

SLOVENSKÁ ŠTATISTIKA a DEMOGRAFIA

SLOVAK STATISTICS
and DEMOGRAPHY

3/2015

ročník/volume 25

Recenzovaný vedecký časopis so zameraním na prezentáciu moderných štatistických a demografických metód a postupov.

Scientific peer-reviewed journal focusing on the presentation of modern statistical and demographic methods and procedures.

Článok/Article: 7

Typ článku/Type of article: vedecký článok/scientific article

Strany/Pages: 85 – 98

Dátum vydania/Publication date: 15. júl 2015/July 15, 2015



Petra DOTLAČILOVÁ, Jitka LANGHAMROVÁ
Fakulta informatiky a statistiky Vysoké školy ekonomické v Praze

**ODHAD PARAMETRŮ GOMPERTZOVY-MAKEHAMOVY FUNKCE:
 KINGOVA-HARDYHO METODA**

**ESTIMATION OF THE PARAMETERS OF GOMPERTZ-MAKEHAM FUNCTION:
 KING-HARDY METHOD**

ABSTRAKT

Úmrtnost nejstarších osob bývá ovlivněna jak náhodnými, tak systematickými odchylkami. Proto je nezbytné ji modelovat. Již po dlouhou dobu se používá pro modelování úmrtnosti nejstarších osob především Gompertzova-Makehamova funkce. Její výhodou je, že k odhadu neznámých parametrů není třeba žádný speciální software. Můžeme také použít Kingovu-Hardyho metodu. Cílem tohoto článku je představit uvedenou metodu a zmínit její výhody a nevýhody. Důležitou součástí také bude následná optimalizace získaných odhadů pomocí řešitele v MS Excel. Pro porovnání budou uvedeny i odhady parametrů získané pomocí softwaru DeRaS. Cílem je zjistit, jak se budou lišit hodnoty získané podle Kingovy-Hardyho metody a odhady získané z DeRaSu. Jako vstupní data budou použita data o úmrtnosti mužů a žen v České republice.

ABSTRACT

Mortality at the highest ages is influenced by systematic and random errors. It is therefore needed to be modelled. For modeling of mortality of the oldest persons is the most frequently used the Gompertz-Makeham function. Its main advantage is that for the estimation of unknown parameters is not necessary to have any special software. The King-Hardy method can also be used. The aim of this paper is to present this methodology and list its advantages or disadvantages. Important part of this paper will also be the subsequent optimization of estimated parameters of the Gompertz-Makeham function by means of MS Excel Solver. The parameter estimations obtained by the DeRaS software will be included for the sake of comparison. The aim will be to identify how will the values obtained by King-Hardy method and the estimates obtained by DeRaS software differ male and female mortality data for the Czech Republic will be used as input data.

KLÍČOVÁ SLOVA

úmrtnost, Gompertzova-Makehamova funkce, odhad neznámých parametrů, Kingova-Hardyho metoda, DeRaS

KEY WORDS

mortality, Gompertz-Makeham function, estimation of unknown parameters, King-Hardy method, DeRaS

1. ÚVOD

V poslední době se stále více diskutuje o prodlužování lidského života a také o stárnutí populace. Prodlužování lidského života znamená, že bude přibývat osob, které se budou dožívat velmi vysokého věku (Gavrilov a Gavrilova, [8]). Pro představu, jak se bude vyvíjet úmrtnost v nejvyšších věcích, bude do budoucna

potřeba co nejpřesněji zachytit úmrtnost osob v nejvyšším věku (Koschin, [12]). K tomuto účelu se již po dlouhou dobu používá Gompertzova-Makehamova funkce (G-M funkce) (Boleslawski a Tabeau, [1], Coelho et al., [3], Gompertz, [9], Makeham, [13] nebo Thatcher et al., [14]). Její výhodou je, že pro odhad neznámých parametrů jsou k dispozici počáteční odhadové vzorce (tzv. Kingova-Hardyho metoda). Výsledky této metody jsou ovlivněny počátečním nastavením šířky věkových intervalů a věku, od kterého budeme provádět vyrovnání. Je důležité uvědomit si, že vyrovnání je dobré provádět od 60 let a naopak horní hranice je dána spolehlivostí empirických dat (tato hranice se zpravidla stanovuje na 85 let – pro použitou Kingovu-Hardyho metodu je stanovena na 83 let). Důležitou součástí článku je i následná optimalizace odhadů neznámých parametrů Gompertzovy-Makehamovy funkce. Tato optimalizace především umožňuje vylepšení získaných odhadů parametrů a také výsledné vyrovnání.

2. METODIKA

Nejprve je třeba uvést, jakým způsobem získáme empirické hodnoty měr úmrtnosti (ze kterých potom vycházíme při modelování úmrtnosti Gompertzovou-Makehamovou funkcí):

$$m_x = \frac{M_x}{S_x}, \quad (1)$$

kde M_x je počet zemřelých v dokončeném věku x a \bar{S}_x je střední stav počtu žijících x -letých.

K modelování úmrtnosti nejstarších osob můžeme použít Gompertzovu-Makehamovu funkci. Mezi její výhody patří i to, že k odhadu neznámých parametrů je možné použít počáteční odhadové vzorce (ČSÚ [4], Gavriloa a Gavriloa [7]). Jde o tzv. Kingovu-Hardyho metodu, která se používá pro odhad neznámých parametrů Gompertzovy-Makehamovy funkce.

$$\text{Gompertzova-Makehamova funkce: } m_x = \mu_{x+0,5} = a + b \cdot c^{x+0,5}, \quad (2)$$

kde μ_x je intenzita úmrtnosti, x je věk, a , b a c jsou neznámé parametry Gompertzovy-Makehamovy funkce.

Kingova-Hardyho metoda vychází z poznatku, že Gompertzova-Makehamova funkce obsahuje tři neznámé parametry. Pro jejich odhad budeme tedy potřebovat tři rovnice. Aby byla soustava rovnic snadněji řešitelná, budou pro odhad použity tři stejně dlouhé na sebe navazující intervaly. Nejprve je třeba zvolit věk, od kterého bude prováděno vyrovnání pomocí Gompertzovy-Makehamovy funkce (x_0), a potom šířku intervalu (k) (Fiala [6], Koschin [11]).

$$G_1 = \sum_{x=x_0}^{x_0+k-1} m_x = \sum_{x=x_0}^{x_0+k-1} (a + b \cdot c^{x+0,5}) = k \cdot a + b \cdot c^{x_0+0,5} \cdot (1 + c + \dots + c^{k-1}) \quad (3)$$

$$G_2 = \sum_{x=x_0+k}^{x_0+2k-1} m_x = k \cdot a + b \cdot c^{x_0+k+0,5} \cdot (1 + c + \dots + c^{k-1}) \quad (4)$$

$$G_3 = \sum_{x=x_0+2k}^{x_0+3k-1} m_x = k \cdot a + b \cdot c^{x_0+2k+0,5} \cdot (1 + c + \dots + c^{k-1}) \quad (5)$$

Pro další výpočty budeme uvažovat, že $k = 8$ a $x_0 = 60$.

Pro řešení rovnic provedeme nejprve jejich odečtení. Nejprve odečteme rovnici (4) od rovnice (5). A potom odečteme rovnici (3) od rovnice (4).

$$G_3 - G_2 = b \cdot c^{x_0+k+0,5} \cdot (c^k - 1) \cdot (1 + c + \dots + c^{k-1})$$

$$G_2 - G_1 = b \cdot c^{x_0+0,5} \cdot (c^k - 1) \cdot (1 + c + \dots + c^{k-1})$$

Tím dojde k vyloučení parametru a . V dalším kroku vyloučíme parametr b a zůstane nám jen k -tá mocnina parametru c (tj. c^k):

$$c^k = \frac{G_3 - G_2}{G_2 - G_1} = \frac{b \cdot c^{x_0+k+0,5} \cdot (c^k - 1) \cdot (1 + c + \dots + c^{k-1})}{b \cdot c^{x_0+0,5} \cdot (c^k - 1) \cdot (1 + c + \dots + c^{k-1})}. \quad (6)$$

Neznámý parametr c potom získáme jako k -tou odmocninou z c^k .

Po dalších úpravách získáme počáteční odhad parametru b jako:

$$b = \frac{G_2 - G_1}{c^{x_0+0,5} \cdot (c^k - 1) \cdot (1 + c + \dots + c^{k-1})}. \quad (7)$$

A počáteční odhad parametru a získáme jako:

$$a = \frac{G_1 - b \cdot c^{x_0+0,5} \cdot (1 + c + \dots + c^{k-1})}{k}. \quad (8)$$

Před prováděním dalších úprav je třeba ověřit, zda pro parametry Gompertzovy-Makehamovy funkce platí:

$$b \neq 0, c \neq 0, c \neq \pm 1. \quad (9)$$

Z rovnice (1) vyplývá, že pokud by byl parametr b roven 0 nebo absolutní hodnota parametru c rovna 1, potom by byly míry úmrtnosti konstantní (tzn. nezávislé na věku). To určitě v nejvyšším věku neplatí.

Dále je třeba ověřit, zda platí, že $b > 0$ a $c > 0$.

Po výpočtu počátečních odhadů parametrů Gompertzovy-Makehamovy funkce je dále možné provést jejich následnou optimalizaci. Optimalizace počátečních odhadů parametrů nám umožňuje získat ještě lepší vyrovnání Gompertzovou-Makehamovou funkcí. Jako optimalizační kritérium můžeme použít minimalizaci součtu vážených čtverců odchylek (Dotlačilová et al [5] nebo Fiala [6]):

$$v.č.o. = \frac{S_{t,x} + S_{t+1,x}}{2 \cdot m_{t,x} \cdot (1 - m_{t,x})} \cdot (m_{t,x} - \tilde{m}_{t,x}^{GM})^2, \quad (10)$$

kde $S_{t,x}$ je počet žijících x -letých na počátku roku t , $m_{t,x}$ je věkově-specifická míra úmrtnosti v roce t , $\tilde{m}_{t,x}^{GM}$ jsou vyrovnané specifické míry úmrtnosti pomocí Gompertzovy-Makehamovy funkce.

Při stanovení optimalizačního kritéria je třeba uvědomit si, že vylepšení počátečních odhadů parametrů Gompertzovy-Makehamovy funkce se často provádí pomocí metody nejmenších čtverců. Je tedy třeba ověřit předpoklady, za kterých je možné metodu použít. Vycházíme z předpokladu, že rezidua mají normální rozdělení s nulovou střední hodnotou a konstantním rozptylem. Potom nejlepším odhadem parametrů regresní funkce je odhad, který minimalizuje součet čtverců odchylek vyrovnaných hodnot od napozorovaných. Pokud mají rezidua různé rozptyly, potom

je třeba veličiny transformovat na stejný rozptyl – vydělíme je směrodatnými odchylkami. Nejlepším odhadem je potom ten, který minimalizuje součet čtverců transformovaných reziduí. Tuto úvahu je možné aplikovat i na specifické míry úmrtnosti, kdy předpokládáme, že jednotlivá úmrtí jsou stejně rozdělená a nezávislá. Celkový počet zemřelých potom pokládáme za náhodnou veličinu s binomickým rozdělením (s parametry: střední stav počtu žijících a vyrovnaná míra úmrtnosti). Za předpokladu, že střední stavy jsou velká čísla, je možné binomické rozdělení aproximovat rozdělením normálním. Nakonec parametry normálního rozdělení vydělíme středním stavem. Výsledné váhy získáme jako převrácené hodnoty rozptylu (Dotlačilová et al., [5]).

Výpočet vážených čtverců odchylek provedeme ve věkovém intervalu $<60; y>$, kde y je nejvyšší věk, ve kterém ještě není míra úmrtnosti rovna 0. Tento způsob volby však není vhodný u malé populace, kde by se mohlo stát, že míra úmrtnosti bude 0 i v relativně nízkém věku. Proto je možné věk y stanovit také jako nevyšší věk, ve kterém se ještě míry úmrtnosti výrazně neodchylují od trendu exponenciálního růstu.

Mezi výhody Kingovy-Hardyho metody patří především její jednoduchost. Také nám poskytne poměrně dobré počáteční odhady neznámých parametrů. Na druhou stranu má i nevýhody. Je důležité uvědomit si, že získané výsledky jsou ovlivněny především tím, jaký je stanoven počáteční věk pro vyrovnání (tj. x_0), a potom také šířkou věkového intervalu (tj. k).

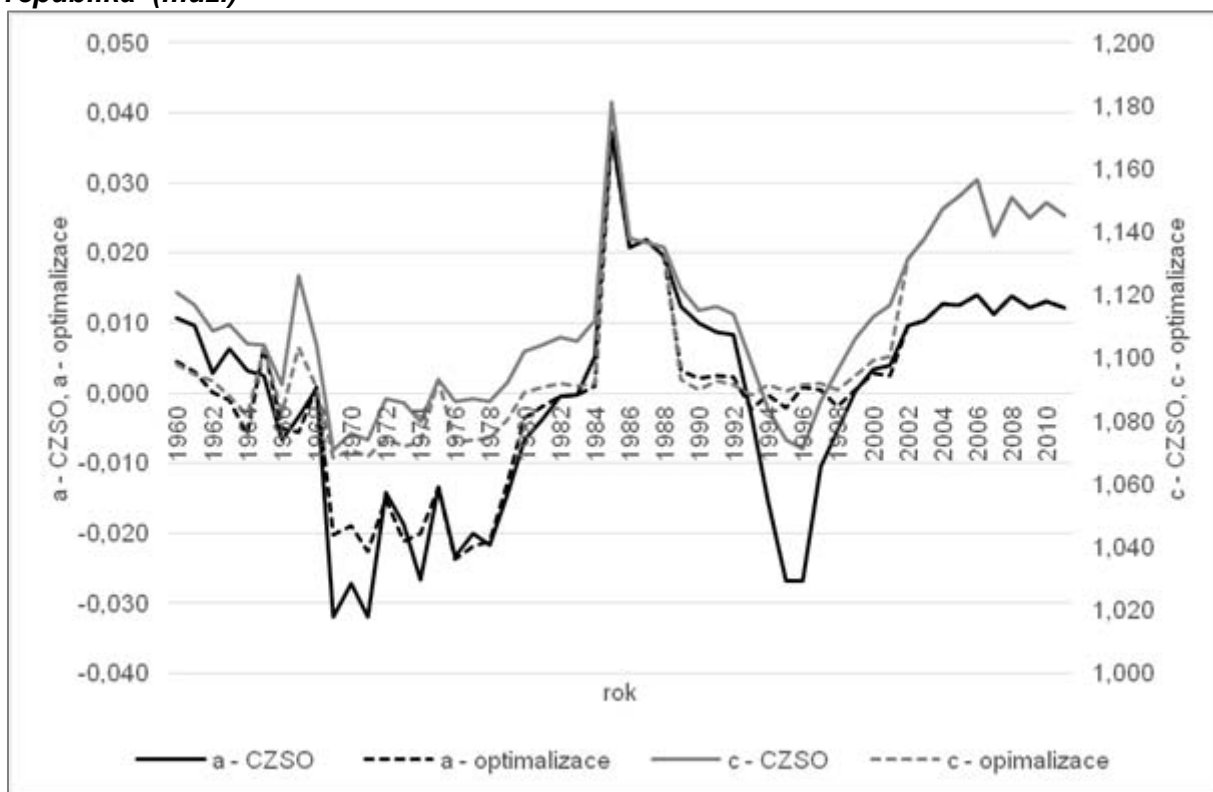
V další části příspěvku bude ukázáno možné vylepšení získaných počátečních odhadů. Jako kritérium pro optimalizaci bude použita minimalizace součtu vážených čtverců odchylek (10). Potřebné výpočty budou provedeny pomocí procedury Řešitel v MS Excel (Fiala, [6]).

Pro porovnání budou uvedeny odhady parametrů získané ze softwaru DeRaS (Burcin et al., [2]). Tento software slouží pro vyrovnávání a následnou extrapolaci křivek úmrtnosti ve vyšším věku. Je zde dostupných několik modelů (např. Gompertzova-Makehamova funkce, Kannistův nebo Thatcherův model). Neznámé parametry jsou odhadovány pomocí nelineární regrese (metoda vážených nelineárních nejmenších čtverců, kdy v každé iteraci dochází k přepočítání vah). Při odhadu neznámých parametrů se vychází z věkového rozpětí od 65 do 90 let.

3. VÝSTUPY

Pro odhady neznámých parametrů Gompertzovy-Makehamovy funkce byla použita data o úmrtnosti české populace od roku 1960 do roku 2011. Nejprve byly odhady získány pomocí Kingovy-Hardyho metody (kde $x_0 = 60$ a $k = 8$). U získaných odhadů byla provedena následná optimalizace.

Graf č. 1: Odhady parametrů G-M funkce a jejich optimalizované hodnoty – Česká republika (muži)



Zdroj: data Human Mortality Database [10], vlastní výpočty

Na grafu č. 1 je znázorněn vývoj odhadů neznámých parametrů získaných pomocí počátečních odhadových vzorců u populace českých mužů. Z grafu jsou pro větší přehlednost vynechány odhady parametru b (ty jsou uvedeny až v tabulce č. 1 resp. v tabulce č. 2).

Ze získaných výsledků je zřejmé, že optimalizované hodnoty parametrů vykazují menší výkyvy v čase (v porovnání s původními odhady).

V tabulce č. 1 jsou uvedeny konkrétní hodnoty parametrů Gompertzovy-Makehamovy funkce získané při použití Kingovy-Hardyho metody. Pokud porovnáme získané odhady se zmíněnými počátečními podmínkami (9), tak je také zřejmé, že ani v jednom z případů nedochází k jejich porušení.

Tabulka č. 1: Odhady parametrů Gompertzovy-Makehamovy funkce získané pomocí Kingovy-Hardyho metody – Česká republika, muži

ČSÚ - muži	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969
a - CZSO	0,010679	0,009663	0,002858	0,006375	0,003194	0,002539	-0,006729	-0,003396	0,000711	-0,031870
b - CZSO	0,000020	0,000026	0,000055	0,000043	0,000067	0,000071	0,000181	0,000012	0,000076	0,000999
c - CZSO	1,120833	1,116857	1,108633	1,110505	1,104763	1,104006	1,091502	1,126194	1,104051	1,070650
ČSÚ - muži	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
a - CZSO	-0,027107	-0,031928	-0,014224	-0,018769	-0,026546	-0,013524	-0,023408	-0,020058	-0,021598	-0,014336
b - CZSO	0,000668	0,000801	0,000270	0,000322	0,000515	0,000185	0,000318	0,000286	0,000309	0,000195
c - CZSO	1,076187	1,073979	1,087246	1,085771	1,079864	1,093044	1,086353	1,087282	1,086258	1,092204
ČSÚ - muži	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
a - CZSO	-0,006639	-0,003733	-0,000475	-0,000195	0,005298	0,037173	0,020806	0,021978	0,019526	0,012417
b - CZSO	0,000098	0,000080	0,000066	0,000072	0,000042	0,000000	0,000006	0,000006	0,000007	0,000019
c - CZSO	1,102033	1,104329	1,106579	1,105434	1,112093	1,180885	1,137819	1,136646	1,134892	1,122123
ČSÚ - muži	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
a - CZSO	0,009984	0,008721	0,008413	-0,003123	-0,015438	-0,026850	-0,026863	-0,010544	-0,005423	0,000366
b - CZSO	0,000031	0,000027	0,000031	0,000097	0,000301	0,000654	0,000764	0,000227	0,000105	0,000049
c - CZSO	1,115029	1,116521	1,114124	1,098622	1,083677	1,073967	1,071112	1,086193	1,096141	1,106137
ČSÚ - muži	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
a - CZSO	0,003397	0,003873	0,009570	0,010440	0,012646	0,012470	0,014075	0,011260	0,013881	0,012124
b - CZSO	0,000029	0,000022	0,000007	0,000005	0,000002	0,000002	0,000001	0,000004	0,000002	0,000002
c - CZSO	1,113250	1,116624	1,131317	1,137968	1,147439	1,151376	1,156449	1,138693	1,151047	1,144421
ČSÚ - muži	2010	2011								
a - CZSO	0,013067	0,012132								
b - CZSO	0,000002	0,000002								
c - CZSO	1,149265	1,145146								

Zdroj: data Human Mortality Database [10], vlastní výpočty

V tabulce č. 2 jsou pro porovnání uvedeny odhady parametrů Gompertzovy-Makehamovy funkce, které byly upraveny následnou optimalizací při použití kritéria minimalizace součtu vážených čtverců odchylek.

Tabulka č. 2: Odhady parametrů Gompertzovy-Makehamovy funkce získané pomocí Kingovy-Hardyho metody a následné optimalizace – Česká republika, muži

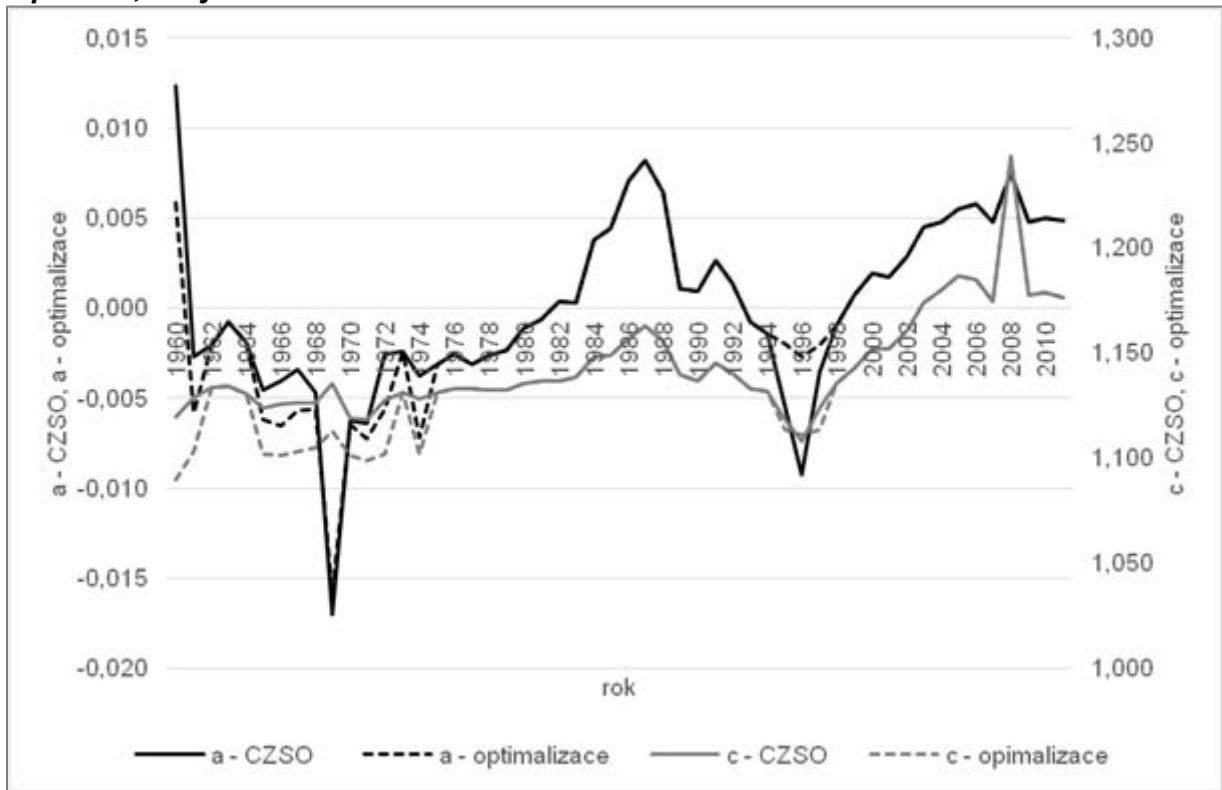
ČSÚ - muži	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969
a - optimalizace	0,010679	0,009663	0,002858	0,006375	0,003194	0,002539	-0,006729	-0,003396	0,000711	-0,031870
b - optimalizace	0,000020	0,000026	0,000055	0,000043	0,000067	0,000071	0,000181	0,000012	0,000076	0,000999
c - optimalizace	1,120833	1,116857	1,108633	1,110505	1,104763	1,104006	1,091502	1,126194	1,104051	1,070650
ČSÚ - muži	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
a - optimalizace	-0,027107	-0,031928	-0,014224	-0,018769	-0,026546	-0,013524	-0,023408	-0,020058	-0,021598	-0,014336
b - optimalizace	0,000668	0,000801	0,000270	0,000322	0,000515	0,000185	0,000318	0,000286	0,000309	0,000195
c - optimalizace	1,076187	1,073979	1,087246	1,085771	1,079864	1,093044	1,086353	1,087282	1,086258	1,092204
ČSÚ - muži	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
a - optimalizace	-0,006639	-0,003733	-0,000475	-0,000195	0,005298	0,037173	0,020806	0,021978	0,019526	0,012417
b - optimalizace	0,000098	0,000080	0,000066	0,000072	0,000042	0,000000	0,000006	0,000006	0,000007	0,000019
c - optimalizace	1,102033	1,104329	1,106579	1,105434	1,112093	1,180885	1,137819	1,136646	1,134892	1,122123
ČSÚ - muži	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
a - optimalizace	0,009984	0,008721	0,008413	-0,003123	-0,015438	-0,026850	-0,026863	-0,010544	-0,005423	0,000366
b - optimalizace	0,000031	0,000027	0,000031	0,000097	0,000301	0,000654	0,000764	0,000227	0,000105	0,000049
c - optimalizace	1,115029	1,116521	1,114124	1,098622	1,083677	1,073967	1,071112	1,086193	1,096141	1,106137
ČSÚ - muži	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
a - optimalizace	0,003397	0,003873	0,009570	0,010440	0,012646	0,012470	0,014075	0,011260	0,013881	0,012124
b - optimalizace	0,000029	0,000022	0,000007	0,000005	0,000002	0,000002	0,000001	0,000004	0,000002	0,000002
c - optimalizace	1,113250	1,116624	1,131317	1,137968	1,147439	1,151376	1,156449	1,138693	1,151047	1,144421
ČSÚ - muži	2010	2011								
a - optimalizace	0,013067	0,012132								
b - optimalizace	0,000002	0,000002								
c - optimalizace	1,149265	1,145146								

Zdroj: data Human Mortality Database [10], vlastní výpočty

Na grafu č. 2 je vyobrazen vývoj odhadů neznámých parametrů Gompertzovy-Makehamovy funkce získaných pomocí Kingovy-Hardyho metody a jejich

optimalizovaných hodnot pro populaci českých žen. I zde jsou pro lepší přehlednost vynechány odhady parametru b . Konkrétní hodnoty jsou uvedeny v tabulce č. 3, resp. v tabulce č. 4. Ze získaných výsledků je patrné, že optimalizované hodnoty vykazují menší výkyvy, než ty získané pomocí Kingovy-Hardyho metody.

Graf č. 2: Odhady parametrů G-M funkce a jejich optimalizované hodnoty – Česká republika, ženy



Zdroj: data Human Mortality Database [10], vlastní výpočty

Tabulka č. 3: Odhady parametrů Gompertzovy-Makehamovy funkce získané pomocí Kingovy-Hardyho metody – Česká republika, ženy

ČSÚ - ženy	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969
parametr a	0,009747	-0,000908	-0,000098	0,001526	-0,000215	-0,002011	-0,004103	-0,003409	-0,002301	-0,013015
parametr b	0,000028	0,000014	0,000010	0,000008	0,000012	0,000017	0,000028	0,000024	0,000016	0,000023
parametr c	1,103595	1,119125	1,124255	1,126437	1,120488	1,115827	1,108808	1,111101	1,117100	1,121877
ČSÚ - ženy	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
parametr a	-0,004023	-0,002683	-0,000226	-0,000123	-0,001235	-0,001183	0,000246	-0,000753	-0,002733	0,000488
parametr b	0,000029	0,000022	0,000012	0,000010	0,000013	0,000011	0,000008	0,000009	0,000016	0,000009
parametr c	1,109762	1,112806	1,120403	1,123437	1,120599	1,121897	1,126992	1,124494	1,117173	1,124991
ČSÚ - ženy	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
parametr a	-0,000604	0,001481	0,002898	0,002580	0,007681	0,007889	0,008899	0,005985	0,003825	0,000998
parametr b	0,000011	0,000007	0,000005	0,000006	0,000001	0,000001	0,000001	0,000002	0,000004	0,000008
parametr c	1,121956	1,128124	1,131851	1,130860	1,151433	1,151625	1,154084	1,143132	1,134795	1,125350
ČSÚ - ženy	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
parametr a	0,000940	0,002622	0,001344	-0,000783	-0,001416	-0,005492	-0,009261	-0,003523	-0,001007	0,000687
parametr b	0,000005	0,000003	0,000004	0,000006	0,000006	0,000017	0,000035	0,000011	0,000004	0,000003
parametr c	1,137017	1,145037	1,140434	1,133240	1,132073	1,118318	1,107913	1,123800	1,135477	1,142859
ČSÚ - ženy	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
parametr a	0,001942	0,001714	0,002846	0,004464	0,004779	0,005493	0,005765	0,004760	0,007472	0,004763
parametr b	0,000001	0,000001	0,000001	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000
parametr c	1,151719	1,151768	1,160446	1,173986	1,180008	1,186472	1,184799	1,174640	1,243617	1,177641
ČSÚ - ženy	2010	2011								
parametr a	0,005014	0,004828								
parametr b	0,000000	0,000000								
parametr c	1,178656	1,176353								

Zdroj: data Human Mortality Database [10], vlastní výpočty

V tabulce č. 3 jsou uvedeny konkrétní hodnoty parametrů Gompertzovy-Makehamovy funkce získané při aplikaci Kingovy-Hardyho metody u populace českých žen během sledovaného období.

Pokud budeme porovnávat odhady získané během sledovaného období, tak zjišťujeme, že i u žen jsou v některých případech hodnoty parametru *a* záporné. V následující tabulce (tabulka č. 4) jsou uvedeny hodnoty neznámých parametrů, které byly následně optimalizovány pomocí kritéria minimalizace součtu vážených čtverců odchylek.

Z grafického zobrazení je také zřejmé, že optimalizované hodnoty parametrů Gompertzovy-Makehamovy funkce vykazují menší výkyvy v čase než původní neoptimalizované hodnoty.

Tabulka č. 4: Odhady parametrů Gompertzovy-Makehamovy funkce získané pomocí Kingovy-Hardyho metody a následné optimalizace – Česká republika, ženy

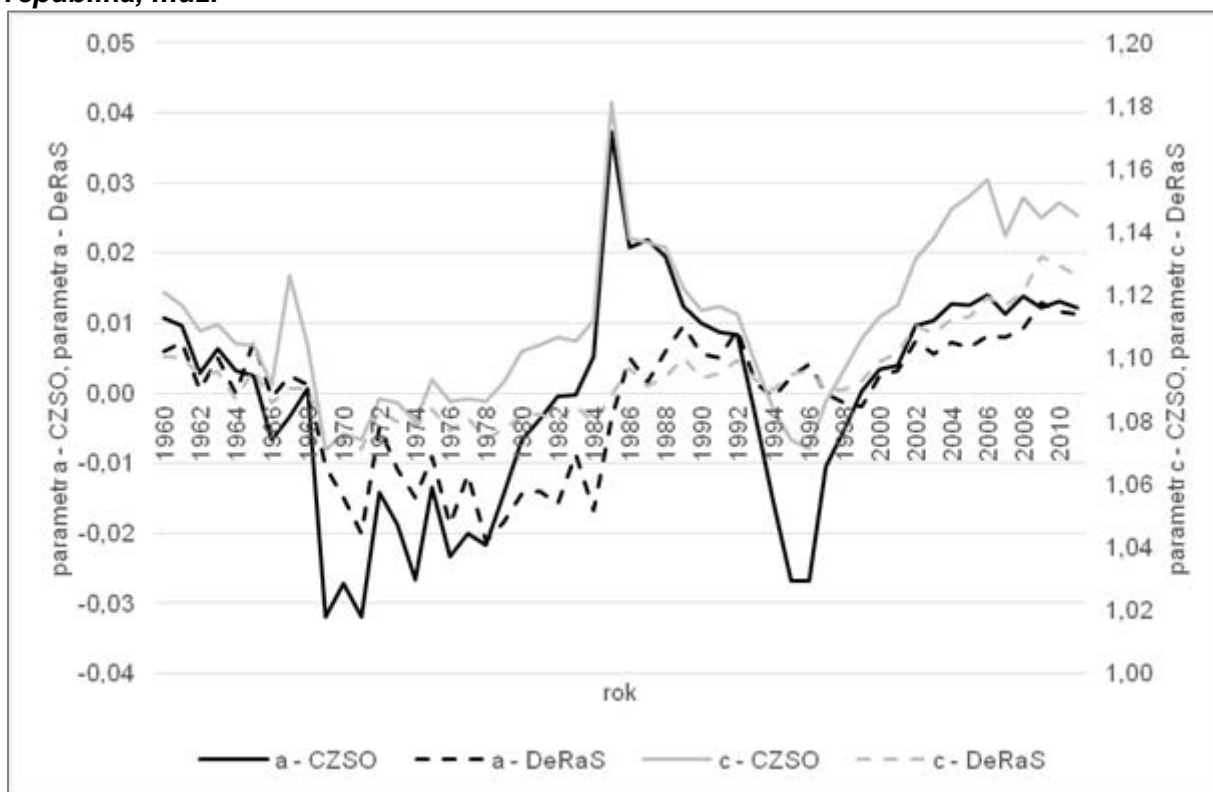
ČSÚ - ženy	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969
a - optimalizace	0,005855	-0,005824	-0,002050	-0,000786	-0,001880	-0,006163	-0,006511	-0,005673	-0,005604	-0,015843
b - optimalizace	0,000076	0,000041	0,000005	0,000004	0,000006	0,000044	0,000046	0,000041	0,000037	0,000039
c - optimalizace	1,089754	1,103254	1,133541	1,134448	1,130364	1,102221	1,101403	1,103013	1,104923	1,113230
ČSÚ - ženy	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
a - optimalizace	-0,006496	-0,007259	-0,005627	-0,002414	-0,007153	-0,003157	-0,002575	-0,003146	-0,002635	-0,002355
b - optimalizace	0,000050	0,000057	0,000044	0,000005	0,000047	0,000005	0,000005	0,000005	0,000005	0,000005
c - optimalizace	1,101162	1,099192	1,101965	1,131216	1,101873	1,131045	1,132979	1,133285	1,132766	1,132568
ČSÚ - ženy	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
a - optimalizace	-0,001097	-0,000665	0,000353	0,000284	0,003783	0,004401	0,007061	0,008189	0,006445	0,001110
b - optimalizace	0,000004	0,000003	0,000003	0,000003	0,000001	0,000001	0,000001	0,000000	0,000001	0,000002
c - optimalizace	1,135464	1,136545	1,136492	1,138548	1,148194	1,148832	1,157013	1,162668	1,155936	1,139535
ČSÚ - ženy	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
a - optimalizace	0,000940	0,002622	0,001344	-0,000783	-0,001416	-0,001988	-0,002701	-0,002058	-0,001007	0,000687
b - optimalizace	0,000003	0,000001	0,000002	0,000004	0,000004	0,000015	0,000017	0,000014	0,000003	0,000002
c - optimalizace	1,137017	1,145037	1,140434	1,133240	1,132073	1,113462	1,111266	1,113823	1,135477	1,142859
ČSÚ - ženy	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
a - optimalizace	0,001942	0,001714	0,002846	0,004464	0,004779	0,005493	0,005765	0,004760	0,007472	0,004763
b - optimalizace	0,000001	0,000001	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000
c - optimalizace	1,151719	1,151768	1,160446	1,173986	1,180008	1,186472	1,184799	1,174640	1,243617	1,177641
ČSÚ - ženy	2010	2011								
a - optimalizace	0,005014	0,004828								
b - optimalizace	0,000000	0,000000								
c - optimalizace	1,178656	1,176353								

Zdroj: data Human Mortality Database [10], vlastní výpočty

V následujících grafech a tabulkách bude vyobrazen vývoj odhadnutých parametrů Gompertzovy-Makehamovy funkce získaný ze softwaru DeRaS (kromě parametru *b*).

Na grafu č. 3 je znázorněn vývoj parametrů Gompertzovy-Makehamovy funkce, který byl získán při výpočtu počátečních odhadových vzorců, a také odhad získaný ze softwaru DeRaS pro populaci českých mužů.

Z grafického výstupu je patrné, že odhady parametrů Gompertzovy-Makehamovy funkce, které byly získány ze softwaru DeRaS, vykazují menší výkyvy v čase.

Graf č. 3: Odhady parametrů G-M funkce získané ze softwaru DeRaS – Česká republika, muži

Zdroj: data Human Mortality Database [10], vlastní výpočty

V tabulce č. 5 jsou uvedeny odhady parametrů Gompertzovy-Makehamovy funkce získané při použití softwaru DeRaS.

Tabulka č. 5: Odhady parametrů Gompertzovy-Makehamovy funkce získané ze softwaru DeRaS – Česká republika, muži

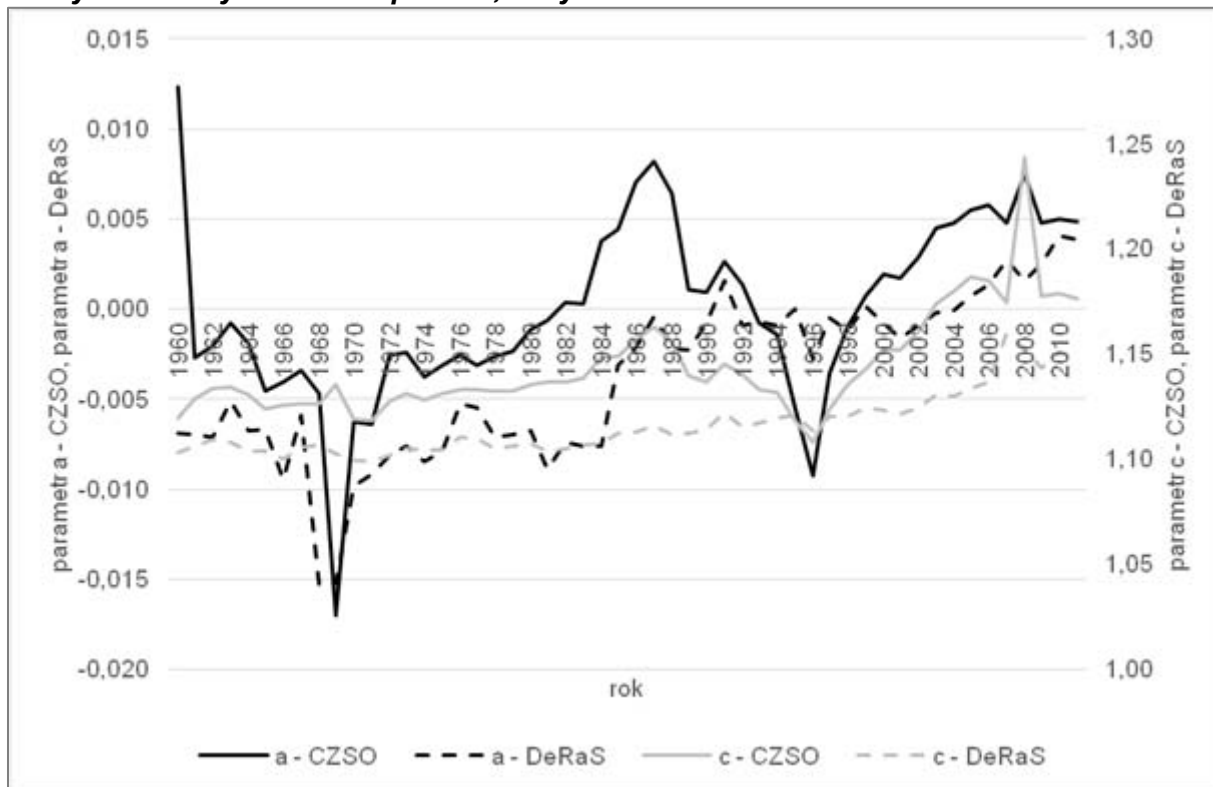
DeRaS - muži	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969
a - DeRaS	0,0059650	0,0071606	0,0005022	0,0050703	-0,0001152	0,0073500	-0,0005525	0,0024168	0,0012044	-0,0104324
b - DeRaS	0,0000574	0,0000585	0,0001032	0,0000854	0,0001652	0,0000796	0,0001796	0,0001297	0,0001376	0,0005045
c - DeRaS	1,1005857	1,1001659	1,0947104	1,0955018	1,0871047	1,0962852	1,0860679	1,0904415	1,0903652	1,0741747
	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
a - DeRaS	-0,0149489	-0,0201113	-0,0048529	-0,0108651	-0,0149482	-0,0091316	-0,0186956	-0,0116982	-0,0211842	-0,0184210
b - DeRaS	0,0005457	0,0006844	0,0002382	0,0003404	0,0004025	0,0002463	0,0004349	0,0003108	0,0005305	0,0004207
c - DeRaS	1,0738046	1,0711395	1,0833734	1,0797697	1,0780307	1,0838344	1,0771151	1,0807858	1,0743371	1,0772821
	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
a - DeRaS	-0,0143769	-0,0140636	-0,0158661	-0,0086450	-0,0166931	-0,0035779	0,0049678	0,0014981	0,0058890	0,0096207
b - DeRaS	0,0003088	0,0002882	0,0003343	0,0002298	0,0003542	0,0001650	0,0000845	0,0001251	0,0000940	0,0000603
c - DeRaS	1,0817500	1,0822076	1,0803180	1,0848524	1,0794535	1,0884785	1,0967256	1,0910426	1,0941283	1,0999593
	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
a - DeRaS	0,0055716	0,0051188	0,0090976	0,0016477	-0,0006700	0,0022849	0,0040477	-0,0001066	-0,0013804	-0,0018239
b - DeRaS	0,0001010	0,0000853	0,0000580	0,0000962	0,0001198	0,0000843	0,0000703	0,0001133	0,0001095	0,0000885
c - DeRaS	1,0934836	1,0953101	1,0994283	1,0930613	1,0900433	1,0946260	1,0959597	1,0897413	1,0900954	1,0928525
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
a - DeRaS	0,0024084	0,0032800	0,0073839	0,0056605	0,0072633	0,0065650	0,0081453	0,0080301	0,0092475	0,0128799
b - DeRaS	0,0000543	0,0000422	0,0000210	0,0000266	0,0000175	0,0000164	0,0000098	0,0000111	0,0000080	0,0000033
c - DeRaS	1,0987289	1,1015842	1,1104236	1,1076480	1,1123836	1,1131706	1,1192185	1,1169795	1,1207721	1,1322424
	2010	2011								
a - DeRaS	0,0115700	0,0112391								
b - DeRaS	0,0000041	0,0000051								
c - DeRaS	1,1292087	1,1259562								

Zdroj: data Human Mortality Database [10], vlastní výpočty

Na první pohled je zřejmé, že ani u jednoho z parametrů nedochází k porušení podmínek (9) v průběhu celého sledovaného období. Pokud se blíže zaměříme na hodnoty parametru a , tak zjišťujeme, že v některých případech dosahuje záporných hodnot (stejně tomu bylo i u Kingovy-Hardyho metody).

Na následujícím grafu je znázorněn vývoj parametrů Gompertzovy-Makehamovy funkce získané při použití Kingovy-Hardyho metody a softwaru DeRaS u populace českých žen.

Graf č. 4: Odhady parametrů G-M funkce získané ze softwaru DeRaS a z Kingovy-Hardyho metody – Česká republika, ženy



Zdroj: data Human Mortality Database [10], vlastní výpočty

Z grafického zobrazení je patrné, že odhady neznámých parametrů Gompertzovy-Makehamovy funkce, které byly získány ze softwaru DeRaS, vykazují o něco menší výkyvy.

Následující tabulka obsahuje odhady parametrů Gompertzovy-Makehamovy funkce pro populaci českých žen, které byly získány ze softwaru DeRaS. Pokud budeme porovnávat hodnoty parametru a získané ze softwaru DeRaS s hodnotami, které byly optimalizovány pomocí kritéria minimalizace součtu vážených čtverců odchylek, zjišťujeme, že u obou metod dostáváme v některých případech záporné hodnoty u parametru a .

Tabuľka č. 6: Odhady parametrov Gompertzovy-Makehamovy funkcie získané ze softwaru DeRaS – Česká republika, ženy

DeRaS - ženy	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969
a - DeRaS	-0,0068803	-0,0070015	-0,0071128	-0,0051109	-0,0067483	-0,0066846	-0,0094260	-0,0058893	-0,0152685	-0,0152540
b - DeRaS	0,0000420	0,0000343	0,0000301	0,0000286	0,0000396	0,0000392	0,0000527	0,0000336	0,0000451	0,0000621
c - DeRaS	1,1031634	1,1060138	1,1089936	1,1081981	1,1037713	1,1040106	1,1000579	1,1057932	1,1069488	1,1023075
	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
a - DeRaS	-0,0098466	-0,0091996	-0,0080882	-0,0075854	-0,0084832	-0,0078857	-0,0052285	-0,0054779	-0,0071344	-0,0069451
b - DeRaS	0,0000581	0,0000596	0,0000448	0,0000375	0,0000408	0,0000386	0,0000242	0,0000252	0,0000363	0,0000330
c - DeRaS	1,0996847	1,0987998	1,1021825	1,1049909	1,1041013	1,1042992	1,1102761	1,1097766	1,1048914	1,1061018
	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
a - DeRaS	-0,0067642	-0,0089006	-0,0074145	-0,0075820	-0,0075779	-0,0031356	-0,0021307	-0,0003924	-0,0021720	-0,0022909
b - DeRaS	0,0000320	0,0000398	0,0000355	0,0000328	0,0000310	0,0000201	0,0000189	0,0000143	0,0000190	0,0000189
c - DeRaS	1,1073381	1,1039585	1,1052282	1,1067340	1,1071614	1,1123514	1,1131018	1,1158912	1,1119625	1,1122880
	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
a - DeRaS	-0,0007873	0,0016030	-0,0009177	-0,0007085	-0,0009294	0,0000149	-0,0027388	-0,0004665	-0,0011712	0,0001781
b - DeRaS	0,0000156	0,0000085	0,0000131	0,0000114	0,0000099	0,0000086	0,0000147	0,0000085	0,0000084	0,0000061
c - DeRaS	1,1143972	1,1218415	1,1155867	1,1174790	1,1193463	1,1210893	1,1134479	1,1203101	1,1202173	1,1243244
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
a - DeRaS	-0,0007642	-0,0016304	-0,0009012	-0,0002238	-0,0000277	0,0007525	0,0013556	0,0027047	0,0015902	0,0025281
b - DeRaS	0,0000064	0,0000074	0,0000058	0,0000038	0,0000037	0,0000027	0,0000021	0,0000005	0,0000008	0,0000012
c - DeRaS	1,1234405	1,1213003	1,1242877	1,1303663	1,1298401	1,1338937	1,1364397	1,1600411	1,1524806	1,1435754
	2010	2011								
a - DeRaS	0,0040614	0,0038196								
b - DeRaS	0,0000006	0,0000007								
c - DeRaS	1,150879372	1,148468033								

Zdroj: data Human Mortality Database [10], vlastní výpočty

4. ZÁVĚR

Jedním z cílů článku bylo představení Kingovy-Hardyho metody používané pro odhad neznámých parametrov Gompertzovy-Makehamovy funkcie a její aplikace na populaci České republiky. Důležitou součástí článku byla i následná optimalizace získaných odhadů, pomocí které docházelo ke zlepšení získaného vyrovnání. Optimalizované hodnoty byly následně porovnány s počátečními odhady. Při porovnání získaných výsledků docházíme jak u mužů, tak u žen k závěru, že Kingova-Hardyho metoda poskytuje poměrně dobré odhady neznámých parametrov. Pokud jsme porovnali počáteční odhady a jejich následnou optimalizaci, tak můžeme říci, že optimalizované hodnoty vykazují menší výkyvy v čase. Jako hodnotící kritérium zlepšení získaného vyrovnání je také možné použít již zmíněný součet vážených čtverců odchylek, který při optimalizaci dává nižší hodnotu než při ponechání původních odhadů, které byly získány pomocí Kingovy-Hardyho metody.

Při použití Kingovy-Hardyho metody je ale důležité uvědomit si, že uvedená metoda má své výhody i nevýhody. Výhodou je její jednoduchost a možnost provedení následné optimalizace počátečních odhadů parametrov pomocí kritéria minimalizace součtu vážených čtverců odchylek. Nevýhodou naopak je, že odhady parametrov jsou ovlivněny zvolením hodnoty x_0 (tj. věku, od kterého provádíme vyrovnání) a šířky intervalu k . Šířku intervalu k je vhodné volit tak, abychom se s poslední horní hranicí věkového intervalu pohybovali okolo 85 let. Tedy věku, který je ještě považován za spolehlivý z hlediska empirických hodnot specifických měř úmrtnosti.

Dále bylo provedeno porovnání odhadů získaných při použití Kingovy-Hardyho metody a ze softwaru DeRaS. Při porovnání získaných výsledků docházíme k závěru, že odhady získané z DeRaSu vykazují menší výkyvy v čase (stejně jako tomu bylo u optimalizovaných hodnot parametrov).

Pokud budeme porovnávat odhady parametrů, které byly optimalizovány pomocí řešitele, a odhady získané z DeRaSu, tak docházíme k závěru, že i mezi nimi jsou rozdíly. To by mohlo být způsobeno tím, že u každé z metod jsou použita odlišná věková rozpětí (Kingova-Hardyho metoda: 60-83 a DeRaS: 65-90). Dalším důvodem by mohla být odlišná konstrukce vah, která je implementována v softwaru DeRaS.

Článek vznikl za podpory projektu GA ČR 15-13283S Projekce populace České republiky podle vzdělání a rodinného stavu.

LITERATÚRA

- [1] BOLESŁAWSKI, L. – TABEAU, E.: Comparing Theoretical Age Patterns of Mortality Beyond the Age of 80. In: TABEAU, E. et al. (Eds.): Forecasting Mortality in Developed Countries: Insights from a Statistical, Demographic and Epidemiological Perspective. Netherlands: Springer, 2001, pp. 127 – 155.
- [2] BURCIN, B. – HULÍKOVÁ TESÁRKOVÁ, K. – KOMÁNEK, D.: DeRaS: software tool for modelling mortality intensities and life table construction. Charles University in Prague, 2012. Dostupný z WWW. <<http://deras.natur.cuni.cz>>
- [3] COELHO, E. – MAGALHÃES, M. G. – BRAVO, J. M.: Mortality and Longevity Projections for the Oldest-Old in Portugal. In: Proceedings of the Joint Eurostat/UNECE (United Nations Statistical Commission and Statistical Office of the Economic Commission for Europe). Work Session on Demographic Projections, Bucharest, Romania, 2007, s. 117 – 132. Dostupné na WWW: <<http://epc2008.princeton.edu/download.aspx?submissionId=80105>>.
- [4] ČSÚ 2014. [cit. 18. 10. 2014]. Dostupný z WWW. <http://www.czso.cz/csu/redakce.nsf/i/umrtnostni_tabulky_metodika>
- [5] DOTLAČILOVÁ, P. – ŠIMPACH, O. – LANGHAMROVÁ, J.: DERAS Versus MS EXCEL Solver in Levelling the Life Expectancy at Birth. In: APLIMAT [CD]. Bratislava, 04. 02. 2014 – 06. 02. 2014. Bratislava: Publishing House of STU, 2014, s. 108 – 114. ISBN 978-80-227-4140-8.
- [6] FIALA, T.: Výpočty aktuárské demografie v tabulkovém procesoru. 1. vyd. Praha: Vysoká škola ekonomická v Praze, 2005. 177 s. ISBN 80-2450821-4.
- [7] GAVRILOV, L. A. – GAVRILOVA, N. S.: Mortality measurement at advanced ages: a study of social security administration death master file. In: North American actuarial journal, vol. 15, no. 3, 2011, s. 432 – 447.
- [8] GAVRILOV, L. A. – GAVRILOVA, N. S.: Stárnutí a dlouhověkost: Zákony a prognózy úmrtnosti pro stárnoucí populace. In: Demografie, 2011, č. 2, s. 109 – 128.
- [9] GOMPERTZ, B.: On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies. In: Philosophical Transactions of the Royal Society of London 115, 1825, s. 513 – 585.
- [10] HUMAN MORTALITY DATABASE 2014. [cit. 18. 3. 2014]. Dostupné z WWW: <www.mortality.org>.
- [11] KOSCHIN, F.: Aktuárská demografie (úmrtnost a životní pojištění). 2. vyd. Praha: Vysoká škola ekonomická v Praze, 2000. 123 s. ISBN 80-245-0022-1.
- [12] KOSCHIN, F.: Jak vysoká je intenzita úmrtnosti na konci lidského života? In: Demografie, 1999, č. 2, s. 105 – 109.
- [13] MAKEHAM, W. M.: On the Law of Mortality and the Construction of Annuity Tables. In: The Assurance Magazine and Journal of the Institute of Actuaries 8, 1860, s. 301 – 310.

[14] THATCHER, R. A. – KANISTÖ, V. – VAUPEL, J. W.: The Force of Mortality at Ages 80 to 120. Odense University Press, 1998. 104 s. ISBN 87-7838-381-1.

RESUMÉ

Článek se zabývá možností alternativního odhadu parametrů Gompertzovy-Makehamovy funkce s využitím Kingovy-Hardyho metody. Výpočty jsou doplněny o možnost vylepšení počátečních odhadů parametrů. Pro porovnání jsou uvedeny odhady parametrů ze softwaru DeRaS. Nejprve byly vypočteny odhady parametrů Gompertzovy-Makehamovy funkce pomocí Kingovy-Hardyho metody. Pro vylepšení získaného vyrovnání byla provedena optimalizace odhadnutých parametrů pomocí kritéria minimalizace součtu vážených čtverců odchylek. Při porovnání získaných výsledků docházíme k závěru, že optimalizované hodnoty parametrů vykazují menší výkyvy v čase.

Odhady parametrů získané pomocí Kingovy-Hardyho metody byly také porovnány s odhady, které byly vypočteny pomocí softwaru DeRaS. I zde docházíme k závěru, že odhady získané z DeRaSu vykazují menší výkyvy v čase.

RESUME

The article deals with an alternative possibility of estimation parameters of the Gompertz-Makeham function by King-Hardy method. The calculations have been supplemented with initial parameter estimates. The parameter estimates obtained by the DeRaS software will be included for the sake of comparison. At first parameter estimates for the Gompertz-Makeham function by King-Hardy method have been calculated. For the improvement of the obtained smoothing the initial parameter estimates were optimized with the help of minimization criterion of the weighted sum of squared deviations. When comparing the obtained results, it can be said that the optimized values are more stable over time.

The estimates from King-Hardy method were also compared with estimates from the DeRaS software. From this can be concluded that values from the DeRaS are more stable over time.

PROFESIJNÝ ŽIVOTOPIS

Ing. Petra Dotlačilová studuje doktorské studium, obor statistika na Fakultě informatiky a statistiky Vysoké školy ekonomické v Praze. Od roku 2014 působí jako asistentka na Katedře matematiky Fakulty informatiky a statistiky. Zabývá se především matematickou demografií (se zaměřením na zachycení úmrtnosti ve vyšším věku).

Doc. Ing. Jitka Langhamrová, CSc., je absolventkou Vysoké školy ekonomické v Praze, obor ekonomická statistika. Je docentkou v oboru statistika. V současné době je vedoucí Katedry demografie na Fakultě informatiky a statistiky Vysoké školy ekonomické v Praze. Je autorkou či spoluautorkou více než 250 publikací. Zabývá se problematikou populačních prognóz, stárnutím populace a jeho důsledky.

KONTAKTY

petra.dotlacilova@vse.cz
langhamj@vse.cz