Tomasz Jurczyk, Denis Grenda, Michał Smoliński[*] [**] [***]

Modeling of a couple lifetime using exponential distribution

April 3, 2019

1. Basic formulas

Exponential distribution Exponential distribution with a parameter λ is a probability measure with a density on $[0, \infty]$:

$$f(x) = \lambda exp(-\lambda).$$

For a exponentially distributed variable $X \sim Exp(\lambda)$ and $z \leq t$, following identities are true:

$$P(X > t) = e^{-\lambda t}$$

$$P(z \leq X < t) = e^{-\lambda z} - e^{-\lambda t}$$

It is memoryless, what means that:

$$\Pr(X > m + n \mid X \geq m) = \Pr(X > n).$$

Intensity functions for Markov processes Consider continuous-time Markov process with discrete state space J, and transition probabilities $P_{ij}(u, t)$ where $i, j \in J$, and $u \leq t$.

Its transition intensities are given by:

$$\mu_{ij}(t) = \lim_{u \to t} \frac{P_{ij}(u, t)}{u - t}$$

and its exit intensities are given by:

$$\mu_i(t) = \sum_{j \neq i} \mu_{ij} = \lim_{u \to t} \frac{P_{ij}(u, t)}{u - t}$$

Kolmogorov equations For a process defined earlier, following differential equations are satisfied

$$\frac{d}{dt} P_{ij}(z, t) = \sum_{k \neq j} P_{ik}(z, t) \mu_{kj}(t) - P_{ij}(z, t) \mu_j(t)$$

$$\frac{d}{dz} P_{ij}(z, t) = P_{ij}(z, t) - \sum_{k \neq i} P_{kj}(z, t) \mu_{ik}(t)$$

The first is called the Kolmogorov Forward Equation, and the latter is the Kolmogorov Backward Equation.

[*] Tomasz Jurczyk - derivation of formulas in 2, 3,6
[**] Denis Grenda - coding
[***] Michał Smolinski - documentation

2. Description of a problem

Consider two people: A and B, who want to buy a joint life insurance. Suppose their lifetimes T_A, T_B are independent and exponentially distributed, such that:

$$T_A \sim Exp(\lambda_A)$$

$$T_B \sim Exp(\lambda_B)$$

Then we get a continuous-time Markov process with four states:
— state 1 - A and B both alive
— state 2 - A died, B alive
— state 3 - B died, A alive
— state 4 - A and B both dead

Let $P_{ij}(x, y)$ be transition probability from state i to state j during the time interval $[x, y]$. Assume that A and B won't both die in the same moment.

Then:

$$P_{11}(z, z+t) = e^{-\lambda_A t} e^{-\lambda_B t}$$

$$P_{12}(z, z+t) = (1 - e^{-\lambda_A t}) e^{-\lambda_B t}$$

$$P_{13}(z, z+t) = e^{-\lambda_A t}(1 - e^{-\lambda_B t})$$

$$P_{14}(z, z+t) = (1 - e^{-\lambda_A t})(1 - e^{-\lambda_B t})$$

$$P_{22}(z, z+t) = e^{-\lambda_B t}$$

$$P_{24}(z, z+t) = 1 - e^{-\lambda_B t}$$

$$P_{33}(z, z+t) = e^{-\lambda_A t}$$

$$P_{34}(z, z+t) = 1 - e^{-\lambda_A t}$$

with the rest of transition probabilities zero everywhere.

Moreover, for every nonnegative z and positive t, following proprety is satisfied:

$$P_{ij}(z, z+t) = P_{ij}(0, t),$$

due to memorylessness of exponential distribution

Derivation of transition probabilities We prove now the identity for $P_{12}(z, z+t)$. The remaining proofs follow a similar pattern.

$$P_{12}(z, z+t) = P(T_A < z+t,\ T_B > z+t \mid T_A,\ T_B > z) = P(T_A < z+t,\ \mid T_A) \cdot P(T_B > z+t \mid T_B > z) =$$
$$= \frac{P(z \leq T_A < z+t)}{P(T_A \geq z)} \cdot \frac{P(T_B > z+t)}{P(T_B \geq z)} = \frac{e^{-\lambda_A z} - e^{-\lambda_A (z+t)}}{e^{-\lambda_A z}} \cdot \frac{e^{-\lambda_B (z+t)}}{e^{-\lambda_B z}} = (1 - e^{-\lambda_A t}) e^{-\lambda_B t}$$

3. Intensity computation

The process introduced in section 2 has constant transition intensities. They are easily derivable from λ_A and λ_B.

We write down derivatives of transition probabilities using Kolmogorov Forward Equation:

Forward derivative of $P_{11}(z,t)$

$$\frac{d}{dt}P_{11}(z,t) = -P_{11}(z,t)(\mu_{12} + \mu_{13} + \mu_{14})$$

$$-(\lambda_A + \lambda_B)e^{-(\lambda_A+\lambda_B)t} = -e^{-(\lambda_A+\lambda_B)t}(\mu_{12} + \mu_{13} + \mu_{14})$$

That results with:

$$\lambda_A + \lambda_B = \mu_{12} + \mu_{13} + \mu_{14}$$

Forward derivative of $P_{22}(z,t)$

$$\frac{d}{dt}P_{22}(z,t) = -P_{22}(z,t)\mu_{24}$$

$$-\lambda_B e^{-\lambda_B t} = -e^{-\lambda_B t}\mu_{24}$$

That results with:

$$\lambda_B = \mu_{24}$$

Forward derivative of $P_{33}(z,t)$

$$\frac{d}{dt}P_{33}(z,t) = -P_{33}(z,t)\mu_{34}$$

$$-\lambda_A e^{-\lambda_A t} = -e^{-\lambda_A t}\mu_{34}$$

That results with:

$$\lambda_A = \mu_{34}$$

Forward derivative of $P_{12}(z,t)$

$$\frac{d}{dt}P_{12}(z,t) = P_{11}(z,t)\mu_{12} - P_{12}(z,t)\mu_{24}$$

$$(\lambda_A + \lambda_B)e^{-(\lambda_A+\lambda_B)t} - \lambda_B e^{-\lambda_B t} = \mu_{12}e^{-(\lambda_A+\lambda_B)t} + \mu_{24}e^{-(\lambda_A+\lambda_B)t} - \mu_{24}e^{-\lambda_B t}$$

That results with:

$$\lambda_A + \lambda_B = \mu_{12} + \mu_{24},\ \lambda_B = \mu_{24}$$

Forward derivative of $P_{13}(z,t)$

$$\frac{d}{dt}P_{13}(z,t) = P_{11}(z,t)\mu_{13} - P_{13}(z,t)\mu_{34}$$

$$(\lambda_A + \lambda_B)e^{-(\lambda_A+\lambda_B)t} - \lambda_A e^{-\lambda_B t} = \mu_{13}e^{-(\lambda_A+\lambda_B)t} + \mu_{34}e^{-(\lambda_A+\lambda_B)t} - \mu_{34}e^{-\lambda_B t}$$

That results with:

$$\lambda_A + \lambda_B = \mu_{13} + \mu_{34},\ \lambda_A = \mu_{34}$$

These results lead to following values of μ_{ij}:

$$\mu_{12} = \mu_{34} = \lambda_A$$

$$\mu_{13} = \mu_{24} = \lambda_B$$

$$\mu_{14} = 0$$

4. Model calibration

First, we fitted densities of λ_A and λ_B to the respective life tables, such as distances:

$$\sum_{i=0}^{100} \left(e^{-\lambda_A t} - P_A(i)\right)^2$$

$$\sum_{i=0}^{100} \left(e^{-\lambda_B t} - P_B(i)\right)^2$$

are minimized, where $P_A(i)$ and $P_B(i)$ are empirical survival rates, derived from life tables.

We made the fit for λ_A using `TTZ-16M` (based on male survival rates) lifetables and fit for λ_B using `TTZ-16K` lifetables (based on female survival rates). Both are included in section 7.[1]

These computations lead to following values of λ_A, λ_B

$$\lambda_A = 0.007755984$$

$$\lambda_B = 0.005310376$$

5. Fitting curve plots

Figures 1, 2, 3, 4 show us the results of model calibration.

Figure 1 reveals transition probabilities according to the model. We can notice that P_{12} and P_{13} graphs are more "convex", while P_{24} and P_{34} graphs are more "flat"

Figure 2 display occupancy probabilities. They graphs look very similar due to low values of λ_A and λ_B probabilities.

Figures 3 and 4 show how modeled probabilities are related to actual survival rates. We can see that exponential function does not fit very well, so the results may be inaccurate.

[1] They are also available at URL `https://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5470/2/11/1/trwanie_zycia_w_2016.zip` and in the repository `https://github.com/ozonowicz/health-exponential-model`

Figure 1. Transition probabilities according to exponential model

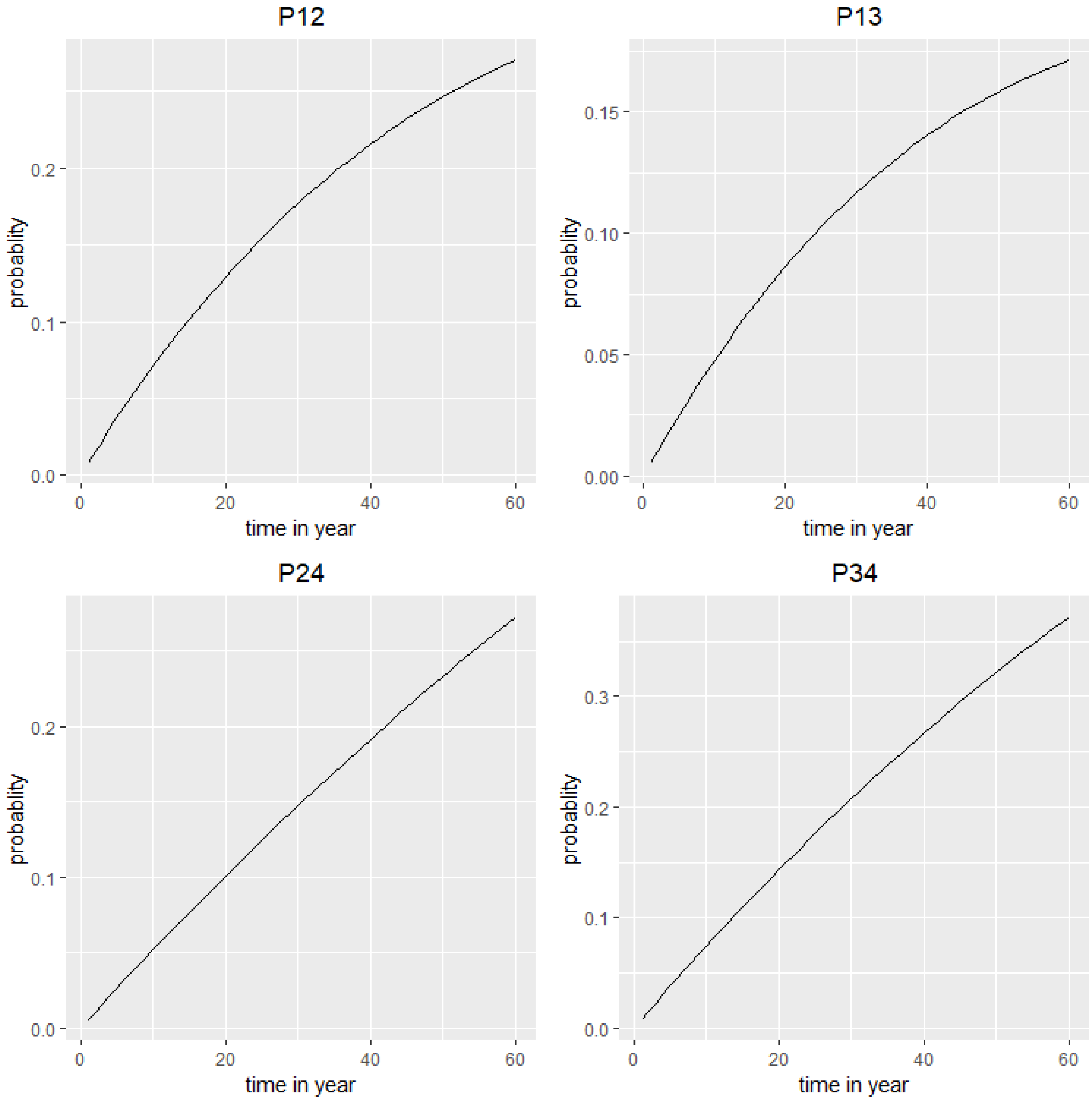


Figure 2. Occupancy probabilities according to exponential model

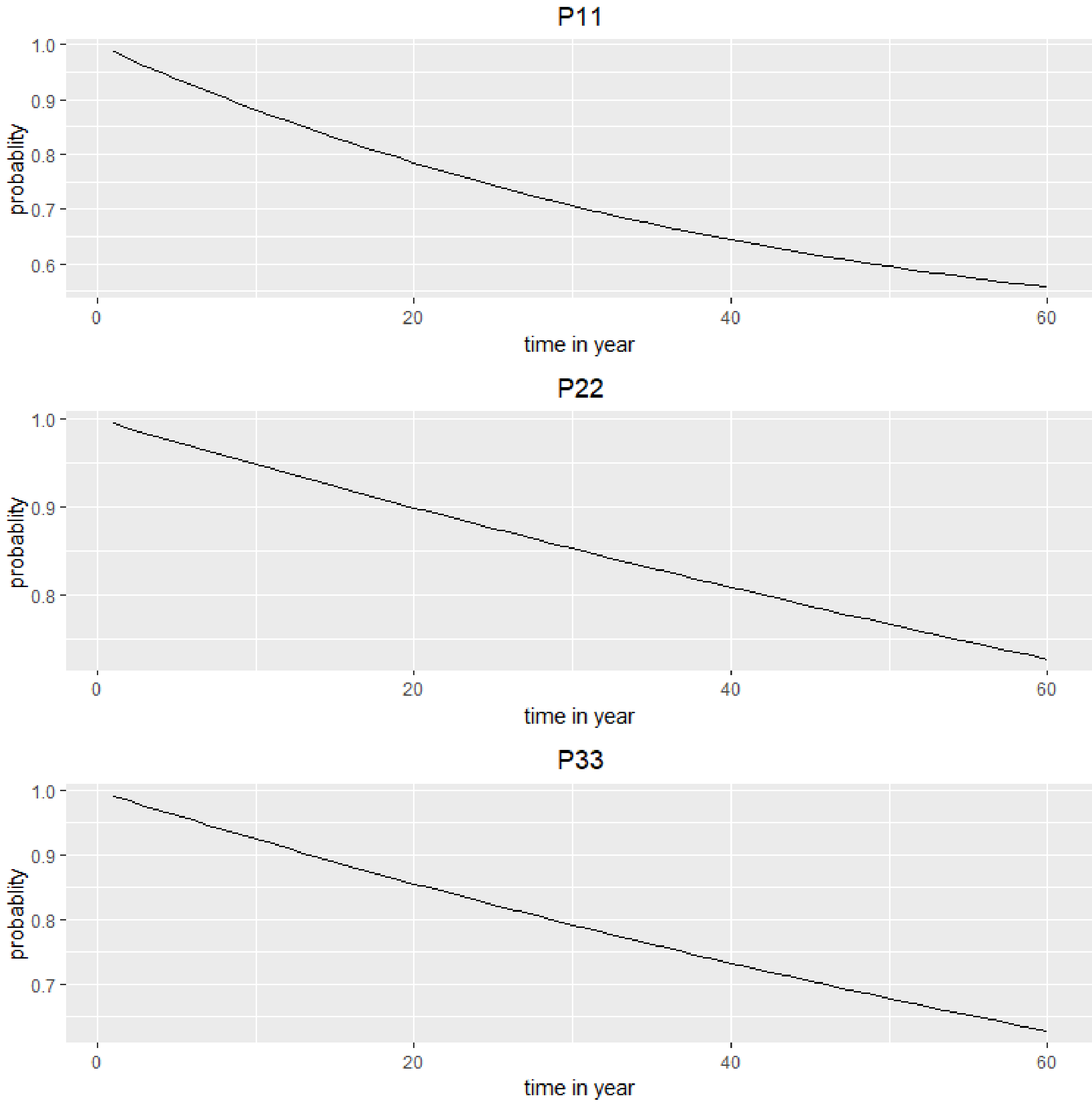


Figure 3. Modeled vs. actual survival rates - male

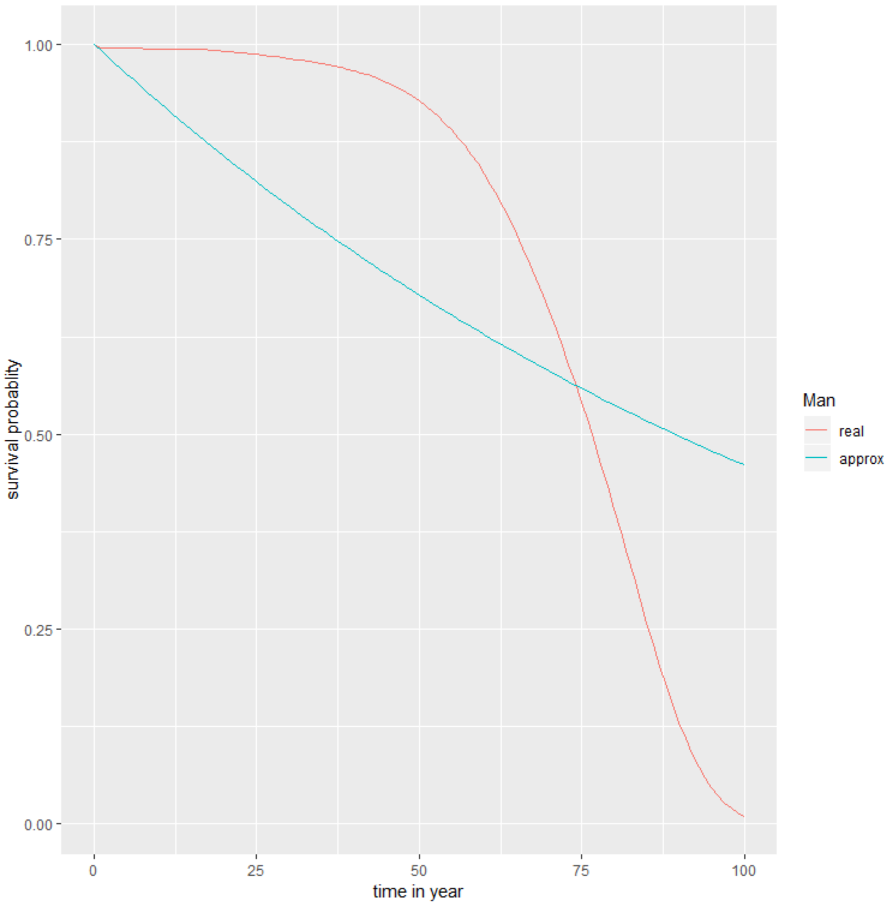
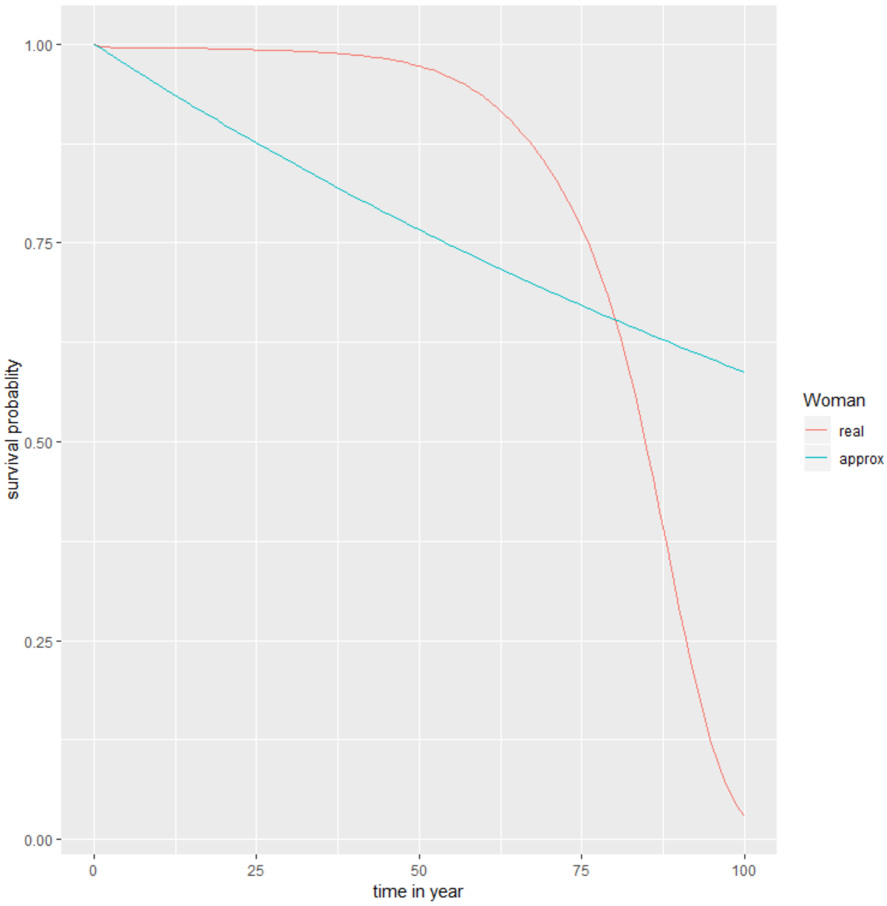


Figure 4. Modeled vs. actual survival rates - female



6. Insurance pricing example using exponential model

Consider an insurance, where a couple pays a yearly premium P as long as both are alive, and their descendants are given a lump sum L when both die in the same year. Assume they buy the insurance when both x-year old

Then present value of the sum of premiums is:

$$PV(premiums) = \sum_{k=0}^{\omega-x} P \cdot \left(\frac{1}{1+i}\right)^k \cdot P_{11}(x, x+k) = P \cdot \sum_{k=0}^{\omega-x} \left(\frac{1}{1+i}\right)^k \cdot e^{-(\lambda_A+\lambda_B)k} =$$

$$= P\frac{(1+i)^{\omega-x+1} - e^{-(\lambda_A+\lambda_B)(\omega-x+1)}}{(1+i-e^{-(\lambda_A+\lambda_B)})(1+i)^{\omega-x}}$$

and the present value of lump sum is :

$$PV(lump\ sum) = \sum_{k=1}^{\omega-x} L \cdot \left(\frac{1}{1+i}\right)^k \cdot P_{14}(k-1, k) = L \cdot (1-e^{-\lambda_A t})(1-e^{-\lambda_B}) \sum_{k=1}^{\omega-x} \left(\frac{1}{1+i}\right)^k =$$

$$= L \cdot (1-e^{-\lambda_A t})(1-e^{-\lambda_B}) \frac{(1+i)((1+i)^{\omega+x}-1)}{i(1+i)^{\omega+x+1}}$$

where ω is the maximum lifetime and i is a yearly interest rate

For $\omega = 100$, $n = 30$, $i = 0.02$, $P = 10$ and $L = 1000$, the respective present values are:

$$PV(premiums) = 292.2383$$

$$PV(lump\ sum) = 1.534411$$

what means that such insurance product is profitable for an insurance company.

7. Appendix - used lifetables

Lifetables used here are available at URL `https://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5470/2/11/1/trwanie_zycia_w_2016.zip` and in the repository `https://github.com/ozonowicz/health-exponential-model`

Table 1. TTZ16-M life tables, used to model T_A lifetime

x	l_x	x	l_x	x	l_x	x	l_x
0	100000	26	98591	51	92126	76	51467
1	99552	27	98494	52	91445	77	48810
2	99529	28	98393	53	90702	78	46076
3	99510	29	98289	54	89892	79	43270
4	99494	30	98179	55	89010	80	40402
5	99481	31	98064	56	88053	81	37488
6	99469	32	97941	57	87017	82	34546
7	99459	33	97810	58	85898	83	31599
8	99449	34	97670	59	84694	84	28671
9	99439	35	97521	60	83401	85	25790
10	99429	36	97361	61	82019	86	22981
11	99418	37	97190	62	80547	87	20274
12	99407	38	97006	63	78986	88	17693
13	99395	39	96806	64	77336	89	15262
14	99380	40	96589	65	75600	90	13001
15	99362	41	96351	66	73781	91	10926
16	99337	42	96089	67	71883	92	9048
17	99301	43	95800	68	69909	93	7376
18	99254	44	95481	69	67862	94	5912
19	99195	45	95128	70	65742	95	4653
20	99124	46	94738	71	63551	96	3592
21	99044	47	94309	72	61288	97	2716
22	98958	48	93836	73	58951	98	2009
23	98869	49	93318	74	56536	99	1451
24	98778	50	92749	75	54042	100	1022
25	98685						

Table 2. TTZ16-K life tables, used to model T_B lifetime

x	l_x	x	l_x	x	l_x	x	l_x
0	100000	26	99281	51	97028	76	75016
1	99645	27	99256	52	96761	77	73012
2	99621	28	99230	53	96464	78	70822
3	99606	29	99203	54	96136	79	68426
4	99597	30	99173	55	95774	80	65807
5	99589	31	99141	56	95375	81	62950
6	99581	32	99107	57	94936	82	59853
7	99573	33	99068	58	94453	83	56521
8	99565	34	99027	59	93925	84	52969
9	99558	35	98982	60	93348	85	49224
10	99552	36	98932	61	92719	86	45323
11	99546	37	98878	62	92036	87	41311
12	99539	38	98818	63	91296	88	37242
13	99531	39	98752	64	90496	89	33174
14	99521	40	98679	65	89632	90	29166
15	99510	41	98597	66	88704	91	25282
16	99496	42	98506	67	87707	92	21577
17	99479	43	98405	68	86640	93	18108
18	99458	44	98291	69	85501	94	14922
19	99436	45	98163	70	84287	95	12057
20	99414	46	98020	71	82993	96	9536
21	99393	47	97862	72	81614	97	7371
22	99373	48	97685	73	80140	98	5558
23	99352	49	97488	74	78561	99	4082
24	99329	50	97270	75	76859	100	2913
25	99306						

8. Appendix - R code used for generating the results

The program is available in the repository `https://github.com/ozonowicz/health-exponential-model`

```
library(tidyr)
library(dplyr)
library(ggplot2)


data <- read.csv("ttz.csv", head=TRUE, dec = ",", sep=";")
x <- data$x
mpx1 <- data$mpx
par1 <- 0.01
kpx1 <- data$kpx
mse <- function(par1, v1, x){
sum((v1 - exp(-x*par1))^2)
}
# mse(par1,x,mpx1)

min_m <- optimize(mse, c(0,1), tol = 0.00001, maximum = FALSE, x = x, v1 = mpx1)$minimum
min_k <- optimize(mse, c(0,1), tol = 0.00001, maximum = FALSE, x = x, v1 = kpx1)$minimum

mpx_est <- exp( -min_m *x)
kpx_est <- exp( -min_k *x)
est <- data.frame(mpx_est, kpx_est)

data <- cbind(data, est)

ggplot(data, aes(x)) +
geom_line(aes(y = mpx, colour = "blue")) +
geom_line(aes(y = mpx_est, colour = "dark green"))+
xlab("time in year") +
ylab("survival probablity")+
scale_color_discrete(name = "Man", labels = c("real", "approx"))

ggplot(data, aes(x)) +
geom_line(aes(y = kpx, colour = "blue")) +
geom_line(aes(y = kpx_est, colour = "dark green"))+
xlab("time in year") +
ylab("survival probablity")+
scale_color_discrete(name = "Woman", labels = c("real", "approx"))

u12 <- min_m
u13 <- min_k
u24 <- min_k
u34 <- min_m


p12 <- function(x,t){
(1-exp(-u12*t))*exp(-u24*t)
}

p13 <- function(x,t){
(1-exp(-u13*t))*exp(-u34*t)
}

p11 <- function(x,t){
1 - p12(x,t) - p13(x,t)
}

p22 <- function(x,t){
exp(-u24*t)
}
p24 <- function(x,t){
```

```
1 - p22(x,t)
}
p33 <- function(x,t){
exp(-u34*t)
}
p34 <- function(x,t){
1 - p33(x,t)
}

y <- 30
t <- 1:60
plot_p12 <- ggplot(data.frame(t, p = p12(y,t)), aes(t,y)) +
geom_line(aes(t,p))+
xlab("time in year") +
ylab("probablity")+
ggtitle("P12")+
theme(plot.title = element_text(hjust = 0.5))


plot_p13 <- ggplot(data.frame(t, p = p13(y,t)), aes(t,y)) +
geom_line(aes(t,p))+
xlab("time in year") +
ylab("probablity")+
ggtitle("P13")+
theme(plot.title = element_text(hjust = 0.5))

plot_p11 <- ggplot(data.frame(t, p = p11(y,t)), aes(t,y)) +
geom_line(aes(t,p))+
xlab("time in year") +
ylab("probablity")+
ggtitle("P11")+
theme(plot.title = element_text(hjust = 0.5))

plot_p22 <- ggplot(data.frame(t, p = p22(y,t)), aes(t,y)) +
geom_line(aes(t,p))+
xlab("time in year") +
ylab("probablity")+
ggtitle("P22")+
theme(plot.title = element_text(hjust = 0.5))

plot_p24 <- ggplot(data.frame(t, p = p24(y,t)), aes(t,y)) +
geom_line(aes(t,p))+
xlab("time in year") +
ylab("probablity")+
ggtitle("P24")+
theme(plot.title = element_text(hjust = 0.5))

plot_p33 <- ggplot(data.frame(t, p = p33(y,t)), aes(t,y)) +
geom_line(aes(t,p))+
xlab("time in year") +
ylab("probablity")+
ggtitle("P33")+
theme(plot.title = element_text(hjust = 0.5))

plot_p34 <- ggplot(data.frame(t, p = p34(y,t)), aes(t,y)) +
geom_line(aes(t,p))+
xlab("time in year") +
ylab("probablity")+
ggtitle("P34")+
theme(plot.title = element_text(hjust = 0.5))


par(mfrow = c(3,2))
plot_p12; plot_p13; plot_p11; plot_p22; plot_p24; plot_p33; plot_p34
```

```


grid.arrange(plot_p12, plot_p13, plot_p24, plot_p34, ncol=2)
grid.arrange(plot_p11, plot_p22, plot_p33)


AV_premiums <- function(x, i, p) #actuarial Value /// x - age, p - premium
{
v = 1/(1+i)
y <- x:100
return(sum(v^(y-30) * p * p11(x,y-x)))
}

AV_premiums(x = 30, i = 0.02, p = 10)


AV_lump_sum<-function(x, i, c)
{
v = 1/(1+i)
y <- 0:(100-(x+1))
return(sum( v^(y+1)*(p24(x,1)*p34(x,1)*c )))
}

AV_lump_sum(x = 30, i =.02, c = 1000)
```