Ocenjevanje v relativnem preživetju

Maja Pohar Perme

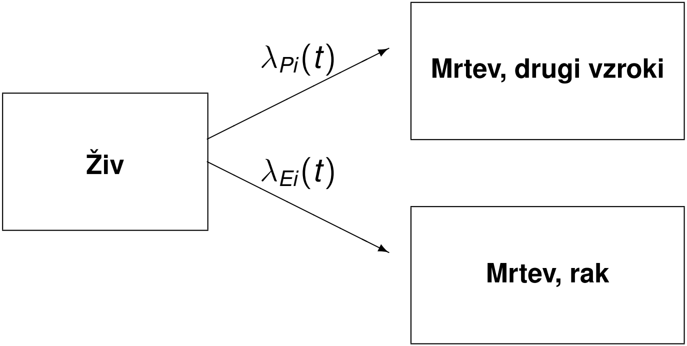
IBMI
Medicinska fakulteta, Univerza v Ljubljani

April 2013

Analiza preživetja

- Bolniki spremljani v dolgih časovnih intervalih
- Možnih je več končnih dogodkov, nas zanima predvsem eden izmed njih
- Vrsta končnega dogodka pogosto ni znana oziroma zanesljiva
- Primer: register raka, smrti zaradi raka ali iz drugih vzrokov, vzrok smrti ni podan

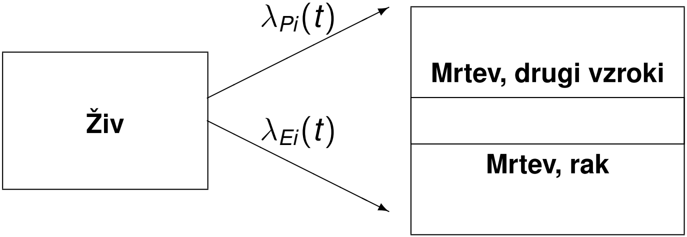
Podatki



Vrste podatkov

- **Dva vzroka: Čas spremljanja, vzrok konca spremljanja (0=izgubljen, 1=mrtev bolezen, 2=mrtev, drugo)**

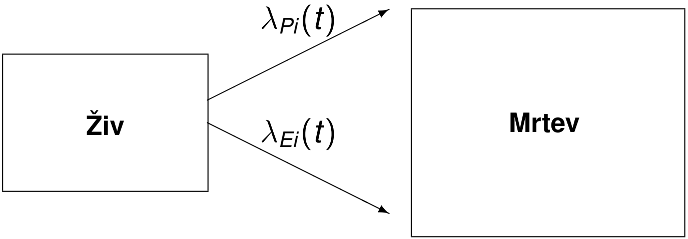
Podatki



Vrste podatkov

- **Dva vzroka: Čas spremljanja, vzrok konca spremljanja (0=izgubljen, 1=mrtev bolezen, 2=mrtev, drugo)**

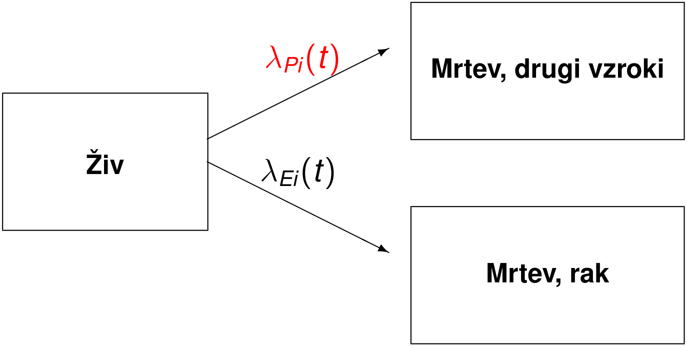
Podatki



Vrste podatkov

- **Dva vzroka: Čas spremljanja, vzrok konca spremljanja (0=izgubljen, 1=mrtev bolezen, 2=mrtev, drugo)**
- **Relativno preživetje: Čas spremljanja, vzrok konca spremljanja (0=izgubljen, 1=mrtev) + podatki o smrtnosti v populaciji**

Podatki



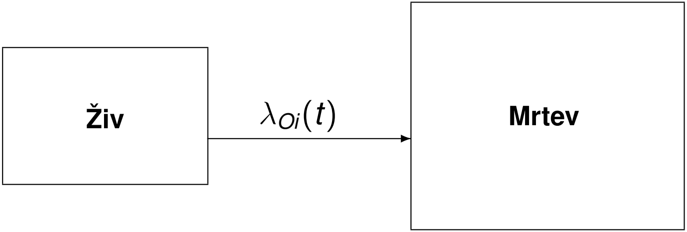
Vrste podatkov

- **Dva vzroka: Čas spremljanja, vzrok konca spremljanja (0=izgubljen, 1=mrtev bolezen, 2=mrtev, drugo)**
- **Relativno preživetje: Čas spremljanja, vzrok konca spremljanja (0=izgubljen, 1=mrtev) + podatki o smrtnosti v populaciji**

Relativno preživetje

Podatki o bolnikih z rakom

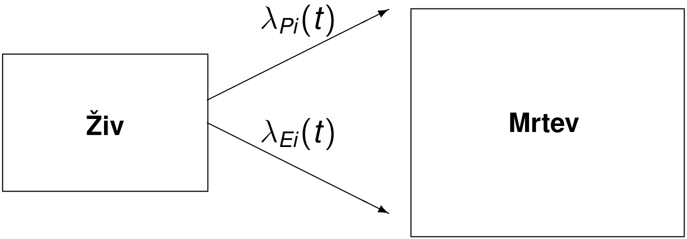
- **Opazovani bolniki: Čas do smrti ali krnjenja; populacija: tveganje za smrt glede na demografske dejavnike**



Relativno preživetje

Podatki o bolnikih z rakom

- **Opazovani bolniki: Čas do smrti ali krnjenja; populacija: tveganje za smrt glede na demografske dejavnike**

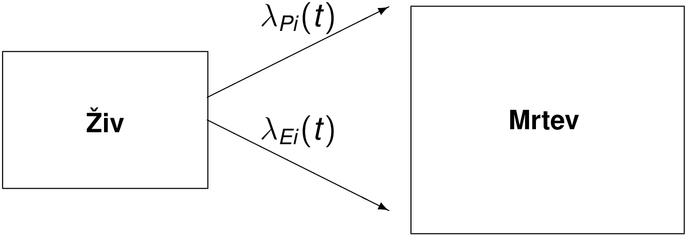


$\lambda_{Pi}(t)$ predstavlja ogroženost, ki bi jo imeli opazovani posamezniki, če ne bi imeli raka

Relativno preživetje

Podatki o bolnikih z rakom

- **Opazovani bolniki: Čas do smrti ali krnjenja; populacija: tveganje za smrt glede na demografske dejavnike**



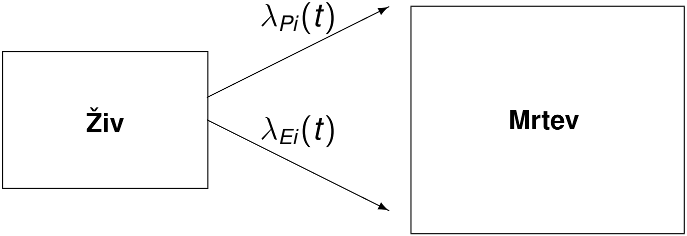
$\lambda_{Pi}(t)$ predstavlja ogroženost, ki bi jo imeli opazovani posamezniki, če ne bi imeli raka

$$\lambda_{Oi}(t) \quad = \quad \lambda_{Pi}(t) + \lambda_{Ei}(t)$$

Relativno preživetje

Podatki o bolnikih z rakom

- **Opazovani bolniki: Čas do smrti ali krnjenja; populacija: tveganje za smrt glede na demografske dejavnike**



$\lambda_{Pi}(t)$ predstavlja ogroženost, ki bi jo imeli opazovani posamezniki, če ne bi imeli raka

$$\begin{aligned}
\lambda_{Oi}(t) &= \lambda_{Pi}(t) + \lambda_{Ei}(t) \\
e^{-\int \lambda_{Oi}(t)dt} &= e^{-\int \lambda_{Pi}(t)dt} e^{-\int \lambda_{Ei}(t)dt} \\
S_{Oi}(t) &= S_{Pi}(t) S_{Ei}(t)
\end{aligned}$$

Kaj želimo oceniti?

Populacijske količine, ki nas zanimajo

- **Opazovano preživetje**
- **Verjetnost posameznega končnega dogodka (crude mortality)**
- **Čisto preživetje**
- **Kvocient opazovanega in pričakovanega preživetja (relative survival ratio)**

Opazovano preživetje

Količina, ki nas zanima

$$
\begin{aligned}
S_{Oi}(t) &= P(T > t) = e^{-\int \lambda_{Oi}(t)dt} = e^{-\int \lambda_{Pi}(t)+\lambda_{Ei}(t)dt} \\
S_{Oi}(t) &= S_{Pi}(t)S_{Ei}(t)
\end{aligned}
$$

Interpretacija

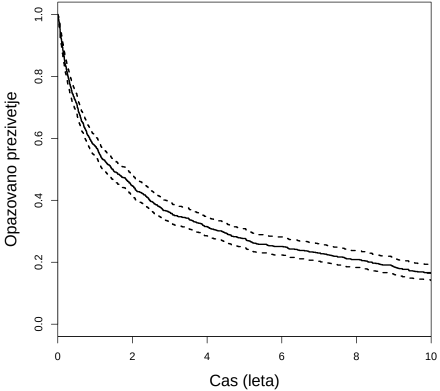
- **Delež še živih ob nekem času**
- **Ne ločimo med vzroki smrti**
- **Rezultat je odvisen od smrtnosti zaradi raka in smrtnosti iz drugih vzrokov**

Metoda ocenjevanja

Kaplan-Meier

Opazovano preživetje, primer

Mehur, diagnoza 1990-2000, starost nad 50 let, 1300 bolnikov



Verjetnost posameznega končnega dogodka (crude mortality)

Količina, ki nas zanima

$$F_C(t) = P(T \leq t, C) = \int_0^t S_O(u-)\lambda_C(u)du$$
$$= \int_0^t \exp\left\{-\int_0^u [\lambda_C(v) + \lambda_P(v)]dv\right\}\lambda_C(u)du$$

Interpretacija

- **Delež posameznikov, ki so do nekega časa umrli zaradi posameznega vzroka**
- **Upoštevamo oba vzroka smrti**
- **Rezultat je odvisen od smrtnosti zaradi raka in smrtnosti iz drugih vzrokov**

Metoda ocenjevanja

Po vzoru metode Aalen Johansen (Cronin, Feuer 2000 za diskretno merjene podatke)

Verjetnost posameznega končnega dogodka (crude mortality) - ocenjevanje

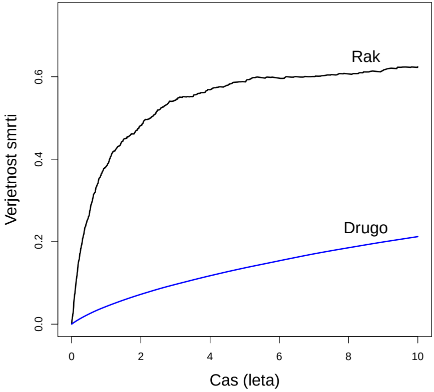
Metoda ocenjevanja

$$\hat{F}_C(t) = \int_s^t \hat{S}_O(u-)d\hat{\Lambda}_C(u)$$

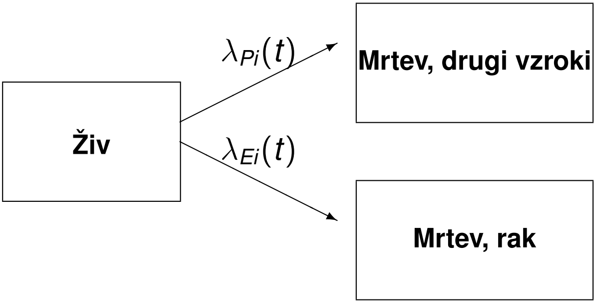
- $\hat{S}_O(u-)$: **opazovano preživetje do trenutka pred tem**
- $d\hat{\Lambda}_C(u)$: **tveganje, da v trenutku u umre zaradi bolezni,** $d\hat{\Lambda}_C(u) = d\hat{\Lambda}_O(u) - d\hat{\Lambda}_P(u)$
- **celotno (opazovano) tveganje** $d\hat{\Lambda}_O(u) = \frac{dN(u)}{Y(u)}$
- **tveganje, da posameznik v trenutku u umre iz drugih vzrokov:** $d\hat{\Lambda}_P(u) = \frac{1}{Y(u)} \sum_{i=1}^{n} Y_i(u)d\Lambda_{Pi}(u)$

Verjetnost posameznega končnega dogodka (crude mortality) - primer

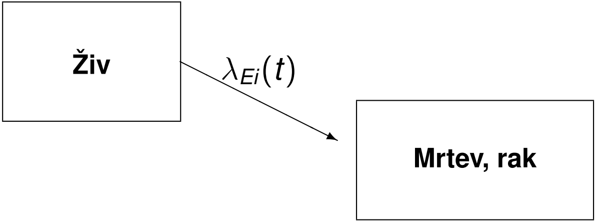
Mehur, diagnoza 1990-2000, starost nad 50 let, 1300 bolnikov



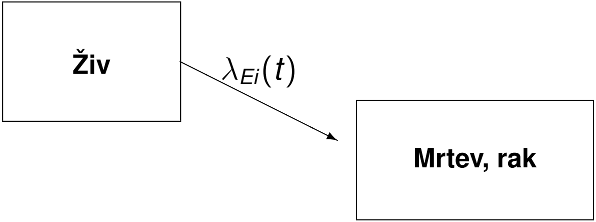
Čisto preživetje



Čisto preživetje



Čisto preživetje



Zakaj želimo ocenjevati tako hipotetično količino?

- **neodvisno od λ_P!**
- **indikator uspešnosti zdravljenja, ki ni odvisen od preživetja splošne populacije**
- **primerjava med populacijami**

Čisto preživetje

Količina, ki nas zanima

$$S_E(t) = \exp\left\{ -\int_0^t \lambda_E(u) \right\} du$$

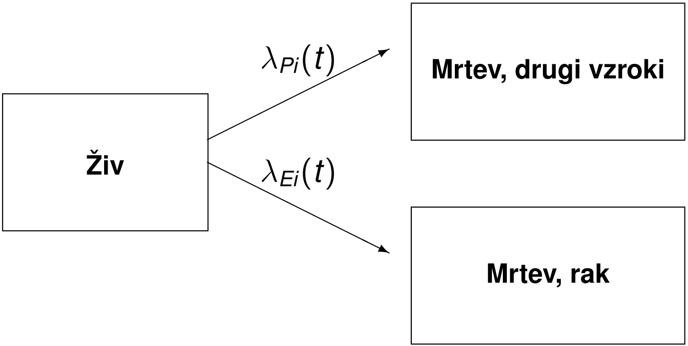
Interpretacija

- **Delež preživelih posameznikov v hipotetičnem svetu, kjer je rak edini možen vzrok smrti**
- **Upoštevamo oba vzroka smrti**
- **Rezultat je odvisen samo od smrtnosti zaradi raka**

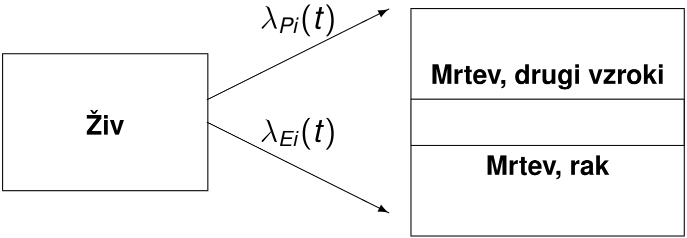
Metoda ocenjevanja

Naš predlog

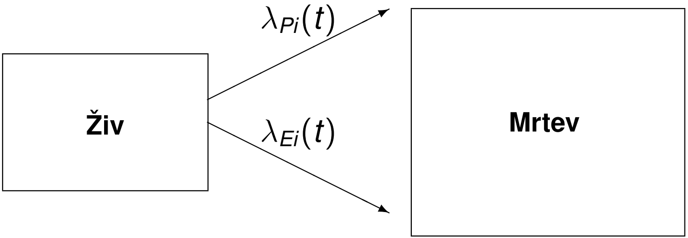
Čisto preživetje, ideja ocenjevanja



Čisto preživetje, ideja ocenjevanja

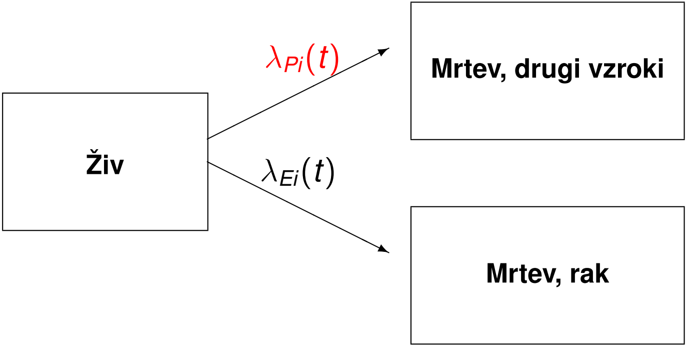


Čisto preživetje, ideja ocenjevanja



- **Dva vzroka smrti - od skupnega števila smrti odštejemo pričakovano število smrti glede na populacijske tabele**

Čisto preživetje, ideja ocenjevanja



- **Dva vzroka smrti - od skupnega števila smrti odštejemo pričakovano število smrti glede na populacijske tabele**
- **Informativno krnjenje - uporabimo uteži, da se zopet približamo pravemu številu (ki bi ga imeli v hipotetičnem svetu)**

Čisto preživetje, metoda ocenjevanja

- **Ocenimo opazovano ogroženost**

Formula

$$\frac{\sum_{i=1}^{n} dN_i(t)}{\sum_{i=1}^{n} Y_i(t)}$$

Oznake

i	**posameznik**
dN	**število dogodkov na kratkem intervalu**
Y	**at risk indikator**
$d\Lambda_P$	**prirastek populacijskega tveganja**
S_P	**populacijsko preživetje**

Čisto preživetje, metoda ocenjevanja

- **Ocenimo opazovano ogroženost**
- **čista = opazovana - pričakovana**

Formula

$$\frac{\sum_{i=1}^{n} dN_i\ (t)}{\sum_{i=1}^{n} Y_i\ (t)} - \frac{\sum_{i=1}^{n} Y_i\ (t) d\Lambda_{Pi}(t)}{\sum_{i=1}^{n} Y_i\ (t)}$$

Oznake

- i **posameznik**
- dN **število dogodkov na kratkem intervalu**
- Y **at risk indikator**
- $d\Lambda_P$ **prirastek populacijskega tveganja**
- S_P **populacijsko preživetje**

Čisto preživetje, metoda ocenjevanja

- **Ocenimo opazovano ogroženost**
- **čista = opazovana - pričakovana**
- **popravimo za tiste, ki so zapustili vzorec zaradi drugih vzrokov**

Formula

$$d\widehat{\Lambda}_E(t) = \frac{\sum_{i=1}^n dN_i^w(t)}{\sum_{i=1}^n Y_i^w(t)} - \frac{\sum_{i=1}^n Y_i^w(t) d\Lambda_{Pi}(t)}{\sum_{i=1}^n Y_i^w(t)}$$

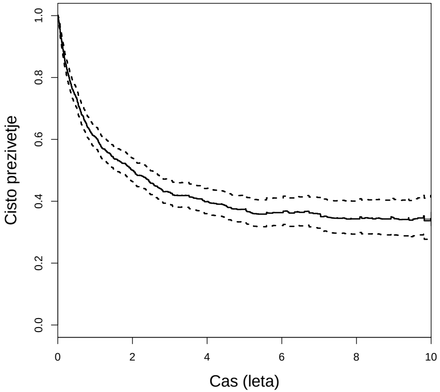
$$dN_i^w(t) = \frac{dN_i(t)}{S_{Pi}(t)}, \quad Y_i^w(t) = \frac{Y_i(t)}{S_{Pi}(t)}$$

Oznake

- i **posameznik**
- dN **število dogodkov na kratkem intervalu**
- Y **at risk indikator**
- $d\Lambda_P$ **prirastek populacijskega tveganja**
- S_P **populacijsko preživetje**

Čisto preživetje, primer

Mehur, diagnoza 1990-2000, starost nad 50 let, 1300 bolnikov



Kvocient opazovanega in pričakovanega preživetja - relative survival ratio

Količina, ki nas zanima

$$S_R(t) = \frac{S_O(t)}{S_P(t)}$$

Interpretacija

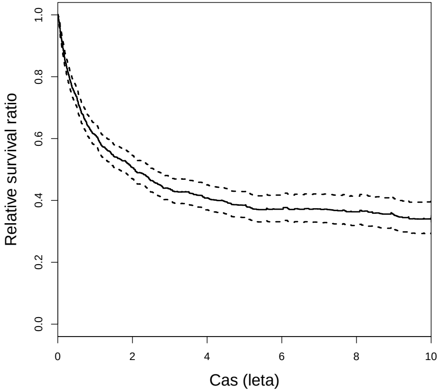
- **Preživetje bolnikov glede na preživetje celotne populacije**
- **Ne ločimo med vzroki smrti**
- **Rezultat je odvisen od smrtnosti zaradi raka in smrtnosti iz drugih vzrokov**

Metoda ocenjevanja

Kaplan-Meier za opazovano, $S_P(t) = \frac{1}{n}\sum S_{Pi}(t)$

Kvocient opazovanega in pričakovanega preživetja, primer

Mehur, diagnoza 1990-2000, starost nad 50 let, 1300 bolnikov



Primerjava konceptov

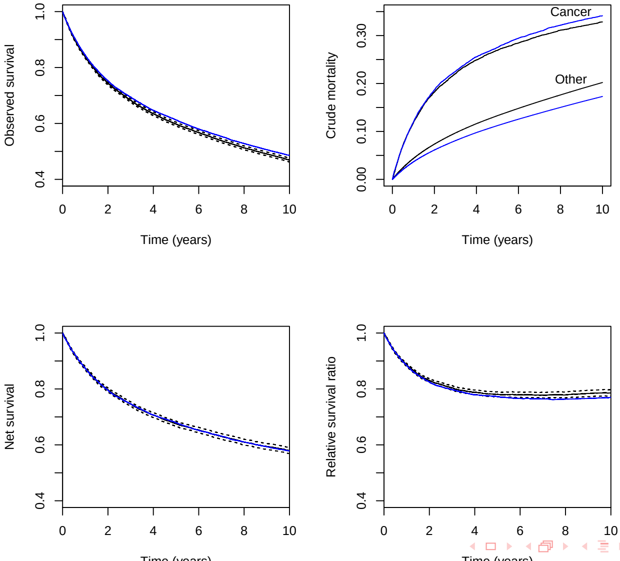
Populacijske količine, ki nas zanimajo

- **Opazovano preživetje**
- **Verjetnost posameznega končnega dogodka (crude mortality)**
- **Čisto preživetje**
- **Kvocient opazovanega in pričakovanega preživetja (relative survival ratio)**

Simulacija

- **Kohorti iz dveh obdobij: 1985-1990 (črna) in 1995-2000 (modra)**
- **Enaka struktura: starost 40-85 let, spol 50/50**
- **Enaka ogroženost za smrt zaradi raka (enak λ_E)**
- **Različna populacijska smrtnost**

Primerjava konceptov



Primerjava konceptov

Populacijske količine, ki nas zanimajo

- **Opazovano preživetje - odvisno od obeh smrtnosti**
- **Verjetnost posameznega končnega dogodka (crude mortality) - v populacijah s slabšim populacijskim preživetjem bo manj smrtnosti zaradi raka**
- **Čisto preživetje - edina mera, ki je neodvisna od populacijskega preživetja. Seveda pa nista neposredno primerljivi populaciji z različno starostno strukturo**
- **Kvocient opazovanega in pričakovanega preživetja (relative survival ratio) - edina krivulja, ki ne predstavlja verjetnosti preživetja neke skupine**

Standardno uporabljane cenilke

- **Ederer I: kvocient opazovanega in pričakovanega preživetja (ni enako kot čisto preživetje)**
- **Hakulinen: popravek Ederer I za primer informativnega krnjenja (potencialni čas opazovanja odvisen od starosti). Kvaliteta popravka vprašljiva, napačen v primeru neinformativnega krnjenja**
- **Ederer II: boljši približek čistemu preživetju kot Ederer I. Enaka ideja kot 'cause specific survival'**
- **'Cause specific survival' na podatkih z znanim vzrokom smrti: krnimo vzroke, ki nas ne zanimajo - napaka: informativno krnjenje**

Uteži

- **Ocenimo opazovano ogroženost**

Formula

$$d\widehat{\Lambda}_E(t) = \frac{\sum_{i=1}^n dN_i^w(t)}{\sum_{i=1}^n Y_i^w(t)} - \frac{\sum_{i=1}^n Y_i^w(t)d\Lambda_{Pi}(t)}{\sum_{i=1}^n Y_i^w(t)}$$

$$dN_i^w(t) = \frac{dN_i(t)}{S_{Pi}(t)}, \quad Y_i^w(t) = \frac{Y_i(t)}{S_{Pi}(t)}$$

Potrebujemo uteži za vsakega posameznika ob vsakem času

Računska zahtevnost, velike uteži

Oznake

- i **posameznik**
- dN **število dogodkov na kratkem intervalu**
- Y **at risk indikator**
- $d\Lambda_P$ **prirastek populacijskega tveganja**
- S_P **populacijsko preživetje**

Uteži

- **Ocenimo opazovano ogroženost**
- **čista = opazovana - pričakovana**

Formula

$$d\widehat{\Lambda}_E(t) = \frac{\sum_{i=1}^n dN_i^w(t)}{\sum_{i=1}^n Y_i^w(t)} - \frac{\sum_{i=1}^n Y_i^w(t) d\Lambda_{Pi}(t)}{\sum_{i=1}^n Y_i^w(t)}$$

$$dN_i^w(t) = \frac{dN_i(t)}{S_{Pi}(t)}, \quad Y_i^w(t) = \frac{Y_i(t)}{S_{Pi}(t)}$$

Oznake

- i **posameznik**
- dN **število dogodkov na kratkem intervalu**
- Y **at risk indikator**
- $d\Lambda_P$ **prirastek populacijskega tveganja**
- S_P **populacijsko preživetje**

Potrebujemo uteži za vsakega posameznika ob vsakem času

Računska zahtevnost, velike uteži

Uteži

- **Ocenimo opazovano ogroženost**
- **čista = opazovana - pričakovana**
- **popravimo za tiste, ki so zapustili vzorec zaradi drugih vzrokov**

Formula

$$d\widehat{\Lambda}_E(t) = \frac{\sum_{i=1}^n dN_i^w(t)}{\sum_{i=1}^n Y_i^w(t)} - \frac{\sum_{i=1}^n Y_i^w(t) d\Lambda_{Pi}(t)}{\sum_{i=1}^n Y_i^w(t)}$$

$$dN_i^w(t) = \frac{dN_i(t)}{S_{Pi}(t)}, \quad Y_i^w(t) = \frac{Y_i(t)}{S_{Pi}(t)}$$

Potrebujemo uteži za vsakega posameznika ob vsakem času

Računska zahtevnost, velike uteži

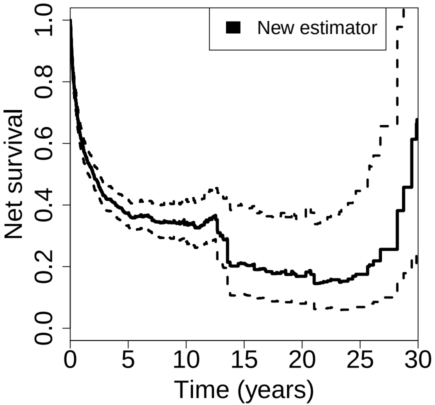
Oznake

i	**posameznik**
dN	**število dogodkov na kratkem intervalu**
Y	**at risk indikator**
$d\Lambda_P$	**prirastek populacijskega tveganja**
S_P	**populacijsko preživetje**

Čisto preživetje - (grd) primer

Mehur, kohorta 1970-1980, spremljana 30 let

- **851 posameznikov, starost 32-96**



Uteži

Definicija čistega preživetja

$$\lambda_{Oi}(t) \quad = \quad \lambda_{Pi}(t) + \lambda_{Ei}(t)$$

Uteži

Definicija čistega preživetja

$$\begin{aligned}\lambda_{Oi}(t) &= \lambda_{Pi}(t) + \lambda_{Ei}(t)\\ S_{Oi}(t) &= S_{Pi}(t)S_{Ei}(t)\end{aligned}$$

Uteži

Definicija čistega preživetja

$$\begin{aligned}
\lambda_{Oi}(t) &= \lambda_{Pi}(t) + \lambda_{Ei}(t) \\
S_{Oi}(t) &= S_{Pi}(t) S_{Ei}(t) \\
S_{Ei}(t) &= \frac{S_{Oi}(t)}{S_{Pi}(t)}
\end{aligned}$$

Uteži

Definicija čistega preživetja

$$\begin{aligned}
\lambda_{Oi}(t) &= \lambda_{Pi}(t) + \lambda_{Ei}(t) \\
S_{Oi}(t) &= S_{Pi}(t) S_{Ei}(t) \\
S_E(t) &= \frac{1}{N} \sum_{i=1}^{N} \frac{S_{Oi}(t)}{S_{Pi}(t)}
\end{aligned}$$

Uteži

Definicija čistega preživetja

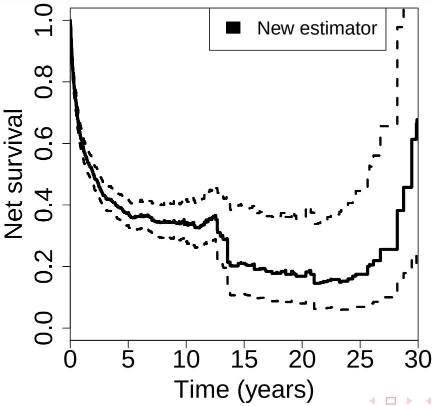
$$\begin{aligned} \lambda_{Oi}(t) &= \lambda_{Pi}(t) + \lambda_{Ei}(t) \\ S_{Oi}(t) &= S_{Pi}(t) S_{Ei}(t) \\ S_E(t) &= \frac{1}{N} \sum_{i=1}^{N} \frac{S_{Oi}(t)}{S_{Pi}(t)} \end{aligned}$$

- **uteži nastopajo že v definiciji**
- **posamezniki z majhnim S_P lahko povzročijo veliko varianco**

Čisto preživetje - primer

Mehur, kohorta 1970-1980, spremljana 30 let

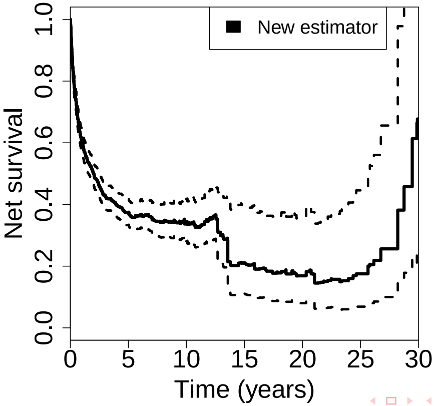
- **851 posameznikov, starost 32-96**
- **Edino krnjenje je konec spremljanja (2011)**



Čisto preživetje - primer

Mehur, kohorta 1970-1980, spremljana 30 let

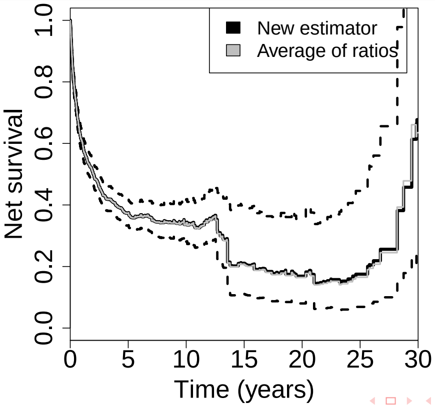
- **851 posameznikov, starost 32-96**
- **Edino krnjenje je konec spremljanja (2011)**
- **Povprečje kvocientov:** $S_E(t) = \frac{1}{N}\sum_{i=1}^{N}\frac{S_{Oi}(t)}{S_{Pi}(t)} \rightarrow \widehat{S}_E(t) = \frac{1}{n}\sum_{i=1}^{n}\frac{Y_i(t)}{S_{Pi}(t)}$



Čisto preživetje - primer

Mehur, kohorta 1970-1980, spremljana 30 let

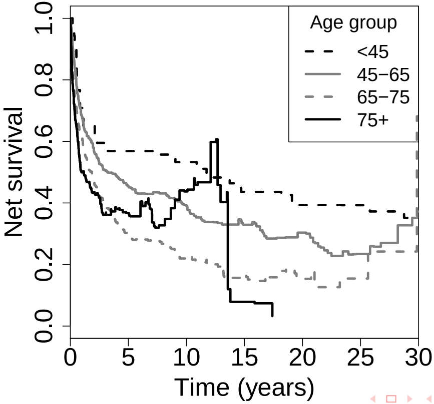
- **851 posameznikov, starost 32-96**
- **Edino krnjenje je konec spremljanja (2011)**
- **Povprečje kvocientov:** $S_E(t) = \frac{1}{N}\sum_{i=1}^{N}\frac{S_{Oi}(t)}{S_{Pi}(t)} \rightarrow \widehat{S}_E(t) = \frac{1}{n}\sum_{i=1}^{n}\frac{Y_i(t)}{S_{Pi}(t)}$



Čisto preživetje - primer

Mehur, kohorta 1970-1980, spremljana 30 let

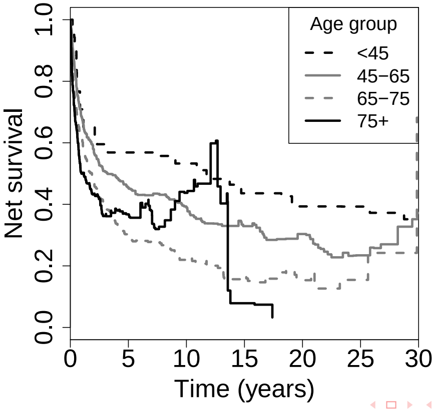
- **851 posameznikov, starost 32-96**
- **Edino krnjenje je konec spremljanja (2011)**
- **Povprečje kvocientov:** $S_E(t) = \frac{1}{N}\sum_{i=1}^{N}\frac{S_{Oi}(t)}{S_{Pi}(t)} \to \widehat{S}_E(t) = \frac{1}{n}\sum_{i=1}^{n}\frac{Y_i(t)}{S_{Pi}(t)}$



Uteži

Vzrok za skoke

- **8 let: v raziskavi še 20 posameznikov 75+, od tega trije 80+**
- **8-12.5: umre manj posameznikov kot pričakovano (v populaciji)**
- **12.5 let: v raziskavi 10 posameznikov 75+, od tega 2 nad 80**
- **12.5 let: najstarejša imata 10x večjo utež kot ostali**



Uteži

Posameznik ima veliko utež:

- **Verjetnost smrti iz drugih vzrokov do sedaj je zanj velika**
- **Živih je le še malo njegovih vrstnikov**
- **Predstavlja vrstnike - njegova utež je velika**
- **Skušamo sklepati na podlagi zelo majhnega števila posameznikov - standardna napaka bo velika**

Dolgoročno preživetje kohorte

Podskupina z majhno verjetnostjo populacijskega preživetja S_P

- **Lahko povzročajo veliko variabilnost ocene dokler so še živi**
- **Kaj pa se zgodi, ko izgubimo vse?**

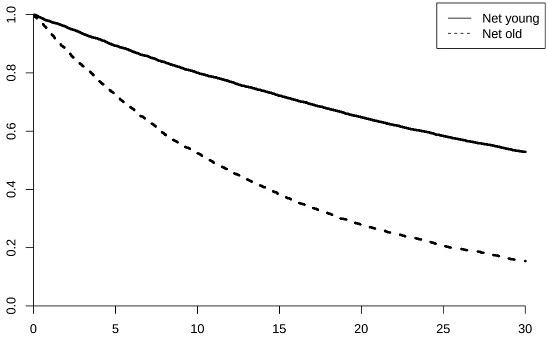
Težava!

- **Čisto preživetje je definirano v hipotetičnem svetu**
- **V našem svetu - ob nekem trenutku vsi umrejo**
- **Ne moremo ocenjevati, kaj se dogaja v hipotetičnem svetu po tem trenutku**
- **Naša ocena: predpostavljamo, da je čisto preživetje te skupine v nadaljevanju enako preživetju ostalih**

Dolgoročno čisto preživetje kohorte

skupina 85+ čez 20 let?

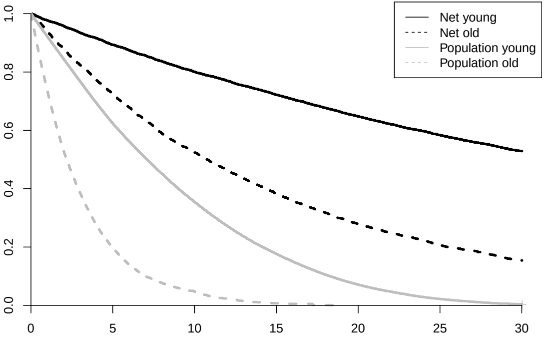
Čisto preživetje cele kohorte je povprečje podskupin



Dolgoročno čisto preživetje kohorte

skupina 85+ čez 20 let?

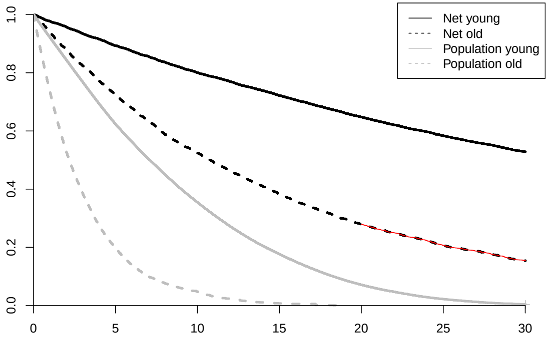
Čisto preživetje cele kohorte je povprečje podskupin



Dolgoročno čisto preživetje kohorte

skupina 85+ čez 20 let?

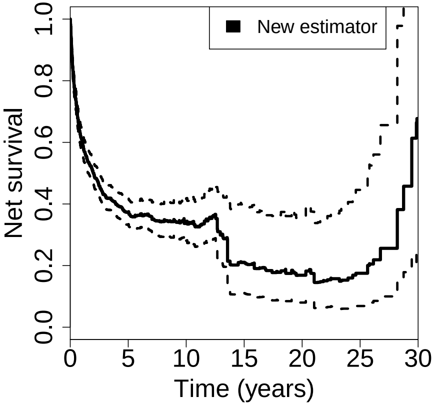
Čisto preživetje cele kohorte je povprečje podskupin



Dolgoročno čisto preživetje kohorte

Poročajmo dolgoročno čisto preživetje le za podskupine

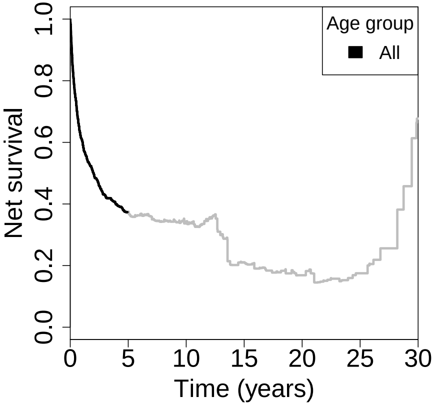
- **Nas res zanima 20-letno čisto preživetje za 80 let stare bolnike?**



Dolgoročno čisto preživetje kohorte

Poročajmo dolgoročno čisto preživetje le za podskupine

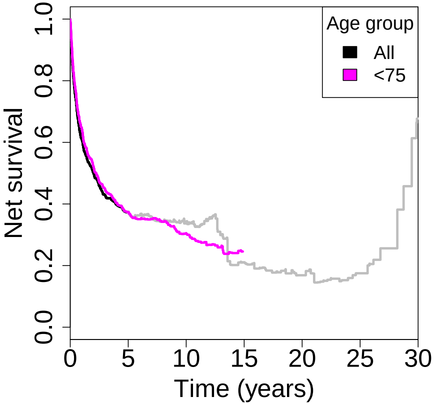
- **Nas res zanima 20-letno čisto preživetje za 80 let stare bolnike?**



Dolgoročno čisto preživetje kohorte

Poročajmo dolgoročno čisto preživetje le za podskupine

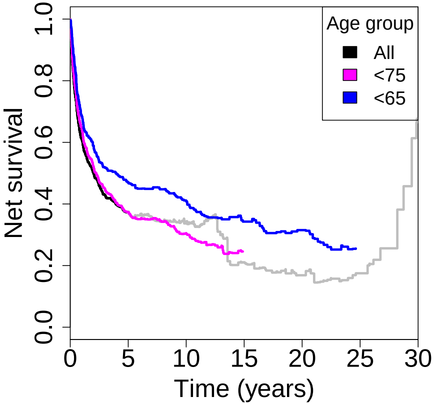
- **Nas res zanima 20-letno čisto preživetje za 80 let stare bolnike?**



Dolgoročno čisto preživetje kohorte

Poročajmo dolgoročno čisto preživetje le za podskupine

- **Nas res zanima 20-letno čisto preživetje za 80 let stare bolnike?**



Zaključki

Zanimajo nas lahko različne količine

- **Opazovano preživetje**
- **Verjetnost posameznega končnega dogodka (crude mortality)**
- **Čisto preživetje**
- **Kvocient opazovanega in pričakovanega preživetja (relative survival ratio)**

Čisto preživetje

- **Velika varianca ni posledica metode ocenjevanja - problematična je definicija čistega preživetja**
- **Težave pri standardiziranem poročanju**
- **Nesmiselna vprašanja lahko dobijo nesmiselne odgovore**

Bibliography

Pohar Perme M., Stare J., Estève J.
On estimation in relative survival
Biometrics, 2012

Danieli C., Remontet L., Bossard N., Roche L., Belot A.
Estimating net survival: the importance of allowing for informative censoring
Statistics in Medicine, 2012

Cronin, K. A., Feuer, E. J.
Cumulative cause-specific mortality for cancer patients in the presence of other causes: a crude analogue of relative survival,
Statistics in Medicine, 2000

Programje

R paket `relsurv`, na CRAN-u, funckiji `rs.surv`, `cmp.rel`