MODELLING OF MORTALITY – KANNISTO AND WEIBULL MODEL

Petra Dotlačilová

Abstract

Several methods can be used for modelling mortality for people over 60 years of age. One of them is the application of analytical functions. One of the oldest is the Gompertz-Makeham function. Logistic functions are the most often used (eg. Kannisto or Thatcher model). We can also mention Weibull model or polynomial functions of various degrees (e.g. polynomial function of second or third degree). As the other models could be mentioned: Gompertz, Cubic model or modified Gompertz – Makeham function. In this paper, Kannisto and Weibull model will be used for modelling mortality. The main goal is the evaluation of both models used on the basis of the proposed criterion (the aim is that this criterion can be applied to the other models). Different weights will be used in the proposed criterion for comparison. The obtained results will also be evaluated by using adjusted R squared (criterion is used in nonlinear regression).

Key words: mortality, mortality models, Kannisto model, Weibull model

JEL Code: J 10, J 11, J 19

Introduction

Over the years, there has been a gradual increase in life expectancy. This leads to changes in the health and social system. In order to make the best possible changes, it is necessary to have an idea of the development of mortality.

The aging of the population also needs to be addressed as part of the reform of the pension system. One possibility is that the increasing number of people living until old age will be offset by the number of children born. Then there is the second (easier to implement) option, namely a gradual increase in the retirement age. This will provide the necessary number of people of working age to feed the aging population. But problem is how to set these changes in retirement age? One of the bases is demographic mortality tables and then also demographic forecasts of population mortality.

During analyzing mortality, it is necessary to get the best idea of its development. Therefore, it is necessary to choose the best approach to its analysis. One of the possible methods is used of analytical functions (especially for people over 60 years of age).

In this paper, two analytical functions: Kannisto and Weibull model will be used for modelling mortality - these will then be compared. The own proposed test criterion (with different weight variants) will be used for comparison, and results of the adjusted *R* squared obtained from the nonlinear regression outputs will be also published for comparison.

1 Methodology

As was already mentioned, several methods can be used for modelling mortality. One of them is the application of analytical functions. These can be divided into exponential (Gompertz-Makeham function), logistic (Kannisto and Thatcher model), power model (Weibull model) or polynomial functions (polynomial functions of various degrees).

Here Kannisto and Weibull models will be used for modelling mortality curve. The obtained results will then be evaluated using both proposed criterion. Adjusted *R* squared will be used as an additional criterion for evaluation.

1.1 Kannisto model

Kannisto model is one of the most widely used analytical functions used in modelling mortality for people over 60 years of age. It is a type of logistic function. This feature is expected to have a slower increase in mortality. It is therefore suitable for population with lower mortality.

In this article, Kannisto model was used in shape (Kannisto et al, 1994 or Thatcher et al. 1998):

$$\mu_x = \frac{ae^{b.x}}{1+ae^{b.x}}\ , \qquad (1)$$

where a and b are unknown parameters of the model, x is the age.

1.2 Weibull model

Another possibility for modelling mortality is to use Weibull model (Boleslawski, Tabeau, 2001 or Burcin et al., 2010):

$$\mu(x) = b \cdot x^a\ , \qquad (2)$$

where a and b are unknown parameters and x is the age.

Another point of this paper is the application of the test criterion which could be used for the evaluation of obtained results (two variants). This give us information about suitability of concrete model. After that it give us information which from these models could be better for modelling of mortality curve.

1.3 Evaluation of expected results

The important aim of this paper is to propose own criterion which could be used for the evaluation of obtained results. Here it could be used weighted squares of deviations (*WSD*) – minimization criterion. In the first one as weight will be used exposure to risk. In the second one as weight will be use exposure to risk and modelled death rates ($m_{t,x}^{(modelled)}$):

$$WSD = \frac{S_{t,x} + S_{t+1,x}}{2} \cdot (m_{t,x} - m_{t,x}^{(\mathrm{mod}elled)})^2 , \qquad (3)$$

where $m_{t,x}^{(modelled)}$ is modelled mortality curve according to Kannisto (*K*) or Weibull model (*W*), $S_{t,x}$ is number of living at the beginning of year t and $S_{t+1,x}$ is number of living at the beginning of year $t + 1$ (or number of living at the end of year t).

$$WSD2 = \frac{S_{t,x} + S_{t+1,x}}{2 \cdot m_{t,x}^{(\mathrm{mod}elled)} \cdot (1 - m_{t,x}^{(\mathrm{mod}elled)})} \cdot (m_{t,x} - m_{t,x}^{(\mathrm{mod}elled)})^2 , \qquad (4)$$

where $m_{t,x}^{(modelled)}$ is modelled mortality curve according to Kannisto (*K*) or Weibull model (*W*), $S_{t,x}$ is number of living at the beginning of year t and $S_{t+1,x}$ is number of living at the beginning of year $t + 1$ (or number of living at the end of year t).

Finally, the sum of WSD has to be calculated $\sum_{60}^{90} WSD$.

Sum of *WSD* is calculated in age interval <60; 90>. The same age interval was used for the estimation of unknown parameters for Kannisto and Weibull model. This criterion could be used for evaluation of analytical function suitability. Information about the analytical function suitability can also be used to select an estimation function to calculate modal length of life.

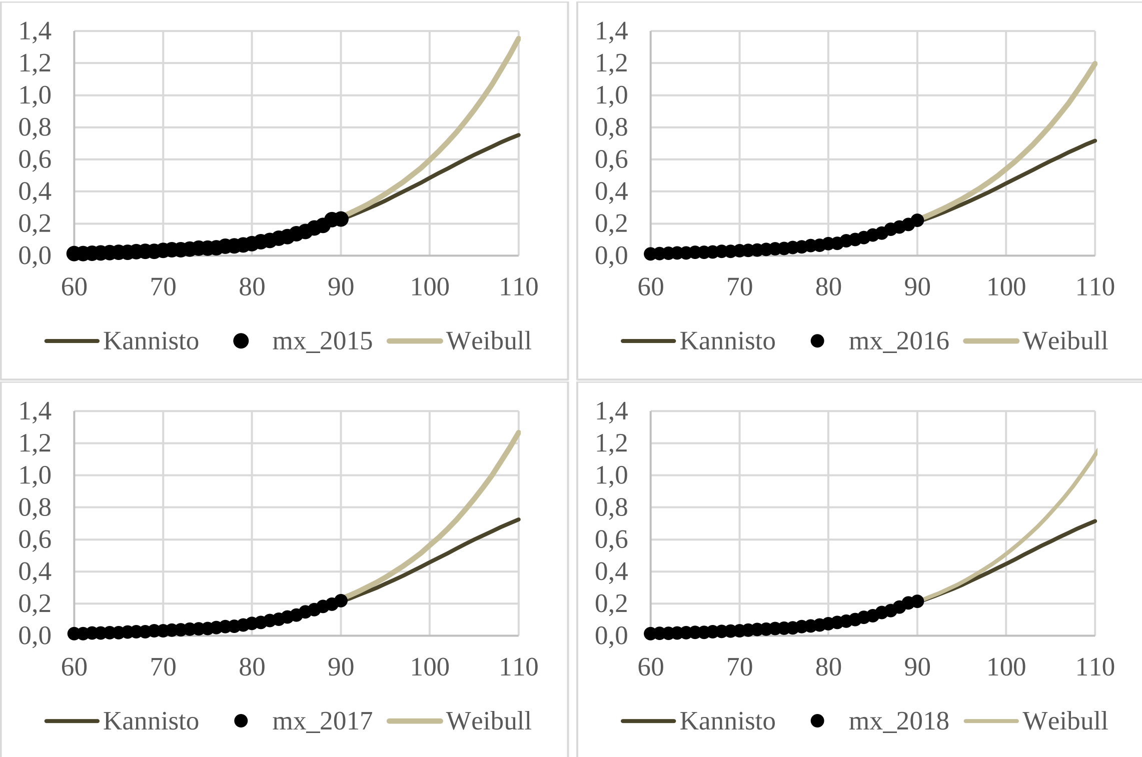
As an additional evaluation criterion will be used adjusted multiple determination coefficient (maximization criterion) ($R^2 adj.$). These values will be used for comparison with values of proposed criterion.

2 Results

The latest data on the mortality of the Czech population available in the Eurostat database were used for the calculations (reporting period: 2015 – 2018).

Figure 1 shows the results of mortality modelling according to both models used for males in the Czech Republic. It is already clear from the graphic representation that Kannisto model is better suited for modelling than Weibull model. We also want to confirm this using the test criterion used.

Fig. 1: Death rates 2015 – 2018 – Czech Republic, males



Source: author’s calculations, Eurostat database

Table 1 shows all values of test criterion. The values of both types of weighted squared deviations show that Kannisto model is more suitable for modelling mortality. This conclusion is also confirmed by the values of adjusted *R* squared.

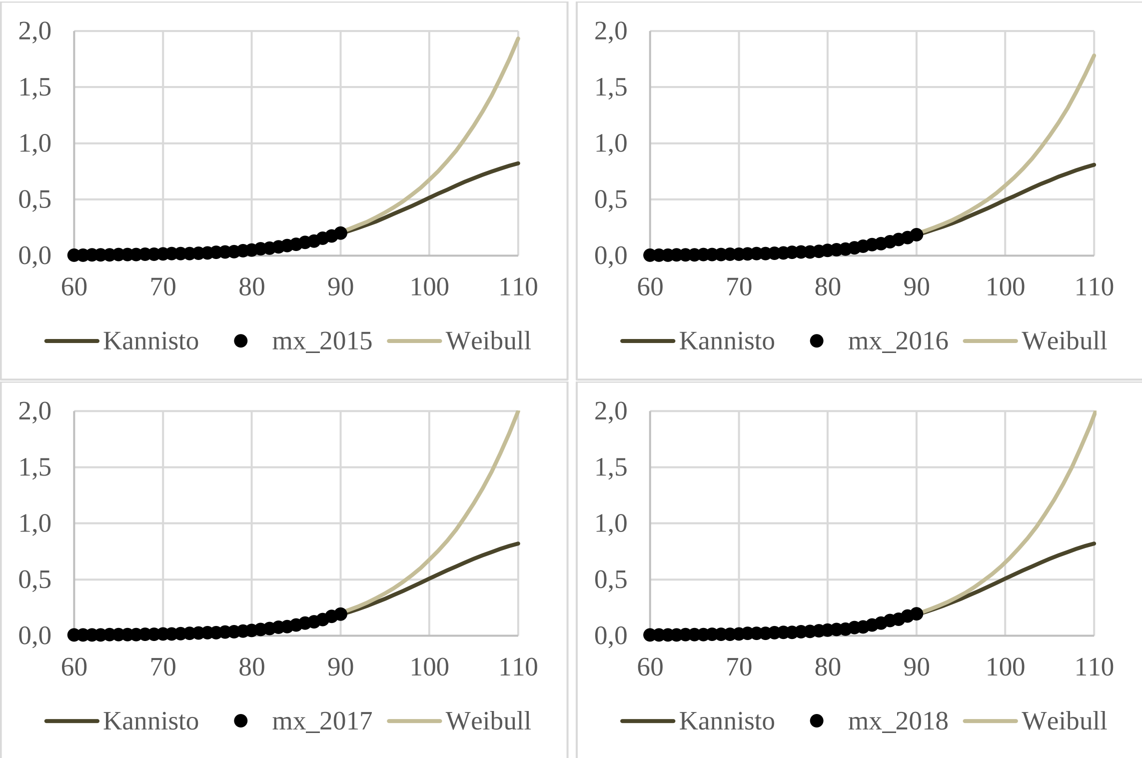
Tab. 1: Values of test criterion – Czech Republic, males

	2015	2016	2017	2018
WSD_Weibull	49,7	38,7	42,9	36,1
WSD_Kannisto	34,1	18,8	21,1	19,0
WSD2_Weibull	2334,0	1667,4	1979,0	1367,3
WSD2_Kannisto	1270,8	1004,4	1009,2	886,5
R^2 adj. Weibull	0,9972	0,9974	0,9971	0,997
R^2 adj. Kannisto	0,9974	0,9979	0,9980	0,998

Source: author’s calculations, Eurostat database

Figure 2 shows the results of mortality modelling according to both models for females' population in the Czech Republic. Here, too, the graphical representation shows that Kannisto model is better suited for modelling mortality of this population. Here, too, the conclusion is based on the values of the test criterion used.

Fig. 2: Death rates 2015 – 2018 – Czech Republic, females



Source: author's calculations, Eurostat database

Table 2 shows the values obtained for both types of weighted squared of deviations. Then results for adjusted *R* squared are also given here. The results obtained by both types of weighted squares of deviations show that the Kannisto model is more suitable for modeling mortality for females over 60 years of age. This conclusion is also confirmed by the results for adjusted *R* squared.

Tab. 2: Values of test criterion – Czech Republic, females

	2015	2016	2017	2018
WSD_Weibull	31,4	28,5	36,3	48,6
WSD_Kannisto	11,7	11,9	14,9	27,8
WSD2_Weibull	2272,9	2204,1	3300,7	3789,5
WSD2_Kannisto	1524,3	1585,9	2096,2	2872,2
R^2 adj. Weibull	0,9996	0,9994	0,9992	0,9980
R^2 adj. Kannisto	0,9996	0,9995	0,9993	0,9980

Source: author's calculations, Eurostat database

The Kannisto model is according to selected evaluation criteria as a more suitable function for modeling mortality in the population of males and females in the Czech Republic. The proposed evaluation criteria thus confirm the conclusion that Kannisto model is more suitable for calculating the mortality tables of the Czech population.

3 Discussion

One way to model mortality is to use analytical functions. Many authors have already dealt with this topic. Logistic functions were used for modeling, for example, by Boleslawski and Tabeau in their article Comparing Theoretical Age Patterns of Mortality Beyond the Age of 80 (Boleslawski and Tabeau, 2001). The authors Burcin, Tesárková and Šídlo also deal with logistic or the other type of functions in their article Nejpoužívanější metody vyrovnávání a extrapolace křivky úmrtnosti a jejich aplikace na českou populaci (Burcin et al., 2010). Mortality modeling using analytical functions is also discussed by Thatcher, Kannistö and Vaupel in The Force of Mortality at Ages 80 to 120 (Thatcher et al., 1998). Koschin deals with this topic in his article Jak vysoká je intenzita úmrtnosti na konci lidského života? (Koschin, 1999).

Analytical functions are commonly used by authors to model mortality curve in ages 60 and higher. However, it is possible to use these analytical functions for calculation of modal length of life (Dotlačilová, 2020).

Analytical functions were also used to model mortality in this article. The benefit is the design of own evaluation criteria that could be used to evaluate and compare the results obtained.

Conclusion

In the first part, this article is focused on mortality modeling based on available analytical functions. Kannisto and Weibull models were chosen for the calculations. It was already clear from the figures that Kannisto model worked better than the second one – it is true for males' and also for females' population.

However, the graphic outputs may not be sufficient. Therefore, the aim of the article was to propose an evaluation criterion that could be used to evaluate the outputs. Two types of sums of weighted squares deviations were used as evaluation criteria. For comparison, results for adjusted *R* squared are published. If we come to the same conclusion when comparing all values of evaluation criteria, it is possible to consider using the proposed criteria for further evaluation.

An important goal of this paper was to design a proper evaluation criterion that would be universally applicable to the evaluation of the obtained results (for each analytical function). Here 2 types of weighted squared deviations (with different weights) were proposed. Based on their values, it is possible to confirm the original conclusion from the figures - Kannisto model is better suited for modeling. This conclusion is also confirmed by the values of adjusted *R* squared. Given the identical conclusions, it is also possible to say that the proposed evaluation criteria can be used to evaluate the results obtained.

Based on the selected evaluation criteria, Kannisto model was chosen as a more suitable model. Based on this model, demographic mortality tables are then calculated, which then serve as one of the bases for changes in the pension system. It is worth mentioning here that the Czech Statistical Office also chose the Kannisto model to calculate mortality tables. On this basis it calculates one of the most important output indicators of mortality tables and that is life expectancy.

For future research it will also be interesting to monitor the effects of the COVID 19 pandemic on mortality. It is likely that the pandemic will also affect mortality (it is likely to increase). Therefore, it will be interesting to observe the suitability of using analytical functions in mortality modelling.

References

BOLESLAWSKI, L., TABEAU, E. Comparing Theoretical Age Patterns of Mortality Beyond the Age of 80. In: Tabeau, E., van den Berg J., A. and Heathcote, Ch. (eds.) 2001. *Forecasting Mortality in Developed Countries: Insights from a Statistical, Demographic and Epidemiological Perspective*. p. 127 – 155. ISBN 978-0-7923-6833-5.

BURCIN, B., TESÁRKOVÁ, K., ŠÍDLO, L. Nejpoužívanější metody vyrovnávání a extrapolace křivky úmrtnosti a jejich aplikace na českou populaci. In: *Demografie* 52, 2010, pp. 77 – 89.

CZECH STATISTICAL OFFICE. Available from: https://www.czso.cz/csu/czso/umrtnostni-tabulky-metodika [cit. 23.11.2019]

DOTLAČILOVÁ, P. Comparison of selected methodologies used for smoothing and extrapolation of mortality curve. In: *The 10th International Days of Statistics and Economics (MSED 2016)* [online]. Prague, 08.09.2016 – 10.09.2016. Slaný: Melandrium, 2016, p. 382–399. ISBN 978-80-87990-10-0. Available from: https://msed.vse.cz/msed_2016/article/234-Dotlacilova-Petra-paper.pdf.

DOTLAČILOVÁ, P. Methods used for the calculations of normal length of life. In: *Aplimat 2019* [flash disk]. Bratislava, 05.02.2019 – 07.02.2019. Bratislava: Publishing house SPEKTRUM STU, 2019, p. 261–265. ISBN 978-80-227-4884-1.

DOTLAČILOVÁ, P. Modal length of life – calculation using Kannisto and Weibull models. In: *Aplimat* [flash disk]. Bratislava, 04.02.2020 – 06.02.2020. Bratislava: Publishing House Spektrum STU, 2020, p. 368–373. ISBN 978-80-227-4983-1.

EUROSTAT. Available from: https://ec.europa.eu/eurostat/data/database [cit. 1.11.2019]

FIALA, T. *Výpočty aktuárské demografie v tabulkovém procesoru*, 1st edition Prague: Oeconomica. 2002. 218 p. ISBN 80-245-0446-4.

GAVRILOV, L., A., GAVRILOVA, N., S. (2011). "Mortality measurement at advanced ages: a study of social security administration death master file." In: *North American actuarial journal* 15 (3): 432 – 447.

KANNISTO, V., LAURISTEN, J., THATCHER, A., R., VAUPEL, J., W. Reductions in mortality at advanced ages - several decades of evidence from 27 countries. In: *Population and development review*, 1994, vol. 20, no. 4, pp. 793-810.

KOSCHIN, F. "Jak vysoká je intenzita úmrtnosti na konci lidského života?" In: *Demografie* 41 (2), 1999: p. 105 – 109.

LANGHAMROVÁ, JA., ARLTOVÁ, M. Life Expectancy and Modal Age at Death in the Czech Republic in 1920-2012. In: *Mathematical Methods in Economics (MME2014)* [CD]. Olomouc, 10.09.2014 – 12.09.2014. Olomouc: Palacký University in Olomouc, 2014, p. 572–577. ISBN 978-80-244-4208-2. CD ISBN 978-80-244-4209-9.

LANGHAMROVÁ, JI., FIALA, T. Ageing of Population of Productive Age in the Czech Republic. In: *International Days of Statistics and Economics* [online]. Prague, 19.09.2013 – 21.09.2013. Slaný: Melandrium, 2013, p. 748–757. 1624 p. ISBN 978-80-86175-87-4. Available from: http://msed.vse.cz/files/2013/105-Langhamrova-Jitka-paper.pdf.

THATCHER, R., A., KANISTÖ, V., VAUPEL, J., W. *The Force of Mortality at Ages 80 to 120*. Odense University Press, 1998. ISBN 87-7838-381-1.

Tab. 3: Parameters' estimates – Kannisto – males (Source: author's calculations, Eurostat database)

	2015		2016		2017		2018	
	value	*p*-value	value	*p*-value	value	*p*-value	value	*p*-value
parameter *a*	0,0000073	0,0000	9,94E-06	0,0000	9,03E-06	0,0000	0,0000101	0,0000
parameter *b*	0,1170459	0,0000	0,112627	0,0000	0,113866	0,0000	0,112394	0,0000

Source: author's calculations, Eurostat database

Tab. 4: Parameters' estimates – Weibull – males (Source: author's calculations, Eurostat database)

	2015		2016		2017		2018	
	value	*p*-value	value	*p*-value	value	*p*-value	value	*p*-value
parameter *a*	8,636289	0,0000	8,389367	0,0000	8,557224	0,0000	8,317772	0,0000
parameter *b*	3,05E-18	0,0000	8,62E-18	0,0000	4,14E-18	0,0000	1,18E-17	0,0000

Source: author's calculations, Eurostat database

Tab. 5: Parameters' estimates – Kannisto – females (Source: author's calculations, Eurostat database)

	2015		2016		2017		2018	
	value	*p*-value	value	*p*-value	value	*p*-value	value	*p*-value
parameter *a*	4,13E-07	0,0000	3,89E-07	0,0000	3,21E-07	0,0000	2,61E-07	0,0000
parameter *b*	0,146879	0,0000	0,146637	0,0000	0,149062	0,0000	0,151558	0,0000

Source: author's calculations, Eurostat database

Tab. 6: Parameters' estimates – Weibull – females (Source: author's calculations, Eurostat database)

	2015		2016		2017		2018	
	value	*p*-value	value	*p*-value	value	*p*-value	value	*p*-value
parameter *a*	1,11E+01	0,0000	1,11E+01	0,0000	1,14E+01	0,0000	1,15E+01	0,0000
parameter *b*	4,14E-23	0,0000	4,41E-23	0,0000	8,1E-24	0,0000	5,33E-24	0,0000

Source: author's calculations, Eurostat database

Contact

Petra Dotlačilová
University of Economics in Prague (Dep. of Mathematics)
W. Churchill Sq. 4, 130 67 Prague 3
petra.dotlacilova@vse.cz