	0.591	-.9264999	.5275166

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
P53m	5.670694	2.308899	4.262	0.000	2.553043	12.59547
P53e	.8191471	.3038457	-0.538	0.591	.3959371	1.694718

- ◆ Obecność mutacji zwiększa śmiertelność ponad 5.5 razy
- ◆ Teoretycznie, model definiuje 4 „grupy ryzyka”:

mutacja (P53m)	ekspresja (P53e)	iloraz hazardów
- (0)	- (0)	1
- (0)	+ (1)	$e^{1.74 \cdot 0 - 0.20 \cdot 1} = 1 \cdot e^{-0.20} = 0.82$
+ (1)	- (0)	$e^{1.74 \cdot 1 - 0.20 \cdot 0} = e^{1.74} \cdot 1 = 5.67$
+ (1)	+ (1)	$e^{1.74 \cdot 1 - 0.20 \cdot 1} = 5.67 \cdot 0.82 = 4.65$

# Model PH: marskość wątroby

- ◆ Christiansen *et al.*, *Gastroenterology* (1985)
- ◆ Próba kliniczna z 248 chorymi na marskość wątroby
- ◆ Czas przeżycia
- ◆ W oparciu o model Coxa, zdefiniowano indeks prognostyczny

# Marskość wątroby: indeks prognostyczny

	Coeff.	Std. Err.	P> z
bilirubin [log (μmol/L)]	2.51	0.32	<0.001
age [exp {(years-20)/10}]	0.0069	0.0016	<0.001
cirrhosis [no=0, yes=1]	0.88	0.22	<0.001
albumin [g/L]	-0.50	0.018	<0.006
cholestiasis [no=0, yes=1]	0.68	0.27	0.01
drug [azathioprine=0, placebo=1]	0.52	0.21	0.01

♦ indeks prognostyczny:

$$\text{PI} = 2.51 \times \text{bilirubin} + 0.0069 \times \text{age} + 0.88 \times \text{cirrhosis} \\ - 0.50 \times \text{albumin} + 0.68 \times \text{cholestiasis} + 0.52 \times \text{drug}$$



# Marskość wątroby: walidacja indeksu

85 chorych z grupy placebo w innej próbie klinicznej

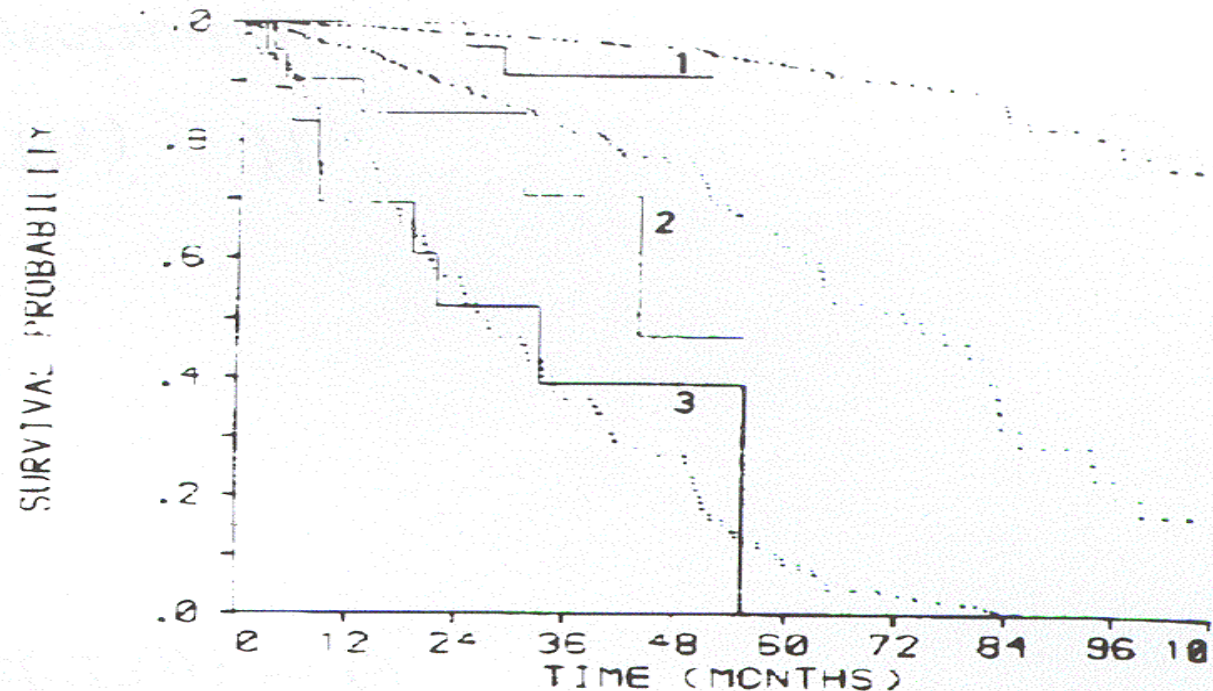


Figure 7. Observed (—) and estimated (----) survival function of three groups of placebo-treated patients from a controlled trial of D-penicillamine (Reference 4): group 1:  $PI < 2.5$ ,  $n = 44$ ,  $O = 2$ ,  $E = 0.9$ ; group 2:  $2.5 < PI < 3.75$ ,  $n = 22$ ,  $O = 5$ ,  $E = 2.7$ ; group 3:  $PI > 3.75$ ,  $n = 19$ ,  $O = 9$ ,  $E = 9.0$ .  $O$  and  $E$  are the observed and expected numbers dying in each group ( $\chi^2_3 = 3.2$ ,  $p = 0.37$ ).

# Warstwowy model PH (1)

- ◆ W niektórych sytuacjach założenie PH może nie być spełnione dla wszystkich zmiennych niezależnych
- ◆ Można wówczas użyć modelu PH dla zmiennych, dla których założenie jest spełnione, przy odpowiedniej modyfikacji bazowych funkcji hazardu

## Warstwowy model PH (2)

- ◆ Załóżmy PH dla zmiennej binarnej  $Y$
- ◆ Przyjmijmy, że założenie nie jest spełnione dla  $X$
- ◆ Możemy wówczas zdefiniować następujący model:

$$\lambda(t) = \lambda_1(t) \cdot e^{\beta \cdot Y} \quad \text{if } X=x_1$$

$$\lambda(t) = \lambda_2(t) \cdot e^{\beta \cdot Y} \quad \text{if } X=x_2$$

- ◆ Jest to **model warstwowy**:
  - bazowe funkcje hazardu są różne dla różnych *warstw*, zdefiniowanych wartościami zmiennej niezależnej  $X$ , ale
  - efekt  $Y$  jest nadal wyrażony jako proporcjonalna zmiana funkcji hazardu (dla ustalonej wartości  $X$ )

# Interpretacja warstwowego modelu PH

$$\lambda(t) = \lambda_1(t) \cdot e^{\beta \cdot Y} \quad \text{jeśli } X=x_1$$

$$\lambda(t) = \lambda_2(t) \cdot e^{\beta \cdot Y} \quad \text{jeśli } X=x_2$$

- ◆ Gdy  $X=x_1$ , wzrost  $Y$  o 1 zmienia funkcję hazardu  $e^\beta$  razy. Podobnie dla  $X=x_2$ .

- ◆ Ale iloraz hazardów dla  $x_2$  vs  $x_1$  wynosi

$$\{\lambda_2(t) \cdot e^{\beta \cdot Y}\} / \{\lambda_1(t) \cdot e^{\beta \cdot Y}\} = \lambda_2(t) / \lambda_1(t)$$

- ◆ Wniosek: zależy od  $t$  !

# Szacowanie parametrów warstwowego modelu PH

- ◆ Częściowa funkcja wiarygodności (s indeksuje warstwy):

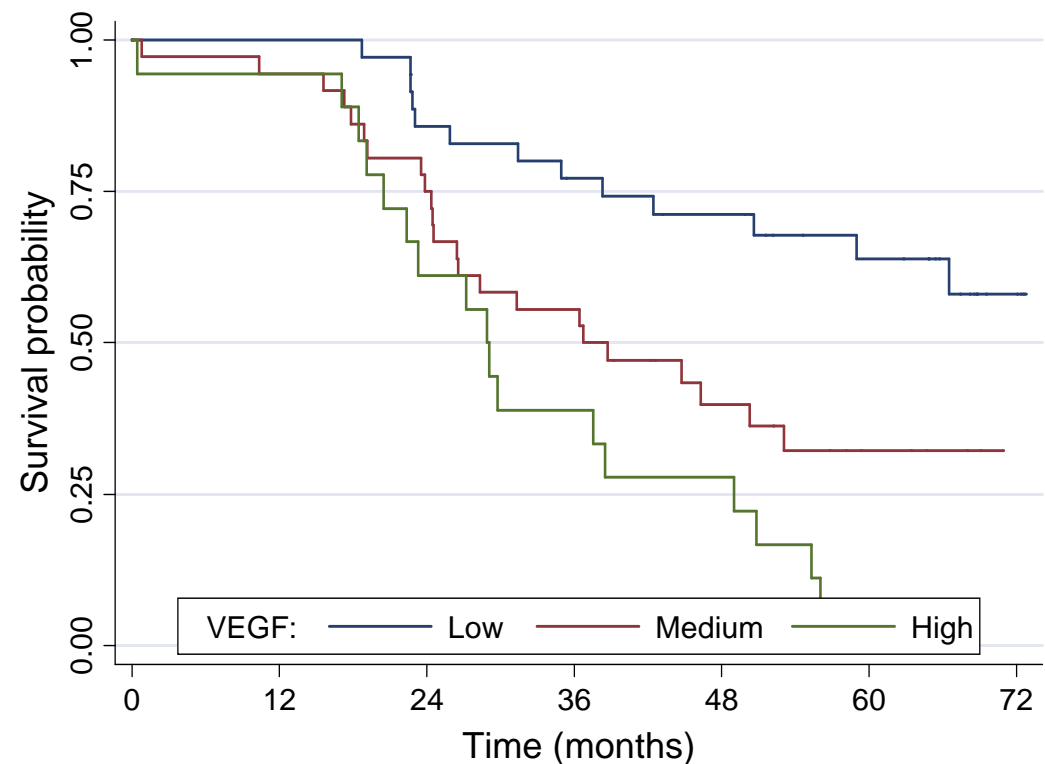
$$L(\beta) = \prod_{s=1}^S \prod_{k=1}^{K_s} \frac{\lambda_{0s}(t_{(k)}) h_k e^{\beta' Z_{(k)}(t_{(k)})}}{\sum_{l=1}^{n_s} Y_l(t_{(k)}) \lambda_{0s}(t_{(k)}) h_k e^{\beta' Z_l(t_{(k)})}} = \prod_{s=1}^S \prod_{k=1}^{K_s} \frac{e^{\beta' Z_{(k)}(t_{(k)})}}{\sum_{l=1}^{n_s} Y_l(t_{(k)}) e^{\beta' Z_l(t_{(k)})}}$$

gdzie  $K_s$  – liczba czasów zdarzeń w warstwie  $s$ ,  $n_s$  – liczba obserwacji w warstwie  $s$

- ◆ Oszacowanie  $\beta$  uzyskujemy rozwiązując  $U(\hat{\beta}) = 0$

# Warstwowy model PH: przykład (1)

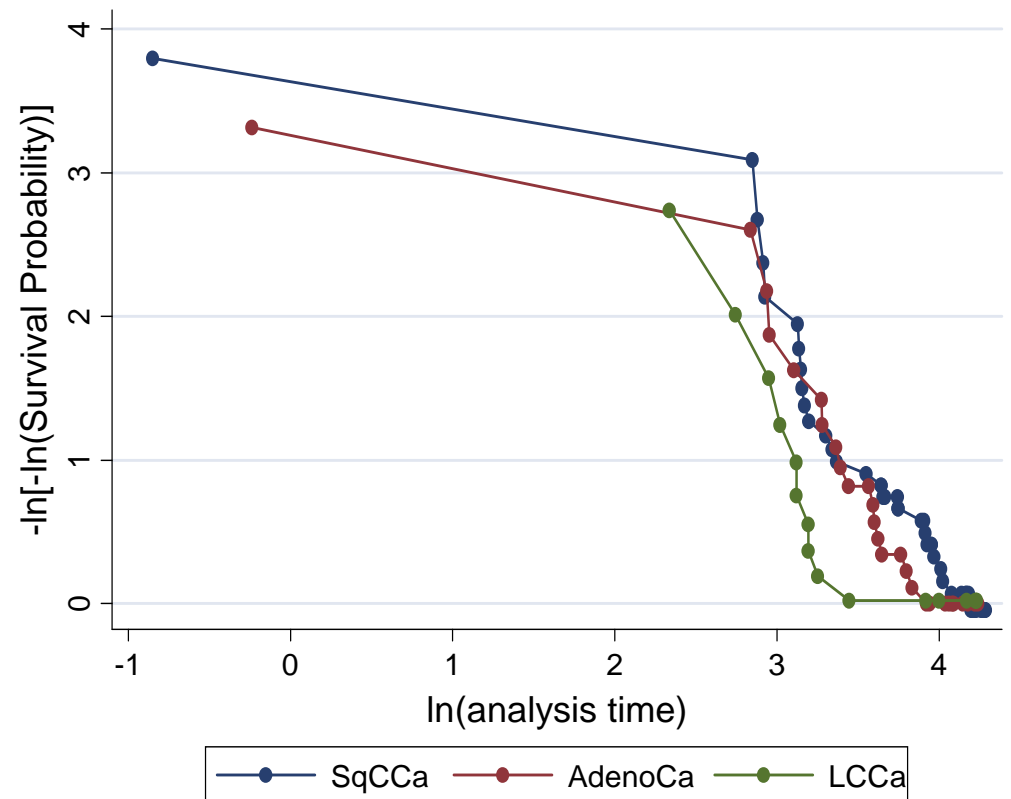
- ◆ Niklińska *et al.*, *Lung Cancer* (2001)
- ◆ Ekspresja czynnika wzrostu VEGF
- ◆ Wpływ na przeżycie 89 chorych leczonych operacyjnie
- ◆ Test logrank (dla trendu):  
 $p < 0.001 \rightarrow$



# Wastwowy model PH: przykład (2)

- ♦ Trzy typy histpat: SqCCa, AdenoCa, LCCa
- ♦ Graficzna ocena założenia proporcjonalnych hazardów:

wykresy powinny być równoległe →



# (Nie-)Warstwowy model PH: przykład (3)

## ◆ Model ze wieloma zmiennymi niezależnymi:

- „vegf” = 1, 2, 3 dla niskiej, średniej, i wysokiej ekspresji VEGF
- „p53m” = 0 jeśli brak mutacji P53, 1 jeśli mutacja
- „sex” = 0 dla kobiet, 1 dla mężczyzn
- „tnm” = 1, 2, 3 dla TNM I, II and IIIA
- „adCC” = 0 dla SqCCa/LCCa, 1 dla AdenoCa
- „lcc” = 0 dla SqCCa/AdenoCa, 1 dla LCCa

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
vegf	1.428507	.3145729	1.619	0.105	.9277663	2.199512
p53m	2.767612	1.039047	2.712	0.007	1.325983	5.776602
sex	1.114986	.471841	0.257	0.797	.4864711	2.555535
tnm	3.208635	.8751451	4.274	0.000	1.88	5.476245
adCC	1.699829	.617637	1.460	0.144	.8339101	3.464906
lcc	3.330073	1.384905	2.893	0.004	1.473864	7.524021



## Warstwowy model PH: przykład (4)

- ◆ Warstwowy model z warstwami zdefiniowanymi typem histopatologicznym:

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
vegf	1.548393	.3493506	1.938	0.053	.99502	2.409519
p53m	2.738736	1.061083	2.600	0.009	1.281636	5.852422
sex	1.00601	.4357343	0.014	0.989	.4304479	2.351171
tnm	3.095804	.8460452	4.135	0.000	1.811968	5.289278

Stratified by histpat

- ◆ Efekt typu histpat nie jest oszacowany
- ◆ Jest uwzględniony poprzez bazowe funkcje hazardu

# Warstwowanie i liczebność próbki (1)

- ◆ Przyjmijmy  $\beta=0$  i  $V(\beta, t) \equiv V$  dla ważonej wariancji  $Z$
- ◆ Oczekiwana wartość oszacowania wariancji wynosi  $V(n-1)/n$
- ◆ Załóżmy brak jednoczesnych zdarzeń i cenzurowania dla  $n=10$
- ◆ Oczekiwana wartość informacji wynosi
$$E\{I(\beta)\} = V(9/10 + 8/9 + \dots + 1/2 + 0/1)$$
  - $I^{-1}$  daje wariancję oszacowania  $\beta$
- ◆ Niech  $n=200$ . Wówczas
$$E\{I(\beta)\} = V(199/200 + 198/199 + \dots + 1/2 + 0/1) \approx 194.1 V$$
- ◆ Rozważmy 4 warstwy, każdą z 50 zdarzeniami. Wówczas
$$E\{I(\beta)\} = 4V(49/50 + 48/49 + \dots + 1/2 + 0/1) \approx 182 V$$
  - $(194.1 - 182) / 194.1 = 6\%$  strata efektywności

# Warstwowanie i liczebność próbki (2)

- ◆ Dla  $n=200$ ,  $E\{I(\beta)\} \approx 194.1 \text{ V}$
- ◆ Dla 4 warstw z 50 zdarzeniami,  $E\{I(\beta)\} \approx 182 \text{ V}$ 
  - $(194.1 - 182) / 194.1 = 6\%$  strata efektywności
- ◆ Dla 50 warstw z 4 zdarzeniami każda,  
 $E\{I(\beta)\} = 50V(3/4 + 2/3 + 1/2 + 0/1) \approx 95.8 \text{ V}$ 
  - $(194.1 - 95.8) / 194.1 = 51\%$  strata efektywności

## Warstwowanie i liczebność próbki (3)

- ◆ Załóżmy  $d=200$  dla  $n=1400$  i obserwacje cenzurowane > czas ostatniego zdarzenia
- ◆ Wówczas

$$E\{I(\beta)\} = V \sum_{i=1}^{200} \frac{1200 + i - 1}{1200 + i} \approx 199.85V$$

- ◆ Dla 50 równolicznych warstw z 4 zdarzeniami każda:

$$E\{I(\beta)\} = 50V \sum_{i=1}^4 \frac{24 + i - 1}{24 + i} = \frac{27}{28} + \frac{26}{27} + \frac{25}{26} + \frac{24}{25} \approx 192.44V$$

- ◆ Strata efektywności:  $(1 - 192.44/199.85) = 3.7\%$

## 4. Diagnostyka dla modelu proporcjonalnych hazardów

Reszty

Ocena graficzna dopasowania modelu

Test założenia proporcjonalnych hazardów

# Model PH: założenia

$$\lambda(t) = \lambda_0(t) \cdot e^{\alpha \cdot X + \beta \cdot Y}$$

## ◆ Główne założenia:

- niezależność obserwacji
- nie-informatywne cenzurowanie
- proporcjonalność hazardów (PH)
- postać funkcjonalna efektu zmiennej niezależnej

## Zmienne zależne od czasu w modelu Coxa (1)

$$\lambda(t) = \lambda_0(t) \cdot e^{\alpha \cdot X + \beta \cdot Z(t)}$$

- ♦ Interpretacja: jeśli w chwili czasu  $t$  dla jednostki A jest  $Z(t)=z$ , a dla B jest  $Z(t)=z+1$  (a wartości pozostałych zmiennych niezależnych są równe), to

$$\lambda_B(t)/\lambda_A(t) = e^{\beta \cdot (z+1) - \beta \cdot z} = e^{\beta \cdot z + \beta - \beta \cdot z} = e^{\beta}$$

- ♦ W rezultacie,  $e^{\beta}$  jest *ilorazem hazardu* związanym ze wzrostem  $Z(t)$  o 1 w dowolnej chwili czasu  $t$

## Zmienne zależne od czasu w modelu Coxa (2)

- ♦ Zmienne „zewnątrzne”: nie związane z obserwowaną jednostką, o wartościach znanych z góry (np. wiek, planowane dawki leku, itd.)
- ♦ Zmienne „wewnętrzne”: mierzalne tylko gdy jednostka pozostaje pod obserwacją; np. wielokrotne pomiary ciśnienia krwi chorego
  - Wymagają ostrożności przy budowie modelu. Ich uwzględnienie może np. spowodować błąd w oszacowaniu efektu zmiennej, którą jesteśmy zainteresowani.



# Współczynniki zależne od czasu w modelu Coxa

- ◆ Możliwe jest również założenie, że *efekt* zmiennej  $Z$  zależy od czasu:

$$\lambda(t) = \lambda_0(t) \cdot e^{\alpha \cdot X + \beta(t) \cdot Z}$$

- ◆ Dla np.  $\beta(t) = \beta + \beta_T \cdot t$  zachodzi

$$\lambda(t) = \lambda_0(t) e^{\alpha \cdot X + \beta \cdot Z + \beta_T \cdot t \cdot Z} = \lambda_0(t) \cdot e^{\alpha \cdot X + \beta \cdot Z + \beta_T \cdot \underbrace{Z \cdot t}_{Z(t)}} = \lambda_0(t) \cdot e^{\alpha \cdot X + \beta \cdot Z + \beta_T \cdot Z(t)}$$

Przyjmując  $Z(t) = Z \cdot t$

- możemy traktować ten model jako model ze zmienną zależną od czasu
- ale efekt zmiennej może być trudny do przedstawienia i interpretacji

# Proste metody oceny założenia PH (1)

- ◆ Iloraz hazardów dla dwóch wartości zmiennej niezależnej  $Z$  jest stały w czasie
- ◆ Proste metody graficzne oceny założenia:
  - Jeśli krzywe przeżycia się krzyżują, funkcje hazardu nie mogą być proporcjonalne
  - Zgodnie z modelem,

$$\ln \{ -\ln S(t) \} = \ln \{ -\ln S_0(t) \} + \beta \cdot Z .$$

Wykres  $\ln\{-\ln S(t)\}$  w funkcji czasu lub jego logarytmu powinny mieć postać równoległych krzywych

## Proste metody oceny założenia PH (2)

- ◆ Metody graficzne mogą być trudne w interpretacji
- ◆ Formalne testy mogą być bardziej przydatne

- ◆ Rozważmy model

$$\lambda(t) = \lambda_0(t) \cdot e^{\beta \cdot Z + \beta_T \cdot Z \cdot t}$$

- można użyć innych funkcji czasu, np.  $\ln t$

- ◆ Dla  $Z=z$  i  $Z=z+1$  otrzymujemy

$$\lambda_{z+1}(t) / \lambda_z(t) = e^{\beta(z+1) + \beta_T(z+1) \cdot t - \beta z - \beta_T z \cdot t} = e^{\beta + \beta_T \cdot t}$$

- ◆ Iloraz hazardów zależy od czasu! Ale tylko jeśli  $\beta_T \neq 0 \dots$
- ◆ ... więc można użyć modelu do testu czy  $\beta_T = 0$

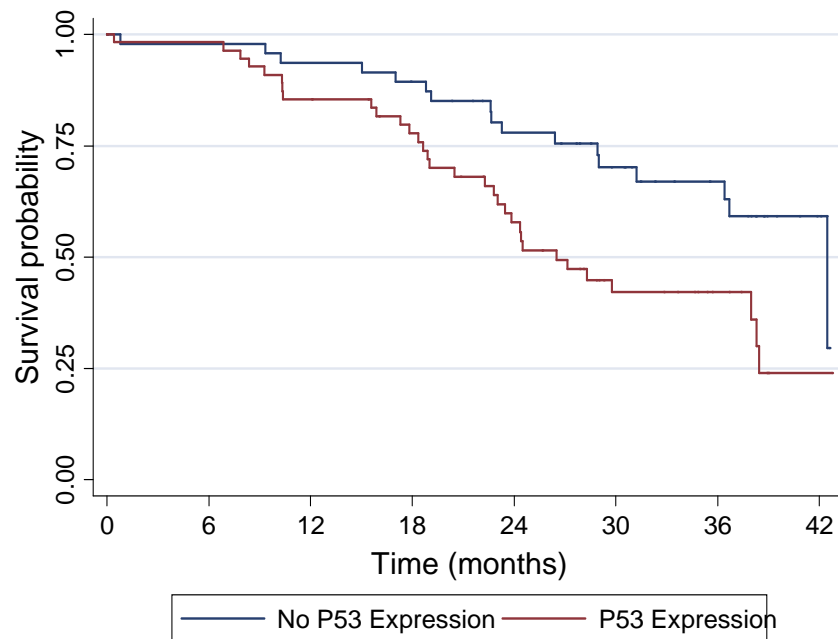
# Model PH: NSCLC, ekspresja P53

- ♦ Zmienna „P53e” = 0 jeśli brak ekspresji białka, 1 jeśli ekspresja

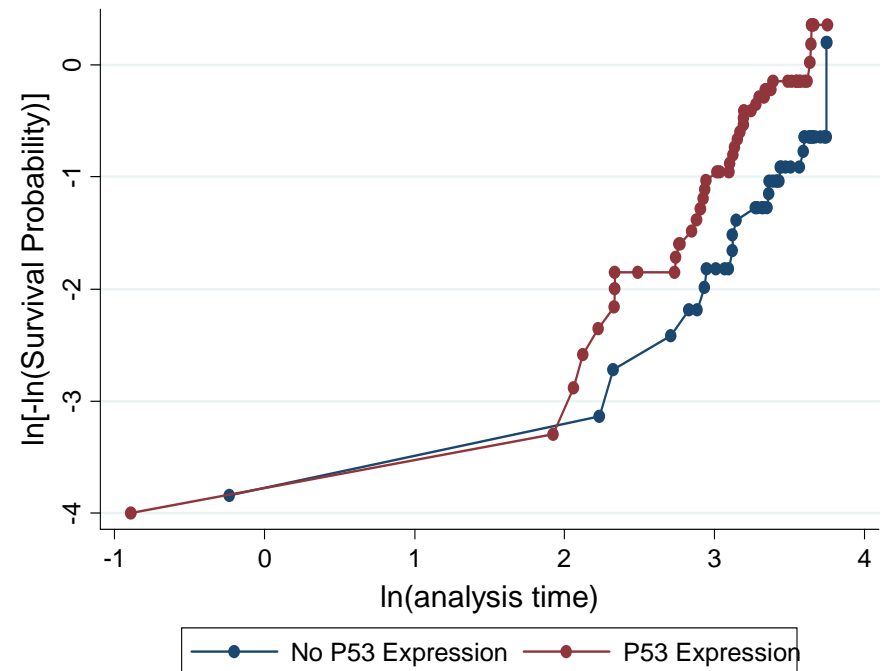
_t	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
P53e	.7859412	.3021812	2.601	0.009	.1936769	1.378205
	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
P53e	2.194471	.663128	2.601	0.009	1.213704	3.967775

# Założenie PH: NSCLC (1)

- ◆ Krzywe przeżycia się nie przecinają



- ◆ Równoległe krzywe?



# Założenie PH: NSCLC (2)

- ♦ Zmienna „P53e” = 0 jeśli brak ekspresji białka, 1 jeśli ekspresja
- ♦ „P53e · time” dodana do modelu

		_t	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
rh		P53e	.9412935	.7483608	1.26	0.208	-.5254666 2.408054
t		P53e	-.0071451	.0313378	-0.23	0.820	-.068566 .0542757

Note: Second equation contains variables that continuously vary with respect to time; variables are interacted with current values of \_t.

- ♦ Wniosek: test dla  $\beta_T = 0$  nie jest istotny statystycznie na poziomie istotności 5%
- ♦ Założenie może być przyjęte dla P53e

# Ocena założeń przy użyciu reszt

- ◆ Dla modelu PH można zdefiniować różne rodzaje reszt:
  - reszty martyngałowe
  - reszty oparte na dewiancji
  - (skalowane) reszty „score”
  - (skalowane) reszty Schoenfelda
  - ...
- ◆ Są używane do oceny różnych aspektów dobroci dopasowania modelu na podstawie wykresów
- ◆ Wykresy nie są łatwe w interpretacji

# Przydatne wykresy reszt

- ◆ Reszty martyngałowe/dewiancji vs lin. komb. zmiennych
  - duże reszty sygnalizują problem z dopasowaniem modelu do danych
- ◆ Reszty martyngałowe vs ciągła zmienna niezależna
  - wygładzony wykres może sugerować postać funkcjonalną zmiennej
- ◆ Reszty Schoenfelda dla zmiennej vs czas
  - wygładzony wykres może być przydatny do oceny założenia PH



# Reszty martyngałowe

$$M(t) = N(t) - \int_0^t Y(u) \lambda_0(u) e^{\beta^T Z(u)} du$$

- ♦  $N(t)$  to proces zliczający zdarzenia,  $Y(t)$  to proces zbioru ekspozycji na ryzyko zdarzenia

- ♦ W szczególności, dla  $t = \infty$  i zmiennych niezależnych stałych w czasie

$$\hat{M}_i = \delta_i - \hat{\Lambda}_0(t_i) \exp(\hat{\beta}^T Z_i) \equiv \delta_i - \hat{E}_i$$

gdzie  $\delta_i$  i  $t_i$  to wskaźnik zdarzenia i czas obserwacji dla  $i$ -tej jednostki

- $M(t)$  = obserwowane – oczekiwane zdarzenia dla chwili czasu  $t$

- ♦ Własności

- $E(M_i) = 0$
- $\text{Cov}(M_i, M_j) = 0$
- $\sum \hat{M}_i = 0$
- $\text{Cov}(\hat{M}_i, \hat{M}_j) < 0$

# Reszty oparte na dewiancji

- ◆ Dla zmiennych niezależnych stałych w czasie

$$\hat{r}_{Di} = \text{sign}(\hat{M}_i) \sqrt{-\hat{M}_i - \delta_i \log(\delta_i - \hat{M}_i)}$$

- Motywacja:  $2\{l(\hat{\beta}^*) - l(\hat{\beta})\} = \sum_i \hat{r}_{Di}^2$

- ◆ Szereg Taylora pierwszego rzędu daje  $\hat{r}_{Di} \approx \frac{\delta_i - \hat{E}_i}{\sqrt{\hat{E}_i}}$ 
  - $\approx$  reszta Pearsona

- ◆ Reszty oparte na dewiancji mają bardziej symetryczny rozkład niż  $M_i$

# Reszty martyngałowe i dewiancji: użyteczne wykresy

- ◆ Reszty martyngałowe/dewiancji vs lin. komb. zmiennych
  - duże reszty sygnalizują problem z dopasowaniem modelu do danych
- ◆ Reszty martyngałowe vs ciągła zmienna niezależna
  - wygładzony wykres może sugerować postać funkcjonalną zmiennej

# Reszty martyngałowe i dewiencji: sumaryczna ocena dopasowania

- Wykres może ujawnić obserwacje, do których model nie pasuje (nie w przykładzie)

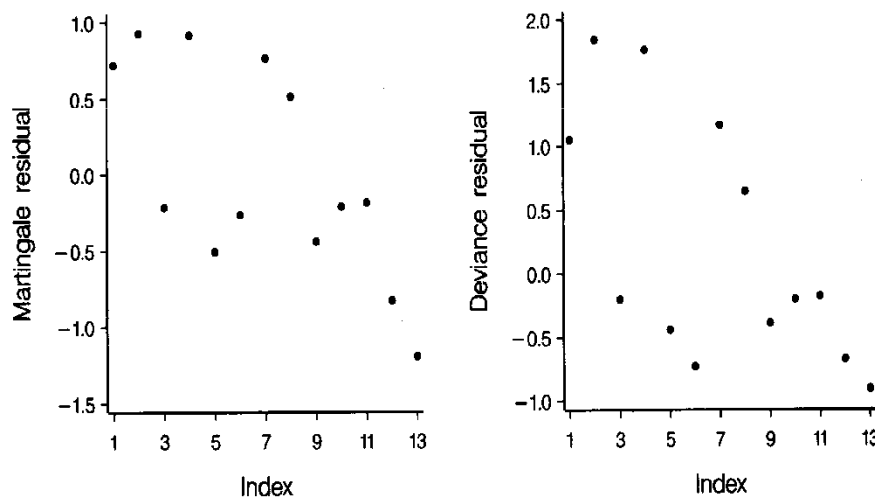


Figure 4.2 Index plots of the martingale and deviance residuals.

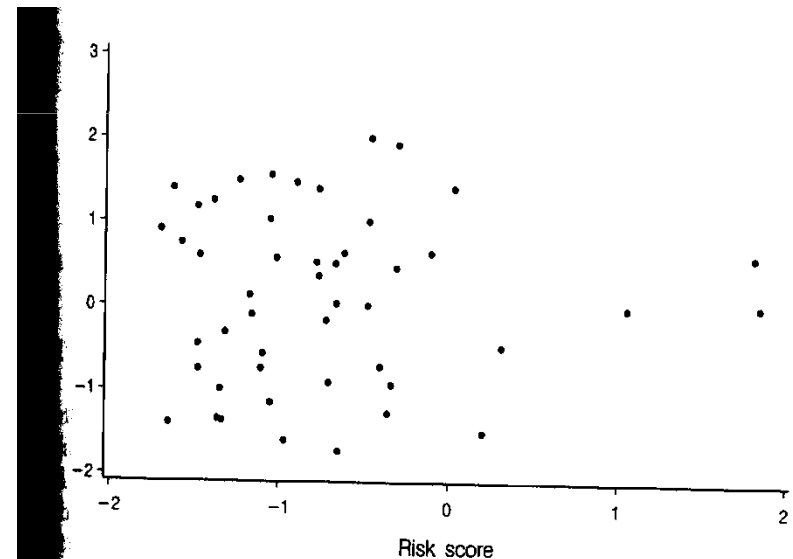


Figure 4.6 Deviance residuals plotted against the risk score.

# Postać funkcjonalna ciągłej zmiennej niezależnej stałej w czasie (1)

$$\lambda(t) = \lambda_0(t) \cdot e^{\beta \cdot f(Z)}$$

## ♦ Jaka $f()$ ?

- Dopasowujemy model „zerowego” (z wartościami  $\beta$  równymi 0)
- Sporządzamy wykres reszt martyngałowych dla danej zmiennej
- „Wygładzamy” wykres dla ustalenia  $f()$

## ♦ Przykład: punkt odcięcia (zmienna binarna) →

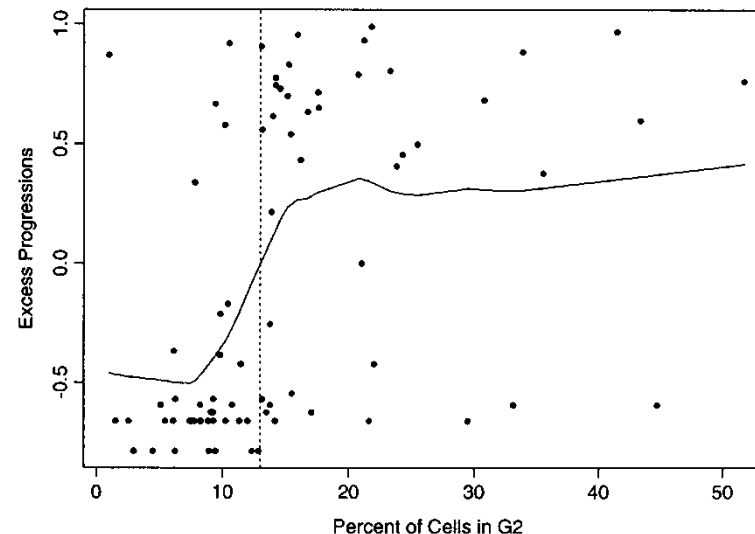


FIGURE 5.1: Stage D1 prostate cancer

# Postać funkcjonalna ciągłej zmiennej niezależnej stałej w czasie (2)

◆ Przykład: trend liniowy i logarytmiczny

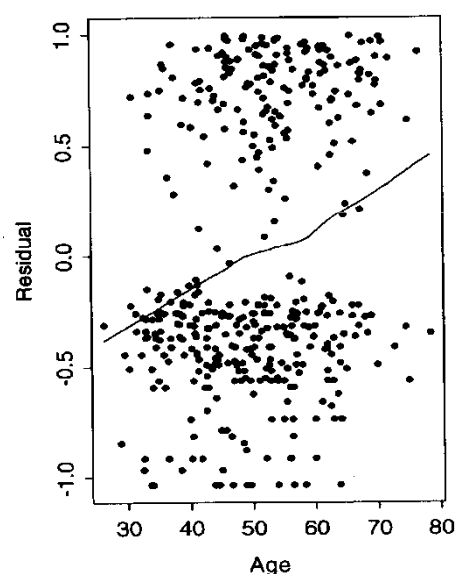


FIGURE 5.3: PBC data, residual plots for age and bilirubin

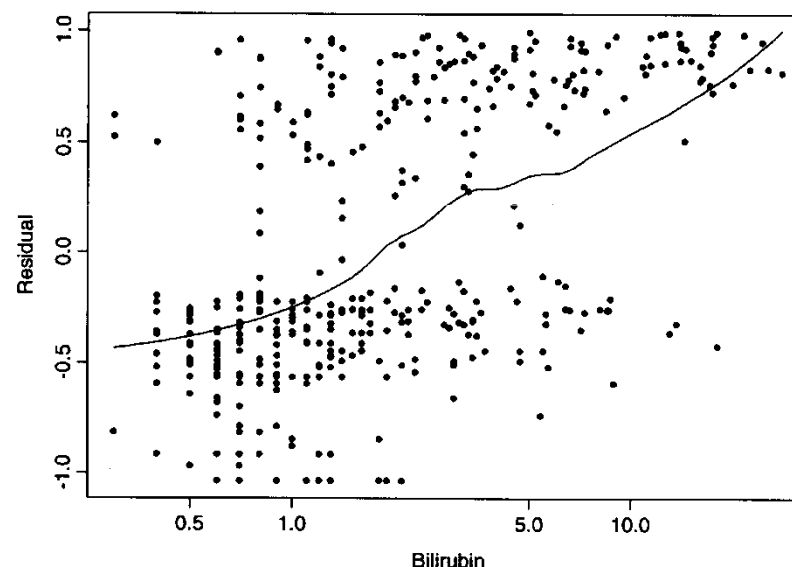
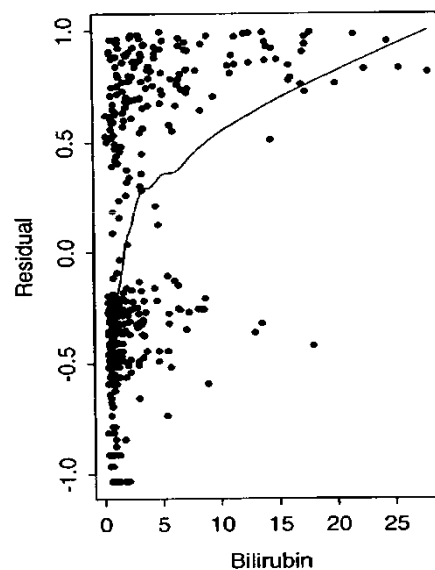


FIGURE 5.4: PBC data, logarithmic plot for bilirubin

◆ Zauważalne „smugi” wynikające z cenzurowania

# Postać funkcjonalna ciągłej zmiennej niezależnej stałej w czasie: model Poissona (1)

- ◆ Przyjmijmy  $\lambda(t) = \lambda_0(t) \cdot e^{\beta f(Z)}$
- ◆ Wówczas  $E(\delta_i|z) = \lambda_0(t_i) \cdot e^{\beta f(z)}$  oraz  $\ln E(\delta_i|z) = \ln \lambda_0(t_i) + \beta f(z)$ 
  - Gdybyśmy znali  $\lambda_0(t_i)$ , byłby to model regresji Poissona
- ◆ Idea: oszacujmy  $\lambda_0(t)$ , potraktujmy wskaźnik zdarzeń jako zliczenie i zastosujmy model regresji Poissona z gładką  $f(z)$ 
  - można użyć modelu PH z oryginalnymi zmiennymi niezależnymi
- ◆ Postać funkcjonalna przyjęta na podstawie wykresu  $f(z)$

# Postać funkcjonalna ciągłej zmiennej niezależnej stałej w czasie: model Poissona (2)

- ◆ Trend logarytmiczny dla bilirubiny, liniowy dla wieku i albuminy

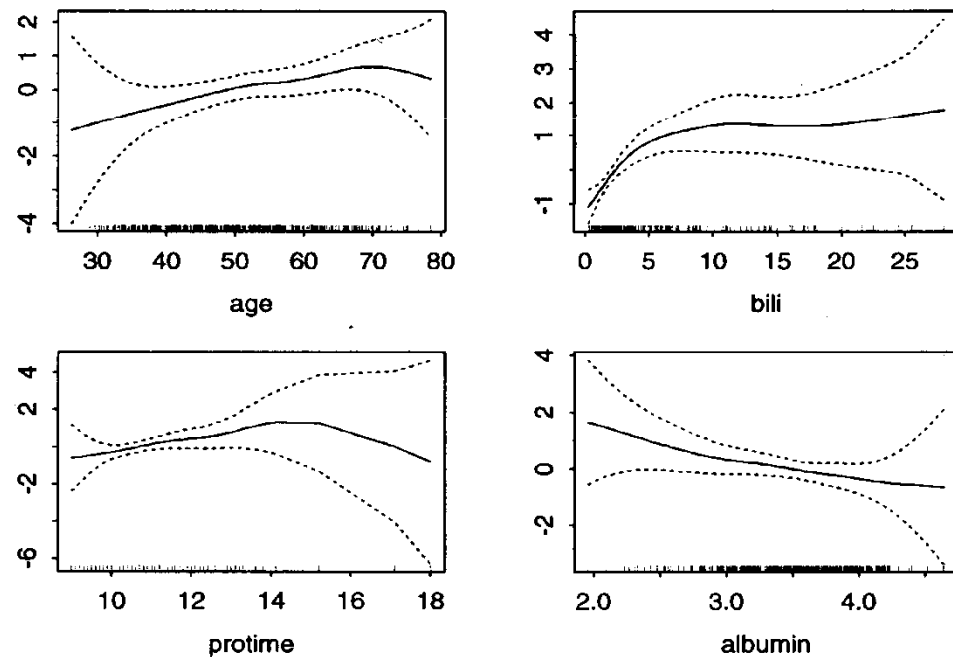


FIGURE 5.11: *PBC Data, functional form using Poisson approach*



# Elastyczne postaci funkcjonalne ciągłej zmiennej niezależnej stałej w czasie

♦ Zamiast „znajdować” postać funkcjonalną, można próbować ją modelować bezpośrednio

- wielomiany „cząstkowe” (fractional polynomials)
- spline’y regresyjne (liniowe, kwadratowe, kubiczne, ...)
- spline’y wygładzające

♦ Spline’y są „lokalne” i dlatego bardziej użyteczne

- spline’y regresyjne wymagają wyboru węzłów
- w spline’ach wygładzających stopień wygładzania może być wybrany „automatycznie”

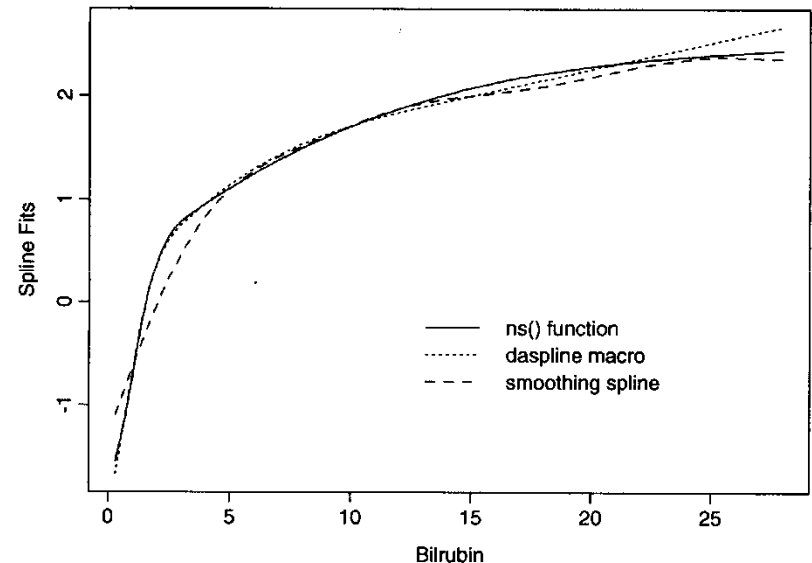


FIGURE 5.14: *Spline fits for bilirubin*

# Reszty „score”

- ◆ *Score process* dla obserwacji  $i$   $U_i(\beta, t) = \int_0^t \{Z_i(u) - \bar{Z}(\beta, u)\}^T dM_i(u)$

- wektor !

- ◆ Dla  $k$ -tego czasu zdarzenia oraz zmiennej  $j$ :

$$U_{ijk}(\beta) = \int_{t_{(k-1)}}^{t_{(k)}} \{Z_{ij}(u) - \bar{Z}_j(\beta, u)\} dM_i(u)$$

- ◆ *Reszta „score”* dla obserwacji  $i$  i zmiennej  $j$ :  $\hat{r}_{Uij} = U_{ij}(\hat{\beta}, \infty)$

- Macierz reszt,  $\mathbf{U}$ , rozmiaru  $n \times p$
- $\sum_i U_{ij} = 0$  (score equation)

# Skalowane reszty „score”

$$\hat{r}_{Ui}^* = \hat{r}_{Ui} I(\hat{\beta})^{-1} = \hat{r}_{Ui} \text{var}(\hat{\beta})$$

◆ Zachodzi  $\hat{r}_{Uij}^* \approx \hat{\beta}_j - \hat{\beta}_{j(i)}$

gdzie  $\hat{\beta}_{j(i)}$  to współczynnik uzyskany dla  $j$ -tej zmiennej po usunięciu  $i$ -tej obserwacji (metoda „syczorykowa”)

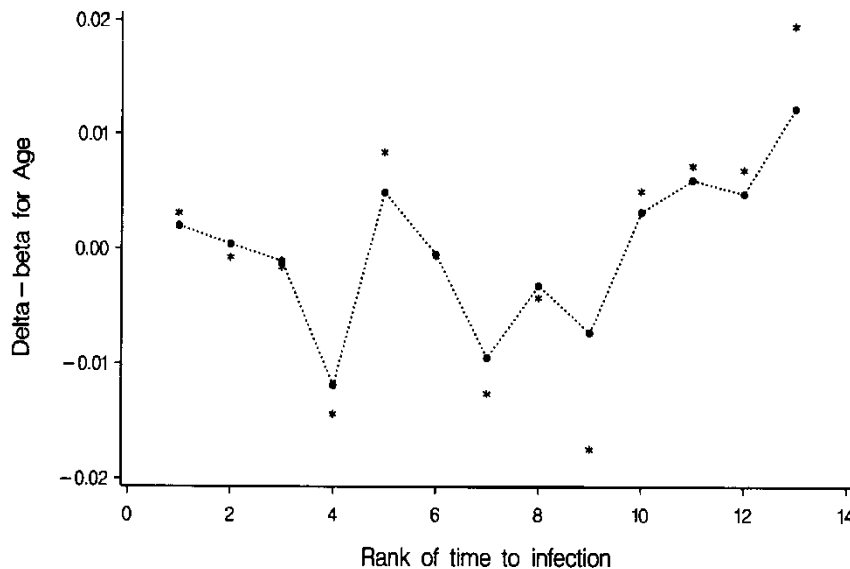
◆ Użyteczne w diagnostyce wpływu obserwacji na model

◆ Ponadto,  $(U^*)^T U^* = \text{var}(\hat{\beta})(U^T U) \text{var}(\hat{\beta})$

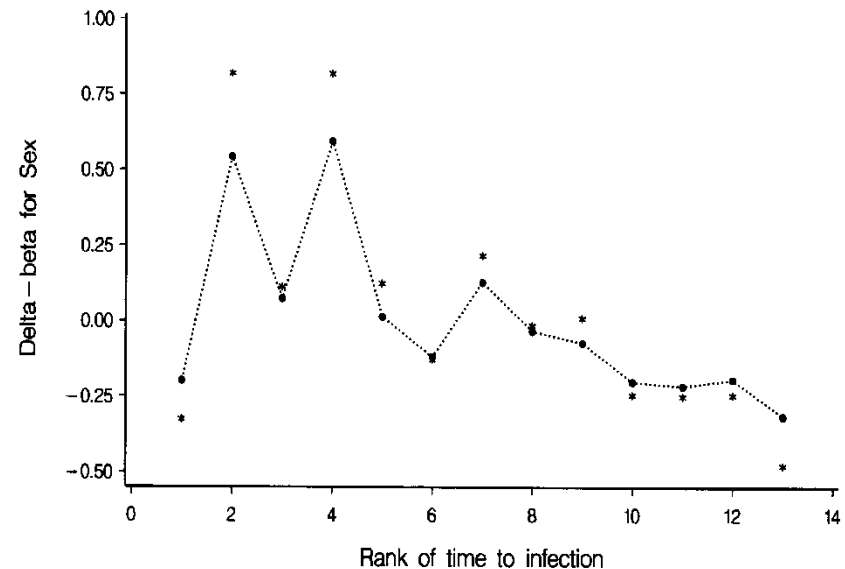
jest „odpornym” estymatorem wariancji

# Skalowane reszty „score”: diagnostyka wpływu na model (1)

- ◆ Nieszacowanie dużych wartości
  - zmiany w wariancji oszacowań współczynników ignorowane
- ◆ Porównujemy wartości z oszacowanymi błędami standardowymi
- ◆ Wykres w funkcji rang czasu może ujawnić związek między czasem a wpływem obserwacji na model



**Figure 4.10** Plot of the exact (\*) and approximate (•) delta-betas for Age against rank order of time to infection.



**Figure 4.11** Plot of the exact (\*) and approximate (•) delta-betas for Sex against rank order of time to infection.

# Skalowane reszty „score”: diagnostyka wpływu obserwacji na model (2)

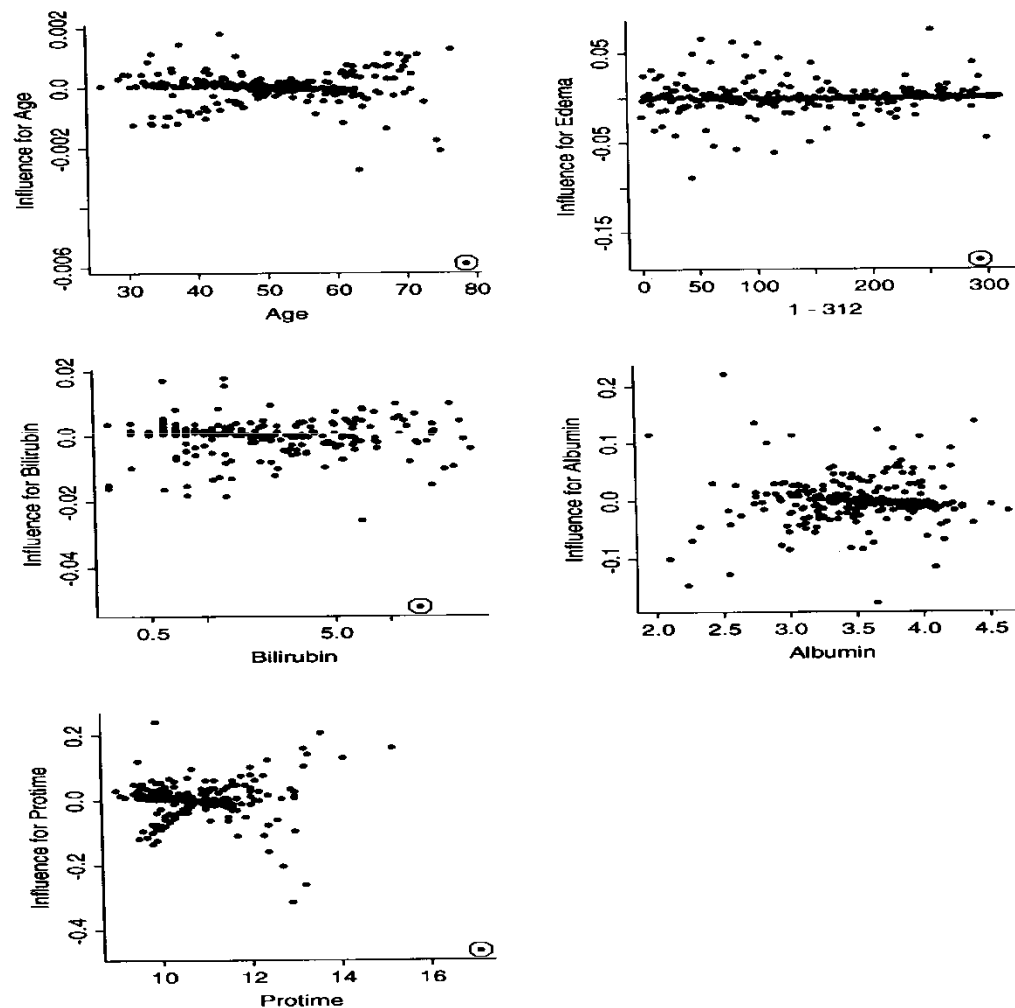


FIGURE 7.3: Influences for the PBC data set

# Reszty Schoenfelda (1)

- ◆ Definiowane dla czasów zdarzeń  $t_{(k)}$ :

$$\begin{aligned}\hat{r}_{Sk} &= \int_{t_{(k-1)}}^{t_{(k)}} \sum_i \left\{ Z_i(u) - \bar{Z}(\hat{\beta}, u) \right\}^T d\hat{M}_i(u) = \int_{t_{(k-1)}}^{t_{(k)}} \sum_i \left\{ Z_i(u) - \bar{Z}(\hat{\beta}, u) \right\}^T dN_i(u) \\ &= \sum_{i: t_i = t_{(k)}} \delta_i \left\{ Z_i(t_{(k)}) - \bar{Z}(\hat{\beta}, t_{(k)}) \right\}^T\end{aligned}$$

- wektor !
- Macierz reszt Schoenfelda **S** rozmiaru  $K \times p$

# Reszty Schoenfelda (2)

- ◆ Jeśli nie ma jednoczesnych zdarzeń  $\hat{r}_{Sk} = \left\{ Z_{(k)}(t_{(k)}) - \bar{Z}(\hat{\beta}, t_{(k)}) \right\}^T$ 
  - $\sum_k r_{kj} = 0$  (score equation)
  - asymptotycznie o średniej równej zero, nieskorelowane
- ◆ Dla  $d_k$  jednoczesnych zdarzeń, programy komp. definiują  $d_k$  osobnych reszt, jak wyżej

# Skalowane reszty Schoenfelda (1)

$$\hat{r}_{Sk}^* = \hat{r}_{Sk} I(\hat{\beta}, t_{(k)}) = \hat{r}_{Sk} \text{var}^{-1}(\hat{\beta}, t_{(k)})$$

♦ Niech  $\lambda(t) = \lambda_0(t) \cdot e^{\beta(t)'Z}$

♦ Wówczas  $E\left\{\left(\hat{r}_{Skj}^*\right)^T\right\} + \hat{\beta}_j \approx \beta_j(t_{(k)})$

gdzie  $\hat{\beta}_j$  jest oszacowaniem współczynnika dla zmiennej stałej w czasie

♦ Wykres  $\left(\hat{r}_{Skj}^*\right)^T + \hat{\beta}_j$  w funkcji czasu może ujawnić problem z PH



# Skalowane reszty Schoenfelda (2)

- ♦ W praktyce,  $V(\hat{\beta}, t_{(k)})$  może być niestabilne dla „późnych”  $t_{(k)}$
- ♦ Z drugiej strony,  $V(\hat{\beta}, t)$  zmienia się powoli
- ♦ Ponadto mamy  $\sum_k V(\hat{\beta}, t_{(k)}) = I(\hat{\beta})$

- ♦ W praktyce, zakłada się stałą  $V$  i używa średniej  $1/d$ , skąd

$$\hat{r}_{Sk}^* = \left( \sum_k d_k \right) \hat{r}_{Sk} \text{var}(\hat{\beta})^{-1} = d \cdot \hat{r}_{Sk} \text{var}(\hat{\beta})^{-1}$$

- Macierz skalowanych reszt Schoenfelda  $\mathbf{S}^*$  rozmiaru  $K \times p$

# Skalowane reszty Schoenfelda (3)

- ◆ Uwaga na modele warstwowe: założenie o stałej  $V$  może nie być spełnione dla warstw o mocno różnych  $Z$
- ◆ Przykład: interakcje zmiennych z warstwami
  - 4 szpitale jako warstwy,  $Z_i=1$  dla leczenia w szpitalu  $i$ , 0 w p.p.
  - Model:  $\lambda(t) = \lambda_{0i}(t) \cdot \exp(\beta_1 Z_1 + \beta_2 Z_2 + \beta_3 Z_3 + \beta_4 Z_4 + \dots)$
  - Problem:  $Z_1=0$  w warstwach 2, 3, 4, więc jej wariancja  $\equiv 0$
- ◆ „Ręczne” rozwiązanie
  1. Dopasuj model do danych
  2. Dopasuj model do każdej z warstw (użyj oszacowań  $\beta$  z pkt. 1 jako wartości startowych, liczba iteracji=0), wylicz skalowane reszty
  3. Użyj reszt do sporządzenia wykresów / testów

# Skalowane reszty Schoenfelda: współczynniki zależne od czasu

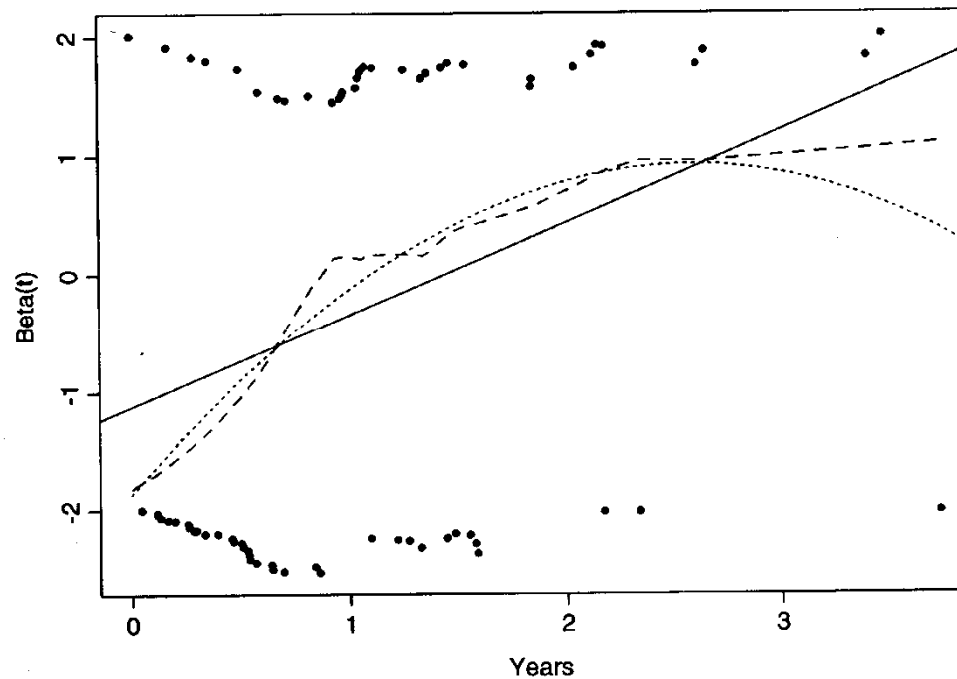


FIGURE 6.4: *Time-dependent coefficient plot for the gastric data*

- ◆ Dla modelu PH oczekujemy poziomej linii

# Wykresy skalowanych reszt Schoenfelda

- ◆ Istnieje szereg testów do testowania założenia PH
- ◆ Większość odpowiada wykresowi skalowanych reszt Schoenfelda dla pewnej funkcji czasu  $g(t)$

# Testy oparte na skalowanych resztach Schoenfelda (1)

- ◆ Ustalamy funkcję  $g(t)$ ; wektor  $g^T = (g(t_{(1)}), \dots, g(t_{(K)})) \equiv (g_1, \dots, g_K)$

- ◆ Test globalny  $\sim \chi^2_p$  
$$T = \frac{(g - \bar{g} \cdot 1_K)^T S^* var^{-1}(\hat{\beta}) S^{*T} (g - \bar{g} \cdot 1_K)}{d \sum_k (g_k - \bar{g})^2}$$

- ◆ Test dla jednej zmiennej 
$$T_j = \frac{\left\{ \sum_k (g_k - \bar{g}) \hat{r}_{skj}^{*T} \right\}^2}{d var(\hat{\beta}_j) \sum_k (g_k - \bar{g})^2}$$

- $\approx$  test dla współczynnika w prostej regresji liniowej dla  $r_{skj}^*$  vs.  $g(t_k)$

# Testy oparte na skalowanych resztach Schoenfelda (2)

- ◆ Oprócz wyniku testu warto zwrócić uwagę na wykres
  - $\beta(t)$  może być funkcją kwadratową, której może nie wykryć test dla funkcji liniowej
- ◆ Problemem jest również moc statystyczna testu

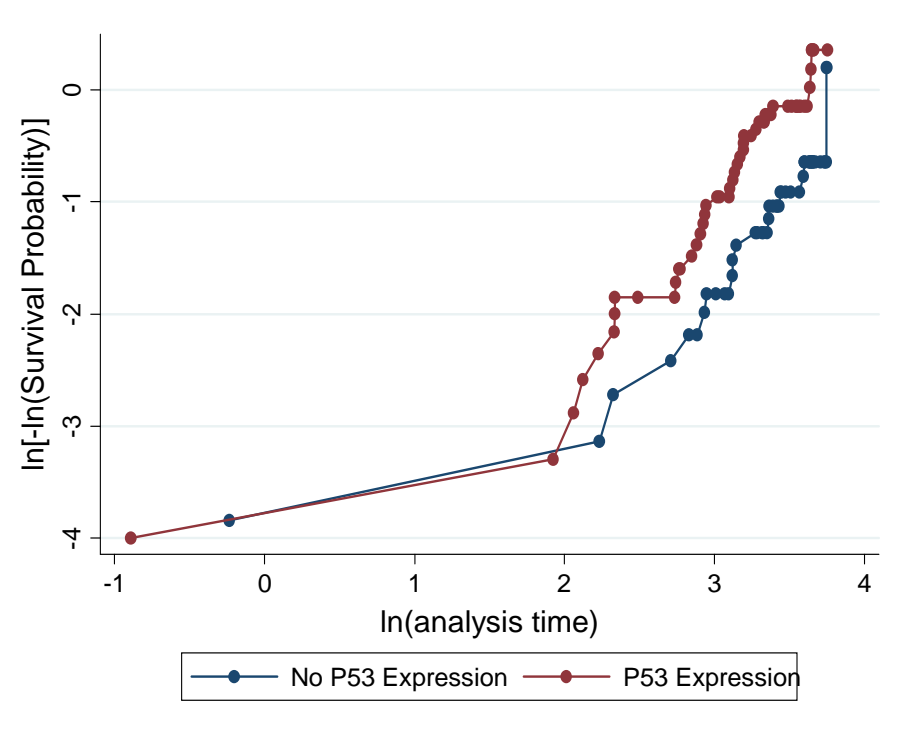
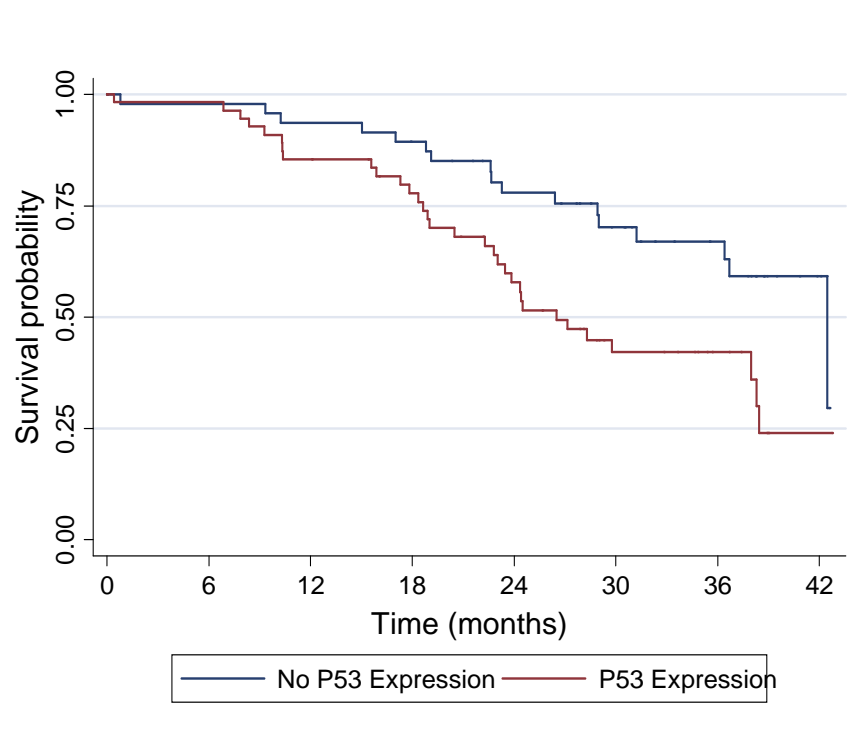
# Model PH: NSCLC, ekspresja P53

- ♦ Zmienna „P53e” = 0 jeśli brak ekspresji, 1 jeśli ekspresja

_t	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
P53e	.7859412	.3021812	2.601	0.009	.1936769	1.378205
	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
P53e	2.194471	.663128	2.601	0.009	1.213704	3.967775

# Założenie PH: NSCLC (1)

- ◆ Krzywe przeżycia się nie przecinają
- ◆ Równoległe?





# Założenie PH: NSCLC (2)

- ♦ Zmienna „P53e” = 0 jeśli brak ekspresji, 1 jeśli ekspresja
- ♦ „P53e · time” dodana do modelu

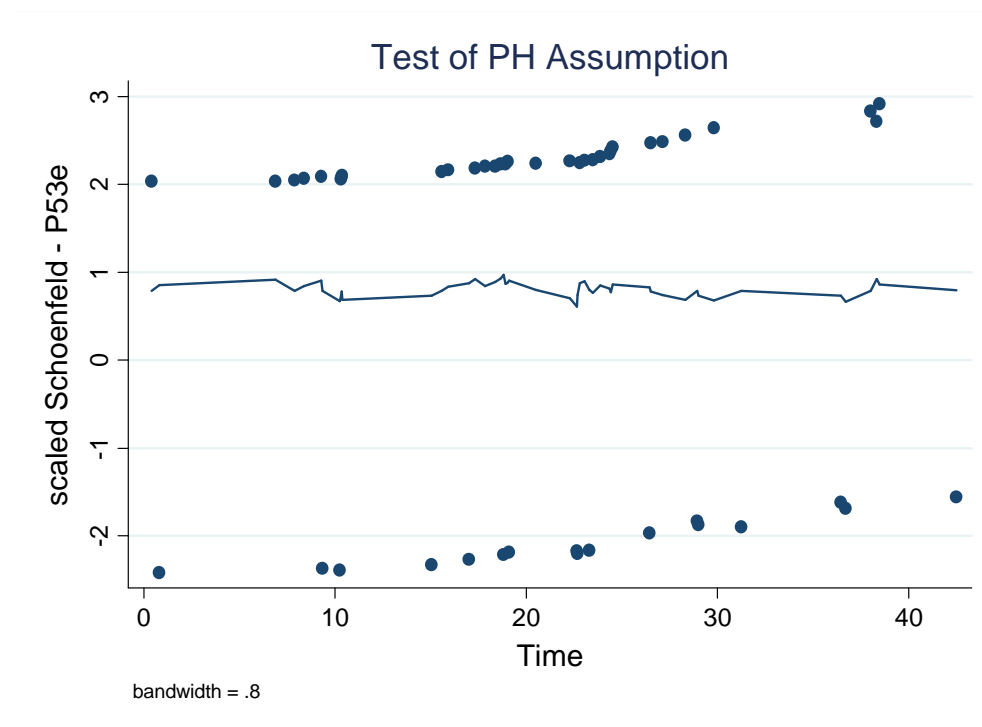
		_t	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
rh		P53e	.9412935	.7483608	1.26	0.208	-.5254666 2.408054
t		P53e	-.0071451	.0313378	-0.23	0.820	-.068566 .0542757

Note: Second equation contains variables that continuously vary with respect to time; variables are interacted with current values of \_t.

- ♦ Wniosek: test dla  $\beta_T = 0$  nie jest statystycznie istotny na poziomie istotności 5%
- ♦ Założenie PH dla ekspresji P53 wydaje się być spełnione

# Założenie PH: NSCLC (3)

- Wygładzony wykres skalowanych reszt Schoenfelda dla P53e



- Wygładzona krzywa bliska poziomej
- Dla testu Schoenfelda  $p=0.819$  – brak trendu w czasie

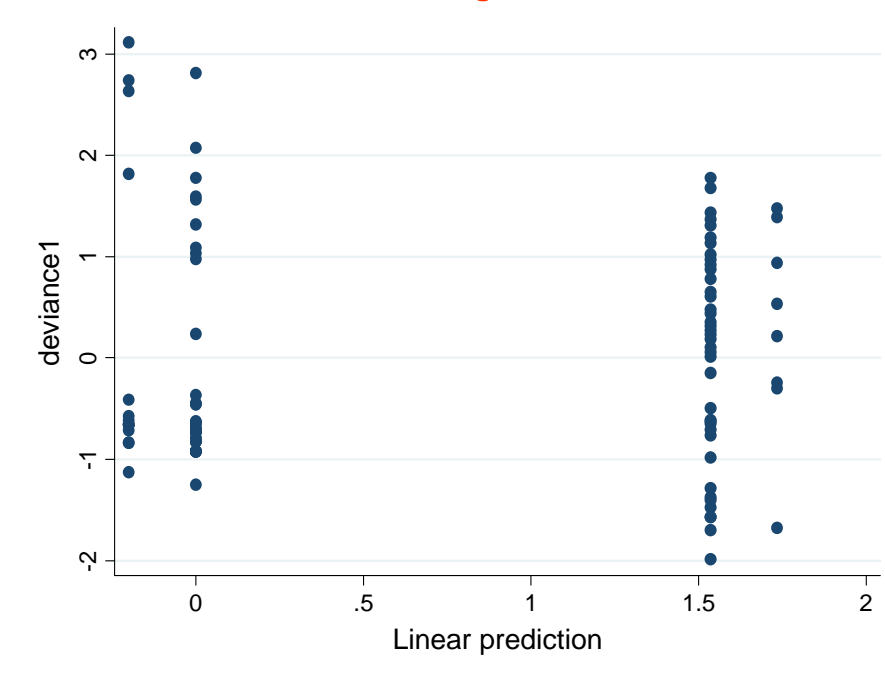
# Model PH: NSCLC, ekspresja i mutacja P53

- ♦ „P53m” = 0 jeśli brak mutacji, 1 jeśli mutacja
- ♦ „P53e” = 0 jeśli brak ekspresji, 1 jeśli ekspresja

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
P53m	5.670694	2.308899	4.262	0.000	2.553043	12.59547
P53e	.8191471	.3038457	-0.538	0.591	.3959371	1.694718

- ♦ Wniosek: skorygowanie ze względu na efekt mutacji „usuwa” efekt ekspresji białka P53 (współliniowość?)
- ♦ Ocena ekspresji białka nie „dodaje informacji” jeśli wiemy czy występuje mutacja

# Dopasowanie modelu: NSCLC, ekspresja i mutacja P53

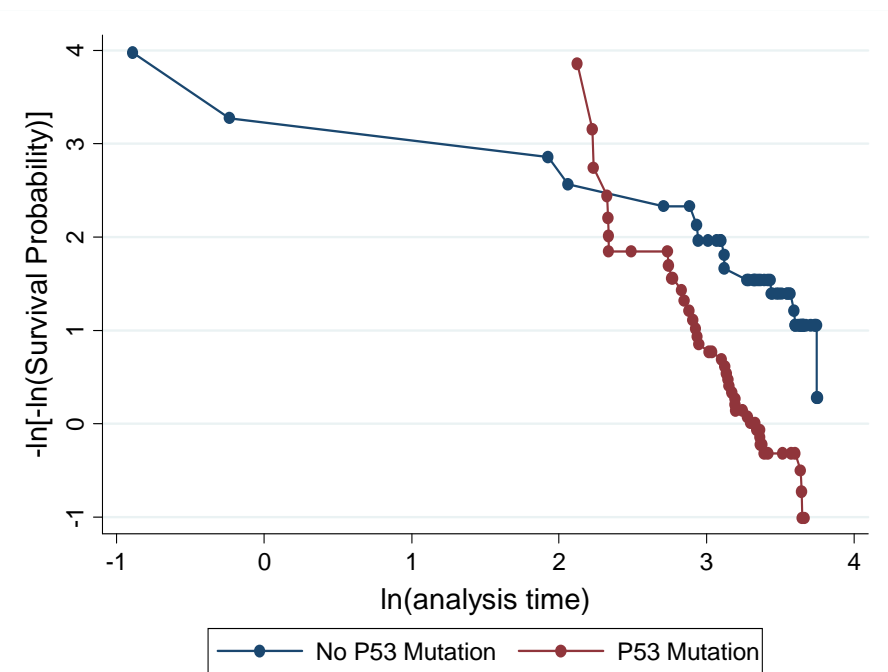
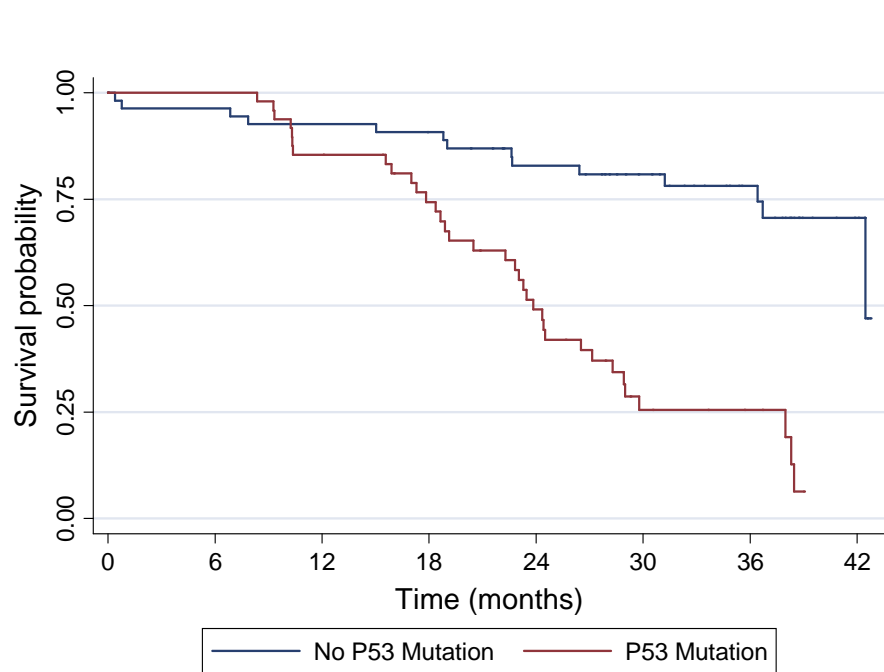


- ◆ Reszty oparte na dewiancji vs lin. kombinacja P53e i P53m
- ◆ Wykres trudny w interpretacji – tylko cztery grupy
- ◆ Powinien być symetryczny, bez wyraźnie odstających obserwacji
  - kilka dużych, dodatnich reszt?

# Założenie PH: NSCLC, mutacja P53 (1)

◆ Krzywe przeżycia się przecinają

◆ Wyraźnie nie równoległe



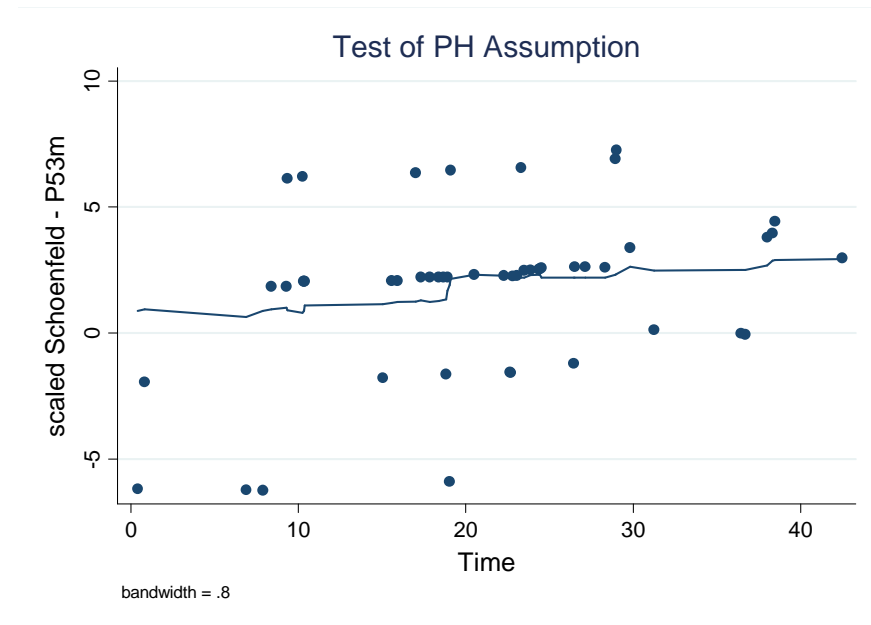
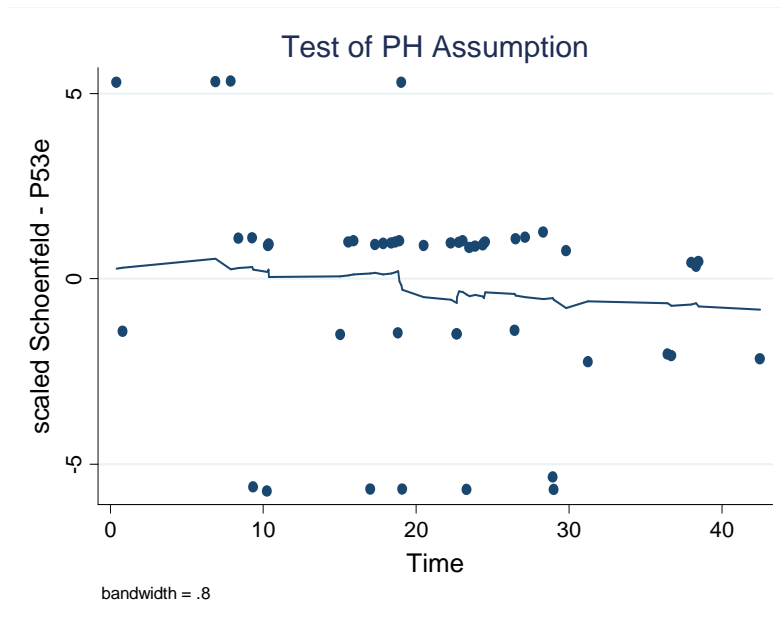
# Założenie PH: NSCLC, mutacja P53 (2)

♦ „P53m · time” added to the model

		_t	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
rh							
	P53e		-.2435042	.3749779	-0.65	0.516	-.9784474 .491439
	P53m		.2350653	.7795496	0.30	0.763	-1.292824 1.762954
t							
	P53m		.0729449	.0344783	2.12	0.034	.0053686 .1405211

Note: Second equation contains variables that continuously vary with respect to time; variables are interacted with current values of \_t.

# Założenie PH: NSCLC, mutacja P53 (3)



## ◆ Test Schoenfelda

	rho	chi2	df	Prob>chi2
P53e	-0.23896	3.29	1	0.0697
P53m	0.34884	7.75	1	0.0054
global test		7.81	2	0.0201

# Warstwowy model PH: NSCLC, ekspresja P53

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
P53m	5.670694	2.308899	4.262	0.000	2.553043	12.59547
P53e	.8191471	.3038457	-0.538	0.591	.3959371	1.694718

- ♦ Model dla ekspresji P53, warstwowy ze względu na mutację P53

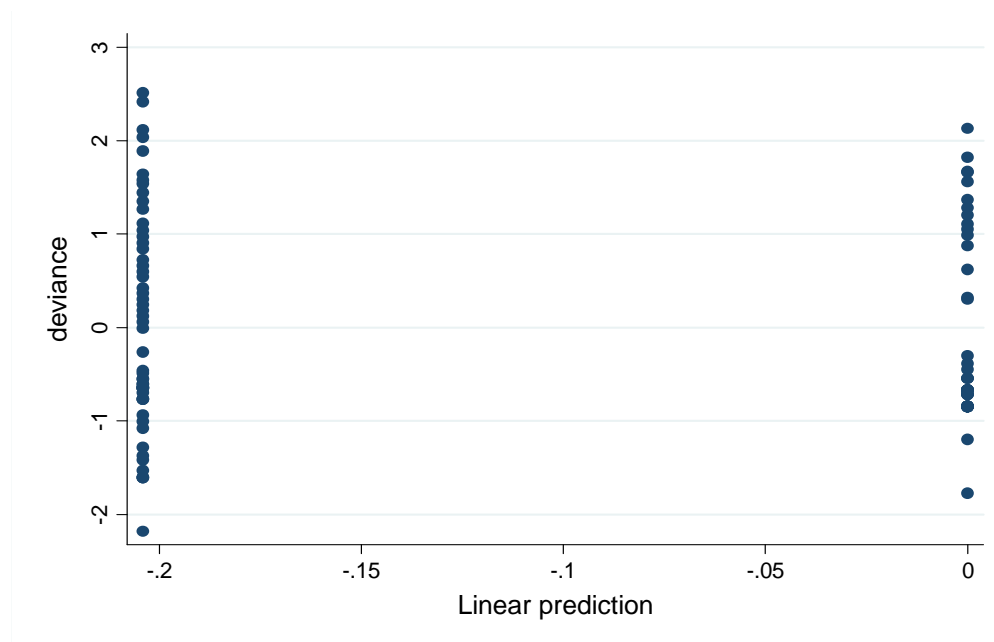
_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
P53e	.8154742	.3033883	-0.55	0.583	.3933054	1.690793

Stratified by P53m

- ♦ Wniosek: efekt ekspresji „znika” po uwzględnieniu mutacji
  - Warstwowy test logrank:  $p=0.58$



# Dopasowanie modelu: NSCLC, ekspresja P53



- ◆ Reszty oparte na dewiancji vs. efekt P53e ( $= -0.20$  dla  $P53e=1$ )
- ◆ Wykres trudny w interpretacji – tylko dwie grupy
- ◆ Bez wyraźnie odstających obserwacji

# Strategie dla nie-proporcjonalnych hazardów (1)

## ◆ „Znaczący efekt”?

- Efekt wieku wydaje się nie być stały w czasie (wzrasta po pierwszych 1-1.5 roku)...
- ale zmiany w czasie są niewielkie w porównaniu do „globalnego” efektu

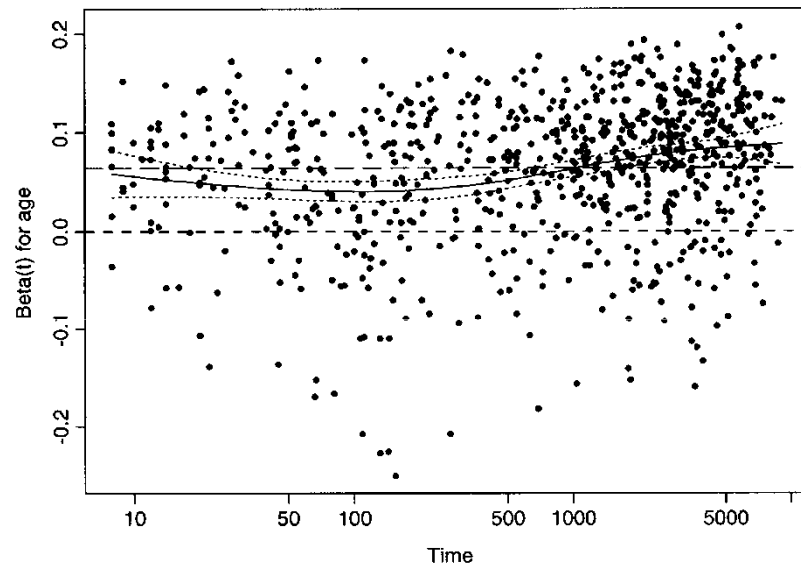


FIGURE 6.9: DVT/PE data, test of PH for age

# Strategie dla nie-proporcjonalnych hazardów (2)

## ◆ „Efekt rzeczywisty”?

- Test dla  $\beta(t)=Zt$  nie jest istotny statystycznie, dla  $\beta(t)=Z \ln t$  jest ...
- ... z powodu odstających reszt

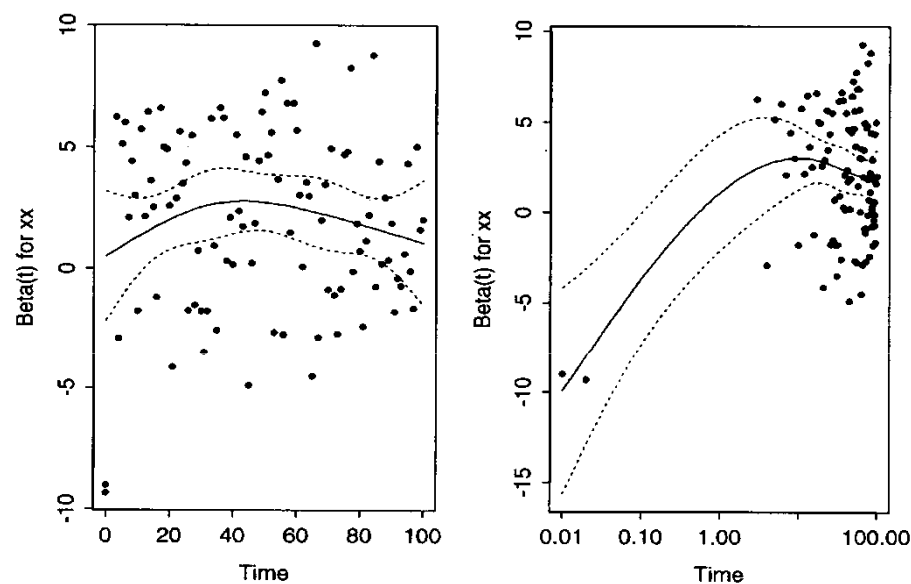


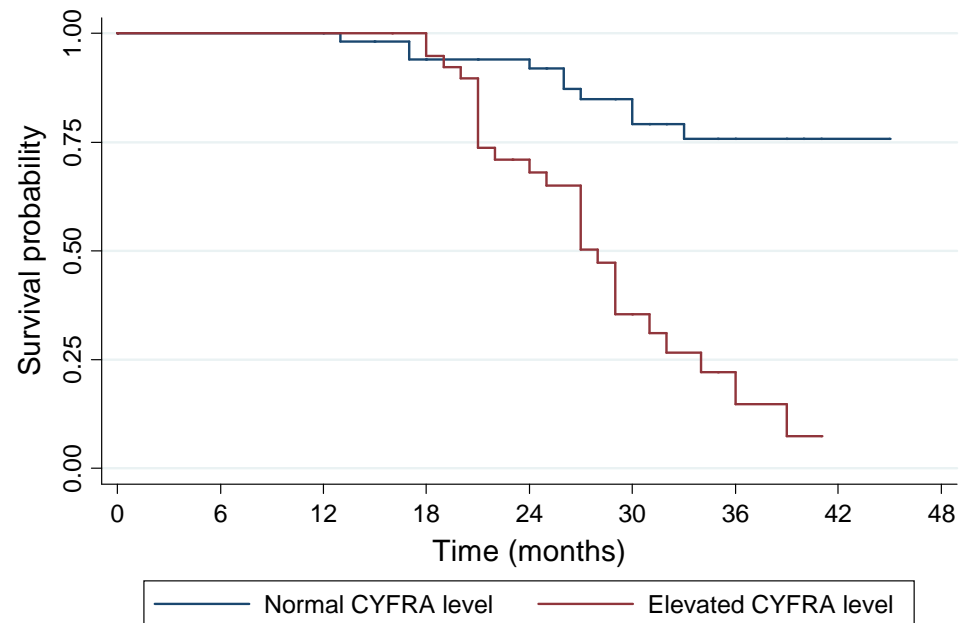
FIGURE 6.10: *Outliers and test for PH*

# Strategie dla nie-proporcjonalnych hazardów (3)

- ◆ Użycie modelu warstwowego
  - efekt zmiennej warstwującej pozostaje nie oszacowany
  - warstwy dla zmiennej ciągłej?
  - strata na efektywności?
- ◆ Użycie modelu PH dla odcinków czasu
- ◆ Użycie modelu ze współczynnikiem zależnym od czasu
- ◆ Użycie innego modelu (sumarycznych hazardów, AFT, ...)

# Współczynnik zależny od czasu w modelu Coxa : przykład (1)

- ♦ Nikliński *et al.*, *Eur Respir J* (1998)
- ♦ Przedoperacyjny poziom CYFRA 21-1 jako czynnik prognostyczny w operowalnym NSCLC
- ♦ Wpływ na czas przeżycia 94 chorych leczonych chirurgicznie
- ♦ Test logrank (trend):  $p < 0.001 \rightarrow$



# Współczynnik zależny od czasu w modelu Coxa : przykład (2)

## ♦ Zmienne objaśniające:

“level” = 1 dla podwyższonego, 0 dla normalnego poziomu CYFRA  
“stage2” = 1 dla TNM II, 0 dla I/IIIA  
“stage3” = 1 dla TNM IIIA, 0 dla I/II  
“adeno” = 1 dla AdenoCa, 0 dla SqCCa/LCCa  
“large” = 1 dla LCCa, 0 dla SqCCa/AdenoCa

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
level	8.793789	3.807352	5.02	0.000	3.763921	20.54526
stage2	1.570952	.8384226	0.85	0.397	.5519169	4.471489
stage3	6.164546	3.181685	3.52	0.000	2.241679	16.9523
adeno	5.018086	2.236916	3.62	0.000	2.094582	12.02206
large	3.530907	1.580739	2.82	0.005	1.468301	8.490974

# Współczynnik zależny od czasu w modelu Coxa : przykład (3)

## ◆ Test założenia PH

	rho	chi2	df	Prob>chi2
level	0.42286	8.24	1	0.0041
stage2	0.01139	0.01	1	0.9418
stage3	-0.09119	0.28	1	0.5977
adeno	0.07211	0.21	1	0.6502
large	0.01765	0.01	1	0.9103
global test		9.98	5	0.0759

_t	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
rh						
level	-2.770215	1.688021	-1.64	0.101	-6.078676	.5382463
stage2	.4993004	.5394972	0.93	0.355	-.5580946	1.556695
stage3	2.037793	.5289788	3.85	0.000	1.001014	3.074573
adeno	1.944006	.492415	3.95	0.000	.9788907	2.909122
large	1.570528	.4876386	3.22	0.001	.6147741	2.526282
t						
level	.1954465	.0676342	2.89	0.004	.0628859	.3280071

Note: Second equation contains variables that continuously vary with respect to time; variables are interacted with current values of \_t.

# Współczynnik zależny od czasu w modelu Coxa : przykład (4)

	_t	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
-----+-----							
rh							
	stage2	.4993004	.5394972	0.93	0.355	-.5580946	1.556695
	stage3	2.037793	.5289788	3.85	0.000	1.001014	3.074573
	adeno	1.944006	.492415	3.95	0.000	.9788907	2.909122
	large	1.570528	.4876386	3.22	0.001	.6147741	2.526282
	level	-2.770215	1.688021	-1.64	0.101	-6.078676	.5382463
-----+-----							
t							
	level	.1954465	.0676342	2.89	0.004	.0628859	.3280071
-----+-----							

- ◆ Dla  $level=1$  dostajemy

$$\lambda(t) = \lambda_0(t) \cdot \exp( 0.499 \cdot stage2 + \dots + 1.571 \cdot large - 2.77 + 0.195 \cdot t )$$

a dla  $level=0$

$$\lambda(t) = \lambda_0(t) \cdot \exp( 0.499 \cdot stage2 + \dots + 1.571 \cdot large )$$

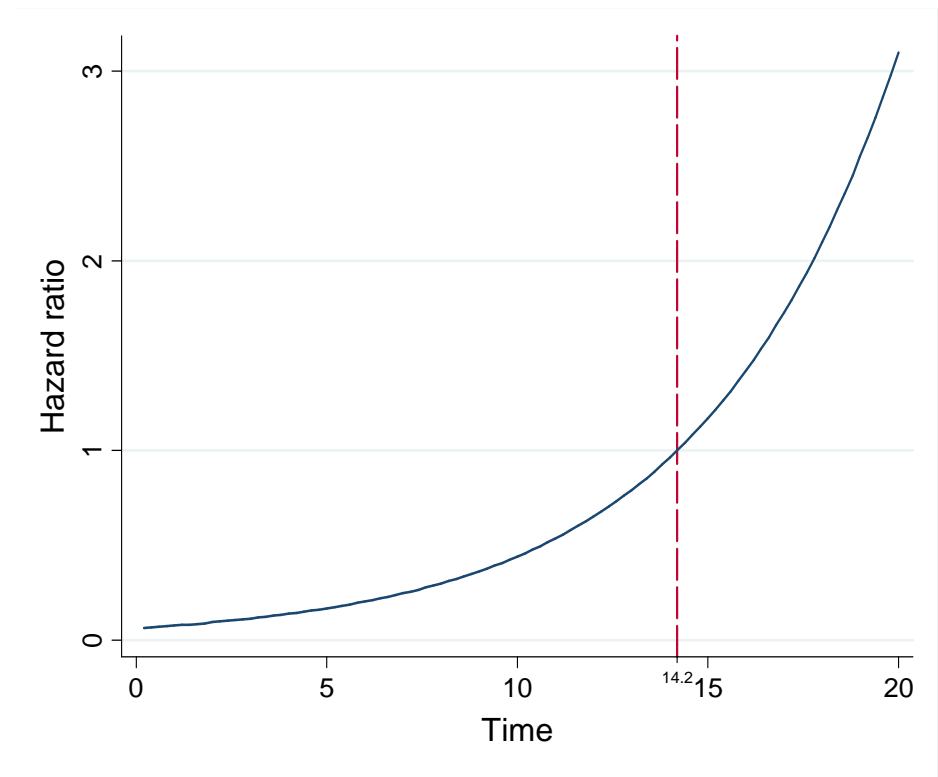
- ◆ Iloraz hazardów wynosi

$$HR = \exp( -2.77 + 0.195 \cdot t )$$



# Współczynnik zależny od czasu w modelu Coxa : przykład (5)

- ♦ Iloraz hazardów dla podwyższonego i normalnego poziomu CYFRA zależy od czasu
  - $HR < 1$  dla  $t < 14.2$  mies.,  $HR > 1$  dla  $t > 14.2$  mies.
- ♦ Trudność w interpretacji: początkowo efekt prewencyjny?
- ♦ Trudność w opisie
  - $HR = \exp(-2.77 + 0.195 \cdot t)$  ?
  - $Z = level \cdot t$  ?



# Model PH: komplikacje (1)

- ♦ Rozważmy dwie zmienne, binarną  $Z_1$  oraz  $Z_2$ .
- ♦ Niech dla obu zachodzi model PH, tj.

$$\lambda(t \mid z_1, z_2) = \lambda_0(t) e^{\beta_1 z_1 + \beta_2 z_2}$$

- ♦ Załóżmy, że używamy modelu tylko z  $Z_1$ .

# Model PH: komplikacje (2)

- ◆ Brzegowa funkcja hazardu dla  $Z_1 = z_1$ :

$$\lambda(t | z_1) = \lambda_0(t) e^{\beta_1 z_1} E\{e^{\beta_2 z_2} | Y(t) = 1, Z_1 = z_1\}$$

gdzie

$$E\{e^{\beta_2 z_2} | Y(t) = 1, Z_1 = z_1\} = \frac{\int e^{\beta_2 z_2} \exp\{\Lambda_0(t) e^{\beta_1 z_1 + \beta_2 z_2}\} C(t; z_2) dG(z_2)}{\int \exp\{\Lambda_0(t) e^{\beta_1 z_1 + \beta_2 z_2}\} C(t; z_2) dG(z_2)}$$

$G(z_2)$ , dystrybuanta  $Z_2$ , a  $C(t, z_2)$  – funkcja przeżycia dla zmiennej cenzurującej.

- ◆ W rezultacie

$$\frac{\lambda(t | z_1 = 1)}{\lambda(t | z_1 = 0)} = e^{\beta_1} \frac{E\{Y(t) e^{\beta_2 z_2} | Z_1 = 1\}}{E\{Y(t) e^{\beta_2 z_2} | Z_1 = 0\}} \equiv e^{\beta(t)}$$

# Model PH: komplikacje (2)

- ◆ Jeśli  $\beta_2 \neq 0$ , pominięcie  $Z_2$  spowoduje obciążenie oszacowania  $\beta_1$ .
  - Nawet jeśli rozkład wartości  $Z_2$  jest zrównoważony dla  $Z_1$
  - Problem w próbach klinicznych!
- ◆ Ponadto spowoduje zależność od czasu HR dla  $Z_1$ .
  - Efekt trudno odróżnić od współczynnika zależnego od czasu