Life Insurance Mathematics

Exam February 2024

Task 1

Markov Model

Definition of a Markov Chain

- Let $(X_t)_{t\in T}$ be a **stochastic process** over (Ω, A, P) with state space S and $T \subset \mathbb{R}$. The process X is called a Markov chain, if for all $n \geq 1$, $t_1 < t_2 < \cdots < t_{n+1} \in T$, $i_1, i_2, \ldots, i_{n+1} \in S$ with $P\left[X_{t_1} = i_1, X_{t_2} = i_2, \ldots, X_{t_n} = i_n\right] > 0$, the following equation holds:

$$P\left[\, X_{t_{n+1}} = i_{n+1} \mid X_{t_k} = i_k \forall k \leq n \,\right] = P[X_{t_{n+1}} = i_{n+1} \mid X_{t_n} = i_n]$$

- This property is called the **Markov property** and states that the probability of transitioning to any particular state depends solely on the current state and time elapsed and is independent of how the system arrived at its current state.

Chapman-Kolmogorov Equation

- Let $(X_t)_{t \in T}$ be a **Markov chain** and let $s \leq t \leq u \in T, i, k \in S$ with $P[X_s = i] > 0$. Furthermore let us denote by $p_{ij}(s,t) \coloneqq P[X_t = j \mid X_s = i]$ the **conditional probability** to change from time s to t from state i to state j, Then we have the following equations:

$$p_{ik}(s,u) = \sum_{j \in S} p_{ij}(s,t) p_{jk}(t,u), \qquad P(s,u) = P(s,t) \times P(t,u)$$

- Hence we can calculate $P(s,u)$ for $s \leq t \leq u$ by matrix multiplication of $P(s,t)$ and $P(t,u)$.

Markov Model

- The **cash flow** $x = (x_k)_{k\in\mathbb{N}}$ in the discrete Markov model is given by

$$x_k = \sum_{i\in S} I_i(k)\times a_i^{\text{Pre}}(k) + \sum_{(i,j)\in S^2} \Delta N_{ij}(k-1)\times a_{ij}^{\text{Post}}(k-1)$$

- In the above, $a_i^{\text{Pre}}(k)$ denotes the **prenumerando** benefits and $a_{ij}^{\text{Post}}(k)$ denotes the **postnumerando** benefits if a transition $i \to j$ happens.

- Furthermore, we have $\mathbb{E}\big[\, \Delta N_{ij}(s) \mid X_t = k \,\big] = p_{ki}(t,s)p_{ij}(s,s+1)$ and $\mathbb{E}[\, I_i(s) \mid X_t = k \,] = p_{ki}(t,s)$ and we assume that $\Delta N_{ij}(-1) = 0$.

Markov Model

- It hence follows from the Chapman-Kolmogorov equations that

$$\mathbb{E}[\, x_s \mid X_t = k \,] = \sum_{i \in S} p_{ki}(t,s) \times a_i^{\mathrm{Pre}}(s) + \sum_{(i,j) \in S^2} p_{ki}(t,s-1) p_{ij}(s-1,s) \times a_{ij}^{\mathrm{Post}}(s-1)$$

- We assume here that $p_{ki}(t,s-1) = 0$ if $t \geq s$.

- The expected cash flow at time s given that we were in state k at time t is given by **two components**:
  - The sum over the **prenumerando benefits** at time step s for all states $\mathrm{i} \in S$ multiplied by the probability of being in state i at time s (given that we were in state k at time t).
  - The sum over the **postnumerando benefits** at time step $s-1$ for all possible transitions from states $\mathrm{i} \in S$ to states $\mathrm{j} \in S$ multiplied by the probability of being in state i at time step $s-1$ (given that we were in state k at time t) and transitioning from i to state j from time step $s-1$ to s.

Thiele Difference Equation

- The **prospective reserve** is defined to be the present value of the future cash flow A given the information at present and we write $V_j^+(t, A) := \mathbb{E}[V\big(t, A \times \chi_{[t,\infty]}\big) \mid X_t = j]$. Re-using what we have derived previously, we can write $\Delta V(t, A) = v(t) x_t$ with $v(t) = \prod_{\tau \le t} \sum_{j \in S} I_j(\tau) v_j(\tau)$.
- The **Thiele difference equation** allows the recursive calculation of the necessary reserves and hence of the necessary single premiums:

$$V_i^+(t) = a_i^{\text{Pre}}(t) + \sum_{j \in S} v_i(t) p_{ij}(t) \{a_{ij}^{\text{Post}}(t) + V_j^+(t+1)\}$$

- The reserve for state i at time t hence consists of the prenumerando benefits corresponding to this state as well as postnumerando benefits for a transition from $i \to j$ and the future necessary reserves, weighted by a discount factor and the probability of this transition happening.

Risk Decomposition

- An insurance premium for a policy can be **decomposed into a risk premium and a savings component**
- The normal subsequent state is defined by a function $\phi: S \rightarrow S, i \rightarrow \phi(i)$, which assigns to each state a state subsequent to it for which no payout is due. We then define the following:
  - The **savings premium** $\Pi_i^{(s)}(t)$ for state i and time interval $]t, t+1]$ is given by

$$\Pi_i^{(s)}(t) = v_t^i V_{\phi(i)}(t+1) - V_i(t)$$

  - The **regular cash flow** or technical pension debt $\Pi_i^{(tr)}(t)$ (which usually coincides with the premium) is given by

$$\Pi_i^{(tr)}(t) = a_i^{\text{Pre}}(t) + v_t^i a_{i\phi(i)}^{\text{Post}}(t)$$

  - The **value at risk** $R_{ij}(t)$ for $i \in S$ and $j \neq \phi(i)$ is given by

$$R_{ij}(t) = V_j(t+1) + a_{ij}^{\text{Post}}(t) - (V_{\phi(i)}(t+1) + a_{i\phi(i)}^{\text{Post}}(t))$$

  - The **total risk premium** $\Pi_{\mathrm{i}}^{r}(t)$ is then given by

$$\Pi_i^{(r)}(t) = \sum_{j \neq \phi(i)} p_{ij}(t) v_t^i R_{ij}(t)$$

Risk Decomposition

- Based on these quantities, a **technical analysis** can be performed by comparing the loss which effectively incurred in the year to the risk premium for the corresponding transaction.
- The regular cash flow is related to the savings premium and the risk premium by the following equation: $-\Pi_i^{(tr)}(t) = \Pi_i^{(r)}(t) + \Pi_i^{(s)}(t)$ (**technical decomposition**). This can be shown using Thiele's difference equations.
- In practice, the gross premium is composed of three parts: savings premiums, risk premium and cost premiums.

Task 2

Stopping to pay premium

Code and calculations can be found here.

Premium for the product – commutation numbers

- Commutation numbers are **auxiliary variables** which simplify the formulas for the present value of insurance policies.
- Discounted numbers of the **living**: $D_x = v^x l_x$, $N_x = \sum_{j=x}^{\infty} D_j$, $S_x = \sum_{j=x}^{\infty} N_j$
- Discounted numbers of the **dead**: $C_x = v^{x+1} d_x$, $M_x = \sum_{j=x}^{\infty} C_j$, $R_x = \sum_{j=x}^{\infty} M_j$
- In our setting, we have the following:
  - The discount factor $v = \frac{1}{1+i_T}$
  - The number of people alive $l_x = \prod_{i=0}^{x-1}(1 - q(i, t_0 + (i - x_0)))$
  - The number of people that died $d_x = l_x - l_{x+1}$

Premium for the product – commutation numbers

- We consider a **mixed endowment** $A_{x:n} = \frac{1}{D_x}(M_x - M_{x+n} + D_{x+n})$ with p**remiums paid prenumerando** $\ddot{a}_{x:n} = \frac{N_x - N_{x+n}}{D_x}$.
- According to the **equivalence principle** at inception, we have $LA_{x:n} = \Pi \ddot{a}_{x:n}$ and hence $\Pi = \frac{LA_{x:n}}{\ddot{a}_{x:n}}$.
- If we put in the numbers from the task description, we arrive at $\Pi = 12302.98$.
- Instead of using commutation numbers, we can also calculate the premium using the equivalence principle at inception directly (see following slides) and we arrive at the same result.

Premium for the product – Markov model

- If we look at the expected payments at time step t, we find the following:
  - The **premiums** paid are equal to $\Pi(t) = p_{**}(x,t)a_{*}^{\text{Pre}}(t)Z^{(t)}$ with $p_{**}(x,t) = \prod_{k\in[x,t-1]} p_{**}(k)$, $a_{*}^{\text{Pre}}(t) = -\Pi$ if $t \in [x, x+n-1]$ and $Z^{(t)}$ is a zero coupon bond with maturity in t years and $Z^{(t)} = 1/(1+i_T)^{t-x}$.
  - The **death benefit** is given by $D(t) = p_{**}(x,t-1)p_{*\dagger}(t-1)a_{*\dagger}^{\text{Post}}(t-1)Z^{(t)}$ with $a_{*\dagger}^{\text{Post}}(t) = L$ if $t \in [x, x+n-1]$.
  - The **life benefit** is given by $L(t) = p_{**}(x,t)a_{**}^{\text{Post}}(t-1)Z^{(t)}$ with $a_{**}^{\text{Post}}(t) = L$ if $t = x+n-1$.
  - According to the equivalence principle, we want to find Π that satisfies the following:

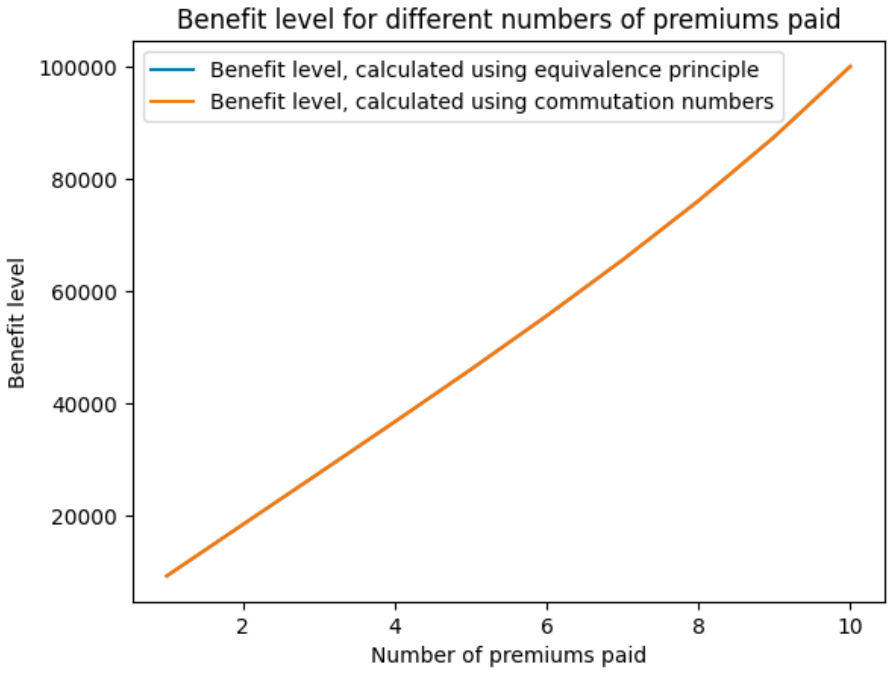
$$\sum_{t\in[x,x+n]} \Pi(t) + D(t) + L(t) = 0$$

Stopping to pay premium – commutation numbers

- According to the task description, the **reduced benefit** if we stop paying premium after p premiums is given by $\tilde{L} = \frac{{}_pV_x}{A_{x+p:n-p}}$.
- For a mixed insurance, we have that the **mathematical reserve** is given by ${}_pV_{x:n} = LA_{x+p:n-kp} - \Pi \ddot{a}_{x+p:n-p}$.
- If we plug in the numbers from the task description, we obtain $\tilde{L} = 9228.77$ for $p = 1$.
- The same result is obtained if we use the equivalence principle directly.

Stopping to pay premium – benefit level

- If we calculate the benefit level $\tilde{L}$ as a **function of the number of premiums** paid, we see the following (there is only one line since the result is equal for both methods, as expected):



Equivalence principle if only $p < n$ premiums are paid

- If we look at the expected payments at time step t, we find the following (changes highlighted):
  - The **premiums** paid are equal to $\widetilde{\Pi}(t) = p_{**}(x,t)a_{*}^{\text{Pre}}(t)Z^{(t)}$ with $p_{**}(x,t) = \prod_{k\in[x,t-1]} p_{**}(k)$, $a_{*}^{\text{Pre}}(t) = -\Pi$ if $t \in [x, x+p-1]$ and $Z^{(t)}$ is a zero coupon bond with maturity in t years and $Z^{(t)} = 1/(1+i)^{t-x}$.
  - The **death benefit** is given by $\widetilde{D}(t) = p_{**}(x,t-1)p_{*\dagger}(t-1)a_{*\dagger}^{\text{Post}}(t-1)Z^{(t)}$ with $a_{*\dagger}^{\text{Post}}(t) = L$ if $t \in [x, x+p-1]$ and $a_{*\dagger}^{\text{Post}}(t) = \tilde{L}$ if $t \in [x+p, x+n-1]$
  - The **life benefit** is given by $\tilde{L}(t) = p_{**}(x,t)a_{**}^{\text{Post}}(t-1)Z^{(t)}$ with $a_{**}^{\text{Post}}(t) = \tilde{L}$ if $t = x+n-1$.
  - According to the equivalence principle, we want to find $\tilde{L}$ that satisfies the following:

$$\sum_{t\in[x,x+n]} \widetilde{\Pi}(t) + \widetilde{D}(t) + \tilde{L}(t) = 0$$

Equivalence principle if only $p < n$ premiums are paid

- Hence, the equivalence principle fulfilled for the **first premium** assuming that only one premium is paid is the normal equivalence principle used to calculate the premium for the product at inception. We hence have $\sum_{t\in[x,x+n]} \Pi(t) + D(t) + L(t) = 0$ and for our example:

$$\sum_{k=80}^{89} \frac{-\Pi \mathrm{p}_{**}(80,\mathrm{k})}{(1+i)^{k-80}} + \sum_{k=80}^{89} \frac{Lp_{**}(80,k)p_{*\dagger}(k)}{(1+i)^{k+1-80}} + \frac{Lp_{**}(80,90)}{(1+i)^{10}} = 0$$

- The equivalence principle fulfilled **after $t = x + 1$ assuming that only one premium** is paid is modified according to the previous slide and we find $\sum_{t\in[x,x+n]} \widetilde{\Pi}(t) + \widetilde{D}(t) + \widetilde{L}(t) = 0$. For the example at hand, this can be written as follows:

$$-\Pi + \frac{\mathrm{L}p_{*\dagger}(80)}{(1+i)} + \sum_{k=81}^{89} \frac{\tilde{L}p_{**}(80,k)p_{*\dagger}(k)}{(1+i)^{k+1-80}} + \frac{\tilde{L}p_{**}(80,90)}{(1+i)^{10}} = 0$$

Task 3

Life Insurance

Code and calculations can be found here.

Premium for the insurance – commutation numbers

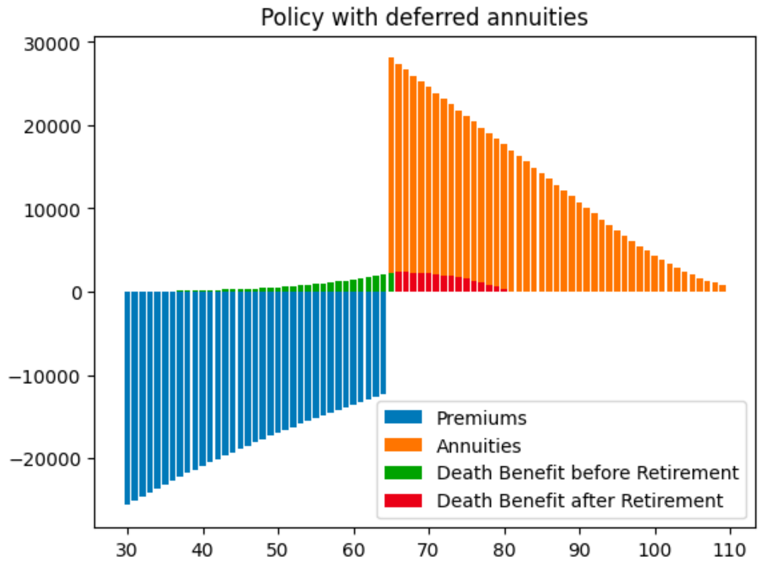
- This insurance policy essentially consists of 4 components:
  - The **premium** paid by the policyholder $\Pi = \mathrm{P}\ddot{a}_{x:m-x} = P\,\frac{N_x - N_m}{D_x}$ (while this is not explicitly mentioned in the task description, we assume that the premium is paid temporarily as in the lecture).
  - The **deferred annuity** $\mathrm{LA} = R_{m-x}\ddot{a}_x = R\,\frac{N_m}{D_x}$.
  - The **death benefit before** payout of annuities $LD_1 = P(IA)_{x:m-x} = P\,\frac{R_x - R_m - (m-x)M_m}{D_x}$.
  - The **death benefit after** the withdrawal has started $LD_2 = \big((m-x)P - nR\big)_m A_{x:n+1} + R_m(DA)_{x:n} = \big((m-x)P - nR\big)\,\frac{M_m - M_{m+n+1}}{D_x} + R\,\frac{nM_m - R_{m+1} + R_{m+n+1}}{D_x}$, where $n = \lfloor (m-x)P/R \rfloor$.
- We get a premium of $P = 25646.58$ by solving $LA + LD_1 + LD_2 = \Pi$.

Premium for the insurance – Markov model

- The **premiums** paid for this policy can be written as $\Pi(t) = p_{**}(x,t)a_{*}^{\mathrm{Pre}}(t)Z^{(t)}$ with $p_{**}(x,t) = \prod_{k\in[x,t-1]} p_{**}(k)$, $a_{*}^{\mathrm{Pre}}(t) = -\mathrm{P}$ if $t \in [x, m-1]$ and $Z^{(t)}$ defined as before.
- The benefit for the policyholder can be split into three components:
  - The **deferred annuity** $LA(t) = p_{**}(x,t)\ a_{*}^{\mathrm{Pre}}(t)Z^{(t)}$ with $a_{*}^{\mathrm{Pre}}(t) = R$ if $t \in [m, \omega]$.
  - The death benefit **before the payout of annuities** $LD_1(t) = p_{**}(x, t-1)p_{*\dagger}(t-1)a_{*\dagger}^{\mathrm{Post}}(t)Z^{(t)}$ with $a_{*\dagger}^{\mathrm{Post}}(t) = (t-x)\Pi$ if $t \in [x+1, m]$.
  - The death benefit **after the withdrawal** has started $LD_2(t) = p_{**}(x, t-1)p_{*\dagger}(t-1)a_{*\dagger}^{\mathrm{Post}}(t)Z^{(t)}$ with $a_{*\dagger}^{\mathrm{Post}}(t) = \max(0, (m-x)\Pi - (\mathrm{t} - \mathrm{m} - 1)\mathrm{R})$ if $t \in [m, \omega]$.
- We can calculate the premium according to the equivalence principle to be 25646.58.

Splitting the premium into the 3 benefits

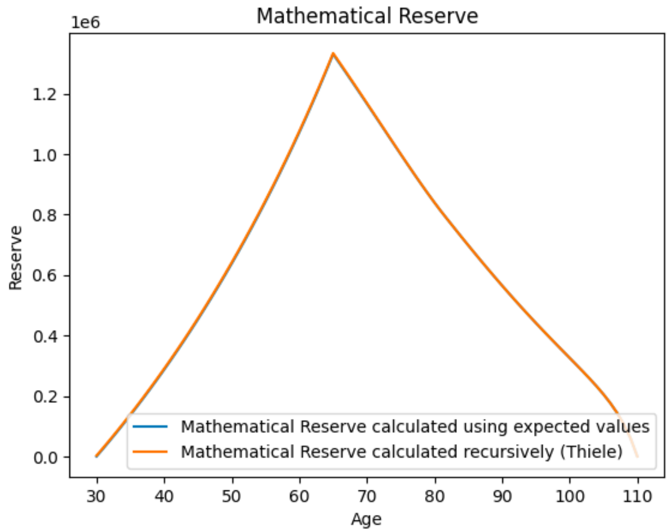
- We can split the premium into the 3 benefits by again using the equality principle and we arrive at
  - The premium for the **annuity** benefit is 23695.72.
  - The premium for the **death benefit before** retirement is 955.98.
  - The premium for the **death benefit after** retirement is 994.89.



Mathematical Reserve

- The **mathematical reserve** is the present value of an insurance policy (expected value of future liabilities). We can calculate the mathematical reserve at $\mathrm{k} = t_r$ with

$$_kV = \sum_{j=0}^{\infty} C_{k+j+1} v^{j+1} {}_jp_{x+k} q_{x+k+j} - \Pi_{j+k} v^j {}_jp_{x+k}$$



Risk and savings premium by type

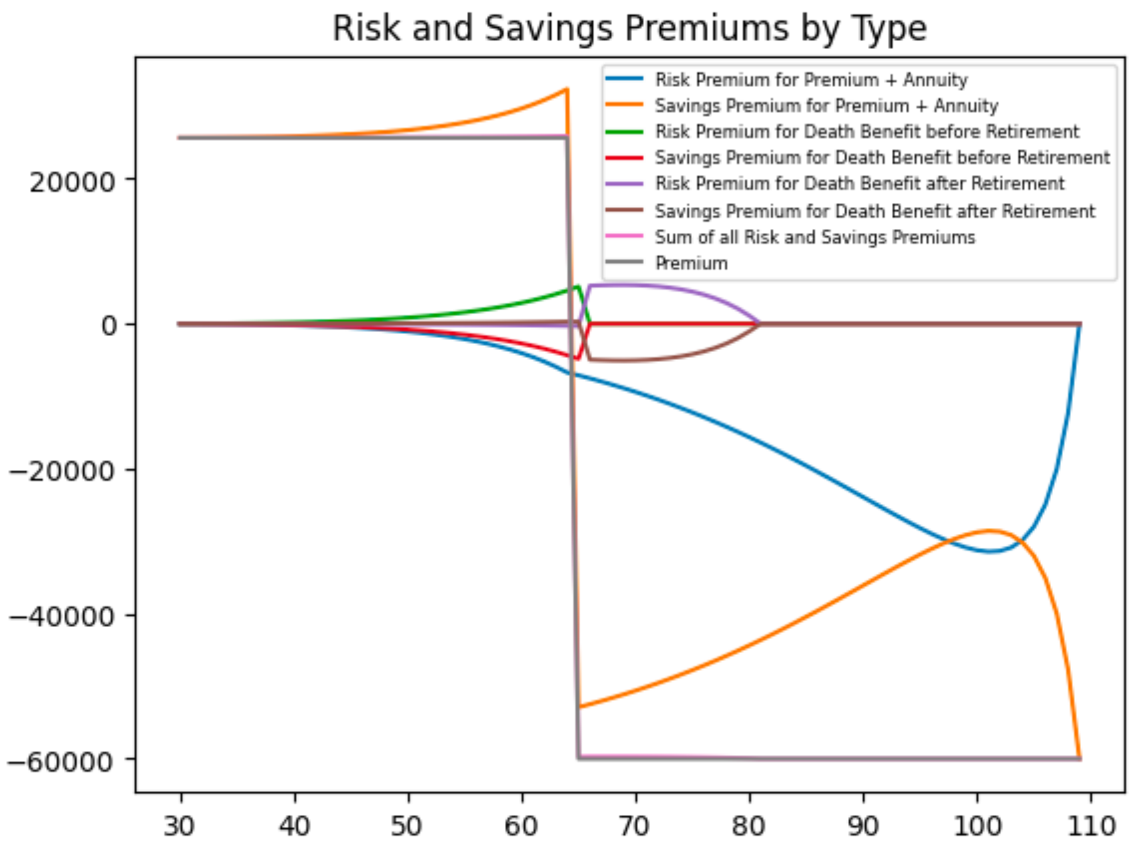
- We have defined the risk premium and the savings premium for task 1. Since we are looking at life insurance, we can define the normal subsequent state as the state where the policyholder is still alive in the next time step.
  - The **savings premium** equals the amount of money that has to be added to the mathematical reserve to still have adequate funds in the next time step if the policyholder is still alive:
  $$\Pi_k^s = {}_{k+1}V v - {}_{k}V$$
  - The **risk premium** corresponds to the loss incurred if the policyholder dies:
  $$\Pi_k^r = (C_{k+1} - {}_{k+1}V) v q_{x+k}$$
- In our example, we define the premium in the next step Π_k as $-P$ if $k < m$ and annuity R afterwards and the C_k is given by the sum of the two death benefits for time step t.

Risk and savings premium by type

Age	$P + LA$ Risk	$P + LA$ Savings	LD_1 Risk	LD_1 Savings	LD_2 Risk	LD_2 Savings	Total Risk	Total Savings	Premium
30	9	25638	-9	9	-10	10	-10	25657	25647
40	-209	25856	184	-173	-25	25	-50	25708	25658
50	-1065	26711	845	-796	-68	68	-287	25983	25695
65	-7112	-52888	5134	-4835	-313	283	-2291	-57439	-59730
70	-9396	-50604	0	0	5334	-5046	-4062	-55650	-59712
75	-12231	-47769	0	0	4427	-4203	-7804	-51972	-59776
85	-19587	-40413	0	0	0	0	-19587	-40413	-60000

- The full code and results can be found on Github.

Risk and savings premium by type



Sum of risk and savings premium

- We have the risk premium and the savings premium defined as on the slide before and as derived in the lecture, we have:

$$ {}_kV + \Pi_k = v(C_{k+1} q_{x+k} + p_{x+k\,k+1}V) $$

- We hence find $\Pi_k^s + \Pi_k^r = {}_{k+1}V v - {}_kV + (C_{k+1} - {}_{k+1}V) v q_{x+k} = v C_{k+1} q_{x+k} + v(1 - q_{x+k})_{k+1}V - {}_kV$ and with $p_{x+k} = 1 - q_{x+k}$, we show that

$$ \Pi_k^s + \Pi_k^r = \Pi_k $$

- Hence, the sum of risk and savings premium equates to the **full premium** at time step k.

Mortality table with 10% margin

- So far, we have been using a **best-estimate** second order mortality $qx(x,t) = e^{a_0 + a_1 x + a_2 x^2 + a_3 t}$.
- For pricing, one typically uses a **first order** mortality table in practice. In such a table, the probability of dying is generally lowered for annuities and increased for life insurance.
- Since we are considering a **deferred annuity** in the current example, we hence have to lower the mortality estimates by a 10% margin which results in an updated mortality function:

$$qx(x,t) = 0.9 * e^{a_0 + a_1 x + a_2 x^2 + a_3 t}$$

- This spread **decreases the risk of default** and covers possible demographic trends.
- If we re-do our calculations from the previous pages with this updated mortality, we find that the policyholder is required to pay a **higher annual premium** for the same annuity payments.

Constructing a first-order mortality table

- To construct a first-order mortality table, one typically relies on **observed mortality data** for a **fixed population of people** in the age range of interest.
- One uses **samples** owned either by the given insurance company or obtained jointly by several insurers.
- The population is **closed** if departures only occur as a result of death of the subject, else it is called **open** (possible additions and departures of subjects).
- To calculate the **mortality rate**, one counts the number of persons at risk at a specific point in time and the number of persons that died over a given period of time. We find that for open populations, we have $\hat{q}_x = d_x / l_x$ and for closed populations, we find $\hat{q}_x = d_x / (l_x + \frac{Z_x - A_x}{2})$.
- The **smoothed mortality** is then obtained by a non-parametric or parametric smoothing algorithm (bias-variance trade-off).
- In general, one also adds a **safety margin** to the calculated numbers to incorporate life expectancy trends or other possible future developments (such as pandemics) that have to be accounted for.

Sources & Code Repository

- Sources:
  - Life Insurance Mathematics lecture at ETH Zurich in Fall 2023 by Prof. Michael Koller (material can be found here)
  - Koller, Michael. (2011). *Life Insurance Risk Management Essentials.* EAA Series. Springer-Verlag Berlin Heidelberg.
  - Koller, Michael. (2012). *Stochastic Models in Life Insurance*. EAA Series. Springer-Verlag Berlin Heidelberg.
- Code repository:
  - Calculations can be found here.