

## KAPITOLA 6

# Štandardizácia údajov

### *Ciele kapitoly*

### *Niečo z histórie*

### *Skreslenie a vychýlenie*

### *Metódy štandardizácie*

### *Očakávaná dĺžka života*

### *Súhrn*

### *Literatúra*

## Ciele kapitoly

Často je potrebné porovnať úmrtnosti v dvoch či viacerých populáciách, povedzme dvoch krajinách, alebo okresoch či mestách. Vieme však, že jedna z nich je podstatne mladšia, je tam viac detí a mladých ľudí, ako druhá, kde je zas málo detí a veľa starých ľudí. Je logické, že tam, kde je viac starých ľudí, bude úmrtnosť vyššia. Hrubá úmrtnosť odráža rovnako rozdiely spôsobené samotnou úmrtnosťou, ako aj rozdiely vo vekovom zložení obyvateľstva. Spôsob, ako riešiť tento problém, sa nazýva štandardizácia a často sa používa pri porovnávaní ukazovateľov zdravia. Porovnávanie úmrtností bolo základnou motiváciou zavedenia štandardizácie. Hoci je možné štandardizovať vzhľadom na rôzne premenné, ktoré môžu vplyvať na jej vývoj, vek sa najčastejšie používa na štandardizáciu úmrtnosti pre jeho vplyv na chorobnosť a úmrtnosť.

## Niečo z histórie

V polovici 19. storočia si odborníci na zdravie verejnosti v Anglicku začali uvedomovať, že hrubé miery sú nevhodné na porovnávanie ukazovateľov zdravia obyvateľov v skupinách, kde vekové rozdelenie bolo výrazne odlišné. Diskusia sa viedla okolo vývoja indexu úmrtnosti, ktorý by nebol ovplyvnený rozdielmi vo veku. V prednáške pre Štatistickú spoločnosť v Londýne pán Edwin Chadwick, ktorého poznáme ako jedeného z prvých reformátorov zdravia verejnosti v Anglicku, navr-

Veková skupina	Európska populácia	SZO populácia
0-4	8.00	8.86
5-9	7.00	8.69
10-14	7.00	8.60
15-19	7.00	8.47
20-24	7.00	8.22
25-29	7.00	7.93
30-34	7.00	7.61
35-39	7.00	7.15
40-44	7.00	6.59
45-49	7.00	6.04
50-54	7.00	5.37
55-59	6.00	4.55
60-64	5.00	3.72
65-69	4.00	2.96
70-74	3.00	2.21
75-79	2.00	1.52
80-84	1.00	0.91
85+	1.00	0.63
<b>Total</b>	<b>100.00</b>	<b>100.00</b>

**Tabuľka 1** Štandardné populácie pre európsku populáciu, ako aj pre všetky krajiny sveta odporúčaný SZO [3]

hol použitie ukazovateľa „stredný vek úmrtia“, ako nástroj porovnania zdravotného stavu rôznych častí Londýna. Tento index podľa neho predstavoval skutočný pohľad na vekovo špecifické riziko umrieť. [1] Neison, ktorý sa zaoberal poistnou matematikou, nesúhlasil s jeho výpočtami a dokázal jeho omyl. Uviedol aj riešenie, ktoré dodnes označujeme tak, ako to on navrhol, teda priama a nepriama štandardizácia, rovnako ako termín štandardná populácia.

Už v roku 1883 v správe o populácii v Anglicku bola použitá metóda priamej štandardizácie podľa Neisona a ako štandard boli použité výsledky sčítania ľudu v Anglicku a Walese z roku 1881. V ďalších správach bol štandard vypočítaný pre každé nové sčítanie ľudu, teda raz za desať rokov. Tieto časté zmeny štandardu boli ťažkopádne, pretože historické bolo potrebné prepočítať vždy s cieľom posúdiť súčasné trendy. Ako riešenie bolo nakoniec prijaté sčítanie ľudu z roku 1901 ako všeobecný štandard v Anglicku a Walese. Ten zostal nezmenený, aj keď už boli k dispozícii novšie sčítania. S cieľom uľahčiť porovnanie s úmrtnosťou v Anglicku a Walese, Spojené štáty prijali v roku 1901 britskú normu. Táto prax pokračovala až do skorých 40. rokov, keď bolo rozhodnuté, že rozdiely medzi počtom obyvateľov v USA v tom čase a anglickou populáciou z roku 1901 boli natoľko významné, že oprávňovali k zmene štandardu. [2] Už v roku 1892 Ogle odporúčal zaviesť medzinárodný štandard na

porovnávanie. Od tých čias boli navrhované viaceré štandardy, ale žiaden z nich sa všeobecne neujal. Únia proti rakovine prišla v roku 1965 s tromi štandardami. SZO používalo všetky tri štandardy, až kým nebol navrhnutý a prijatý jeden nový. [3] V tabuľke 1 sú hodnoty štandardu pre európsku populáciu, ako aj pre všetky krajiny sveta. Pri štúdiu zdravia v populácii sa stretáme so situáciami, ktoré môžu skreslovať vzťahy medzi premennými, či zanášať systematickú chybu. V oboch prípadoch je jedným z riešení použitie niektorej z metód štandardizácie. V ďalšom si tieto situácie preberieme podrobnejšie.

## Skreslenie a vychýlenie

Niekedy sa stane, že pozorujeme príčinný vzťah medzi premennými, ktorý v skutočnosti nie je prítomný. Premenné, ktoré skresľujú vzťahy medzi ďalšími premennými, ktoré sú predmetom nášho skúmania, sa nazývajú *mätúce premenné (confounders)* [4]. Aby mohla byť premenná mätúcou musí byť vo vzťahu k ochoreniu alebo k stavu, o ktorý sa zaujímame a tiež k faktoru rizika, ktorý študujeme. [5] Pokiaľ sa vychýlenie neodstráni, sú výsledky vychýlené. Preberme si podrobnejšie tieto procesy, keďže sa s nimi často môžeme stretnúť v odbornej literatúre a nie vždy je jasný aj ich plný obsah.

### Bias, Vychýlenie, Skreslenie

V štatistike považujeme proces za vychýlený, ak nie sú všetky možné výstupy rovnako pravdepodobné; v takom prípade hovoríme, že máme vychýlku od najpravdepodobnejšieho výstupu. Keď hádzeme kockou, automaticky predpokladáme, že pri veľkom počte hodov je rovnaká pravdepodobnosť padnutia každej strany kocky. Hod kockou je nespravodlivý, ak nie sú všetky dosiahnuté skóre rovnako možné. Ak po mnohých hodoch kockou zaznamenáme viac trojok alebo päťok a veľmi málo iných výsledkov, môžeme predpokladať, že kocka bola vychýlená – biased – alebo sfaľšovaná. Slovník epidemiológie definuje bias spolu s príčinami vychýlenia: *Odchýlka výsledkov alebo záverov od skutočnosti alebo procesy, ktoré vedú k takejto odchýlke. Lubovoľný trend pri zbere, analýze, interpretácii, publikácii alebo overovaní údajov, ktoré môžu viesť k záverom, ktoré sa systematicky líšia od skutočnosti. Medzi postupmi, v ktorých sa môže vyskytnúť odchýlka od skutočnosti, patria nasledujúce: 1. Systematická (jednostranná) odlišnosť merania od skutočnej hodnoty (synonymum systematická chyba). 2. Variácie štatistických súhrnných mier (priemery, proporcie, miery asociácie atď.), od ich skutočných hodnôt v dôsledku systematickej odchýlky meraní, iné slabiny v zbere dát, alebo nedostatky v návrhu štúdie alebo analýzy. 3. Odchýlka záverov od skutočnosti v dôsledku nedostatkov v návrhu štúdie, zbere údajov alebo v analýze alebo interpretácii výsledkov. 4. Tendencia postupov (návrh štúdie, zber dát, analýza, interpretácia, hodno-*

tenie, alebo publikácia) prinášať výsledky alebo závery, ktoré sa odchyľujú od skutočnosti. 5. Zaujatosť vedúca k vedomému alebo nevedomému výberu postupov štúdia, ktoré sa líšia od skutočnosti v určitom smere alebo smerujú k jednostrannosti pri interpretácii výsledkov. Výraz vychýlenie, skreslenie, nemusí nutne niesť predpoklad zaujatosti alebo iného subjektívneho faktora, akým by mohlo byť želanie konkrétneho výsledku experimentátorom. Tým sa líši od hovorového použitia, v ktorom zaujatosť odkazuje na hľadisko (uprednostnenie) určitej strany. [6]

Slovník uvádza veľké množstvo zdrojov a príčin vychýlenia. V oblasti klinického výskumu sa nimi podrobne zaoberá klasická práca Sacketta. [7] V oblasti epidemiológie ich popisuje napríklad Rodriguez a Llorca [8]. História spoznávania skreslenia v epidemiológii je veľmi zaujímavá a dobre ilustruje vývoj myslenia ľudstva. [9] Pre potreby tejto publikácie si uvedieme len niekoľko z nich, tie, ktoré musíme mať na pamäti pri práci s údajmi z rutinne zbieraných štatistík.

### **Skreslenie detekciou**

Skreslenie detekciou je spôsobené systematickou chybou v metódach zisťovania, diagnostiky alebo overenia prípadov v epidemiologickej štúdií. Príkladom je overenie diagnózy pomocou laboratórnych testov v nemocnici a mimo nej v dôsledku neuplatnenia rovnakých testov. Môže to byť uvádzanie počtov chorých na tuberkulózu na základe skriningového testu, ktorý sa používal v teréne a zároveň sa započítali prípady v nemocnici, ktoré boli overené laboratórne.

### **Skreslenie dostupnosťou**

Dostupnosť zdravotnej a zdravotníckych služieb nie je zvyčajne teritoriálne rovnaká. Prístup môže byť ovplyvnený napríklad záujmom zdravotníckych pracovníkov o určitý druh prípadov (skreslenie popularitou), keď si zariadenie vyberá určitých pacientov, napríklad pre skúšania lieku. Niekedy dostupnosť ovplyvní prestíž niektorých pracovišť či lekárov alebo sestier (skreslenie dostredivé). Iným prípadom je situácia, keď ťažké prípady sú odsunuté na hierarchicky vyššiu úroveň starostlivosti (skreslenie filtrom) a podobne.

### **Skreslenie hlásením**

Skreslenie pri hlásení nejakého stavu môže vzniknúť napríklad tým, že pýtaní sa snažia „vyhovieť“ výskumníkovi a odpovedať v zmysle, aký si predstavujú, že je pre pýtajúceho sa vhodný či očakávaný (skreslenie z podlízavosti). Citlivé otázky, ktoré môžu priviesť pýtaného do rozpakov alebo ho sociálne zraniť môžu odmietnuť (neprijateľné choroby/expozície), napríklad otázka na alkohol alebo intímne správanie sa.

### **Skreslenie klasifikáciou**

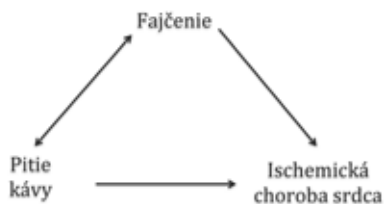
Skreslenie klasifikáciou vzniká často pri nejednotnej interpretácii klasifikačných

kritérií. Napríklad keď neskúsený lekár alebo sestra posudzujú stav pacienta a klasifikujú ho podľa určitej schémy, napríklad stav vedomia či stupeň poranenia. Ich hodnotenie sa môže líšiť od hlásenia skúseného kolegu či kolegyně. Táto situácia vzniká aj vtedy, keď nie je dohoda (alebo sa nedodriava) o pravidlách klasifikácie. Napríklad najčastejšie sa ako príčina smrti uvádza Srdcové zlyhanie (MKCH I50) a nie jeho najčastejšia príčina Hypertenzná choroba (MKCH I11), hoci samotná klasifikácia na to upozorňuje. Výsledkom je potom podcenený výskyt hypertenzie, ktorá viedla k úmrtiu.

### Confounding – zmätenie

Slovník epidemiológie definuje confounding (z latinského *confundere*, miešať) ako 1. Stav, keď nie sú účinky dvoch procesov oddelené. Zmiešanie zdanlivého vplyvu expozície na riziká, ktoré prináša spojenie s inými faktormi, ktoré môžu ovplyvniť výsledok. 2. Vzťah medzi účinkami dvoch alebo viacerých príčinných faktorov, ktoré sa pozorovali v súbore dát takým spôsobom, že nie je logicky možné oddeliť príspevok k účinku každého z jednotlivých kauzálnych faktorov. 3. Situácia, keď je miera vplyvu expozície na riziko skreslená vzťahom expozície na iný faktor ( $y$ ), ktoré ovplyvňujú skúmaný výsledok. [6] Na ilustráciu tejto koncepcie uvedieme príklad, publikovaný ako súčasť prehľadového článku o štandardizácii mier v epidemiológii. [10]

V štúdií o vplyve pitia kávy ako rizikového faktoru rozvoja ischemickej choroby srdca sa musel prirodzene zobrať do úvahy fakt, že subjekt bol alebo nebol fajčiar. Tabakizmus je všeobecne uznávaný faktor s výrazným vplyvom na rozvoj tohto ochorenia. Zároveň sa vie, že tabak a káva sa vo veľkej väčšine prípadov konzumujú súčasne, ale nemôžeme povedať, že fajčenie je následkom pitia kávy. Preto fajčenie nemôžeme považovať za premennú, ktorá by stála medzi kávou a ischemickou chorobou srdca. Jednoduchý obrázok 1 ukazuje na vzťah medzi spomínanými tromi premennými.



**Obrázok 1** Vzťah medzi pitím kávy, fajčením a ischemickou chorobou srdca, podľa [10]

Spojovníky naznačujú, že keď skúmame vplyv pitia kávy na rozvoj ischemickej choroby srdca, potom musíme predpokladať vzájomné pôsobenie medzi kávou a tabakom, ale pôsobenie tabaku na ochorenie je len jedným smerom. Ak by sme úlohu

tabaku vynechali, potom dôjde k pochybeniu. Takúto situáciu nazývame zmätením a samotný faktor označujeme ako mäťúci. Hoci v slovenčine nie je tento výraz dostatočne zaužívaný, budeme ho používať. Často sa však používa aj jeho slovný ekvivalent v angličtine.

### **Confounding variable, confounder, mäťúca premenná**

Mäťúca premenná, pokiaľ sa nezoberie do úvahy, môže viesť k falošnému výsledku a musí mať vzťah k faktoru, ktorý je predmetom vyšetrovania. Ak by sa nebrala do úvahy, potom dôjde k skresleniu výsledku. [6] Takýto faktor (mäťúca premenná) musí splniť nasledujúce kritériá: musí byť známym rizikovým faktorom vo vzťahu k študovanému výsledku, musí byť faktorom súvisiacim s expozíciou, ale nie jej výsledkom a tiež nesmie stáť medzi dvoma primárne študovanými premennými (expozícia a účinok). [11]

Confounding, alebo zmätenie, nie je chyba v postupe, ale skôr je to fenomén, ktorý je identifikovaný skúmaním a je potrebné s ním rátať. Skreslenie je dôsledkom chyby v postupe, akým bola štúdia vykonaná, ale zmätenie je platný nález, ktorý opisuje povahu vzťahu medzi niekoľkými faktormi a rizikom ochorenia. Avšak neschopnosť prijať mäťúcu premennú do úvahy pri interpretácii výsledkov skúmania je naozaj chyba v realizácii štúdie a môže ovplyvniť jej závery. [11]

Vzhľadom na snahu čo najviac vylúčiť vplyv mäťúcich premenných, a tak odstrániť či znížiť ich vplyv, používame štandardizáciu mier ako základnú metódu. V praxi sa často štandardizuje na vek, najmä pri porovnávaní mier či ukazovateľov odvodených od úmrtnosti. Menej často sa štandardizuje na iné mäťúce premenné. Rovnako ako u akejkolvek miery, ktorá vznikla na základe kombinácie viacerých premenných, môže štandardizovaná premenná prekryť veľké rozdiely medzi skupinami, ktoré môžu byť dôležité pre vysvetlenie zmien mier v dôsledku alebo v súvislosti s premennou, ktorú chceme upraviť. Preto vždy, keď je to možné, je dôležité analyzovať štandardizované miery spolu s hrubými (teda pred štandardizáciou).

### **Metódy štandardizácie**

V podstate sa používajú dve metódy na odstránenie rozdielov v demografickom zložení dvoch populácií: priama a nepriama metóda. Pokiaľ sa použijú s rovnakými vstupnými údajmi, mali by viesť k rovnakým výsledkom. Ak by sme získali dva rozličné výsledky, je nevyhnutné hľadať príčiny hlbším skúmaním vlastností vstupných údajov. Bez ohľadu na použitú metódu potrebujeme stanoviť, akú štandardnú populáciu použijeme. Môže ňou byť jedna z porovnávaných populácií, alebo úplne iná či umelo vytvorená teoretická populácia.

### Priama metóda štandardizácie

Priamu metódu štandardizácie budeme demonštrovať na príklade porovnania úmrtnosti v dvoch častiach Bratislavy: Bratislava I, teda Staré mesto, kde je býva pomerne málo mladých ľudí a Petržalka, teda Bratislava V, kde naopak je mladých ľudí veľa.

Veková skupina (i)	Bratislava I (a)			Bratislava V (b)			B V/BI
	Počet úmrtí $x_{ia}$	Úmrtnosť /1000 $p_{ia}$	Populácia $n_{ia}$	Počet úmrtí $x_{ib}$	Úmrtnosť /1000 $p_{ib}$	Populácia $n_{ib}$	Populácia $n_{ib}/n_{ia}$
0--14	6	1.3	4 636	5	0.4	12 617	2.7
15--59	63	2.7	23 595	241	2.9	83 491	3.5
60+	490	38.9	12 597	283	13.2	21 464	1.7
Spolu	559	13.7	40 828	529	4.5	117 572	2.9

**Tabuľka 2** Počty zomretých a počet obyvateľov v Bratislava I a Bratislava IV v roku 2009, ženy a muži spolu. Zdroj: ŠÚ SR

Počtom zomretých sa obe populácie veľmi nelíšia, keď sa však pozrieme na pomer medzi počtom obyvateľov v Bratislave I-Staré mesto a Bratislave V-Petržalka vidíme zaujímavý rozdiel (Tabuľka 2). Populácia v Starom meste Bratislava má menej mladých ľudí ako Petržalka. Štandardizujeme na nejakú spoločnú populáciu, ktorá je rovnaká pre obe populácie. V podstate je jedno, na akú populáciu budeme štandardizovať, ale pre ilustráciu postupu zvolíme populáciu, ktorá vznikne sčítaním počtov obyvateľov v Bratislave Starom meste a Petržalke v každej vekovej skupine. Na túto novú populáciu prepočítame počet zomretých využijúc úmrtnosť na 1000 obyvateľov v každej komunite. Zhrnutím výsledkov získame výsledné miery úmrtnosti pre každú komunitu (Tabuľka 3).

Vek (i)	Populácia BA I + BA V $P_i = n_{ia} + n_{ib}$	Bratislava I			Bratislava V		
		Úmrtnosť /1000 $p_{ia}$	Očakávaný počet úmrtí $e_{ia}$	Štandardizovaná úmrtnosť $sdr_{ia}$	Úmrtnosť /1000 $p_{ib}$	Očakávaný počet úmrtí $e_{ib}$	Štandardizovaná úmrtnosť $sdr_{ib}$
0-14	17 253	1.3	22.3		0.4	6.8	
15-59	107 086	2.7	285.9		2.9	309.1	
60+	34 061	38.9	1 324.9		13.2	449.1	
Spolu	158 400		1 633.2	0.01031	4.499	712.7	0.00450

**Tabuľka 3** Výpočet štandardizovanej úmrtnosti metódou priamej štandardizácie

Preberieme si výpočet krok po kroku. V prvom rade je potrebné povedať, že pre zjednodušenie sme zvolili len tri vekové skupiny, ktoré označujeme ako  $i$ . Oblasť Bratislava I označujeme písmenom  $a$ , Bratislava V písmenom  $b$ . Počet úmrtí v jednotlivých vekových skupinách označíme ako  $x_p$ , kde budeme mať dve rôzne premenné, jednu pre prvú oblasť  $x_{ia}$  a ďalšiu pre druhú  $x_{ib}$ . Analogicky bude označená populácia ako  $n_{ia}$  a  $n_{ib}$ . Hrubú mieru úmrtnosti v príslušnej vekovej skupine a oblasti označíme ako  $p_{ia}$  a  $p_{ib}$ . Jej výpočet je jednoduchým pomerom počtu zomretých v danej vekovej skupine k strednému stavu obyvateľov danej oblasti a vekovej skupine:

$$\text{pre oblasť } a: p_{ia} = \frac{x_{ia}}{n_{ia}} 1000,$$

$$\text{pre oblasť } b: p_{ib} = \frac{x_{ib}}{n_{ib}} 1000.$$

Súčtom obyvateľov z oboch oblastí podľa vekových skupín sme získali spoločnú, alebo štandardnú populáciu  $P_i$ , na ktorú prepočítame počty úmrtí z jednotlivých oblastí. Tak získame hodnotu očakávaného počtu úmrtí v jednotlivých oblastiach a vekových skupinách  $e_{ia}$  a  $e_{ib}$ :

$$\text{pre oblasť } a: e_{ia} = \frac{p_{ia}}{P_i} 1000,$$

$$\text{pre oblasť } b: e_{ib} = \frac{p_{ib}}{P_i} 1000.$$

Štandardizovaná úmrtnosť podľa veku, ktorú označíme ako  $sdr_i$  potom vypočítame podobne ako pomer očakávanej úmrtnosti a stredného počtu obyvateľov danej oblasti:

$$sdr_{ia} = \frac{e_{ia}}{n_{ia}} 1000,$$

$$sdr_b = \frac{e_{ib}}{n_{ib}} 1000.$$

Interpretácia výsledku vychádza z porovnania štandardizovanej úmrtnosti podľa veku pre obe oblasti. Vidíme, že tá je v Starom meste Bratislava (I) výrazne vyššia ako v Petržalke (Bratislava V). Vychádzajúc z koncepcie relatívneho rizika, môžeme tieto úmrtnosti interpretovať ako pravdepodobnosť umrieť a ich pomer ako relatívne riziko úmrtnia:



$$RR = \frac{sdr_a}{sdr_b} = \frac{0,01}{0,0045} = 2,29.$$

Teda môžeme konštatovať, že relatívne riziko úmrtia je pre obyvateľa Starého mesta Bratislava 2,3-krát väčšie ako u obyvateľa Petržalky.

Nie vždy je výhodné používať kombináciu zúčastnených populácií ako štandard. V štúdiu zdravia a choroby sa zaužívala štandardná populácia definovaná Svetovou zdravotníckou organizáciou. Je podstatne jednoduchšia a dnes predstavuje určitý celosvetový štandard. Využíva sa prakticky vždy pri medzinárodných porovnaníach. Pri skúmaní úmrtnosti či iného ukazovateľa, ktorý súvisí so zdravím pomocou porovnania medzi dvoma a viacerými geografickými či administratívnymi celkami sa často použije populácia vyššieho celku, napríklad VÚC pri skúmaní miest či okresov alebo celého štátu. Výber štandardnej populácie totiž výsledok významne neovplyvní.

### Epitools

Na uľahčenie výpočtu štandardizovaných mier pomocou priamej i nepriamej štandardizácie ponúka projekt {R} knižnicu epidemiologických metód nazvanú *epitools*. [12] Na predchádzajúcom príklade si demonštrujeme použitie funkcie *ageadjust.direct(count, pop, rate = NULL, stdpop, conf.level = 0.95)* z tejto knižnice, ktorá vypočíta štandardizovanú úmrtnosť aj s intervalom istoty. Postup prípravy údajov v prostredí RExcel a získané výsledky sú zrejme z nasledujúceho obrázku (Obrázok 2).

---

```
# údaje sme uložili do premenných výberom z tabuľky v RExcel pomocou Put R Var
# počet úmrtí v oblasti A, Bratislava I
> pocetA
[1] 6 63 490
# počet úmrtí v oblasti B, Bratislava V
> pocetB
[1] 5 241 283
# počet obyvateľov v oblasti A, Bratislava I
> popA
[1] 4636 23595 12597
# počet obyvateľov v oblasti B, Bratislava V
> popB
[1] 12617 83491 21464
# štandardná populácia
> stdpop
[1] 17253 107086 34061
# volanie funkcie pre oblasť A výsledok
# premenná crude.rate obsahuje hodnotu hrubej úmrtnosti, premenná adj.rate hodnotu štandardizovanej
úmrtnosti, premenná lci prináša dolný interval istoty, kým premenná uci horný
```

## Štandardizácia údajov

```
> ageadjust.direct(pocetA, popA, rate = NULL, stdpop, conf.level = 0.95)
crude.rate adj.rate lci uci
0.013691584 0.010310384 0.009456982 0.011230608
> ageadjust.direct(pocetB, popB, rate = NULL, stdpop, conf.level = 0.95)
crude.rate adj.rate lci uci
0.004499371 0.004829766 0.004424762 0.005262736
```

Oblasť	Hrubá miera	Štandardizovaná	Dolný interval	Horný interval
Bratislava I	0,014	0,010	0,009	0,011
Bratislava V	0,004	0,005	0,004	0,005

**Obrázok 2** Postup prípravy údajov a výsledky priamej štandardizácie

Teraz sa pozrieme, čo sa stane, ak použijeme inú štandardnú populáciu. Začneme s populáciou, ktorú používa Svetová zdravotnícka organizácia. [3]

```
# do premennej stdWHO sme zadali jednotlivé vekové skupiny
> stdWHO
[1] 26.15 61.93 11.95
# Volaním funkcie ageadjust.direct() vypočítame hodnoty štandardizácie
> ageadjust.direct(pocetB, popB, rate = NULL, stdWHO, conf.level = 0.95)
crude.rate adj.rate lci uci
0.004499371 0.003465815 0.003167792 0.003800535

> ageadjust.direct(pocetA, popA, rate = NULL, stdWHO, conf.level = 0.95)
crude.rate adj.rate lci uci
0.013691584 0.006638342 0.006013983 0.007359279
```

Oblasť	Hrubá miera	Štandardizovaná	Dolný interval	Horný interval
Bratislava I	0,014	0,007	0,006	0,007
Bratislava V	0,004	0,003	0,003	0,004

**Obrázok 3** Štandardizácia úmrtnosti v dvoch častiach Bratislavy za použitia populačného štandardu SZO

Z výsledku (Obrázok 3) je vidieť, že použitie štandardnej populácie SZO zásadne nezmenilo výsledok štandardizácie.

### Nepriama metóda štandardizácie

Nepriama štandardizácia sa zvolí vtedy, keď nie je dostatok údajov pre priamu, používa sa menej často. Nepriama vychádza zo štandardnej množiny vekovo špecifických úmrtností, ktoré sa vzťahujú na celkovú úmrtnosť skúmanej populácie. Ve-

kovo špecifické úmrtnosti sa získajú z referenčnej populácie. Týmto spôsobom sa získa „očakávaný“ počet úmrtí v pozorovanej populácii očakávajúc, že vekovo špecifické úmrtnosti štandardnej populácie sa prepočítajú na sledovanú populáciu. Najčastejšie sa uvádza pomer medzi pozorovanými a očakávanými úmrtiami, ktorý sa nazýva *index štandardizovanej úmrtnosti* (Standardized Mortality Ratio SMR). [13]

Skúsime vykonať proces nepriamej štandardizácie na učebnicovom príklade prebranom z [14]. Vstupné údaje a výsledok kalkulácie je v dvoch tabuľkách na obrázku 4. Očakávaný počet úmrtí v každej zo študovaných populácií sme získali vynásobením stredného stavu populácie skúmanej v danej vekovej skupine štandardnou úmrtnosťou. Súčin sme prepočítali na 1000 obyvateľov.

$$e_{ia} = \frac{n_{ia} p_i}{1000}$$

Sumárny ukazovateľ pomeru očakávaných úmrtí k celkovej populácii sme získali vydelením očakávaného počtu úmrtí v celej skúmanej populácii A a B a vydelením celkovým stredným stavom obyvateľstva v každej z oblastí

$$A_a = \frac{e_a}{n_a} 1000 \qquad B_a = \frac{e_b}{n_b} 1000$$

Potom sme prišli k nepriamej štandardizácii, kde sme hrubú mieru úmrtnosti vynásobili štandardnou úmrtnosťou pre celú populáciu a následne vydělili pomerom očakávaných úmrtí k celkovej populácii

$$imdr_a = \frac{p^* mdr}{A_a} \qquad imdr_b = \frac{p^* mdr}{A_b}$$

Index štandardizovanej úmrtnosti je potom jednoduchý pomer medzi nepriamo štandardizovanými úmrtnosťami

$$imdr = \frac{imdr_a}{imdr_b}$$

Index *imdr* (Indirect Method Death Rate) hovorí o tom, že nepriamo štandardizovaná úmrtnosť v populácii A je menšia ako v populácii B približne o 10 %. Ak by boli totiž rovnaké, potom by sa *imdr* muselo rovnať jednej. Tolko k samotnému výpočtu a teraz si zjednodušíme život použitím funkcie z EPITOOLS, ktorá vypočíta nielen jednotlivé úmrtnosti, ale zároveň uvedie aj intervaly istoty. Na výpočet použijeme údaje z príkladu porovnania dvoch častí Bratislavy, avšak s tým rozdielom, že štandardnú populáciu a príslušné počty zomretých budeme brať zo štatistiky celej Slovenskej republiky v roku 2010 (Obrázok 5).

Štandardizácia údajov

Veková skupina	Štandardná populácia			Populácia v komunite	
	Populácia	Úmrtia	Úmrtnosť /1000	A	B
0-14	35 000	330	9.43	10 000	25 000
15-59	30 000	345	11.50	15 000	15 000
60+	35 000	535	15.29	25 000	10 000
Spolu	100 000	1 210	12.10	50 000	50 000
<b>Hrubá miera úmrtnosti/1000</b>	<i>mdr</i>	<b>12.40</b>	<b>11.80</b>		

Štandardná úmrtnosť/1000 $P_i$	Populácia $n_{ia}$	Očakávaný počet úmrtí $e_{ia}$	Populácia $n_{ib}$	Očakávaný počet úmrtí $e_{ib}$
9.43	10 000	94.29	25 000	235.71
11.50	15 000	172.50	15 000	172.50
15.29	25 000	382.14	10 000	152.86
12.10	50 000	648.93	50 000	561.07
<b>pomer očakávaných úmrtí k celkovej populácii</b>	$A_a$	<b>12.98</b>	$A_b$	<b>11.22</b>
<b>nepriamo štandardizovaná úmrtnosť</b>	$imdr_a$	<b>11.56</b>	$imdr_b$	<b>12.72</b>
<b>Index štandardizovanej úmrtnosti</b>	$imdr$	<b>0.91</b>		

Obrázok 4 Príklad výpočtu indexu štandardizovanej úmrtnosti metódou nepriamej štandardizácie

Veková skupina	Bratislava I A		Bratislava V B		Slovensko	
	Počet úmrtí <i>countA</i>	Populácia <i>popA</i>	Počet úmrtí <i>countB</i>	Populácia <i>popB</i>	Počet úmrtí <i>stdcount</i>	Populácia <i>stdpop</i>
0-14	6	4 636	5	12 617	505	831 246
15-59	63	23 595	241	83 491	10 133	3 636 173
60+	490	12 597	283	21 464	42 807	963 605
Spolu	559	40 828	529	117 572	53 445	5 431 024

Obrázok 5 Vstupné údaje príkladu nepriamej štandardizácie. Názvy premenných, ktoré sa nachádzajú ako argumenty volanej funkcie, sú vyznačené kurzívou.

Volaním funkcie `ageadjust.indirect(count, pop, stdcount, stdpop, stdrate=NULL, conf.level = 0.95)` z EPITOOLS štandardizujeme úmrtnosti najprv pre Bratislava I (označené ako A) a následne pre Bratislava V (označené B).

```
> ageadjust.indirect(countA, popA, stdcount, stdpop, stdrate=NULL, conf.level = 0.95)
```

```
$sir
  observed   exp   sir   lci   uci
559.0000000 628.1758184 0.8898783 0.8190843 0.9667910

$rate
  crude.rate adj.rate   lci   uci
0.013691584 0.008757012 0.008060351 0.009513886
```

SMR populácie A

Nepriamo štandardizovaná úmrtnosť populácie A

```
> ageadjust.indirect(countB, popB, stdcount, stdpop, stdrate=NULL, conf.level = 0.95)
```

```
$sir
  observed   exp   sir   lci   uci
529.0000000 1193.8437692 0.4431066 0.4069110 0.4825218

$rate
  crude.rate adj.rate   lci   uci
0.004499371 0.004360472 0.004004283 0.004748345
```

SMR populácie B

Nepriamo štandardizovaná úmrtnosť populácie B

Oblasť	Hrubá miera	Štandardizovaná	Dolný interval	Horný interval
<b>Bratislava I</b>	0,014	0,009	0,008	0,009
<b>Bratislava V</b>	0,004	0,004	0,004	0,005

**Obrázok 6** Štandardizácia úmrtností na vek v dvoch komunitách za použitia metódy nepriamej štandardizácie.

Z výsledkov nepriamej metódy štandardizácie, použitých na údaje dvoch komunit rovnakých ako v prípade priamej štandardizácie len za použitia inej štandardnej populácie (Obrázok 6) je vidieť, že sme sa dopracovali k podobným číslam.

Použitie štandardizovaných úmrtností môže byť sporné. Akákoľvek súhrnná miera môže prekryť faktory, ktoré by mohli mať významné dôsledky pre zdravie verejnosti. Napríklad pri štandardizácii na vek je možné prehliadnúť vekovo špecifické rozdiely v riziku v čase alebo mieste. Takáto situácia môže nastať vtedy, keď výskyt nových prípadov ochorenia zdanlivo narastá v dôsledku efektu kohorty narodení (ľudia v mladšom veku môžu mať vyššie riziko ochorenia v porovnaní so staršími). Vekovo štandardizovaná miera môže skrývať tieto trendy. Aj napriek tomuto riziku štandardizované miery poskytujú užitočné informácie, najmä v prípadoch zriedkavých ochorení a pri širokej variabilite špecifických mier.

## Očakávaná dĺžka života

Ľudia si od pradávna kládli otázku o dĺžke života, ktorá ich ešte čaká. Len nedávno získala táto problematika pozornosť odborníkov a impulzom bol vznik poisťovníctva. Poisťovalo sa všeličo – lode, domy, fabriky. Keď sa začali poisťovať ľudia na to, že sa dožijú určitého veku, vznikli nielen detektívne romány, ale aj potreba vypočítať, koho sa oplatí poistiť a koho nie. Veď zoberme si príklad, že bohatý starý pán sa rozhodne poistiť na predčasnú smrť a dať vyplatiť vdove či deťom poistné, ak do určitého času zomrie. Poisťovňa je podnikateľský subjekt, ktorý musí mať zisk. Teda platby poistného by mali byť väčšie ako skutočne vyplatená poistka proti neočakávanej situácii. Ak by šlo o poistenie jediného človeka, potom by poisťovňa bola postavená pred riziko, že ňou poistený jedinec umrie pred dohodnutou lehotou a suma, ktorú sa zaviazala vyplatiť, prevýši sumu, ktorú dovtedy vybrala. Poisťovne však konajú tak, aby prípadov, keď majú vyplatiť poistku, mali čo najmenej, a tým mali čo najväčší zisk. Riešia otázku, koľko sa ešte môže dožiť daný človek, v danom veku, daného pohlavia a podobne. Na základe odhadu veku dožitia potom sú schopní vypočítať poistnú sumu i splátky. Je možné, že občas človek zomrie skôr a vtedy musia poistku vyplatiť, ale počet takýchto prípadov pri správnom výpočte očakávanej dĺžky života je malý a poisťovne sú vysoko ziskové.

### Aktuariálna matematika

Ako odhadnúť, či sa niekto dožije, alebo nedožije určitého veku? Ak máme k dispozícii dostatočne veľkú skupinu ľudí a poznáme pravdepodobnosť úmrtia v jednotlivých vekových skupinách, môžeme predpovedať počet tých, ktorí prežijú s dostatočne veľkou istotou. Táto matematika sa nazýva *aktuariálna*. Na základe takéhoto výpočtu môžeme vypočítať priemerný počet celkových rokov života, čo sa bežne nazýva *očakávaná dĺžka života*. Vypočítame pravdepodobnosť prežiť na základe pravdepodobnosti umrieť v danej vekovej kategórii. Ale tieto výpočty boli vykonané na údajoch o ľuďoch, ktorí sú starší a často je rozdiel niekoľko generácií. Inými slovami, keď chceme zisťovať pravdepodobnosť dožitia sa 60 rokov pre dnes narodené dieťa, tak to robíme na základe údajov o starých rodičoch, narodených na začiatku 50. rokov minulého storočia. Teda do očakávanej dĺžky života premietame skutočnosti, ktoré ovplyvnili život ľudí tejto vekovej kategórie a tie sa môžu výrazne odlišovať od skutočností, ktoré čakajú novorodenca v jeho živote a ktoré nevieme dostatočne predvídať. Preto actuariálna štatistika poskytuje vlastne historický obraz skutočnosti a jej interpretácia je založená na predpoklade, že odhady (napríklad pre dnešného novorodenca) budú správne len pri predpoklade rovnakého správania sa vekových skupín. Vieme, že tento predpoklad je vždy len hypotetický a je zatažený štatistickou chybou. Taktiež nemôžeme povedať, že dieťa dnes narodené sa dožije presne toľko rokov, ako by sa dalo predpokladať na základe očakávanej dĺžky života.

Výhodou *očakávanej dĺžky života* (LE) je, že predstavuje sumárnu mieru úmrtnosti a nevyžaduje použitie štandardnej populácie. Preto postup jej výpočtu môžeme považovať za štandardizovanie na vek a porovnávať populácie medzi sebou. Preto ju možno ľahko porovnávať medzi územnými celkami a krajinami. Zároveň je pomerne všeobecne zrozumiteľná. Prirodzene, jej výpovedná hodnota o stave zdravia populácie je obmedzená, rovnako ako sú údaje o vekovo štandardizovanej úmrtnosti. Jednoduchá definícia *strednej dĺžky života pri narodení* (Life expectancy at birth, expectation of life at birth) hovorí, že je to *počet rokov, ktoré v priemere ešte prežije práve narodená osoba za predpokladu, že sa úmrtnostné pomery nezmenia*. Podobne sa definuje *stredná dĺžka života v určitom veku* (Life expectancy at given age) ako *počet rokov, ktoré v priemere ešte prežije osoba v príslušnom veku za predpokladu, že sa úmrtnostné pomery nezmenia*. Zvykne sa používať aj názov nádej na dožitie alebo očakávaná dĺžka života. [15]

LE je najčastejšie používanou globálnou charakteristikou na hodnotenie úmrtnosti. Vzhľadom na časté použitie tejto miery si ukážeme nielen jej výpočet, ale najmä myšlienkový postup pri ňom. Zároveň ukážeme, ako správne interpretovať výsledky a ako ich použiť pri úvahách o faktoroch, ktoré k nim viedli. Tiež si ukážeme, ako uvažovať o kvalite života a zdravia pri porovnávaní rôznych skupín v populácii.

### História

Úmrtnostné tabuľky majú dlhú históriu. John Graunt publikoval knihu *Natural and Political Observations Upon the Bills of Mortality* v roku 1662 na základe ním zozbieraných Bills of Mortality (Výkazy úmrtnosti). [16] Podobne Edmund Halley (1656 – 1742), známy objaviteľ kométy, publikoval tabuľky z nemeckého mesta Breslau (teraz poľské mesto Vroclav) v roku 1693, ktoré zostavil ctihodný Caspar Neumann (1648 – 1715). Tieto tabuľky zahŕňali päťročné obdobie rokov 1687 – 1691 a na ich základe Halley odvodil predpokladaný čas trvania života ľudí. Svoj postup publikoval v eseji *An estimate of the degrees of the mortality of mankind, drawn from the curious tables of the birth and funerals at the city of Breslaw, with an attempt to ascertain the price of annuities upon lives* (Odhad úmrtností človeka odvodený z tabuliek narodenia a pohrebov v meste Breslaw s pokusom určiť cenu anuity<sup>21</sup> počas života). Význam úmrtnostných tabuliek ako nástroja na výpočty našiel menšiu odozvu u štatistikov ako u aktuariálnych matematikov. [17]

### Úmrtnostné tabuľky

Úmrtnostné tabuľky (Mortality tables, life tables) alebo tiež tabuľky života popisujú proces zomierania obyvateľstva. Princíp konštrukcie je založený na určení pravdepodobností úmrtia osôb podľa veku v určitom období a na zmenšovaní

21 Anuita (anuitná splátka) je splátka úveru zahŕňajúca splátku istiny úveru aj splátku úrokov.

tabuľkového počtu žijúcich osôb na základe týchto pravdepodobností. Počítajú sa oddelene podľa pohlavia pre vek od 0 do 104 rokov. [15] Sú nástrojom na zisťovanie úmrtnostných pomerov danej populácie. [18] Poskytujú pohľad podľa vekovej skupiny a pohlavia na pravdepodobnosť úmrtia a na nádej na dožitie. Úmrtnostné tabuľky sa niekedy nazývajú aj tabuľkami ďalej žijúcich. Predstavujú model (teda nie skutočnosť) o úmrtnostnom správaní uzavretej populácie, kde sa nepočíta s narodením a jediné možné ukončenie života je smrť jedinca. Teda model či abstrakcia, na ktorej je založený, neuvažuje s migráciou, novonarodenými jedincami a pod.

### **Generačné (kohortové) úmrtnostné tabuľky**

Pokiaľ tabuľka zachytáva záznam priebehu života konkrétnej populácie súčasne narodených jedincov od ich narodenia až do smrti posledného z nich, hovoríme o *generačnej úmrtnostnej tabuľke*. Vytvoriť takúto tabuľku je náročné, pretože je potrebné sledovať populáciu počas jej života. Prirodzene, pri takomto sledovaní ľudskej kohorty by určite došlo k stratám, napríklad migráciou, a tiež k strate follow-up. Z uvedených dôvodov sa preto využívajú najmä pri sledovaní populácie zvierat, hmyzu či baktérií, kde je dĺžka života kratšia ako u človeka. V ostatnom čase nachádzajú využitie aj v epidemiologickom výskume.

### **Bežné (prierezové) úmrtnostné tabuľky**

Tie popisujú úmrtnostnú skúsenosť určitej populácie počas časového obdobia. Na základe úmrtnostných mier podľa jednotlivých vekových skupín sa potom konštruuje obraz života hypotetickej populácie súčasne narodených jedincov. Podľa toho, ako sa určí sledovaný vekový interval, sa rozdeľujú na úplné a skrátené. Úplné (podrobné) úmrtnostné tabuľky (complete life tables) počítajú funkcie pre každý rok. Skrátené úmrtnostné tabuľky využívajú vekový interval väčší ako jeden rok, okrem prvého roku života. [14] Sú základom výpočtu očakávanej dĺžky života.

### **Výpočet úmrtnostných tabuliek**

Vstup pre výpočet úmrtnostných tabuliek tvoria počty zomretých v danom veku a počet žijúcich v danom veku (stredný stav), všetko za sledovaný rok. Postup ich výpočtu uvádza napríklad INFOSTAT. [18]

*Skrátené úmrtnostné tabuľky (abridged life tables)* sú odvodené z podrobných úmrtnostných tabuliek a počítajú sa pre vekové skupiny. Ako sme už uviedli v definícii, je šírka vekového intervalu  $x$  väčšia ako 1, okrem vekovej skupiny do jedného roku.

Výpočet očakávanej dĺžky života a skrátených úmrtnostných tabuliek je pomerne jednoduchý, ale vyžaduje dosť úsilia, aby sa všetky pravidlá výpočtu dodržali. Je to možné realizovať aj v programe Excel. Väčšinou sa dnes už nepočítajú, ale preberú sa hotové zo zdrojov, ktoré sa demografiou profesionálne zaoberajú, alebo sa použije hotový program na ich výpočet. My si uvedieme a vysvetlíme metodiku výpočtu jed-



noduchými slovami pri použití hotových úmrtnostných tabuliek pre rok 2010, ktoré pripravil INFOSTAT a sú voľne dostupné na jeho portáli.<sup>22</sup> Jediné vstupné údaje potrebné na konštrukciu úmrtnostnej tabuľky sú počet zomretých v danom veku a počet žijúcich v danom veku (stredný stav), obe zo sledovaného roku. Obrázok 7 predstavuje takúto úmrtnostnú tabuľku (jej časť od 0 do 10 rokov a od 90 do 100 a viac rokov) pre populáciu Slovenskej republiky v roku 2010. Pozrime sa na ňu bližšie.

Podrobné úmrtnostné tabuľky  
Územie: Slovenská republika

Obdobie:2012  
spolu

Vek	Zomrelí	Žijúci	$q_x$	$l_x$	$d_x$	$L_x$	$T_x$	$e_x$
0	321	58213	0,005499	100000	550	99505	7606521	76,07
1	31	59402	0,000522	99450	52	99424	7507016	75,49
2	15	58813	0,000255	99398	25	99386	7407591	74,52
3	8	58441	0,000137	99373	14	99366	7308206	73,54
4	13	55790	0,000146	99359	14	99352	7208840	72,55
5	6	54236	0,000093	99345	9	99340	7109488	71,56
6	6	54271	0,000102	99336	10	99330	7010148	70,57
7	3	54268	0,000110	99325	11	99320	6910817	69,58
8	11	52849	0,000170	99314	17	99306	6811497	68,59
9	8	51294	0,000162	99298	16	99289	6712191	67,60
10	10	51218	0,000151	99281	15	99274	6612902	66,61
90	1125	5209	0,222650	12550	2794	11153	37844	3,02
91	891	3700	0,248887	9756	2428	8542	26691	2,74
92	703	2561	0,277747	7328	2035	6310	18149	2,48
93	368	1459	0,309326	5293	1637	4474	11839	2,24
94	202	744	0,343678	3655	1256	3027	7365	2,01
95	183	581	0,380795	2399	914	1942	4338	1,81
96	140	501	0,420592	1486	625	1173	2395	1,61
97	161	437	0,462890	861	398	662	1222	1,42
98	96	336	0,507397	462	235	345	561	1,21
99	88	249	0,553697	228	126	165	216	0,95
100+	109	383	0,601238	102	102	71	71	0,70
<b>Spolu</b>	<b>52 437</b>	<b>5 407 580</b>						

**Obrázok 7** Časť z podrobných úmrtnostných tabuliek pre Slovensko 2012. Zdroj: Štatistický úrad SR

Prvý stĺpec obsahuje vek, v ktorom ľudia zomreli, v druhom je počet zomrelých a v treťom je počet tých, ktorí žijú. Nasledujúci stĺpec označený ako  $q_x$  predstavuje pravdepodobnosť umrieť v danej vekovej skupine. Je to odhad pravdepodobnosti, že jednotlivec, ktorý žije vo veku presne rovnom  $x$ , umrie počas tohto roku. Odhad sa vykoná na základe vekovo špecifických úmrtností podľa mierne komplikovaného

22 [http://portal.statistics.sk/files/Sekcie/sek\\_600/Demografia/Obyvatelstvo/tabulky/UT/2012/ut-2012.pdf](http://portal.statistics.sk/files/Sekcie/sek_600/Demografia/Obyvatelstvo/tabulky/UT/2012/ut-2012.pdf)

vzorca, ktorý možno nájsť, pokiaľ to niekoho zaujme, v literatúre. [14, 18] Ďalším stĺpcom začína model populácie, kde vychádzame z predstavy, že sa narodí stotisíc detí (tabuľkový počet dožívajúcich pre vek 0) a tie budú umierať postupne, podľa pravdepodobnosti umrieť v danom veku. Tak dostaneme postupne sa znižujúci počet žijúcich v umelej populácii označovaný ako  $l_x$ , tabuľkový počet dožívajúcich. Nasledujúci stĺpec nadpísaný  $L_x$  je tabuľkový počet zomretých. Všimnite si, že rozdiel dvoch nasledujúcich počtov dožívajúcich v tabuľke sa rovná tabuľkovému počtu zomretých pre prvý z rokov rozdielu. Nasledujúci stĺpec nadpísaný  $d_x$  predstavuje tabuľkový priemerný počet žijúcich (stacionárna populácia). Jeho výpočet je zas o niečo zložitejší, ale v podstate ide o to, že každý člen kohorty, ktorý sa dožije konca vekového intervalu, pridá jeden rok k hodnote  $L_x$ , zatiaľ čo každý, ktorý umrie, prispeje len čiastkou, ktorú v intervale dožil. Často sa pre zjednodušenie predpokladá, že zomrelý umrel v polovici intervalu. Predposledný stĺpec tabuľky nadpísaný  $T_x$  udáva počet rokov života, ktoré má tabuľková generácia (nie jednotlivec) v danom veku pred sebou. Pre každý vek sa vypočíta ako počet zostávajúcich rokov života, teda  $T_x = L_x + L_{x+1} + \dots + L_{100}$ . Nakoniec sa vypočítajú hodnoty nádeje na dožitie v danom veku ako pomeru tabuľkového priemerného počtu žijúcich a tabuľkového počtu dožívajúcich

$$e_x = \frac{T_x}{l_x}.$$

Skrátené úmrtnostné tabuľky  
Územie: Slovenská republika

Obdobie:2012  
spolu

Vek	Zomrelí	Žijúci	$q_x$	$l_x$	$d_x$	$L_x$	$T_x$	$e_x$
0	321	58213	0,005499	100000	550	99505	7606521	76,07
1-4	67	232445	0,000288	99450	105	99397	7499265	75,41
5-9	34	266918	0,000127	99345	63	99313	7101675	71,49
10-14	43	273999	0,000157	99281	80	99242	6605110	66,53
15-19	128	324444	0,000394	99202	190	99107	6108902	61,58
20-24	220	392065	0,000561	99012	284	98870	5613368	56,69
25-29	278	431779	0,000644	98728	310	98573	5119018	51,85
30-34	309	455289	0,000678	98418	343	98247	4626151	47,00
35-39	567	444111	0,001276	98075	613	97769	4134918	42,16
40-44	772	364976	0,002113	97462	1037	96944	3646074	37,41
45-49	1334	369345	0,003605	96426	1733	95559	3161353	32,79
50-54	2309	375688	0,006127	94693	2841	93272	2683558	28,34
55-59	3596	387971	0,009226	91851	4170	89766	2217199	24,14
60-64	4723	329898	0,014215	87681	6142	84610	1768368	20,17
65-69	4698	227726	0,020419	81539	8048	77515	1345317	16,50
70-74	5542	179183	0,030456	73491	10651	68166	957742	13,03
75-79	6978	134370	0,050606	62840	14518	55581	616914	9,82
80-84	8888	95577	0,088800	48322	18190	39227	339008	7,02
85+	11630	63587	0,167148	30132	30132	142873	142873	4,74
<b>Spolu</b>	<b>52 437</b>	<b>5 407 580</b>						

**Obrázok 8** Skrátená úmrtnostná tabuľka pre Slovenskú republiku 2012. Zdroj: Štatistický úrad SR

Výpočet skrátených úmrtnostných tabuliek sa vykoná odvodením z podrobných. Výsledná tabuľka pre Slovensko 2010 je na obrázku 8.

Pokúsme sa interpretovať túto tabuľku. Vyjdime z Mésárosovho [19] zoznamu možných použití úmrtnostných tabuliek (Obrázok 9).

$l_y/l_0$	pravdepodobnosť prežitia od narodenia $l_0$ (po presný vek $y$ ( $L_y$ ))
$l_y/l_x$	pravdepodobnosť prežitia z veku $x$ do veku $y$
$1 - l_y/l_x$	pravdepodobnosť úmrtia medzi vekmi $x$ a $y$
$L_x - l_y/l_0$	pravdepodobnosť, že novorodenec zomrie medzi vekmi $x$ a $y$
$T_x - T_y/l_0$	počet rokov, ktoré novorodenec prežije medzi vekmi $x$ a $y$
normálna dĺžka života	modus tabuľkových zomrelých $d_x$ , je to vek, v ktorom ľudia najčastejšie zomierajú (okrem veku 0)
pravdepodobná dĺžka života	medián tabuľkových dožívajúcich $l_x$ , je to vek, ktorého by sa za danej úmrtnosti dožila práve polovica narodených, respektíve je to vek $x$ pre ktorý platí $l_x = 1/2l_0$
stratené roky života zomrelého	ak sa človek dožije presného veku $x$ , pravdepodobne má ešte pred sebou $e_x$ rokov. Ak zomrie vo veku $x$ , tak $e_x$ rokov stratil. Samozrejme nie každý zomrie v presnom veku, a tak stratené roky počítame ako priemer dvoch po sebe idúcich rokov: $v_x = \frac{e_x + e_{x+1}}{2}$
stratené roky života pre celú populáciu	$V = \sum_{x=0}^{\omega} D_x^{\text{muži}} * v_x^{\text{muži}} + \sum_{x=0}^{\omega} D_x^{\text{ženy}} * v_x^{\text{ženy}}$

**Obrázok 9** Možnosti interpretácie úmrtnostných tabuliek. Mésáros [19]

V prvom rade uvádza pravdepodobnosti prežitia od narodenia po určitý vek. Ak sa pozrieme do podrobnej tabuľky prežitia pre Slovensko v roku 2010, potom by bola pre obyvateľa Slovenska pravdepodobnosť dožiť sa 50 rokov rovná 0,94, teda 94 % narodených detí by sa mohlo dožiť päťdesiatky.

Z tabuľky (Tabuľka 4) je vidieť, že pravdepodobne 92 % mužov, narodených v roku 2010 sa dožije oslavy svojej päťdesiatky, kým žien bude viac. Tento rozdiel sa stupňuje v ďalších dekádach života až napokon deväťdesiatku bude oslavovať viac ako dvojnásobok žien. Toto hľadisko sa však zriedkavo používa v odbornej literatúre, určite sa však hodí ako základ úvah o stave skúmanej populácie.

Na ilustráciu druhého spomínaného využitia urobíme prepočet na základe predchádzajúcich údajov a pokúsime sa zistiť, aká je pravdepodobnosť dožiť sa vyšších dekád života pre súčasných päťdesiatnikov a sedemdesiatnikov (Tabuľka 5).

Vek dožitia $y$	$l_y$		Pravdepodobnosť prežitia	
	Muži	Ženy	Muži	Ženy
50	92353	96664	92%	97%
60	81778	92403	82%	92%
70	61644	82522	62%	83%
80	32537	57527	33%	58%
90	5875	14022	6%	14%
100	43	35	0,04%	0,03%

**Tabuľka 4** Pravdepodobnosť dožiť sa určitého veku na základe podrobnej úmrtnostnej tabuľky pre Slovensko, 2010.  $l_0 = 100\ 000$ . Zdroj: POPIN

Východzí vek $x$	Vek dožitia $y$	$l_x$		$l_y$		Pravdepodobnosť prežitia	
		Muži	Ženy	Muži	Ženy	Muži	Ženy
50	60	92353	96664	81778	92403	89%	96%
	70			61644	82522	67%	85%
	80			32537	57527	35%	60%
	90			5875	14022	6%	15%
	100			43	35	0,05%	0,04%
70	80	61644	82522	32537	57527	53%	70%
	90			5875	14022	10%	17%
	100			43	35	0,0007	0,0004

**Tabuľka 5** Pravdepodobnosť prežitia z veku  $x$  do veku  $y$  na základe podrobnej úmrtnostnej tabuľky pre Slovensko, 2010. Zdroj: POPIN

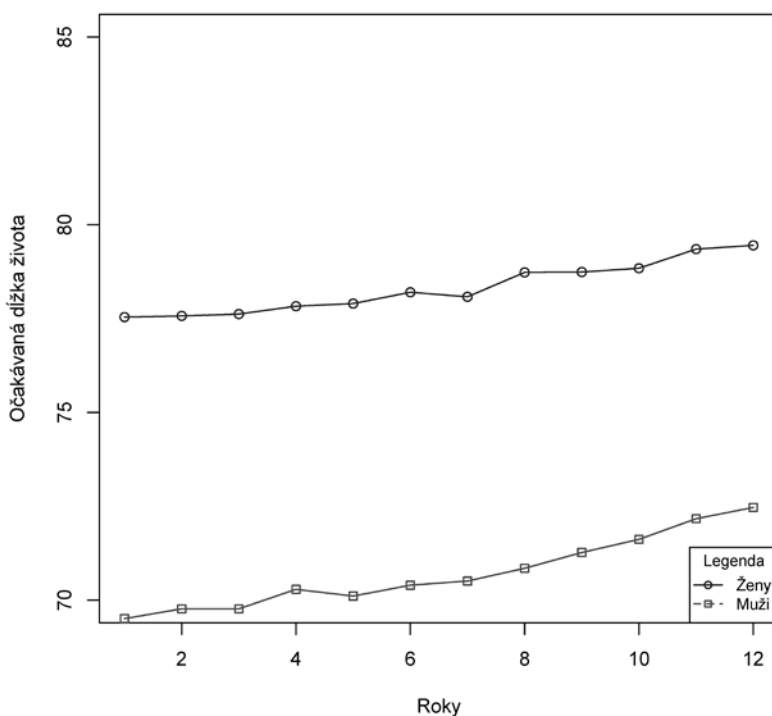
Vidieť, že muži, ktorí sa dožili päťdesiatky majú menej ako 90 % pravdepodobnosť sa dožiť nasledujúcej dekády. Sedemdesiatky sa dožijú len s pravdepodobnosťou menšou ako 70 %, pritom pre ženy je pravdepodobnosť dožiť sa rovnakého veku o 8 % väčšia.

### Výhody a nevýhody

Napriek relatívne zložitému výpočtu prináša výpočet očakávanej dĺžky života výsledok, ktorý je intuitívne ľahko pochopiteľný. Táto vlastnosť je však spojená s nebezpečenstvom nesprávnej interpretácie, ktorá neberie do úvahy retrospektívny základ výpočtu či predpoklad nemennosti trendu úmrtnosti. Významnou výhodou oproti úmrtnosti je, že porovnanie očakávanej dĺžky života medzi dvoma a viacerými populáciami sa obíde bez štandardizácie na vek. Napriek tomu je táto hodnota veľmi ovplyvnená úmrtnosťou vo včasnom detskom veku. V krajinách s vysokou

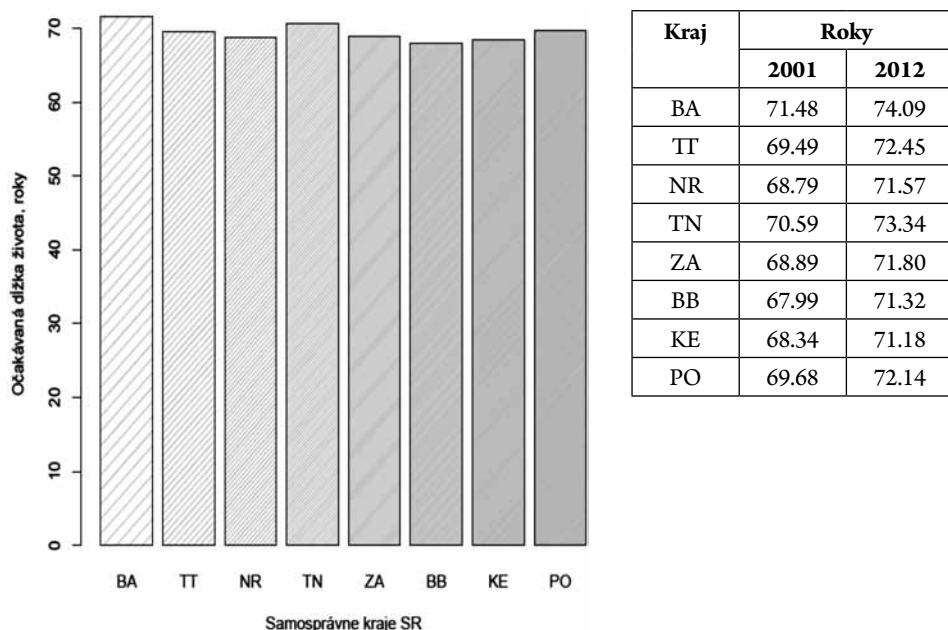
mierou dojčenskej úmrtnosti je očakávaná dĺžka života pri narodení veľmi citlivá na úmrtnosť v prvých rokoch života. Preto môže výrazne skresľovať úvahy o celkovej nádeji na dožitie v zmysle očakávania, že populácia s vysokou mierou dojčenskej úmrtnosti bude mať nutne malý podiel starších ľudí. Ďalšou pomernou nevýhodou, respektíve častou chybou pri interpretácii očakávanej dĺžky života je precenenie jej vzťahu k zdravotnému stavu obyvateľstva a najmä k funkčnosti systému zdravotníckej starostlivosti. Pokiaľ by sme vylúčili úmrtia zdravých ľudí, napr. samovraždami, utopením, úrazmi či z iných príčin, potom by sme boli blízko pravde. V demografii sa zvyčajne pracuje s úmrtnosťou bez ohľadu na príčiny smrti.

Z obrázku 10 je vidieť, že očakávaná dĺžka života pri narodení sa postupne predlžuje, najmä u mužov, kde za pozorované obdobie vzrástla zo 69,5 roka na 72,5 roka,



```
>plot(LE_SR$rok[1:12], type = "n", col="red", ylim=c(70,85), xlab = "Roky", ylab = "Očakávaná dĺžka života") # prázdna plocha
>lines(LE_SR$LE[13:24], type="o", pch=21, col="blue") # hodnoty pre mužov
>lines(LE_SR$LE[1:12], type="o", pch=22, col="red") # hodnoty pre ženy
>legend("bottomright", c("Ženy","Muži"), cex=0.8, col=c("blue","red"), pch=21:22, lty=1:2, title="Legenda") # legenda
```

**Obrázok 10** Hodnoty očakávanej dĺžky života v SR za roky 2001 až 2012 pre mužov a ženy. Zdroj: ŠÚSR [20]



**Obrázok 11** Rozdiely v očakávanej dĺžke života pri narodení mužov podľa krajov SR v roku 2001 a 2012. Zdroj ŠÚSR [21]

teda o 3 roky. U žien bol pozorovaný nárast menší, zo 77,5 roka v roku 2001 na 79,5 roka v 2012, teda rozdiel 2 rokov. Zároveň diferencia medzi mužmi a ženami v ostatnom pozorovanom roku zostala 7 rokov v zmysle kratšieho očakávaného prežitia pre mužskú populáciu. Zaujímavý je aj pohľad na regionálne rozdiely v očakávanej dĺžke života. Hodnoty očakávanej dĺžky života pri narodení podľa krajov a okresov v SR stiahneme z databázy regionálnych štatistických dát. [21] Výsledok porovnania vo forme stĺpcových grafov pre mužov z jednotlivých krajov Slovenska vidíme na nasledujúcom obrázku (Obrázok 11).

Hoci sú na prvý pohľad rozdiely medzi kraji minimálne, predsa je to niekoľko rokov. Pokiaľ v roku 2001 bol Banskobystrický kraj s najkratšou očakávanou dĺžkou života, o 11 rokov je ním kraj Košický. Rozdiel medzi Bratislavským a Košickým krajom je 3 roky, čo predstavuje výrazný rozdiel v dožívaní sa mužov medzi oboma kraji. Tiež je potrebné si uvedomiť, o koľko stúpla LE za 11 rokov. Prírodzene, o príčinách tohto javu môžeme diskutovať, ich poznanie si určite zaslúži našu pozornosť.

## Súhrn

Táto kapitola bola venovaná problému skreslenia, ktoré riešime aj pomocou štandardizácie údajov či používaním očakávanej dĺžky života. Popísali sme dve me-

tódy, z ktorých najmä prvá sa bežne používa pri štandardizácii mier najmä na vekové zloženie porovnávaných populácií. Uviedli sme postup za použitia knižnice prostredia {R}. Podrobnejšie sme sa zaoberali otázkami skreslenia a mäťúcich premenných, ktoré sú významné pre výskum v mnohých oblastiach skúmania príčinných súvislostí nielen vo vedách o zdraví. Pripomenuli sme aj koncepciu štandardnej populácie, ako ju uvádza SZO. Tiež sme urobili krátky náhľad na problematiku charakteristiky dynamiky populácie na základe úmrtnosti. LE očakávaná dĺžka života je jedným z častých charakteristík populácie, na základe ktorej sa možno zhruba orientovať pri usudzovaní a porovnávaní zdravia medzi viacerými populáciami. Jej významnou výhodou je jednoduchá konštrukcia a možnosť populáciu charakterizovať jedným číslom. To je zároveň jej obmedzením. Za týmto číslom sa totiž skrýva história života, respektíve smrti generácie, ktorá tu už nie je. Preto úvahy, ktoré sa na základe tohto indikátoru zvyknú robiť, je potrebné konfrontovať s ďalšími ukazovateľmi, najmä zo štúdií stavu zdravia žijúcej populácie. Napriek všetkému poskytuje veľmi užitočnú orientáciu a do veľkej miery dynamika jej vývoja charakterizuje zdravotné správanie populácie. Čitateľ by po oboznámení sa so spomínanými koncepciami mal byť schopný chápať a interpretovať údaje, ktoré poskytujú mnohé organizácie.

## Literatúra

1. Hanley, J. *Edwin Chadwick and the poverty of statistics*. Med Hist, 2002. 46(1): p. 21 – 40.
2. Choi, B. C., de Guia, N. A., Walsh, P. *Look before you leap: stratify before you standardize*. Am J Epidemiol, 1999. 149(12): p. 1087 – 1096.
3. Ahmad, O. et al. *Age standardization of rates: a new WHO standard*. WHO EIP/GPE/EBD, 2001, 14.
4. Kahn, H. A., Sempos, C. T. *Statistical methods in epidemiology*. 1989: Oxford University Press.
5. Miettinen, O. S. *Estimation of relative risk from individually matched series*. Biometrics, 1970. 26(1): p. 75 – 86.
6. Last, J. M. *A Dictionary of Epidemiology*. 1995, Oxford: Oxford University Press. 196.
7. Sackett, D. L. *Bias in analytic research*. J Chron Dis, 1979. 32: p. 51 – 63.
8. Delgado-Rodriguez, M., Llorca, J. *Bias*. J Epidemiol Community Health, 2004. 58: p. 635 – 641.
9. Vineis, P. *History of bias*, in *History of epidemiology methods and concepts*, B. Morabia, Birkhauser, Editor. 2004, Springer. p. 405.
10. *Standardization: a classic epidemiological method for the comparison of rates*. Epidemiol Bull/PAHO, 2002. 23(3): p. 9 – 12.
11. Gordis, L. *Epidemiology*. 4th ed. 2008: Saunders.
12. Aragon, T. J. *Epidemiology Tools*. 2010.
13. Anderson, R. N., Rosenberg HM. *Age standardization of death rates: implementation of the year 2000 standard*. Natl Vital Stat Rep, 1998. 47(3): p. 1 – 16, 20.
14. Chiang, C. L. *The Life Table and its Applications*. 1984, MALABAR, FLORIDA: Robert E. Krieger Publishing Company. 289.
15. Jurčová, D. *Slovník demografických pojmov*. 2005, INFOSTAT – INŠTITÚT INFORMATIKY A ŠTATISTIKY. Výskumné demografické centrum: Bratislava.

16. Wynder, E. L. *A corner of history: John Graunt, 1620 – 1674, the father of demography*. *Prev Med*, 1975. 4(1): p. 85 – 88.
17. Haas, F. *Edmond Halley (1656 – 1742) – Astronomer and Vital Statistician*. *JAMA*, 1965. 193: p. 66.
18. Mészáros, J. (2000) *Výpočet úmrtnostných tabuliek a Výpočet stratených rokov života úmrtím*, s. 14.
19. Mészáros, J. (2006) *Konštrukcia úmrtnostných tabuliek*. *Forum Statisticum Slovaca*, s. 45 – 52.
20. Štatistický úrad SR. *Demografia*. 2010 [cited 2010 28. 08.]; Available from: <http://portal.statistics.sk/showdoc.do?docid=2434>.
21. Štatistický úrad Slovenskej republiky. *Databáza regionálnej štatistiky RegDat*. 2013 [cited 2013 22.5.]; Available from: <http://px-web.statistics.sk/PXWebSlovak/>.