9. Metode izglađivanja

9.1 Izglađivanje parametarskom formulom

Sirovim procjenama prilagođavamo *funkcije dobi* koje ovise o *konačno* mnogo *parametara*.

Uobičajeni modeli:

$$\mu_x = B \cdot C^x \qquad \text{(Gompertzov)}$$
$$\mu_x = A + B \cdot C^x \qquad \text{(Makehamov)}$$

Primjer. Tablice smrtnosti osiguranika životnih osiguranja u UK (CMIB data 1979-82) koriste formulu:

$$\mu_x = \text{Polinom}_1(x;\mathbf{a}) + \exp(\text{Polinom}_2(x;\mathbf{b}))$$

za koje su Gompertzov i Makehamov model specijalni slučajevi.

Pretpostavka: Sirove stope su dobivene pomoću Poissonovog modela, a parametri $\mathbf{a}$ i $\mathbf{b}$ su procijenjeni metodom maksimane vjerodostojnosti.

Napomena 1. Za praktičnu upotrebu izglađenih vrijednosti statistički kriteriji odabira formule nisu dovoljni
$\rightarrow$ potrebno je ispitati rezultat i sa stanovišta *poznatih činjenica o ponašanju smrtnosti*

Npr. da je:
– smrtnost muškaraca veća nego žena,
– smrtnost osiguranika životnih osiguranja manja nego u općoj populaciji,
– smrtnost osoba koje su nedavno kupile policu životnog osiguranja manja od onih koji ju dugo imaju

$\rightarrow$ potrebno je ispitati rezultat i sa stanovišta *mogućeg financijskog rizika*

(problem potcijenjene ili precijenjene smrtnosti)

Napomena 2. Izglađene tablice smrtnosti koriste se za primjenu na *buduću* smrtnost, a baziraju se na *prošlom* iskustvu smrtnosti.
$\rightarrow$ Dakle, potrebno je znati *trend* smrtnosti.

Npr., ako se smrtnost smanjuje (u stabilnim i ekonomski razvijenim društvima), tada je nužno *projicirati* buduće smanjenje smrtnosti (to je problem za mirovinska osiguranja).

Koraci u procesu odabira modela/izglađivanja:

1. Odabir jedne ili više različitih formula za izglađivanje (predanalizom podataka, iz prethodnog/očekivanog iskustva,...)

2. Odabir statističkog modela i metode procjene parametara (ML, LS, minimu χ^2,...)

3. Testirati prilagođeni model pomoću:
– statističkih testova
– poželjnih svojstava obzirom na primjenu

$\rightarrow$ odabire se najbolji model

9.2 Izglađivanje pomoću standardnih tablica

– kada imamo podataka iz nekoliko (*malo*) različitih dobi x_1,..., x_m
– kada je razumno pretpostaviti da se smrtnost izučavane grupe *ponaša kao* neka populacija za koju postaje *standardne tablice smrtnosti*

$\rightarrow$ tada sirove stope možemo izgladiti pomoću tih standardnih tablica smrtnosti

– *koja formula*?

9.3 Grafičko izglađivanje

– radi se o *gotovo prostoručnom* povlačenju odgovarajuće glatke krivulje kroz točke $(x, \widehat{q}_x)$ ili $(x, \widehat{\mu}_x)$ ili u ovisnosti o nekim standardnim veličinama

$\rightarrow$ mana: nepreciznost, prednost: brzina

Napomene.

(a) stope smrtnosti treba prikazivati u logaritamskoj skali (jer se kreću u rasponu od $< 10^{-3}$ do $> 10^{-1}$)

(b) ako su podaci rijetki, treba ih grupirati u dobne razrede

(c) glatka krivulja mora prolaziti kroz barem $0.95m$ intervala pouzdanosti 95% za m veličina koje izglađujemo

10. Izloženost riziku

Problem:
Kako izračunati E_x^c (i E_x) iz nepotpunih podataka?
Posebno iz *cenzusnih* podataka?

10.2 Homogenost

– do sada: pret. *jednake distribuiranosti* podataka
$\rightarrow$ u praksi nemoguće postići
$\rightarrow$ podjela u podgrupe ovisno o **faktorima** koji utječu na smrtnost

Problem:
– dostupnost informacija za klasifikaciju
– podjelom se podaci razrjeđuju

Neki od faktora koji utječu na smrtnost:

1. spol
2. dob (jedino smo do sada i koristili)
3. tip police
4. pušač?
5. koliko je dugo polica na snazi (tablice s odabirom)
6. razina preuzetog rizika
7. način prodaje police
8. veličina police
9. zanimanje osiguranika
10. poznati štetni faktori

10.3 Princip suglasnosti

– bitno za određivanje jesu li D_x i E_x^c opažani *obzirom na isti dobni interval*

Princip suglasnosti: Za osobu koju opažamo živu u trenutku t kažemo da je dobi x akko bi se njena smrt u trenutku t uračunala u statistiku D_x.

10.3 Egzaktno računanje E_x^c

Potrebni podaci o osiguraniku:

1. datum rođenja
2. datum ulaska u razdoblje opažanja
3. datum izlaska iz opažanja
4. razlog izlaska iz opažanja (za izračun D_x)

Što ako se neki od podataka ne bilježe?

10.5 Cenzusna aproksimacija E_x^c

Npr. podaci su (CMIB):

d_x...ukupan broj umrlih dobi x na *zadnji* rođendan u kalendarskim godinama $K, K+1, \ldots, K+N$

$P_x(t)$...broj osoba koje se opažaju i koje su žive u dobi x na *zadnji* rođendan na dan t ($t =$ 1.1. u god. $K, K+1, \ldots, K+N, K+N+1$)

$\rightarrow$ t je datum popisa (ili cenzusa)

Aproksimacija *trapeznom formulom*:

$$\begin{aligned} E_x^c &= \int_{1.1.K.}^{1.1.(K+N+1).} P_x(t)\,dt \approx \\ &\approx \sum_{t=1.1.K.}^{1.1.(K+N).} \frac{1}{2}(P_x(t) + P_x(t+1)) = \\ &= \frac{1}{2}P_x(t_0) + \sum_{i=1}^{N} P_x(t_i) + \frac{1}{2}P_x(t_{N+1}) \end{aligned}$$

$t_i \equiv 1.1.(K+i)$ za $i = 0, 1, \ldots, N+1$

10.6 Definicije dobi

def. dobi x	interval stope	$\widehat{q}_{x+f}$	$\widehat{\mu}_{x+f}$
na zadnji rođ.	$[x, x+1]$	$\widehat{q}_x$	$\widehat{\mu}_{x+\frac{1}{2}}$
na najbliži rođ.	$[x-\frac{1}{2}, x+\frac{1}{2}]$	$\widehat{q}_{x-\frac{1}{2}}$	$\widehat{\mu}_x$
na sljedeći rođ.	$[x-1, x]$	$\widehat{q}_{x-1}$	$\widehat{\mu}_{x-\frac{1}{2}}$

interval stope...dobni interval na koji se def. dobi x odnosi

Ako se definicije dobi za d_x i $P_x(t)$ razlikuju, **popisne podatke $P_x(t)$ treba prilagoditi – d_x se nikada ne prilagođava!**

10.7 Interval stope obzirom na kalendarsku godinu

– kada u podacima imamo *samo* godinu rođenja i godinu smrti

d_x... broj umrlih koji su bili dobi x u kalendarskoj godini smrti

$P_x(t)$... broj opažanih osoba koje su dobi x na sljedeći rođendan na datum cenzusa $t = 1.1.$

$$E_x^c \approx \sum_{t=1.1.K.}^{1.1.(K+N).} \frac{1}{2}(P_x(t) + P_{x+1}(t+1))$$

Procjena: $\widehat{q}_{x-\frac{1}{2}}$, $\widehat{\mu}_x$

10.8 Interval stope obzirom na osigurateljnu godinu

osigurateljne godine = cjelobrojna trajanja police

– kada u podacima imamo *samo* dob na zadnji rođendan kada je polica izdana i trajanje police na dan smrti (u god.)

$d_x \ldots$ broj umrlih dobi x na zadnji rođendan u osig. godini koja prethodi ili se podudara s datumom smrti

$P_x(t) \ldots$ broj opažanih osoba koje su dobi x na zadnji rođ. na godišnjicu police koja prethodi ili se podudara s datumom cenzusa $t = d.m.$

$$E_x^c \approx \sum_{t=d.m.K.}^{d.m.(K+N).} \frac{1}{2}(P_x(t) + P_x(t+1))$$

Procjena: $\widehat{q}_{x+\frac{1}{2}}$, $\widehat{\mu}_{x+1}$

Primjer. U tablici se nalaze podaci o osiguranicima dobi x na zadnji rođendan, na dane 1.1.1999., 1.1.2000. i 1.1.2001. Broj umrlih je klasificiran obzirom na dob na najbliži rođendan u trenutku smrti. Prikazani podaci su dio istraživanja smrtnosti osiguranika osiguranih za slučaj smrti nekog OD.

Neka su: $t_0 = 1.1.1999.$, $t_1 = 1.1.2000.$, $t_2 = 1.1.2001.$

x	$P_x(t_0)$	$P_x(t_1)$	$P_x(t_2)$	$d_x(1999)$	$d_x(2000)$
40	473	512	491	17	18
41	450	470	482	20	18
42	490	460	480	21	19

Pretpostavljamo da su intenziteti smrtnosti za sve dobi konstantni u razdoblju $[t_0, t_2]$.

(i) Izvedite formulu za procjenu intenziteta smrtnosti za osobe dobi x na najbliži rođendan.
(ii) Pomoću dobivene formule procijenite μ_{41} i μ_{42}.

Rješenje:

(i):

D_x... broj umrlih dobi x na 'najbliži rođendan'

$P_x(t)$... broj o. dobi x na 'zadnji rođendan' na dan t

$\Rightarrow$ popisne podatke treba prilagoditi

$\rightarrow$

$P'_x(t)$... broj o. dobi x na 'najbliži rođ.' na dan t

$\Rightarrow$ uz pretp.: rođ. su uniformno distr. duž godine:

$$P'_x(t) = \frac{1}{2}(P_{x-1}(t) + P_x(t))$$

Po trapeznoj formuli:

$$
\begin{aligned}
E_x^c &= \frac{1}{2}P_x'(t_0) + P_x'(t_1) + \frac{1}{2}P_x'(t_2) = \\
&= \frac{1}{4}(P_{x-1}(t_0) + P_x(t_0)) + \frac{1}{2}(P_{x-1}(t_1) + P_x(t_1)) + \\
&\quad + \frac{1}{4}(P_{x-1}(t_2) + P_x(t_3))
\end{aligned}
$$

$\Rightarrow$

$$
\widehat{\mu}_x = \frac{d_x(1999) + d_x(2000)}{E_x^c}
$$

Interval stope je $[x - \frac{1}{2}, x + \frac{1}{2}]$.

(ii): Iz tablice slijedi:

$$\begin{aligned} E^c_{41} &= \frac{1}{4}(473+450)+\frac{1}{2}(512+470)+ \\ &\quad +\frac{1}{4}(491+482)=965 \\ E^c_{42} &= \frac{1}{4}(450+490)+\frac{1}{2}(470+460)+ \\ &\quad +\frac{1}{4}(482+480)=940.5 \end{aligned}$$

Dakle,

$$\widehat{\mu}_{41}=\frac{20+18}{965}=0.0394,\quad \widehat{\mu}_{42}=\frac{21+19}{940.5}=0.0425$$

11. Definicija i procjena funkcija smrtnosti s odabirom

– intenzitet smrtnosti ovisi o dobi i *vremenu članstva u grupi*

$\mu_{[x]+r}$... intenzitet smrtnosti osoba koje su bile dobi x na dan pristupa grupi i u grupi su r god.

Privremena početna selekcija:

$s \ldots$ *period odabira*

Tablica:

$$\begin{array}{lllll}
\mu_{[x]} & \mu_{[x]+1} & \cdots & \mu_{[x]+s-1} & \mu_{x+s} \\
\mu_{[x+1]} & \mu_{[x+1]+1} & \cdots & \mu_{[x+1]+s-1} & \mu_{x+1+s} \\
\vdots & \vdots & & \vdots & \vdots
\end{array}$$

Procjena veličina s odabirom:

$d_{[x],r}$... broj umrlih osoba pristupne dobi x na najbliži rođendan, a u trenutku smrti su bile članovi grupe punih r godina

$E^c_{[x],r}$...ukupno rezidualno vrijeme čekanja osoba pristupne dobi x koje su punih r godina bile članovi grupe

$$\widehat{\mu}_{[x]+r} = \frac{d_{[x],r}}{E^c_{[x],r}}$$

– nakon što se odredi period odabira s:

$$\widehat{\mu}_{x+s} = \frac{d_{[x],s} + d_{[x-1],s+1} + \cdots}{E^c_{[x],s} + E^c_{[x-1],s+1} + \cdots}$$

11.3 Konstrukcija tablica s odabirom i krajnjih tablica

Izglađene stope:

$$\begin{array}{lllll}
q_{[x]} & q_{[x]+1} & \cdots & q_{[x]+s-1} & q_{x+s} \\
q_{[x+1]} & q_{[x+1]+1} & \cdots & q_{[x+1]+s-1} & q_{x+1+s} \\
\vdots & \vdots & & \vdots & \vdots
\end{array}$$

za dobi $x = \alpha, \alpha + 1, \ldots$

1. odaberemo korijen $\ell_{[\alpha]}$
2. rekurzivno računamo prvi redak tablice:

$$\left.\begin{array}{lcl} \ell_{[\alpha]+r} & = & \ell_{[\alpha]+r-1}(1-q_{[\alpha]+r-1}) \\ d_{[\alpha]+r-1} & = & \ell_{[\alpha]+r-1}-\ell_{[\alpha]+r} \end{array}\right\}, r=1,2,\ldots,s-1$$

$$\begin{array}{lcl} \ell_{\alpha+s} & = & \ell_{[\alpha]+s-1}(1-q_{[\alpha]+s-1}) \\ d_{[\alpha]+s-1} & = & \ell_{[\alpha]+s-1}-\ell_{\alpha+s} \end{array}$$

3. rekurzivno računamo zadnji (krajnji) stupac:

$$\ell_{x+1+s}=\ell_{x+s}(1-q_{x+s}),\;\; x=\alpha,\alpha+1,\ldots$$

4. rekurzivno se računaju svi stupci prema naprijed:

$$\left.\begin{array}{rcl} \ell_{[x]+s-1} & = & \frac{\ell_{x+s}}{1-q_{[x]+s-1}} \\ d_{[x]+s-1} & = & \ell_{[x]+s-1} - \ell_{x+s} \end{array}\right\}, \quad x = \alpha+1, \alpha+2, \ldots$$

pa za $r = s-1, s-2, \ldots, 1$,

$$\left.\begin{array}{rcl} \ell_{[x]+r-1} & = & \frac{\ell_{[x]+r}}{1-q_{[x]+r-1}} \\ d_{[x]+r-1} & = & \ell_{[x]+r-1} - \ell_{[x]+r} \end{array}\right\}, \quad x = \alpha+1, \alpha+2, \ldots$$

11.4 Upotreba tablica smrtnosti s odabirom

$$
\begin{aligned}
{}_{n|m}q_{[x]+r} &= \frac{\ell_{[x]+r+n} - \ell[x] + r + n + m}{\ell_{[x]+r}} \\
{}_{n|}q_{[x]+r} &\equiv {}_{n|1}q_{[x]+r} \\
{}_{n}q_{[x]+r} &= \frac{\ell_{[x]+r} - \ell[x] + r + n}{\ell_{[x]+r}} = 1 - {}_{n}p_{[x]+r}
\end{aligned}
$$

Primjer. Tablica A 1967/70 ($s = 2$)

Uvjetna vjerojatnost da osoba pristupne dobi 50 ne doživi dob 51 je:

$$
\begin{array}{rlcll}
(i) & q_{[50]} & = & 0.00286243 & \text{za } \textit{novog} \text{ pristupnika} \\
(ii) & q_{[49]+1} & = & 0.00351979 & \text{za } \textit{lanjskog} \text{ pristupnika} \\
(ii) & q_{48+2} & = & 0.00478880 & \text{za } \textit{starog} \text{ osiguranika.}
\end{array}
$$