Predikcia úmrtnosti na Slovensku

Ján Barlák[1], Národná banka Slovenska

Už dlhšie obdobie pozorujeme, že trend vývoja úmrtnosti na Slovensku má, s troškou zjednodušenia, klesajúci charakter. Pre väčšinu z nás je pozitívom, že si tu požijeme v priemere dlhšie ako naši rodičia a že naše deti, ktorým chceme len to najlepšie, budú žiť ešte dlhšie. Pre finančné inštitúcie, kde dĺžka dožitia človeka hrá prim pri oceňovaní časti ich produktov, je tento klesajúci trend úmrtnosti skôr výzvou ako pozitívom. Predstavme si napríklad, že dnes poisťovňa uzavrie s klientom produkt, v ktorom sa zaväzuje vyplácať mu mesačnú rentu od začatia jeho dôchodkového veku až po jeho smrť. Môže len odhadovať, koľko rokov sa klient dožije, a aj to iba za pomoci dnes už starej úmrtnostnej tabuľky. Na vytvorenie novej a úplnej je pre získanie pravdepodobnosti úmrtia totiž potrebné mať údaje o úmrtiach za ešte stále plynúci rok. Vieme len, že priemerná dĺžka života sa predlžuje a že jej lepší odhad by mal viesť k presnejšiemu oceňovaniu spomínaných produktov.

1 Ak nie je uvedené inak, zdrojom grafov v tomto príspevku sú štatistické tabuľky úmrtnosti v SR a výpočty autora.

2 Viac o SVD nájdeme napríklad na http://www.ling.ohiostate.edu/~kbaker/pubs/Singular_Value_Decomposition_Tutorial.pdf

Na predikciu dĺžky života použijeme slovenské úmrtnostné tabuľky od roku 1980 po rok 2012 (www.infostat.sk) ako vstup do modelu. Nástrojom na fitovanie úmrtnostných kriviek, ako aj na predikciu ich ďalšieho pokračovania bude dnes už hádam aj klasický Lee-Carterov (ďalej len „*LC*") model z roku 1992. Na začiatok si uvedieme základné dôvody nášho rozhodnutia. Keďže úmrtnosť klesá, je zrejmé, že na regresiu potrebujeme klesajúcu funkciu. Našej pozornosti ale neunikne fakt, že jej pokles sa musí spomaľovať, inak by sme sa skôr či neskôr s jej hodnotami dostali nielenže na nulu, ale aj do záporných čísel, čo je nereálne. Tento predpoklad spĺňa exponenciálna funkcia, ktorú používa práve *LC* model. Ďalším jeho pozitívom je celkom pragmatický prístup k predikcii úmrtnosti, ale o tom si povieme až ďalej.

Regresia

Ako sme už naznačili, vstupom do *LC* modelu je úmrtnostná tabuľka, a to pre mužskú aj pre ženskú časť populácie. Jej maticový tvar je takýto:

$$M_{x,t} = \begin{pmatrix} m_{0,1980} & \cdots & m_{0,2012} \\ \vdots & \ddots & \vdots \\ m_{100+,1980} & \cdots & m_{100+,2012} \end{pmatrix},$$

kde *x* označuje vekové skupiny od 0 po 100+, *t* označuje roky a $m_{x,t}$ je teda pravdepodobnosť toho, že sa veková skupina *x* v roku *t* nedožije ďalšieho roku.

Ako prvé bude našou úlohou nájsť funkciu, ktorá vhodne opisuje vstupné dáta. Intuícia a rovnako aj *LC* model nám hovoria, že tento tvar bude exponenciálny. Preto bude vhodnejšie modelovať nie priamo úmrtnosť, ale jej prirodzený logaritmus, aby sme mali jednoduchšiu lineárnu závislosť namiesto exponenciálnej. Naša regresná funkcia (pre pedantov sústava funkcií) teda bude:

$$ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t} \qquad (1)$$

pre všetky *x* a *t*, kde $m_{x,t}$ už poznáme a a_x, b_x a k_t sú parametre, ktoré potrebujeme odhadnúť. Podľa (Lee a Carter, 1992) a (King a Girosi, 2007) takáto parametrizácia nie je jednoznačná. Ak vektory *a*, *b*, *k* sú riešením, potom pre ľubovoľné *c* platí, že *a*, *b/c*, *k*c* alebo *a – bc*, *b*, *k + c* sú tiež riešením. Nie je to však prekážka, keďže každé z týchto riešení nám dá tú istú predikciu. Preto v záujme jedinečného riešenia nech platia tieto obmedzenia:

$$\sum_x b_x = 1\,, \sum_t k_t = 0$$

teda že súčet prvkov vektora *b* bude 1 a pre vektor *k* bude súčet 0. Z toho aj a_x bude jednoducho dané priemerom logaritmov úmrtnosti cez všetky časy *t* pre vekovú skupinu *x*. Preto označme

$$\bar{a}_x = \frac{1}{počet\ regresných\ rokov} \sum_t \ln(m_{x,t})$$

V prehľadnejšom maticovom zápise potom máme:

$$\ln\begin{pmatrix} m_{0,1980} & \cdots & m_{0,2012} \\ \vdots & \ddots & \vdots \\ m_{100+,1980} & \cdots & m_{100+,2012} \end{pmatrix} = \begin{pmatrix} \bar{a}_0 \\ \vdots \\ \bar{a}_{100+} \end{pmatrix} + \\ + \begin{pmatrix} b_0 \\ \vdots \\ b_{100+} \end{pmatrix} \begin{pmatrix} k_{1980} & \cdots & k_{2012} \end{pmatrix} + \\ + \begin{pmatrix} \varepsilon_{0,1980} & \cdots & \varepsilon_{0,2012} \\ \vdots & \ddots & \vdots \\ \varepsilon_{100+,1980} & \cdots & \varepsilon_{100+,2012} \end{pmatrix},$$

kde $\ln(M_{x,t})$ a $\bar{a}$ poznáme. Ostáva nám teda odhadnúť vektory *b* a *k*. Podarí sa nám to za pomoci metódy rozkladu matice. Metóda má názov *Singular Value Decomposition* (ďalej len „*SVD*").[2]

Označme:

$$\widetilde{M}_{x,t} = \ln\begin{pmatrix} m_{0,1980} & \cdots & m_{0,2012} \\ \vdots & \ddots & \vdots \\ m_{100+,1980} & \cdots & m_{100+,2012} \end{pmatrix} - \\ - \begin{pmatrix} \bar{a}_0 \\ \vdots \\ \bar{a}_{100+} \end{pmatrix} \begin{pmatrix} 1 & \ldots & 1 \end{pmatrix},$$

teda od každého riadku v matici odrátame jeho priemernú hodnotu.

Podľa (Wang 2007), ak platí, že $SVD(\tilde{M}_{x,t})=ULV^T$, potom vektor b je prvý stĺpec matice U a vektor k bude násobok prvého riadku matice V^T a prvého prvku matice L. Lee a Carter a aj iní ešte vylepšujú odhad vektora k ďalším postupom, často je ale tento krok reestimácie vektora k úplne vynechaný (King & Girosi, 2007) a my, v súlade s úmyslom nekomplikovaného modelu, ho tiež vynecháme.

Predikcia

Našli sme teda všetky potrebné parametre funkcie, ktorá nám opisuje zatiaľ pozorovaný vývoj úmrtnosti na Slovensku. My by sme ale boli radi, keby sme vedeli túto funkciu „predĺžiť" aj do budúcnosti a vďaka tomu získať odhady pre budúce hodnoty úmrtnosti. Len jediný z parametrov, ktorý sme odhadovali, bol závislý od času, a to bol parameter t. Podľa Wanga (2007) Lee a Carter (1992) predikujú budúce hodnoty vektora k veľmi jednoducho a to lineárnou funkciou, kde priamku určuje iba prvý a posledný bod vektora k. Model vyzerá takto:

$$\hat{k}_{t+1} = \hat{k}_t + \theta + \varepsilon_{t+1},$$

kde ε_t je vzdialenosť medzi reálnou a odhadovanou hodnotou k_t a konštanta θ (drift) je určená jednoducho pomocou jeho prvej a poslednej hodnoty ako:

$$\theta = \frac{k_T - k_{T_0}}{T - T_0}, \qquad (2)$$

kde $[T_0, T]$ je obdobie, za ktoré máme vstupné dáta. V našom prípade, s použitím všetkých rokov, sa bude $T - T_0$ rovnať 33. Vektor k teda môžeme predikovať na „akékoľvek" dlhé obdobie a k nemu jednoducho vypočítať aj odhady pre úmrtnosti.

Lee a Carter (1992) presne neuvádzajú spôsob výpočtu konfidenčných intervalov pre ich odhady, ale jednoduchou úvahou vieme rozumné odhady urobiť aj bez nich. Ako prvé si vypočítame štandardnú odchýlku našich hodnôt vektora k od priamky, ktorou sme ju aproximovali na intervale $[T_0, T]$:

$$s = \sqrt{\frac{1}{(T - T_0) - 2} \sum_{t=T_0+1}^{T-1} (k_t - \hat{k}_t)^2},$$

keďže $k_{T_0} = \hat{k}_{T_0}$ a $k_T = \hat{k}_T$. Teraz, keď vieme, ako sa v priemere odlišovala reálna hodnota k_t od odhadovanej, budeme štandardnú odchýlku pre budúce hodnoty k_t (pre $t>T$) odhadovať ako $s_{T+n} = s\sqrt{n}$. 95-percentné konfidenčné intervaly pre k_t, kde $t>T$, teda odhadneme ako $k_{T+n} \pm 1{,}96\, s_{T+n}$ a z toho pre samotné úmrtnosti ako:

$$e^{[a_x + b_x(k_{T+n} \pm 1{,}96\, s_{T+n})]} = e^{(a_x + b_x k_{T+n})} e^{\pm(b_x 1{,}96\, s_{T+n})} = m_{x,T+n}\, e^{\pm(b_x 1{,}96\, s_{T+n})} \cong m_{x,T+n} . e^{\pm(2 b_x s_{T+n})},$$

čo už je formula pre 95-percentné konfidenčné intervaly uvedená v (Lee & Carter, 1992).

Týmto sme ukončili viac-menej teoretickú časť o *LC* modeli a mohli by sme sa rýchlo pozrieť na jeho výsledky pre slovenské dáta. Niektorým však ešte môže napadnúť celkom jednoduchá otázka. Keď už sa snažíme vyrobiť jednoduchý a použiteľný model, prečo potom namiesto akejsi takmer exponenciálnej funkcie s nie práve intuitívnou premennou k_t nepoužívame obyčajnú exponenciálnu funkciu, kde premennou by bol jednoducho čas? Naša regresná funkcia by bola

$$\ln(m_{x,t}) = a_x + b_x t + \varepsilon_{x,t}$$

s rovnakým označením ako v pôvodnej regresnej funkcii (1). Tento priam stredoškolský model (ďalej len „*EXP*") sa elegantne zbavil najzložitejších častí *LC* modelu a pre čitateľa, ktorý sa s porozumením dostal až sem, zrejme nie je nutné ho podrobne opisovať, keďže je jeho zjednodušením. Azda len spomeňme výpočet jeho štandardnej odchýlky:

$$s_{EXP} = \sqrt{\frac{1}{T - T_0} \sum_{t=T_0}^{T} \varepsilon_{x,t}^2},$$

kde $\varepsilon_{x,t}$ je vektor rozdielov medzi skutočnou a modelovanou hodnotou úmrtnosti na intervale $[T_0, T]$. Vytvorili sme teda dva modely na predikciu úmrtnosti a na to, ako sa obom modelom darí na slovenských dátach, sa pozrieme v nasledujúcej časti.

Výsledky a testy

My muži a ženy sme síce v mnohom rovnakí, ale ak ide o úmrtnosť a veci s tým súvisiace, aspoň zatiaľ to tak nie je. Preto si výsledky modelov ukážeme oddelene pre obe pohlavia. Začneme najprv vstupmi zo štatistického úradu. Grafmi (1 a, b) si znázorníme úmrtnosti pre niektoré vekové kategórie (0, 10, 20, 30, 40, 50, 60 a 70).

V oboch prípadoch vidíme, že úmrtnosť má klesajúci trend. Navyše môžeme vidieť aj fakt, že u žien je úmrtnosť v zodpovedajúcich vekových skupinách nižšia ako u mužov. Prejdime však teraz k číslam. Z *LC* modelu si uvedieme odhad vektorov k a b a pre prvý spomínaný aj jeho budúce hodnoty a k nim prislúchajúce štandardné odchýlky. Pre úspornosť textu a zároveň jednoduchosť výpočtu hodnôt vektora a nebudeme tento vektor uvádzať. Rovnako, pre výrazne jednoduchšie výpočty pri *EXP* modeli, si jeho parametre taktiež neuvedieme.

Máme viacero možností, čo si ukážeme ako výsledok modelu a aby to nebolo jednotvárne, namiesto úmrtnosti sa teraz pozrieme na odhadovanú priemernú dĺžku života, a to už pre oba modely (graf 2 a, b).

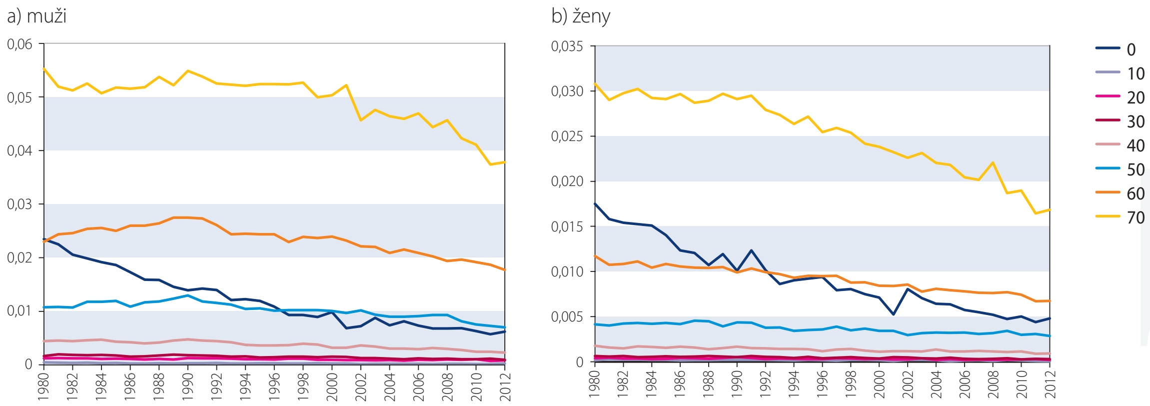
Pre oba grafy platí nasledovné. Čiara Real0 zobrazuje odhady strednej dĺžky života pri narodení (teda 0-ročných) z dostupných údajov za roky 1980

Tabuľka 1 Odhad parametra k, jeho predikcia a štandardná odchýlka, a odhad parametra b.

		Muži		Ženy		b			b		
	Rok	k	s	k	s	Vek	Muži	Ženy	Vek	Muži	Ženy
Fit	1980	-2,2752		-2,0723		0	-0,2471	-0,2295	53	-0,0734	-0,0809
	1981	-2,2728		-2,3211		1	-0,1578	-0,1764	54	-0,0733	-0,0814
	1982	-2,0537		-1,7827		2	-0,1859	-0,1547	55	-0,0757	-0,0820
	1983	-2,1638		-1,9652		3	-0,1964	-0,0694	56	-0,0763	-0,0803
	1984	-1,7292		-1,5156		4	-0,1719	-0,1182	57	-0,0722	-0,0825
	1985	-1,8092		-1,9213		5	-0,1877	-0,1532	58	-0,0692	-0,0850
	1986	-1,3266		-1,6200		6	-0,1804	-0,1918	59	-0,0652	-0,0861
	1987	-1,0294		-1,7007		7	-0,1627	-0,1731	60	-0,0640	-0,0883
	1988	-1,1482		-1,0695		8	-0,1349	-0,1456	61	-0,0604	-0,0900
	1989	-1,3390		-1,5783		9	-0,1272	-0,1625	62	-0,0587	-0,0939
	1990	-1,3956		-1,5925		10	-0,1552	-0,1627	63	-0,0580	-0,0933
	1991	-1,7758		-1,6548		11	-0,1787	-0,1360	64	-0,0576	-0,0937
	1992	-1,0569		-1,0042		12	-0,1637	-0,1193	65	-0,0586	-0,0905
	1993	-0,3569		-0,7258		13	-0,1334	-0,1050	66	-0,0588	-0,0901
	1994	-0,3354		-0,2462		14	-0,1117	-0,1115	67	-0,0569	-0,0939
	1995	-0,1643		-0,3292		15	-0,1111	-0,1037	68	-0,0554	-0,0996
	1996	0,2498		0,1006		16	-0,1205	-0,0971	69	-0,0549	-0,1035
	1997	-0,1286		-0,1632		17	-0,1246	-0,0877	70	-0,0539	-0,1013
	1998	-0,3013		-0,3428		18	-0,1044	-0,0848	71	-0,0523	-0,0985
	1999	0,0940		0,0304		19	-0,0787	-0,1017	72	-0,0534	-0,0934
	2000	0,4739		0,9588		20	-0,0735	-0,1139	73	-0,0552	-0,0897
	2001	0,3512		1,0657		21	-0,0852	-0,1198	74	-0,0539	-0,0866
	2002	1,2847		0,8987		22	-0,1059	-0,1077	75	-0,0529	-0,0831
	2003	1,0334		1,1893		23	-0,1014	-0,1015	76	-0,0524	-0,0799
	2004	1,2957		1,0264		24	-0,0939	-0,0833	77	-0,0531	-0,0762
	2005	1,5773		1,8653		25	-0,0900	-0,0776	78	-0,0539	-0,0731
	2006	1,4703		2,1192		26	-0,0970	-0,0839	79	-0,0518	-0,0686
	2007	1,6834		1,0686		27	-0,0985	-0,1106	80	-0,0492	-0,0640
	2008	1,7618		2,6543		28	-0,1066	-0,1365	81	-0,0470	-0,0586
	2009	2,2108		2,0411		29	-0,1093	-0,1443	82	-0,0461	-0,0539
	2010	3,0107		2,3858		30	-0,1218	-0,1341	83	-0,0457	-0,0484
	2011	2,6745		2,6387		31	-0,1215	-0,1303	84	-0,0449	-0,0430
	2012	3,4903		3,5629		32	-0,1287	-0,1376	85	-0,0441	-0,0376
Predikcia	2013	3,6704	0,7436	3,7390	0,8843	33	-0,1256	-0,1336	86	-0,0432	-0,0322
	2014	3,8506	1,0517	3,9151	1,2505	34	-0,1340	-0,1274	87	-0,0424	-0,0268
	2015	4,0308	1,2880	4,0912	1,5316	35	-0,1252	-0,1123	88	-0,0415	-0,0213
	2016	4,2109	1,4873	4,2673	1,7685	36	-0,1222	-0,1259	89	-0,0405	-0,0160
	2017	4,3911	1,6628	4,4434	1,9773	37	-0,1167	-0,1232	90	-0,0396	-0,0107
	2018	4,5713	1,8216	4,6195	2,1660	38	-0,1177	-0,1196	91	-0,0386	-0,0055
	2019	4,7515	1,9675	4,7956	2,3395	39	-0,1174	-0,1086	92	-0,0375	-0,0006
	2020	4,9316	2,1033	4,9717	2,5011	40	-0,1221	-0,0993	93	-0,0364	0,0042
	2021	5,1118	2,2309	5,1478	2,6528	41	-0,1207	-0,0983	94	-0,0353	0,0088
	2022	5,2920	2,3516	5,3238	2,7963	42	-0,1159	-0,0925	95	-0,0341	0,0130
	2023	5,4721	2,4664	5,4999	2,9328	43	-0,1080	-0,0949	96	-0,0329	0,0168
	2024	5,6523	2,5761	5,6760	3,0632	44	-0,1070	-0,0902	97	-0,0316	0,0203
	2025	5,8325	2,6813	5,8521	3,1882	45	-0,1051	-0,0849	98	-0,0303	0,0232
	2026	6,0126	2,7825	6,0282	3,3086	46	-0,1036	-0,0850	99	-0,0289	0,0256
	2027	6,1928	2,8801	6,2043	3,4247	47	-0,0992	-0,0863	100+	-0,0274	0,0273
	2028	6,3730	2,9746	6,3804	3,5370	48	-0,0947	-0,0919			
	2029	6,5531	3,0661	6,5565	3,6459	49	-0,0896	-0,0804			
	2030	6,7333	3,1550	6,7326	3,7516	50	-0,0832	-0,0742			
	2031	6,9135	3,2415	6,9087	3,8544	51	-0,0788	-0,0711			
	2032	7,0937	3,3257	7,0848	3,9545	52	-0,0740	-0,0756			

Zdroj: Vlastné výpočty.

Graf 1 Vývoj úmrtnosti podľa vekových skupín



až 2012. Čiary LC0, resp. EXP0, znázorňujú tieto dĺžky života modelované *LC* modelom, resp. *EXP* modelom. Takisto je znázornená aj predpoveď na roky 2013 – 2032 pomocou spomínaných modelov.

Uvedieme si niektoré hodnoty v číselnej podobe a porovnáme ich aj s údajmi zo Správy o dlhodobej udržateľnosti verejných financií (Rada pre rozpočtovú zodpovednosť – RRZ, 2012), ktorá sa tomuto sčasti venovala (tab. 2).

Vidíme, že predpoveď oboch našich modelov pre odhadovanú strednú dĺžku života je nižšia ako predpoveď Rady pre rozpočtovú zodpovednosť. Navyše, ak si spomenieme na (2), pri *LC* modeli dosť závisí, na akých dátach kalibrujeme model, keďže odhady vektora *k* závisia iba od dvoch bodov. Ak by sme napríklad namiesto všetkých 33 rokov, ktoré máme k dispozícii, vzali o jeden menej, celá predikcia by sa posunula u mužov zhruba o rok nadol. Predikciu *EXP* modelu by to zmenilo minimálne. Napriek tomu však určiť, ktorý z modelov je ten správny, dnes nevieme a ani vedieť nemôžeme. Čo však ale môžeme, je spätne otestovať, ako si naše modely počínajú, keď ich nakalibrujeme na rokoch 1980 – 1999 a predpovedáme roky 2000 – 2012, a približne tým zachováme pomer vstupných rokov k odhadovaným. Skúsme sa teda pozrieť, či nám reálne hodnoty úmrtnosti „padli" do 95-percentných intervalov spoľahlivosti pre konkrétne modely. Pre názornosť sa pozrime napríklad na skupiny 30-ročných (grafy 3 a 5).

Plnou čiarou sú znázornené reálne hodnoty úmrtnosti a čiarkovanou zasa horný (U) a dolný (L) 95-percentný interval spoľahlivosti našej predpovede.

Zaujímavý pohľad na všetky vekové kategórie je v grafe 4. Stratíme síce ucelenú predstavu o pohybe reálnych dát v medziach intervalu spoľahlivosti, ale zároveň získame celkový vizuálny prehľad pre celú vzorku populácie.

Na zvislej osi sú vekové kategórie a na horizontálnej osi sú roky testovaného obdobia. Svetložltou je znázornené spadnutie do intervalu spoľahlivosti pre konkrétnu vekovú kategóriu a konkrétny rok a červenou zrejme opak. Vidíme napríklad, že úmrtnosti 85-ročných mužov boli v intervale spoľahlivosti *LC* modelu, kým úmrtnosti 100- a viacročných žien vôbec nie, a to počas celej dĺžky testovacieho obdobia. Pozrime sa teraz na obdobné výstupy pre exponenciálny model.

Na prvý pohľad vidieť, že *EXP* model si počína lepšie. Všimnime si ale, že veľkosť konfidenčných intervalov pre oba modely je rozdielna. Nie je to prekvapujúce, keďže už vizuálne pre vekovú skupinu 0 vieme posúdiť, že *LC* model lepšie kopíruje vstupné úmrtnostné krivky a v priemere to inak nie je ani v prípade ostatných vekových skupín. Pre zhruba 75 % vekových skupín je R^2 (miera indikujúca, ako dobre model opisuje dáta) v prípade *LC* modelu vyššia ako v prípade *EXP* modelu. Aj keď vo všeobecnosti neplatí nepriama úmernosť medzi výškou R^2 a štandardnou odchýlkou. Výsledok teda nehovorí, že by *LC* model bol akokoľvek horší, ale že v tomto prípade – a dá sa predpokladať, že aj vo všeobecnosti – *EXP* model má menej ambiciózne intervaly spoľahlivosti. A teda budúce hodnoty úmrtnosti do nich padnú ľahšie.

Okrem iného, pre poisťovne je zaujímavé aj vedieť, koľko sa v priemere dožijú ľudia práve vstupujúci do dôchodkového veku. Pre zjednodušenie si zvoľme za tento vek práve 62 rokov a pozrime sa, ako si v spätnom testovaní počínali naše modely, opäť u mužov a u žien.

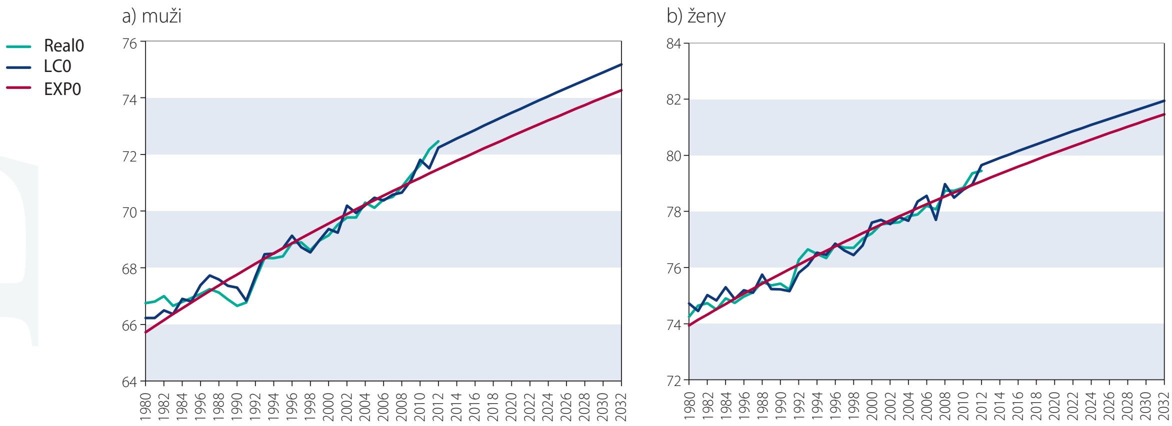
V grafe 7 vidíme, ako sa pohybovala stredná dĺžka života z reálnych dát a odhadnuté predpovede (plné čiary) a ich 95-percentné intervaly spoľahlivosti (čiarkované čiary). Kým u mužov oba modely podcenili našu rastúcu dlhovekosť, u žien sa exponenciálnemu modelu podarilo celkom dobre

Tabuľka 2 priemerná dĺžka dožitia pri narodení

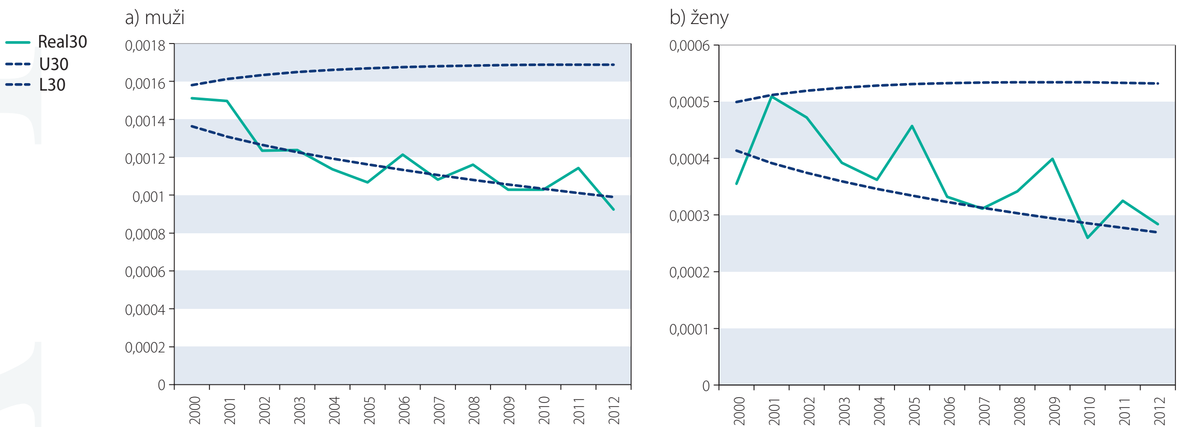
Muži	Skutočnosť	Predpoveď	
	2011	2020	2030
LC	72,2	73,5	74,9
EXP	72,2	72,6	74,0
RRZ	72,2	74,0	76,2
Ženy			
LC	79,4	80,6	81,7
EXP	79,4	80,1	81,2
RRZ	79,4	81,0	82,7

Zdroj: RRZ a vlastné výpočty.

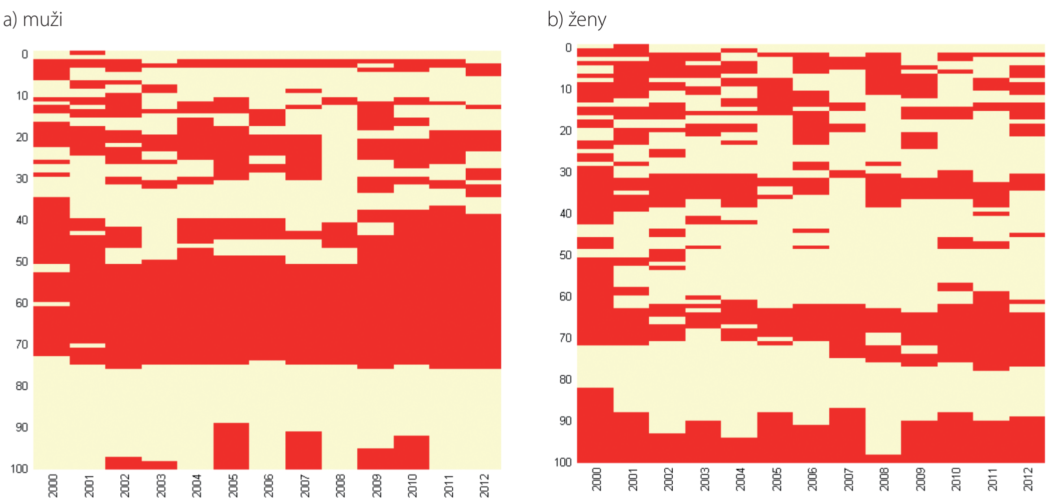
Graf 2 Porovnanie odhadovanej dĺžky života pri narodení



Graf 3 Úmrtnosť 30-ročných podľa spätného testu (LC model)



Graf 4 Vizuálny prehľad úmrtnosti pre všetky vekové kategórie podľa spätného testu (LC model)

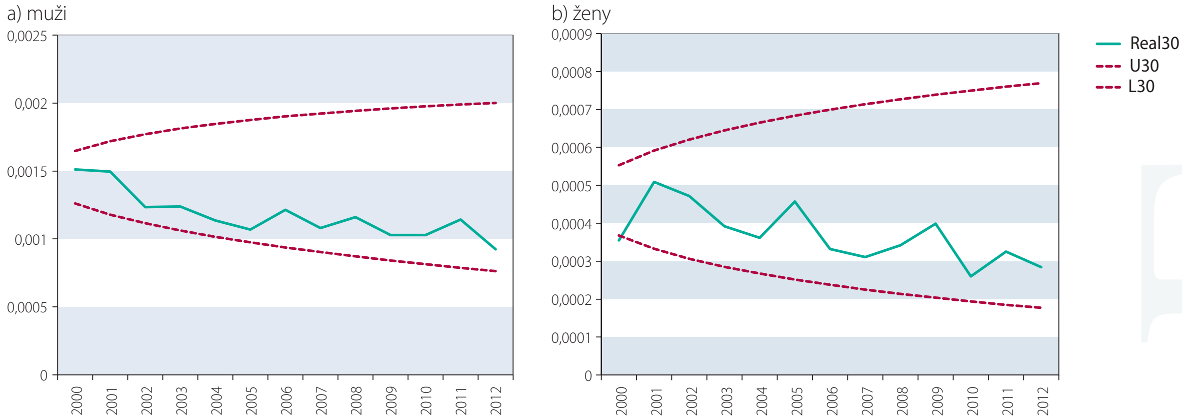


odhadnúť realitu. Všimnime si ale, že pri mužoch bola dĺžka života na zhruba tej istej úrovni až do roku 1992, čo malo, samozrejme, za následok horšiu predpovedaciu schopnosť oboch modelov. V tomto príklade bol pre obe pohlavia bližšie k realite *EXP* model. Ako sme už spomenuli, rozdiely vo veľkosti intervalov spoľahlivosti sú zrejmé, ale je potešiteľné, že krivka reálnych dát sa viac-menej zmestila do oboch intervalov spoľahlivosti.

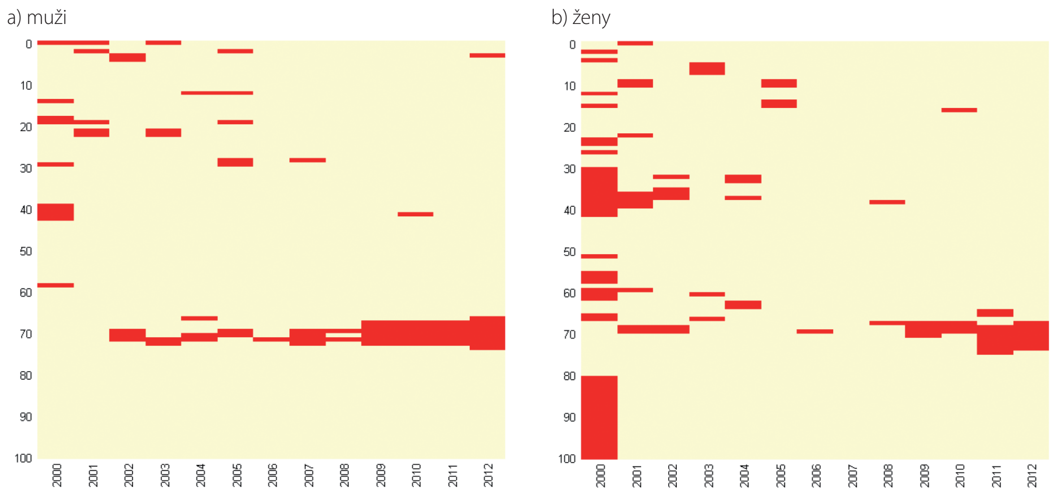
Záver

Pôvodným zámerom tohto textu bolo ukázať možnosť používať Lee-Carterov model aj na Slovensku, a to v jeho najjednoduchšej forme.

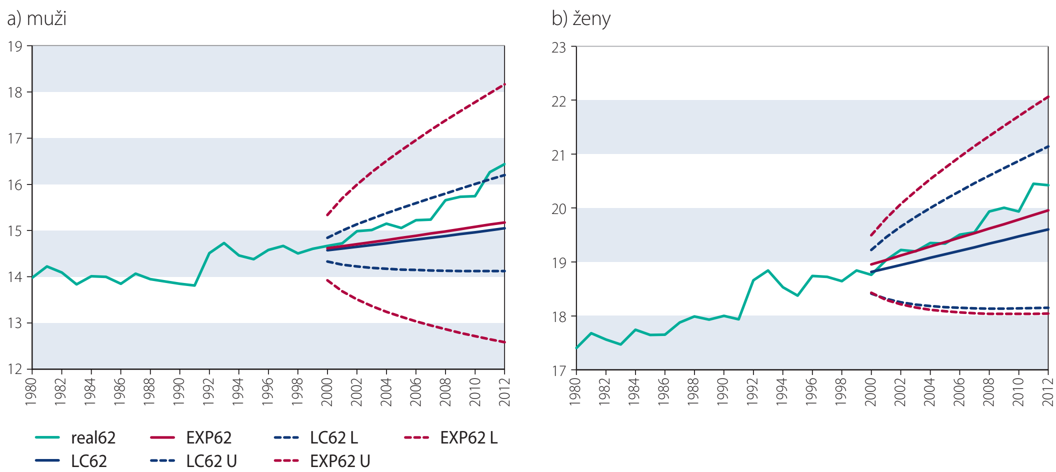
Graf 5 Úmrtnosť 30-ročných podľa spätného testu (EXP model)



Graf 6 Vizuálny prehľad úmrtnosti pre všetky vekové kategórie podľa spätného testu (EXP model)



Graf 7 Stredná dĺžka dožitia 62-ročných



Ďalej sa ale ukázalo, že sa možno dá ušetriť množstvo práce použitím jednoduchšieho modelu a to bez výraznej ujmy na výsledku, ak vôbec nejakej. A je tu ešte aspoň jedno pozitívum. Je dobré, keď ľudia rozumejú tomu, čo používajú. Veľa zaujímavého sa do tohto príspevku z priestorových dôvodov nedostalo, ale iste by bola aj škoda, keby sa čitateľ nemusel trochu ponamáhať aj sám. Neurčili sme, ktorý z modelov je lepší. Vo všeobecnosti to vopred aj tak nikdy nevieme. Jednoznačne však vieme určiť ten jednoduchší. Matematika je silný nástroj, ale aby sme nakoniec zbytočne nešli (ako sa hovorí) s bubnom na zajace...

Literatúra:

1. King, G. a Girosi, F. (14. september 2007): *Understanding the Lee-Carter Mortality Forecasting Method.*
2. Lee, R. D. a Carter, L. R. (September 1992): *Modeling and Forecasting U. S. Mortality.* Dostupné na http://www.jstor.org/stable/2290201
3. Rada pre rozpočtovú zodpovednosť(december 2012): *Správa o dlhodobej udržateľnosti verejných financií.* Dostupné na http://www.rozpoctovarada.sk/download2/sustainability_report_dec2012.pdf
4. Wang, J. Z. (marec 2007): *Fitting and Forecasting Mortality for Sweden: Applying the Lee-Carter Model.*
5. Úmrtnostné tabuľky SR, dostupné na http://www.infostat.sk/vdc/sk/index.php?option=com_wrapper&view=wrapper&Itemid=50