7. Konkurujące ryzyka

Definicje/dane

Interesujące pytania

Podstawowe koncepcje

Szacowanie/testowanie

Modele

Definicja

- *Konkurujące ryzyko* to

  zdarzenie, które uniemożliwia obserwację lub zmienia prawdopodobieństwo wystąpienia interesującego nas zdarzenia

- Przykład:
  - zdarzenie nas interesujące: wznowa nowotworu
  - konkurujące ryzyko (CR): zgon z przyczyny niezależnych od choroby
  - cenzurowanie: koniec próby klinicznej

- Problem: jak uwzględnić CR w analizie?

Dane

- Obserwujemy
  - czas do zdarzenia *T*
  - wskaźnik zdarzenia $\delta = 0$ (cenzurowanie), 1, 2, ... , *J*

Pytania

- Q1: Jaki jest związek między zmiennymi niezależnymi a intensywnością różnych zdarzeń?
  - Np. wpływ leczenia na p-stwo zgonu z powodu wznowy białaczki lub odrzucenia przeszczepu (GVHD) u chorych z przeszczepem szpiku
- Q2: Jaka jest zależność różnych typów zdarzeń?
  - Np. zgonu z powodu wznowy białaczki i GVHD
- Q3: Jaka jest intensywność zdarzeń określonego typu pod warunkiem *wyeliminowania* niektórych (lub wszystkich) innych typów zdarzeń?
  - Np. intenstywność zgonów z powodu wznowy białaczki gdyby GVHD zostało wyeliminowane

Kwestie związane z pytaniem 1

- Q1: Jaki jest związek między zmiennymi niezależnymi a intensywnością różnych zdarzeń?

- Ponieważ inne zdarzenia mogą utrudnić obserwację lub zmienić ryzyko interesującego nas zdarzenia, musimy uwzględnić ich wpływ w analizie

Kwestie związane z pytaniem 2

- Q2: Jaka jest zależność różnych typów zdarzeń?

- Informacja może być użyteczna np. przy leczeniu
  - “łagodne” GVHD może przyczyniać się redukcji ryzyka wznowy białaczki, podczas gdy “ostre” GVHD może wskazywać na nieskuteczność przeszczepu i wyższe ryzyko wznowy

- W oparciu o dane typu (T,δ), nie można odpowiedzieć na pytanie 2

Kwestie związane z pytaniem 3

- Q3: Jaka jest intensywność zdarzeń określonego typu pod warunkiem *wyeliminowania* niektórych (lub wszystkich) innych typów zdarzeń?
  - Często uważane za *główne* pytanie analizy CR

- Ale czy realistyczne?
  - Wyeliminowanie GVHD np. przy pomocy leków immunosupresyjnych lub przeszczepów od bliźniaków może zmienić ryzyko wznowy białaczki

- W oparciu o dane typu (T,δ), nie można odpowiedzieć na pytanie 3

Podstawowy formalizm dla pojedyńczego czasu do zdarzenia, bez CR

- Funkcja przeżycia $S(t) = P(T > t)$
- Dystrybuanta $F(t) = P(T \leq t) = 1 - S(t)$
- Funkcja gęstości $f(t) = \partial F(t) / \partial t$
- Funkcja hazardu $\lambda(t) = \lim_{h \to 0} \frac{P\{T \in [t, t+h) \mid T > t\}}{h}$

Podstawowy formalizm, CR (1)

- Całkowita funkcja przeżycia $S(t) = P(T>t)$
- Całkowita dystrybuanta $F(t) = P(T \leq t)$
- Całkowita funkcja gęstości $f(t) = \partial F(t) \,/\, \partial t$
- Całkowita funkcja hazardu $\lambda(t) = f(t)/S(t)$

- Sub-dystrybuanta $F_j(t) = P(T \leq t,\ \delta = j);\ \lim_{t \to \infty} F_j(t) = P(\delta = j)$
  - „*skumulowana częstość*” (*cumulative incidence*)
  - $F(t) = F_1(t) + F_2(t) + \ldots + F_J(t)$

- Sub-gęstość $f_j(t) = \partial F_j(t) \,/\, \partial t$
  - $f(t) = f_1(t) + f_2(t) + \ldots + f_J(t)$

- Funkcja sub-przeżycia $S_j(t) = P(T > t,\ \delta = j)$
  - „*surowa*” funkcja przeżycia (*crude survival function*)
  - $F_j(t) + S_j(t) = P(\delta = j) < 1$

Podstawowy formalizm, CR (2)

- Sub-dystrybuanta $F_j(t) = P(T \leq t,\ \delta = j)$
- Sub-gęstość $f_j(t) = \partial F_j(t) / \partial t$

- Funkcja sub-hazardu

$$\lambda_j(t) = \lim_{h \to 0} \frac{P\{T \in [t, t+h), \delta = j \mid T > t\}}{h} = \frac{f_j(t)}{1 - F(t)} = \frac{f_j(t)}{S(t)}$$

  - $\lambda(t) = \lambda_1(t) + \ldots + \lambda_J(t)$
  - *„surowa”*, *„specyficzna dla typu” funkcja hazardu* (*crude, cause-specific hazard*)

- Funkcja hazardu dla sub-dystrybuanty

$$h_j(t) = \lim_{h \to 0} \frac{P\{T \in [t, t+h), \delta = j \mid T > t \text{ or } (T \leq t \,\&\, \delta \neq j)\}}{h} = \frac{f_j(t)}{1 - F_j(t)}$$

Podstawowy formalizm, CR (3)

- $F_j(t) = {}_0\int^t f_j(u)du = {}_0\int^t \lambda_j(u)S(u)du = 1 - \exp\{ -{}_0\int^t h_j(u)du \}$

- Związek „specyficznej dla typu” funkcji hazardu z funkcją hazardu dla sub-dystrybuanty:

$$\lambda_j(t) = \left\{ 1 + \frac{\sum_{k \neq j} F_k(t)}{S(t)} \right\} h_j(t)$$

CR, formalizm dla „ukrytych” czasów (1)

- Dla każdego typu zdarzenia j, przyjmijmy istnienie czasu T_j
- Obserwujemy $T = \min(T_1, T_2, \ldots, T_J)$ oraz $\delta = j$ jeśli $T_j = T$
- Łączna funkcja przeżycia $Q(t_1, \ldots, t_J) = P(T_1 > t_1, \ldots, T_J > t_J)$
  - $S(t) = Q(t, \ldots, t)$
- Sub-gęstość $f_j(t) = [-\partial Q(t_1, \ldots, t_J) / \partial t_j]_{(t, \ldots, t)}$
- Sub-dystrybuanta $F_j(t) = {}_0\!\int^t f_j(t)dt$
- Funkcja sub-hazardu $\lambda_j(t) = [-\partial \ln Q(t_1, \ldots, t_J) / \partial t_j]_{(t, \ldots, t)}$

CR, formalizm dla „ukrytych” czasów (2)

- Łączna funkcja przeżycia $Q(t_1, \ldots, t_J) = P(T_1 > t_1, \ldots, T_J > t_J)$

- Brzegowa funkcja przeżycia $S_j^*(t) = P(T_j > t) = Q(0, \ldots, t, \ldots, 0)$
  - *funkcja przeżycia „netto“* (*net survival function*) (bez innych CR)

- Hazard dla rozkładu brzegowego: $\lambda_j^*(t) = -d \ln S_j^*(t)/dt = f_j(t)/S_j^*(t)$
  - *funkcja hazardu „netto”* (*net hazard function*)

Podstawowy formalizm, CR, przykład (1)

- Dwywymiarowy (niezal.) rozkład wykładniczy $Q(t_1, \dots, t_J) = \exp(-\lambda_1 t_1 - \lambda_2 t_2)$
- Całkowita funkcja przeżycia $S(t) = \exp\{-(\lambda_1+\lambda_2)t\}$
- Sub-gęstość $f_1 = \lambda_1 S(t)$
- Funkcja sub-hazardu $\lambda_1(t) = \lambda_1$
- Funkcja sub-przeżycia $S_1(t) = S(t)\,\lambda_1 / (\lambda_1+\lambda_2)$
- Sub-dystrybuanta $F_1(t) = \lambda_1 \{1-S(t)\} / (\lambda_1+\lambda_2)$
  - $S_1(t) + F_1(t) = \lambda_1 / (\lambda_1+\lambda_2)$
- F. hazardu sub-dystrybuanty $h_1(t) = f_1(t)/\{1-F_1(t)\}$
- Brzegowa funkcja przeżycia $S_1^*(t) = \exp(-\lambda_1 t)$
- Brzegowa funkcja hazardu $\lambda_1^*(t) = \lambda_1$

Podstawowy formalizm, CR, przykład (2)

- Dwywymiarowy (zal.) rozkład wykładniczy $Q(t_1, \dots, t_J) = \exp(-\lambda_1 t_1 - \lambda_2 t_2 - \mu t_1 t_2)$
  - $0 \le \mu < \lambda_1 \lambda_2$, $\mu = 0 \rightarrow$ niezależność

- Całkowita funkcja przeżycia $S(t) = \exp\{-(\lambda_1+\lambda_2)t - \mu t^2\}$

- Sub-gęstość $f_1 = (\lambda_1 + \mu t) S(t)$
- Funkcja sub-hazardu $\lambda_1(t) = \lambda_1 + \mu t$
- Funkcja sub-przeżycia $S_1(t) = 0.5 \exp\{-(\lambda_1+\lambda_2)t - \mu t^2\}$

$$+0.5\,(\lambda_1-\lambda_2)\,(\pi/\mu)^{1/2} \exp\{(\lambda_1+\lambda_2)^2/4\mu\}\ \Phi[-(2\mu)^{1/2}\{t+(\lambda_1+\lambda_2)/(2\mu)\}]$$

- Sub-dystrybuanta $F_1(t) = 0.5\,[1 - \exp\{-(\lambda_1+\lambda_2)t - \mu t^2\}]$

$$+0.5(\lambda_1-\lambda_2)(\pi/\mu)^{1/2}\exp\{(\lambda_1+\lambda_2)^2/4\mu\}[\Phi\{(2\mu)^{1/2}t+(\lambda_1+\lambda_2)/(2\mu)^{1/2}\}-\Phi\{(\lambda_1+\lambda_2)/(2\mu)^{1/2}\}]$$

- F. hazardu sub-dystrybuanty $h_1(t) = f_1(t) / \{1 - F_1(t)\}$

- Brzegowa funkcja przeżycia $S_1^*(t) = \exp(-\lambda_1 t)$
- Brzegowa funkcja hazardu $\lambda_1^*(t) = \lambda_1$

Nie-identyfikowalność rozkładów brzegowych dla „ukrytych” czasów

- Jeśli mamy tylko dane typu (T, $\delta = j$)
  - zobacz poprzednie przykłady
- Identyfikowalność przy założeniu niezależności:
$$Q(t_1, \dots , t_J) = \prod_j Q_j(t_j)$$
- Wówczas $\lambda_j^*(t) \equiv \lambda_j(t)$

Estymator Kaplana-Meiera i konkurujące ryzyka

Table 4.1 Discharge or death data.

Patient ID	Time to hospital discharge or death (days)	Type of event[a]
1	1	2
2	2	2
3	3	2
4	4	1
5	5	2
6	7	2
7	8	1
8	10	2
9	13	1
10	14	2
11	15	1
12	17	2
13	20	1
14	30	1
15	42	1
16	50	1
17	65	1
18	80	2
19	84	1
20	90	1

[a] 1 = Discharge from hospital, 2 = Death.

- Oszacowanie $F_1(30)=P(T \leq 30, \delta=1)$:
  6 wypisów / 20 obserwacji = 0.30

- Estymator K-M: zgony jako cenzurowanie:
  $1\times(16/17)\times(13/14)\times(11/12)\times(9/10)\times(7/8)\times(6/7)= .54$
  $1 - 0.54 = 0.46 > 0.30$

- Dopełnienie K-M szacuje $_0\int^t \lambda_j(u)S_j^*(u)du$

- Chcemy szacować $F_j(t) = {}_0\int^t \lambda_j(u)S(u)du$

- Jako że $S(t) \leq S_j^*(t)$, dopełnienie estymatora K-M przeszacowuje $F_j(t)$

Szacowanie funkcji skumulowanej częstości

$$F_j(t) = \int_0^t \lambda_j(u) S(u) du \longrightarrow \hat{F}_j(t) = \sum_{t_l \le t} \frac{d_{lj}}{n_l} \hat{S}(t_{l-1})$$

gdzie

- d_{lj} to liczba zdarzeń typu *j* w chwili (dowolnego) zdarzenia t_l
- n_l *to* liczebność zbioru ryzyka w chwili (dowolnego) zdarzenia t_l
- $\hat{S}(t_l)$ to oszacowanie Kaplana-Meiera $S(t_l)$

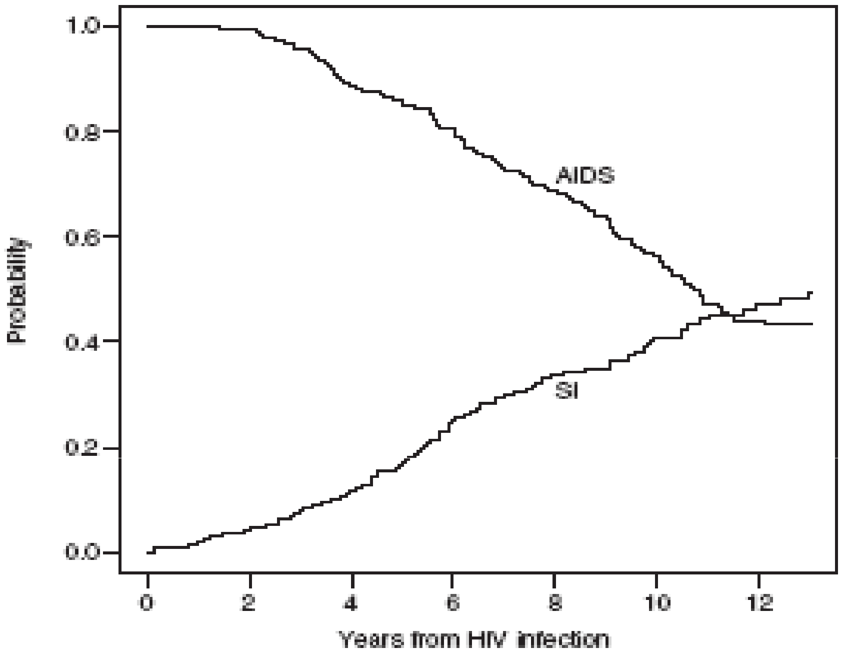
Przykład

- Dane dla 329 homoseksualnych mężczyzn z badania kohortowego infekcji HIV i AIDS
  - Putter, Fiocco, Geskus (2007), *Statist Med*, 26, 2389-2430
- U niektórych mężczyzn występuje fenotyp SI-HIV
  - Komórki zarażone wirusem tworzą *syncytia* (twory wielojądrowe), które nie posiadają właściwości immunologicznych, w związku z czym odporność organizmu zarażonego jest obniżona
- Wystąpienie AIDS jest konkurującym ryzykiem
- Zmienna niezależna: delecja genu C-C chemokine receptor 5
  - obniżona podatność na HIV i opóźniona progresja AIDS
  - 259/324 mężczyzn (80%) bez delecji (genotyp „wild-wild")

Przykład: “Naiwny” estymator

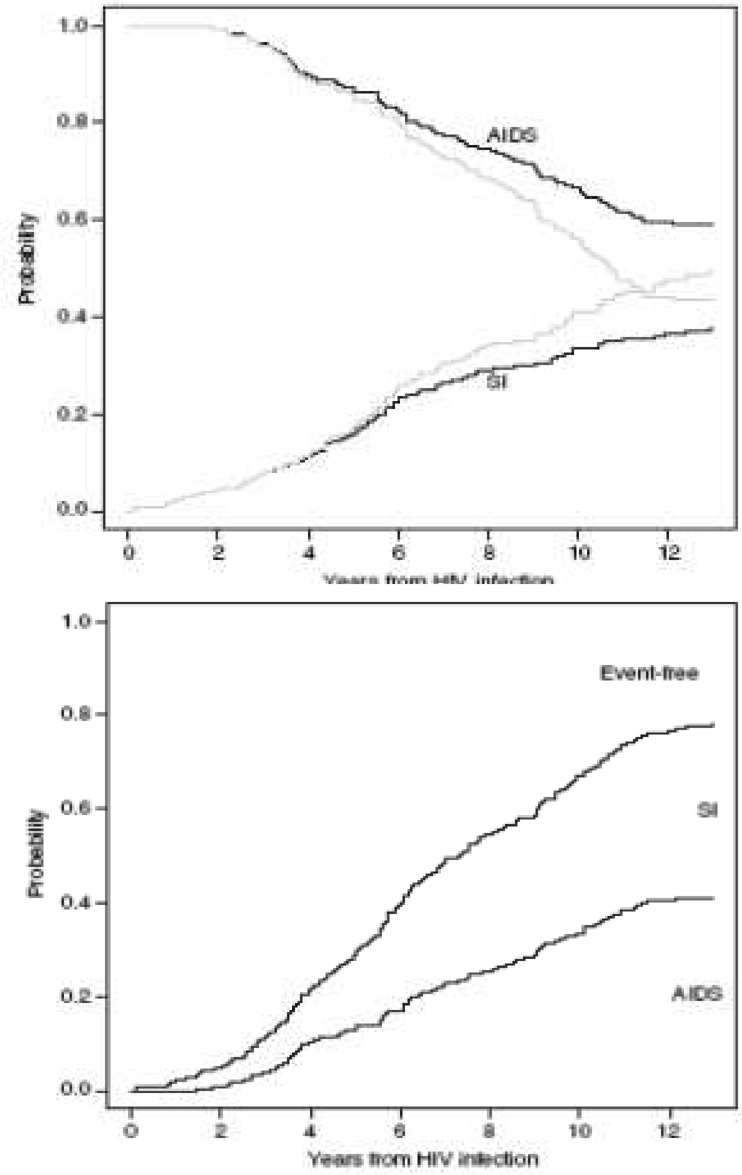
- Krzywa przeżycia dla AIDS, 1-krzywa przeżycia dla SI

- Dla 13 lat, oszacowane p-stwa zdarzeń wynoszą 0.567 (AIDS) & 0.496 (SI)
  - suma > 1 !



Przykład: „Poprawne” szacowanie

- Przy użyciu funkcji skumulowanych częstości (sub-dystrybuant)
- Dla 13 lat, oszacowane p-stwa wynoszą 0.408 (AIDS) & 0.375 (SI)
  - “naiwne”: 0.567 & 0.496
- Wykres „nakładkowy”: $F_{AIDS}(t)$, oraz $F_{AIDS}(t) + F_{SI}(t) = F(t)$
  - różnica = $F_{SI}(t)$



Porównywanie funkcji skumulowanej częstości (sub-dystrybuant)

- Są różne testy, najczęściej używany/wspominany: test Gray'a (Gray 1988, *Ann Stat*)
- Dwa typy zdarzeń, dwie gupy
- Statystyka testowa dla *j*-tego typu zdarzenia:

$$U = \sum_{t_l} \left(d_{1lj} - R_{1lj} \frac{d_{lj}}{R_{lj}} \right) \quad gdzie \quad R_{1lj} = n_{1l} \frac{1 - \hat{F}_{1j}(t_{l-1})}{\hat{S}_1(t_{l-1})}$$

d_{1lj} (d_{lj}) – zdarzenia typu *j* w groupie 1 (łącznie) w chwili zdarzenia t_l

n_{1l} – liczebność zbioru ryzyka w grupie 1 w chwili zdarzenia t_l

$R_{lj} = R_{1lj} + R_{2lj}$

- Jak test log-rank, ale z modyfikowanym zbiorem ryzyka R_{lj}

Częściowa funkcja wiarogodności dla funkcji sub-hazardu

- Zakładając niezależne/nieinformatywne cenzurowanie:

$$L = \prod_{i=1}^{n} \left[\lambda_{j_i}(t_i; x_i)^{\delta_i} \prod_{j=1}^{J} \exp\left\{ -\int_0^{t_i} \lambda_j(u; x_i) du \right\} \right]$$

$$= \prod_{j=1}^{J} \left[\prod_{i=1}^{n} \lambda_j(t_i; x_i)^{\delta_{ji}} \exp\left\{ -\int_0^{+\infty} \sum_{i=1}^{n} Y_i(u) \lambda_j(u; x_i) du \right\} \right]$$

- δ_{ji}=1 jeśli zdarzenie *j-tego* typu dla jednostki *i*, 0 w p.p.

- Iloczyn osobnych czynników dla każdego typu zdarzenia
  - w terminach sub-hazardów „specyficznych dla typu” (cause-specific)
  - dla każdego czynnika, zdarzenia innych typów są cenzurowane

Model proporcjonalnych funkcji sub-hazardów

$$\lambda_j(t) = \lambda_{0j}(t)\exp(Z'(t)\beta_j)$$

- Częściowa funkcja wiarogodności: $L = \prod_{j=1}^{J} \prod_{i=1}^{n_j} \frac{e^{Z_{ij}^{'}(t_{(ij)})\beta_j}}{\sum_{l \in R_{ij}} e^{Z_l^{'}(t_{(ij)})\beta_j}}$

  $t_{(1j)} < ... < t_{(nj,j)}$ to czasy zdarzeń *j*-tego typu

  Z_{ij} to wektor zmiennych niezależnych dla jednostki ze zdarzeniem w $t_{(ij)}$

  R_{ij} to zbiór ryzyka dla czsu zdarzenia $t_{(ij)}$

- Może być szacowany przy użyciu „standardowych” programów
  - cenzurujemy zdarzenia inne niż nas interesujące
    - zbiór ryzyka maleje dla czasów innych typów zdarzeń

Model proporcjonalnych funkcji sub-hazardów

$$\lambda_j(t) = \lambda_{0j}(t)\exp(Z'(t)\beta_j)$$

- Problem: $F_j(t)={}_0\int^t \lambda_j(u)S(u)du$, czyli funkcja skumulowanej częstości (sub-dystrybuanta) zależy od $S(t)$
  - tzn. zależy od hazardów specyficznych dla innych typów zdarzeń
  - efekt zmiennych niezależnych jest trudny w interpretacji

Model proporcjonalnych funkcji hazardu sub-dystrybuanty

$$h_j(t) = h_{0j}(t)exp(Z'\beta_j)$$

- Zaproponowany przez Fine & Gray (1999, *JASA*)
  - Mniej dopracowany/elastyczny niż model PH dla funkcji sub-hazardu
- Wymaga specjalnego oprogramowania
  - jednostki z innymi typami zdarzeń pozostają w zbiorze ryzyka
- Ponieważ $F_j(t) = 1 - \exp\{ -\int_0^t h_j(u)du \}$
  - funkcja skumulowanej częstości zależy jedynie od $h_j(t)$
  - efekt zmiennych niezależnych łatwiejszy do interpretacji

Częściowa funkcja wiarogodności dla funkcji hazardu sub-dystrybuanty

$$L = \prod_{i=1}^{n_j} \frac{e^{Z'_{ij}\beta_j}}{\sum_{l \in R_{ij}} w_{lj} e^{Z'_l \beta_j}}$$

R_{ij}= {l: $T_l \geq t_{(ij)}$ lub ($T_l < t_{(ij)}$ ale zdarzenie innego typu niż j)}

Wagi $$w_{lj} = \frac{\hat{G}\{t_{(ij)}\}}{\hat{G}\{\min(t_{(ij)}, t_l)\}}$$

gdzie G – oszacowanie K-M funkcji przeżycia dla cenzurowania

- 1 jeśli jednostka bez zdarzenia przed $t_{(ij)}$
- ≤1 jeśli jednostka z konkurującym zdarzeniem przed $t_{(ij)}$
  - im wcześniejsze zdarzenie w stosunku do $t_{(ij)}$, tym mniejsza waga

Częściowa funkcja wiarogodności dla funkcji hazardu sub-dystrybuanty: przykład

Table 6.1 Hypothetical example dataset.

Individual	Time	Type of event	X	$\hat{G}$
SN1	1	0	12	0.9
SN2	2	2	10	0.9
SN3	3	1	9	0.9
SN4	4	1	13	0.9
SN5	5	0	8	0.75
SN6	6	2	9	0.75
SN7	7	1	12	0.75
SN8	8	0	10	0.5
SN9	9	1	11	0.5
SN10	10	0	8	0

X = covariate; $\hat{G}$ = estimator of the censoring distribution.
Type of event: 1 = Event of interest, 2 = Competing risk event, 0 = Censored.

R_{ij}= {l: $T_l \geq t_{(ij)}$ lub
($T_l < t_{(ij)}$ ale zdarzenie innego typu niż j)}

$$w_{lj} = \frac{\hat{G}\left\{t_{(ij)}\right\}}{\hat{G}\left\{\min\left(t_{(ij)}, t_l\right)\right\}}$$

Table 6.2 Calculation of the weights for example in Table 6.1.

Time	SN1	SN2	SN3	SN4	SN5	SN6	SN7	SN8	SN9	SN10
3		$\frac{\hat{G}(3)}{\hat{G}(2)} = 1$	1	1	1	1	1	1	1	1
4		$\frac{\hat{G}(4)}{\hat{G}(2)} = 1$		1	1	1	1	1	1	1
7		$\frac{\hat{G}(7)}{\hat{G}(2)} = 0.83$				$\frac{\hat{G}(7)}{\hat{G}(6)} = 1$	1	1	1	1
9		$\frac{\hat{G}(9)}{\hat{G}(2)} = 0.56$				$\frac{\hat{G}(9)}{\hat{G}(6)} = 0.67$			1	1

Związek między funkcjami hazardu

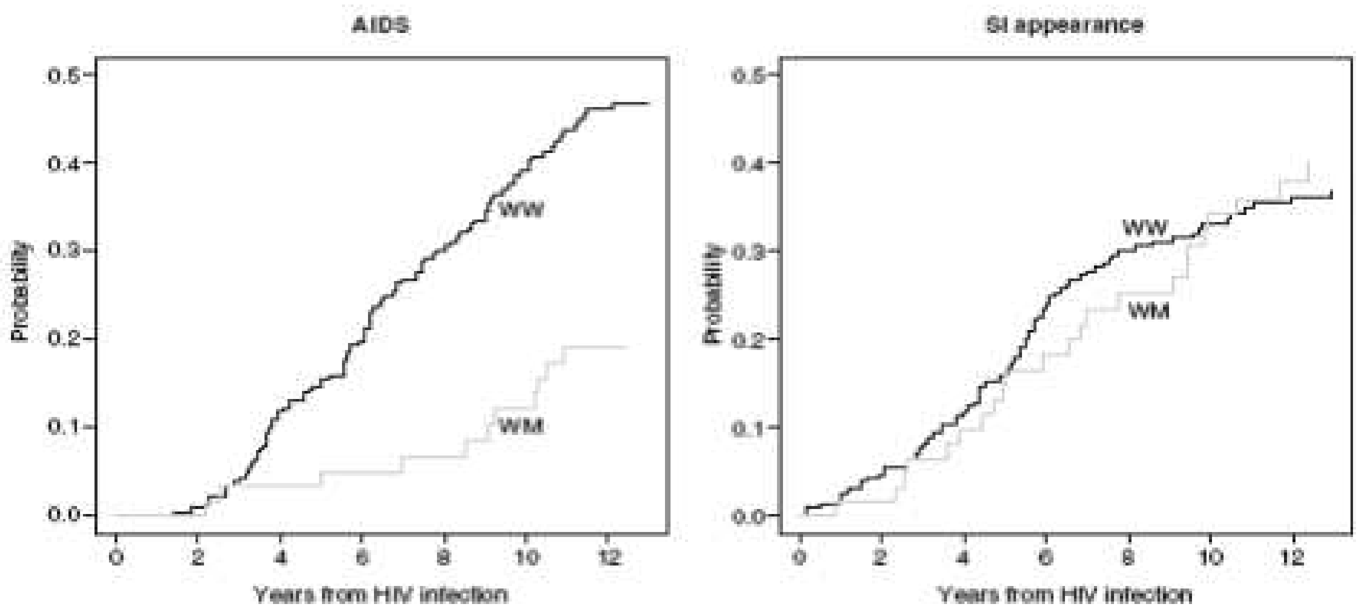
- Mamy

$$\lambda_j(t) = \left\{ 1 + \frac{\sum_{k \neq j} F_k(t)}{S(t)} \right\} h_j(t)$$

- Model PH dla $h_j(t)$ nie daje modelu PH dla $\lambda_j(t)$ i *vice versa*
- Efekt zmiennych niezależnych dla $h_j(t)$ i $\lambda_j(t)$ może być różny

Przykład: Efekt zmiennej niezależnej

- Nie-parametryczne oszacowania funkcji skumulowanych częstości (sub-dystrybuant) dla genotypów bez delecji (WW) i z delecją (WM)



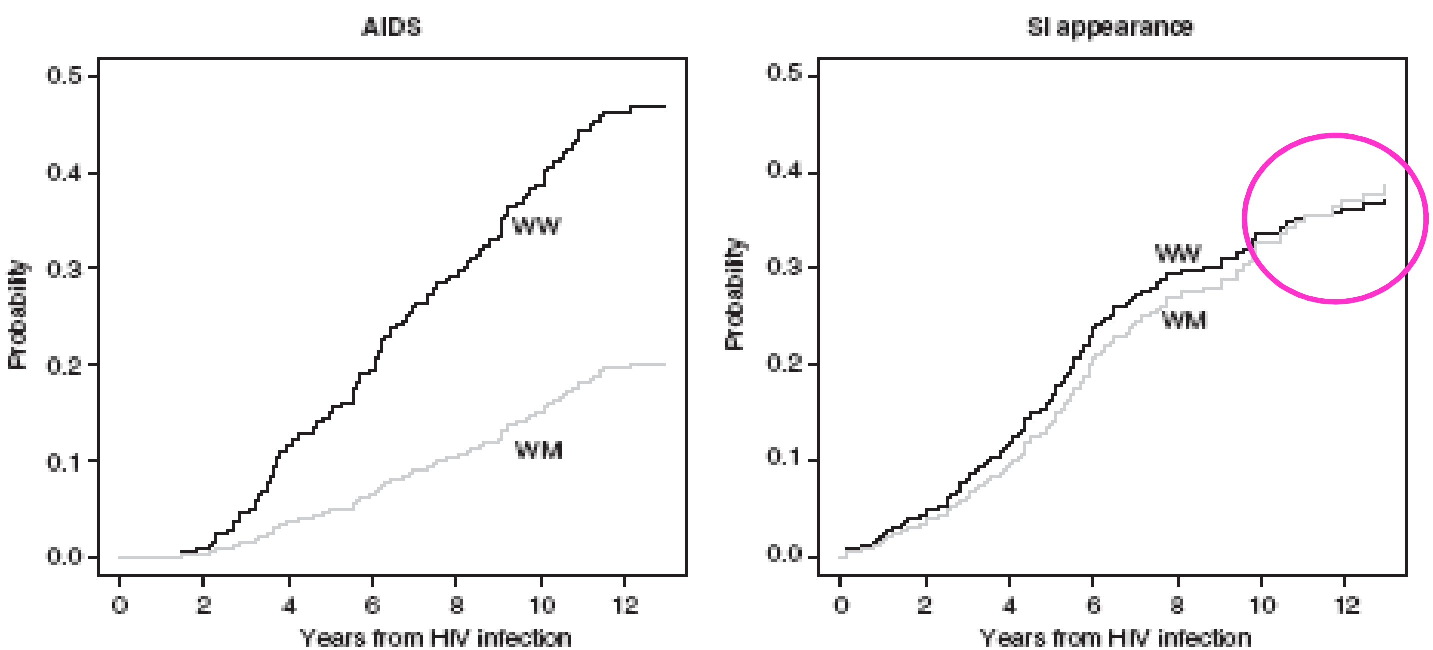
Przykład: Model PH dla sub-hazardów

- Oszacowania współczynników (WM *vs.* WW):

  AIDS: -1.24 (SE 0.31), HR=0.29, $p<10^{-4}$

  SI: -0.24 (SE 0.24), HR=0.78, $p=0.29$

- Oszacowania funkcji skumulowanych częstości z modelu:

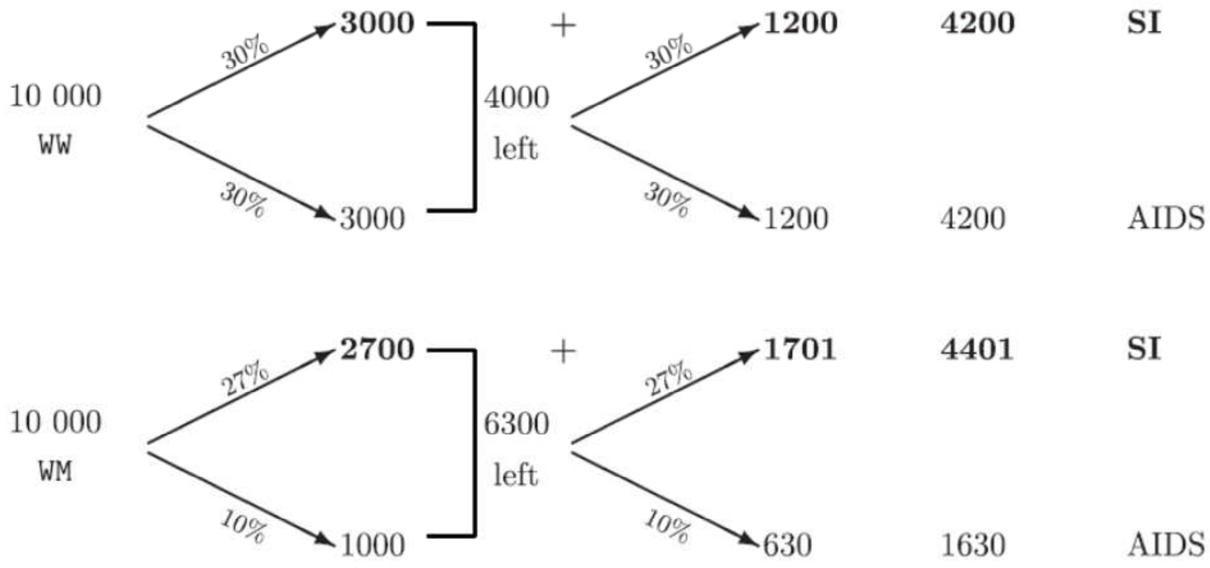


Przecięcie dla modelu PH?!

Przykład: Model PH dla funkcji sub-hazardu, symulacja (1)

- Załóżmy następujące intensywności zdarzeń (stałe w czasie):

	WW	WM	HR
SI	0.30	0.27	0.90
AIDS	0.30	0.10	0.33



- Większy zbiór ryzyka w WM po 1ej chwili czasu z racji efektu dla AIDS
- Częstość SI wzrasta dla WM w 2ej chwili czasu (choć HR=0.90)

Przykład: Model PH dla funkcji sub-hazardu, symulacja (2)

- Przyjmijmy HR z modelu (AIDS: 0.29, SI: 0.78) i bazową funkcję hazardu dla SI
- Zwiększamy bazową funkcję hazardu dla AIDS
- Otrzymane funkcje skumulowanej częstości dla SI:

Bez CR

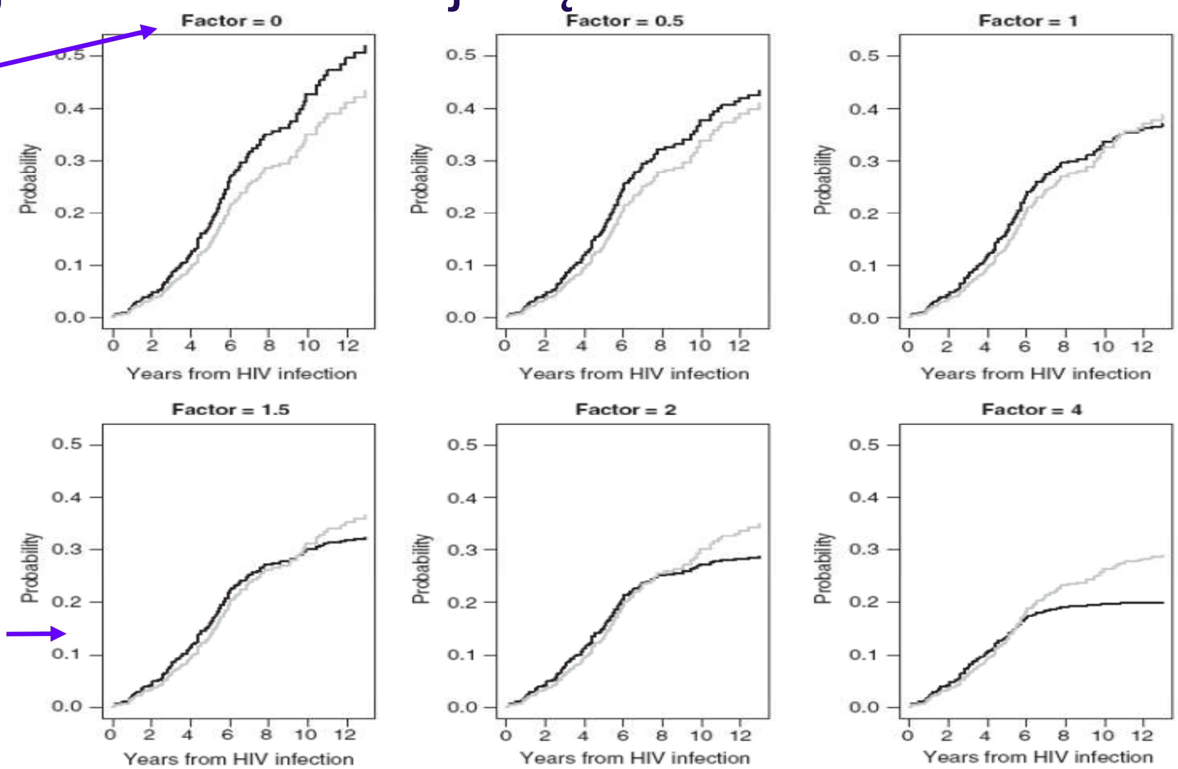
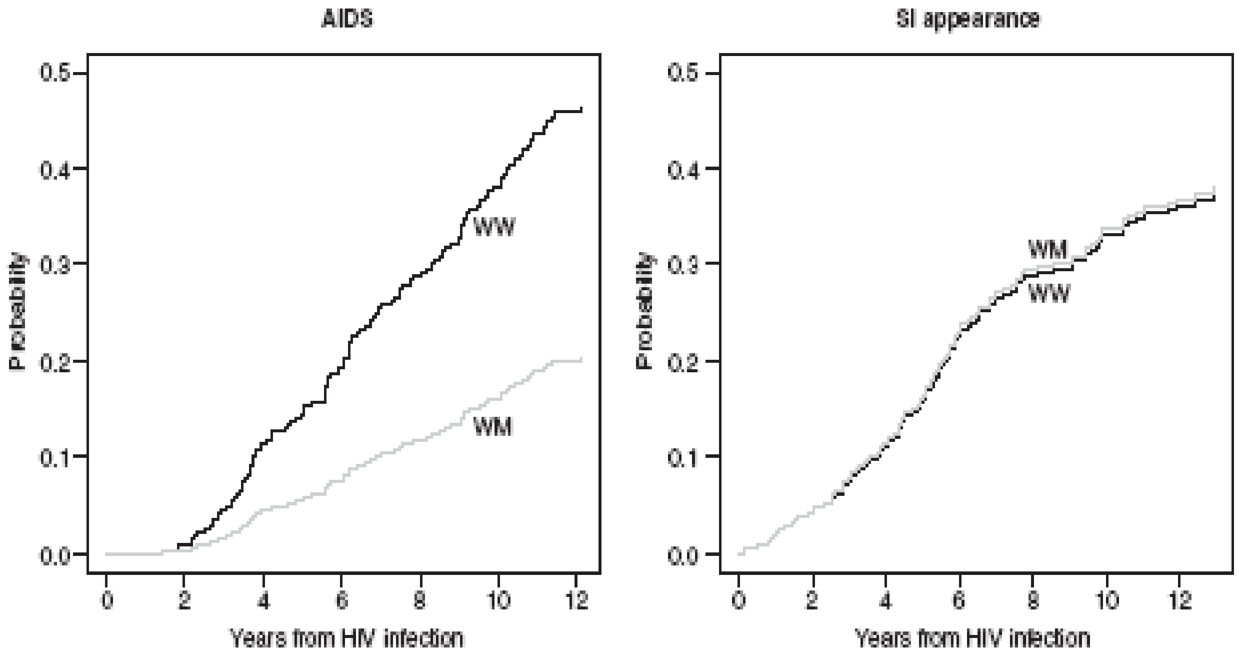


Figure 7. Cumulative incidence functions for SI appearance, for CCR5 wild-type WW (black) and mutant WM (grey). The baseline hazard of AIDS was multiplied with different factors, while keeping everything else the same.

- Zwiększona „konkurencja” ze strony AIDS ogranicza zbiór ryzyka w WW
- W efekcie, częstość SI rośnie dla WM (choć HR=0.78)

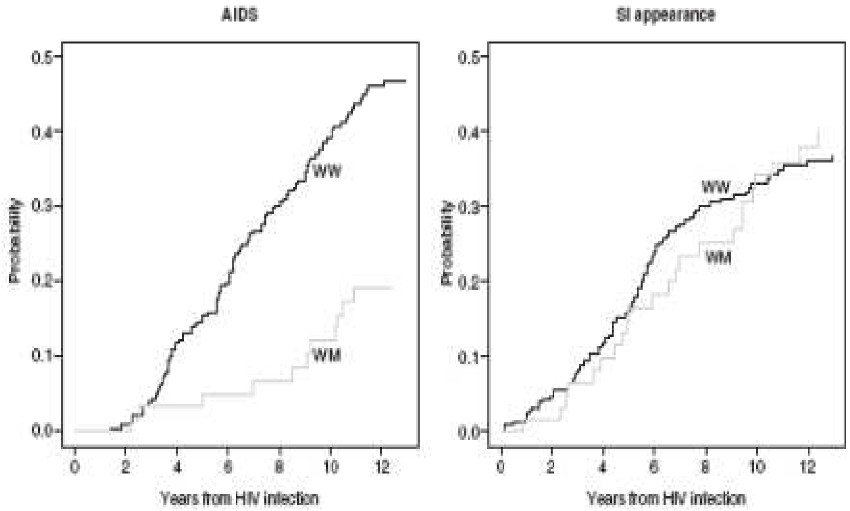
Przykład: Model PH dla funkcji hazardu sub-dystrybuanty

- Oszacowane współczynniki (WM *vs.* WW):
  AIDS: -1.004 (SE 0.295), HR=0.366, p=0.0007
  SI: 0.024 (SE 0.227), HR=1.024, p=0.92
  - HR_{SI}=0.78 dla modelu PH dla funkcji sub-hazardu
- Oszacowania funkcji skumulowanych częstości dla modelu:

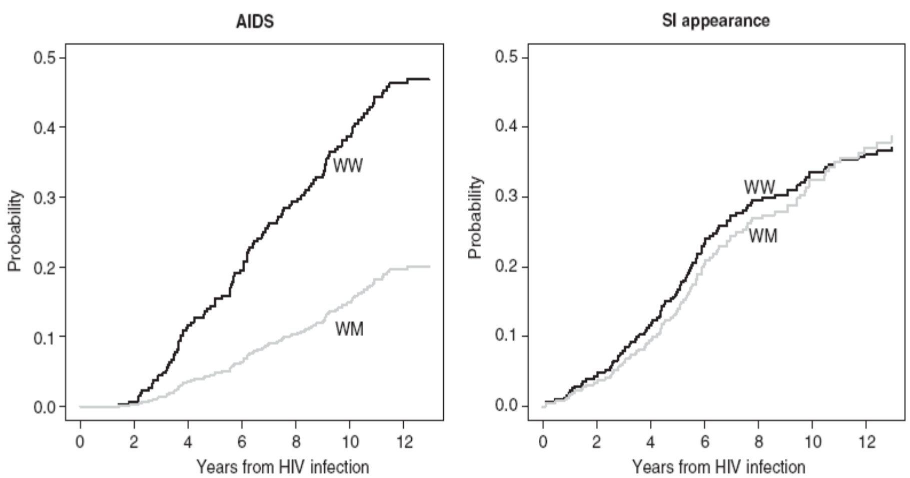


Przykład: Dopasowanie do danych?

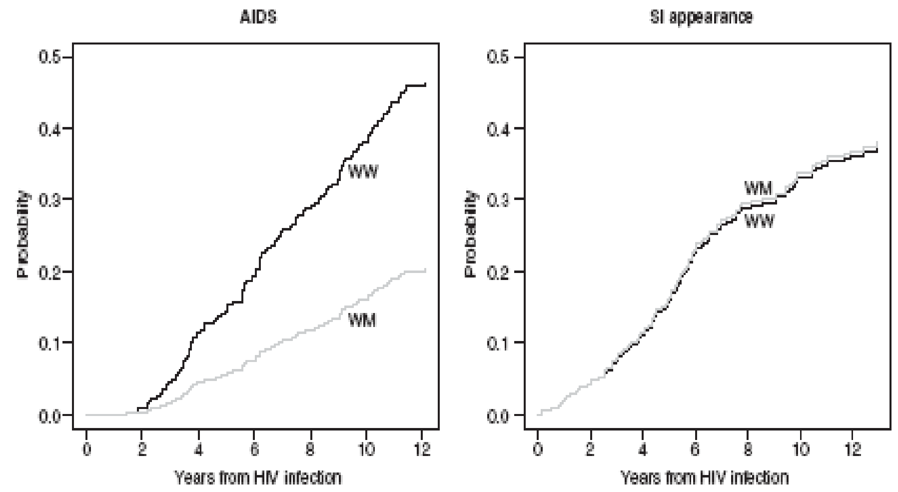
- Oszacowania nie-parametryczne



- Model dla „specyficznych” hazardów



- Model dla hazardów sub-dystrybuanty



Inne modele

- Modele addytywnej funkcji hazardu, np.

  $$\lambda_j(t) = \lambda_{0j}(t) + Z(t)'\beta_j \quad \text{lub} \quad h_j(t) = h_{0j}(t) + Z(t)'\beta_j$$

- Modele przyśpieszonego czasu do niepowodzenia:

  $$\lambda_j(t) = \lambda_{0j}\{t\ \mathrm{exp}(Z'\beta_j)\}\mathrm{exp}(Z'\beta_j)$$

- Modele addytywnej funkcji skumulowanej częstości:

  $$F_j(t) = F_{0j}(t) + Z(t)'\beta_j$$

- Modele wielu stanów (multi-state)

- ...

Bibliografia

- Crowder (2001), Chapman & Hall/CRC
- Kalbfleisch & Prentice (2002), Chapter 8, Wiley
- Pintilie (2006), Wiley

- Klein (2006), *Statist Med*, 25, 1015-1034
- Putter, Fiocco, Geskus (2007), *Statist Med*, 26, 2389-2430
- Beyersmann *et al.* (2007), *Statist Med*, 26, 5360-5369

Software

- R: *cmprsk*, *surv2samp*
- STATA: *stcompet.ado*
- Melanie Pintilie's website:
  http://www.uhnres.utoronto.ca/labs/hill/People_Pintilie.htm