DEPENDENT STRUCTURE INDUCED BY MARKOV CHAIN IN THE MULTIPLE LIFE INSURANCE

STANISŁAW HEILPERN
University of Economics, Department of Statistics,
Komandorska 118-120, 54-345 Wrocław
email: Stanislaw.heilpern@ue.wroc.pl

Abstract
The paper is devoted to the modeling of the dependent structure. The multiple life insurance of married couple is investigated. The lives of spouses may be dependent in this model, so it differs from classical approach, which assumes the independent lives. The survival probabilities are estimated by Makeham approximation. The dependent structure is described by the Markov chain and the join distribution function of lives of spouses are obtain using this Markov chain. The copula describing the depending structure of these lifetimes is estimated based on the obtained empirical cumulative distribution function and Kendall's tau coefficient of correlation is derived in this paper. The Clayton, Gumbel, Frank, AMH and FGM copulas are investigated.

Key words: *multiple life insurance, Markov chain, copula, Kendall's tau.*

1. Introduction

We will study the dependent structure of the lifetimes of insured spouses. This paper is a continuation and extension of the previous papers of the author (Heilpern, 2011, 2014). We still allow, that the lifetimes of spouses may be dependent on the contrary the classical approach, which assumes the independence. This approach is more realistic, because in real life these lifetimes are often dependent. We assume, that the dependent structure is described by the Markov model based on the stationary Markov chain.

We introduce the basic notion connected with the lifetimes analysis and Makeham estimation of the survival probabilities. Next we estimate the parameters describing the Markovian forces of mortalities based on Lower Silesian data during 2011. Makeham estimators and Markovian forces of mortalities,taking into account dependencies, let us to derive the joint cumulative distribution function of the lifetimes of spouses. We compute the Kendall's tau coefficient of correlation based on this joint distribution function. The end of the paper is devoted to selection of copula which the best described such dependent structure. First we select the five families of copulas: Clayton, Gumbel, Frank, Ali-Mikhail-Haq (AMH) and Farlie-Gumbel-Morgenstern (FGM) and next we choose the copulas connected with Kendall's tau form these families. The best copula satisfies the criterion based on the distance between the joint distribution function and theoretical distribution function induced by copula. All calculations were performed using Mathematica 7.

2. General notation

Now, we present the general notions connected with the lifetimes of spouses. Let T_x^M and T_y^W be the remaining lifetimes of x-year-old man and y-year-old woman. These lifetimes take values in [0, w_x^M] and [0, w_y^W], where w_x^M (resp. w_y^W) denotes the difference between the border age of the man (resp. woman) and *x* (resp. *y*).

We denote by symbol ${}_t p_x^M$ a survival function of the lifetime T_x^M, i.e.

$$ {}_t p_x^M = P(T_x^M > t) $$

and by μ_x^M a force of mortality. It satisfies the known formula

$$ {}_t p_x^M = \exp\left(-\int_0^t \mu_{x+r}^M dr\right). \quad \textbf{(1)} $$

The survival function ${}_t p_y^W$ and the force of mortality μ_y^W for women are defined in the similar way.

We can estimate ${}_t p_x^M$ using the life tables. Let l_x^M (l_y^W) be the number of living *x*-years-old men (*y*-years-old women) from initial l_0^M = 10 000 men (women). These values are done in the life tables for *x* = 0, 1, … , 100 (POLISH CSO, 2014). Then

$$ {}_t p_x^M = \frac{l_{x+t}^M}{l_x^M}. $$

We will study the dependence of the lifetimes T_x^M and T_y^W, so we must know their joint distribution.

$$ F_{xy}(t,s) = P(T_x^M \le t, T_y^W \le s). $$

If we know the survival joint function

$$ S_{xy}(t,s) = P(T_x^M > t, T_y^W > s). $$

and survival marginal probabilities ${}_t p_x^M$ and ${}_s p_y^W$ we may compute joint distribution using the following formula

$$ F_{xy}(t,s) = 1 + S_{xy}(t,s) - {}_t p_x^M - {}_s p_y^W. \quad \textbf{(2)} $$

3. Makeham estimation

If we compute the survival probabilities ${}_t p_x^M$ and ${}_t p_y^W$ using the life tables, then we have the values for the integer *x* or *y* only. The Makeham estimation allows us to obtain the remaining values. We assume, that (Denuit et al., 2001)

$$ {}_{t}p_{x}^{M} = s_{M}^{t}\, g_{M}^{c_M^{x+t} - c_M^{x}}, $$

where $c_M > 1$, s_M, $g_M \in [0, 1]$. Hence, we obtain

$$ \mu_{x+t}^{M} = A_M + B_M c_M^{x+t}, \tag{3} $$

where $A_M = -\ln(s_M)$ and $B_M = -\ln(c_M)\ln(g_M)$. In similar way we receive ${}_{t}p_{y}^{W}$ and μ_{y+t}^{W}. We will use (3) in the section 5.

Denuit et al. (2001) decomposed the age range $[v_1, v_2]$ into two parts $[v_1, 40]$ and $[41, v_2]$ and obtain the following estimators of the parameters A_M, c_M and $\beta_M = -(c_M - 1)\ln(g_M)$:

$$ \ln(c_M) = \frac{\sum_{x=41}^{v_2} x \sum_{x=41}^{v_2} \ln(\alpha_x) - (v_2 - 40)\sum_{x=41}^{v_2} x \ln(\alpha_x)}{\left(\sum_{x=41}^{v_2} x\right)^2 - (v_2 - 40)\sum_{x=41}^{v_2} x^2}, $$

$$ \ln(\beta_M) = \frac{1}{v_2 - 40}\left(\sum_{x=41}^{v_2} \ln(\alpha_x) - \ln(c_M) \sum_{x=41}^{v_2} x \right), $$

$$ A_M = \frac{1}{41 - v_1} \sum_{x=v_1}^{40} (\alpha_x - \beta_M c_M^{x}), $$

where $\alpha_x = -\ln({}_{1}p_{x}^{M})$. We can estimate the values of the survival probabilities ${}_{1}p_{x}^{M}$ and ${}_{1}p_{y}^{W}$ using live tables.

Now, we estimate the parameters s_M, s_W, g_M, g_W, c_M and c_W based on the life tables from Polish Central Statistical Office (CSO) contained the data from Lower Silesia (Poland)during 2011 year (Polish General Census). The values of these estimators are presented in table 1. There are the values based on the data from Belgium during 1991 (Denuit et al., 2001) and from Lower Silesia during 2011.

Table 1. The values of the Makeham parameters

	s_M	g_M	c_M	s_W	g_W	c_W
L. Silesia	0.99997	0.99840	1.08329	0.99982	0.99985	1.10569
Belgium	0.99941	0.99960	1.10290	0.99977	0.99983	1.10673

Source: own elaboration.

4. Markov model

Now, we investigate the Markov model based on the stationary Markov chain, see (Wolthuis and Van Hoeck, 1986),(Norberg, 1989); (Denuitet al., 2001). This model let us to derive the joint distribution of the lifetimes of spouses, i.e. (T_x^M, T_y^W), for fixed ages x and y. We have four states in this model (see fig. 1).

Let, $p_{ij}(t, s)$ be the transition probabilities, i.e. the conditional probability that the couple is in state j at time s, given that it was in state i at time t. The Markovian forces of mortality $\mu_{ij}(t)$ from the state i to state j at time t is done by formula

$$\mu_{ij}(t) = \lim_{\Delta t \to 0} \frac{p_{ij}(t, t+\Delta t)}{\Delta t}.$$

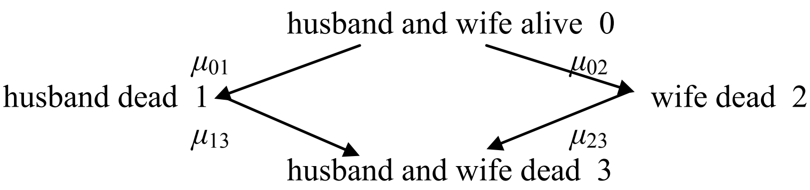

Figure 1. The space of states of Markov model
Source: Denuit et al.(2001).

Let, $p_{ij}(t, s)$ be the transition probabilities, i.e. the conditional probability that the couple is in state j at time s, given that it was in state i at time t. The Markovian forces of mortality $\mu_{ij}(t)$ from the state i to state j at time t is done by formula

$$\mu_{ij}(t) = \lim_{\Delta t \to 0} \frac{p_{ij}(t, t+\Delta t)}{\Delta t}.$$

The following formulas reflect the relation between the transition probabilities and the Markovian forces of mortalities (Denuitet al., 2001):

$$p_{00}(t,s) = \exp\left(-\int_t^s (\mu_{01}(u) + \mu_{02}(u))du \right), \qquad \textbf{(4)}$$

$$p_{ii}(t,s) = \exp\left(-\int_t^s \mu_{i3}(u)du \right), \qquad \textbf{(5)}$$

$$p_{0i}(t,s) = \int_t^s p_{00}(t,u)\mu_{0i}(u)p_{ii}(u,s)du, \qquad \textbf{(6)}$$

where$i$ = 1, 2. The joint and marginal survival functions of the lifetimes T_x^M and T_y^W are equal to

$$P(T_x^M > t, T_y^W > s) = \begin{cases} p_{00}(0,s) + p_{00}(0,t)p_{01}(t,s) & 0 \le t \le s \\ p_{00}(0,t) + p_{00}(0,s)p_{02}(s,t) & 0 \le s \le t \end{cases}, \qquad \textbf{(7)}$$

$$P(T_x^M > t) = p_{00}(0,t) + p_{02}(0,t),$$

$$P(T_y^W > t) = p_{00}(0,t) + p_{01}(0,t).$$

The lifetimes are independent iff$\mu_{01}(t) = \mu_{23}(t)$, $\mu_{02}(t) = \mu_{13}(t)$ (Norberg, 1989). Denuit et al. (2001) made the following simply relations between the forces of mortalities and Markovian forces of mortalities

$$\mu_{01}(t) = (1 + \alpha_{01})\, \mu^M_{x+t} \qquad \mu_{23}(t) = (1 + \alpha_{23})\, \mu^M_{x+t} \quad \textbf{(8)}$$

$$\mu_{02}(t) = (1 + \alpha_{02})\, \mu^W_{y+t} \qquad \mu_{23}(t) = (1 + \alpha_{13})\, \mu^W_{y+t} \quad \textbf{(9)}$$

for fixed x and y. They presented the procedure of the estimation of the parameters α_{01}, α_{02}, α_{23} and α_{13}, see also (Jones, 1997),(Heilpern, 2014). The table 2 contains the values of these estimators. The estimators based on the data from Belgium during 1991 (Denuit et al., 2001) from Poland during 2002 (Heilpern, 2011), from Poland during 2011 (Heilpern, 2014, the values of α_{01} and α_{02} only) and from Lower Silesia during 2011.

Table 2. The values of estimators of the parameters α_{01}, α_{02}, α_{23} and α_{13}

	α_{01}	α_{02}	α_{23}	α_{13}
L. Silesia	–0.0427	–0.0792	0.3603	0.0580
Poland 2011	–0.1257	–0.2009	–	–
Poland 2002	–0.0706	–0.1155	0.2817	–0.0212
Belgium	-0.0929	-0.1217	0.2410	0.0413

Source: own elaboration.

5. Analysis of the joint distribution of lifetimes

5.1. Joint distribution of lifetimes

In this section we estimate the joint distribution of lifetime of spouses T_x^M and T_y^W using the data from Lower Silesia in 2011. We assume, that the husband and wife are is $x = y = 60$ years old. First, we derivethe forces of mortalities μ^M_{60+t} , μ^W_{60+t} using the Makeham approximation and Markovian forces of mortalities $\mu_{ij}(t)$ trading on formulas (8) and (9).

The transition probabilities $p_{ij}(t, s)$ are computed using formulas (4)-(6). These values and formula (7) let us to derive the joint survival function$S_{60,60}(t, s)$. We estimate the joint distribution function $F_{60,60}(t, s)$ using equation (2).

In Table 3, one can find the selected values of the joint distribution function $F_{60,60}(t, s)$ for Lower Silesian men and woman.

Table 3. Selected values of the joint distribution function

	y							
x	65	70	75	80	85	90	95	100
65	0.0042	0.0102	0.0190	0.0315	0.0473	0.0645	0.0786	0.0863
70	0.0103	0.0249	0.0456	0.0748	0.1119	0.1520	0.1851	0.2030
75	0.0173	0.0435	0.0808	0.1312	0.1953	0.2647	0.3220	0.3529
80	0.0247	0.0629	0.1197	0.1973	0.2935	0.3974	0.4834	0.5297
85	0.0311	0.0799	0.1538	0.2588	0.3913	0.5319	0.6481	0.7107
90	0.0355	0.0914	0.1770	0.3007	0.4619	0.6372	0.7808	0.8581
95	0.0376	0.0969	0.1879	0.3204	0.4951	0.6899	0.8534	0.9412
100	0.0381	0.0984	0.1909	0.3258	0.5042	0.7043	0.8747	0.9678

Source: own elaboration.

5.2. Copulas

The copula$C$ is a link between the joint and marginal distributions (Genesta and McKay, 1986),(Nelsen, 1999):

$$F_{xy}(t,s) = P(T_x^M \le t, T_y^W \le s) = C(P(T_x^M \le t), P(T_y^W \le s)).$$

When the random variables are independent, then the corresponding copula takes the form:

$$C_I(u, v) = uv.$$

and for the strict positive (comonotonic) C_M and strict negative (countermonotonic) C_W dependence we obtain:

$$C_M(u, v) = \min\{u, v\}, \qquad C_W(u, v) = \max\{u + v - 1, 0\}.$$

We have

$$C_W(u, v) \le C(u, v) \le C_M(u, v)$$

for every copula C (Nelsen, 1999).

The Archimedean copula is a copula described by the simple, quasi-additive formula (Nelsen, 1999):

$$C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v)),$$

whereφ: $[0, 1] \rightarrow R_+$ is decreasing function, satisfying condition $\varphi(1) = 0$. This function is called a generator.

Archimedean copulas form the families of copulas characterized by some parameter. This parameter described the degree of dependence. The Kendall's tau coefficient of correlation τ,popular measure of dependence, is done by its generator:

$$\tau = 1 + 4\int_0^1 \frac{\varphi(t)}{\varphi'(t)} dt .$$

Now, we introduce the four families of Archimedean copulas, which we will use in the next section. These are the Clayton, Gumbel, Frank and AMH families. The main characteristics: the formula, values of parameters, generator and corresponding Kendall's tau of them is done in Table 4, where $D_1(x) = \frac{1}{x}\int_0^x \frac{t}{e^t - 1} dt$ is Debye'a function.

For Clayton copula in the limit case, when parameter α tend to 0, we have the independence and when α goes to ∞, the strict, positive dependence. We obtain the independence when α tend to 1 for the Gumbel copula, when α tend to 0 for Frank copula and when $\alpha = 0$ for AMH copula. For Frank copula when α goes to $-\infty$ we have strict, positive dependence and negative dependence when α tend to ∞.

Table 4. The characteristics of selected Archimedean copulas

	$C(u, v)$	α	$\varphi(t)$	τ
Clayton	$(u^{-\alpha} + v^{-\alpha} - 1)^{-1/\alpha}$	$\alpha > 0$	$t^{-\alpha} - 1$	$\frac{\alpha}{\alpha+2}$
Gumbel	$\exp(-((-\ln u)^{\alpha} + (-\ln v)^{\alpha})^{1/\alpha})$	$\alpha > 1$	$(-\ln t)^{\alpha}$	$1-\frac{1}{\alpha}$
Frank	$\ln\left(1+\frac{(e^{\alpha u}-1)(e^{\alpha v}-1)}{(e^{\alpha}-1)}\right)$	R\{0}	$-\ln\frac{e^{\alpha t}-1}{e^{\alpha}-1}$	$1-\frac{4}{\alpha}(D_1(-\alpha)-1)$
AMH	$\frac{uv}{1-\alpha(1-u)(1-v)}$	$-1 \le \alpha \le 1$	$\ln\frac{1-\alpha(1-t)}{t}$	$1-\frac{2(\alpha+(\alpha-1)^2\ln(1-\alpha))}{3\alpha^2}$

Source: own elaboration based on Nelsen (1999).

We will use the FGM copula, not Archimedean, also. It is described by formula $C(u, v) = uv + \alpha uv(1 - u)(1 - u)$, where $-1 \le \alpha \le 1$ and it reflects the small degree of dependence only, because $\tau = 2\alpha/9$ in this case. For $\alpha = 0$ we obtain the dependence.

5.3. Copula selection

First, we can compute the Kendall's tau coefficient of correlation for lifetimes T_x^M and T_y^W using the following formula (Nelsen, 1996)

$$\tau = 4\int_x^{\infty}\int_y^{\infty} F_{xy}(t,s)dF_{xy}(t,s) - 1.$$

The coefficient Kendall's tau for data from Lower Silesia during 2011 is equal $\tau = 0.073$. It is small positive dependence in this case. It is smaller than Kendall's tau based on data from cemeteries in Brussels (Denuit et al., 2001), which is equal $\tau = 0.092$. This coefficient based on similar data from Wrocław (Heilpern, 2011) is equal $\tau = 0.156$.

Then, we select the copula connected with the coefficient Kendall's tau $\tau = 0.073$ for the families of copulas mentioned in the previous section. Next, we choose the optimal copula using criterion based on the distance between the distribution functions of lifetimes F_{xy} and the theoretical distribution function induced by copula C(Durrleman et al., 2000):

$$d_2(F_{xy}, F_C) = \left(\int_0^{\infty}\int_0^{\infty}\left|F_{xy}(t,s) - C(F_x(t), F_y(s))\right|^2\right)^{1/2} dxdy.$$

We can see another methods of selection of copula, mainly Archimedean, which best fit to the data in Genest's andRivest's (1993), Frees's et al. (1996), Shemyakin's andYoun's (2006) and Spreeuw's(2006) papers.

We present the values of this distances for these five families of copulas and the values of the parameters corresponding to the Kendall's tau $\tau = 0.073$ in Table 5.
We see, that the Gumbel copula, with parameter $\alpha = 1.0786$ gives us the smallest value of this criterion. So, we can treat it as the best copula, which describe dependent structure of the lifetimes T_x^M and T_y^W in our Lower Silesian case. But, the AMH and FGM copula give us the similar results. The Frank copula is the worst copula.

Table 5. The values of criterion d2 for the selected families of copulas.

	α	d_2
Clayton	0.1572	0.03765
Gumbel	**1.0786**	**0.01252**
Frank	0.0743	0.08964
AMH	0.3019	0.01806
FGM	0.3280	0.01603

Source: own elaboration.

The Gumbel copula with α = 1.1015 was chosen by Denuit et al. (2001) in Belgium, too. Heilpern (2011) chose the AMH copula for the data from Wrocław.

6. Conclusion

The modeling of the dependent structure of the lifetimes of insured spouses is the aim of the paper. We used the Markov model based on the stationary Markov chain to this end. Next we obtained the joint cumulative distribution function describing the joint distribution of such lifetimes. We used the data from Lower Silesia during 2011. Based on this joint distribution function we derived Kendall's tau coefficient of correlation and we chose the optimal copula from five families of copulas using this measure of dependence. The knowledge of the joint distribution function and copula enables us to better investigate such joint distribution of the lifetimes of spouses, derive the insurance premiums and reversionary annuities.

Acknowledgements

The project was funded by the National Science Centre allocated on the basis of decision No DEC-2013/09/B/HS4/00490.

References

1. DENUIT, et al. 2001. Measuring the impact of a dependence among insured lifelengths. In Belgian Actuarial Bulletin, 2001, vol. 1, iss1, pp. 18-39.
2. DURRLEMAN V., NIKEGHBALI, A., RONCALLI T. 2000. Which copula is right now? Groupe de RechercheOperationnelle, Credit Lyonnais, Working Papers 2000.
3. FREES, E. W., CARRIERE J. F., VALDEZ, E. 1996. Annuity valuation with dependent mortality. In The Journal of Risk and Insurance, 1996, vol. 63, pp. 229-261.
4. GENEST, C., MACKAY, R. J. 1986. The joy of copulae: bivariate distributions with uniform margins. In The American Statistician, 1986, vol. 40, pp. 280-283.
5. GENEST, C., RIVEST, L.-P. 1993. Statistical inference procedures for bivariate Archimedean copulas. In JASA, 1993, vol. 88, pp. 1034-1043.
6. HEILPERN, S. 2011. Wyznaczanie wielkości renty w zależnych grupowych ubezpieczeniach na życie. In Prace Naukowe UE Wrocław, 2011, vol. 230, pp. 30-48.
7. HEILPERN, S. 2014. Multiple life insurance – pension calculation. In Proceedings of 17th AMSE, Poland, Jerzmanowice 2014, pp. 114-121, http://amse.ue.wroc.pl/proceedings.html.
8. JONES, B. L. 1997. Methods for the analysis of CCRC data. In North American Actuarial Journal, 1997, vol. 1, pp. 40-54.

9. NELSEN, R. B. 1996. Nonparametric measures of multivariate association. In NELSEN, R. B. (eds.) Distribution with Fixed Marginals and Related Topics. IMS Lecture Notes – Monograph Series, vol. 28, pp. 223-232. Hayward, CA: Instituteof Mathematical Statistics. 1996.
10. NELSEN, R. B. 1999. An Introduction to Copulas. New York: Springer. 1999. ISBN 0-387-98623-5.
11. NORBERG, R. 1989. Actuarial analysis of dependent lives. In Bulletin de l'Association Suisse des Actuaries, 1989, vol. 40, pp. 243-254.
12. POLISH CENTRAL STATISTICAL OFFICE: [cit. 20-02-2014] www.stat.gov.pl.
13. SHEMYAKIN A. E., YOUN H. 2006. Copula models for joint last survival analysis. Applied Stochastic Models in Business and Industry, 2006, vol. 22, pp. 211-225.
14. SPREEW J. 2006. Types of dependence and time-dependent association between two lifetimes in single parameter copula models. Scandinavian Actuarial Journal, 2005, vol. 5, pp. 286-309.
15. WOLTHUIS, H., VAN HOECK L. 1986. Stochastic models for life contingencies. In Insurance: Mathematics and Economics, 1986, vol. 5, pp. 217-254.