Computational Models with Applications in Econometrics and Actuarial Science

Costel Bălcău

2017

Contents

Bibliography

[1] C. Bălcău, P. Radovici, R. Georgescu, *Matematică aplicată în economie*, Ed. Univ. Piteşti, 2010.

[2] B.H. Baltagi, *Econometrics*, Springer, 2011.

[3] N. Boboc, *Analiză matematică. Partea I*, Tipografia Universităţii din Bucureşti, Bucureşti, 1988.

[4] D.N. Gujarati, *Basic Econometrics*, McGraw-Hill, 2004.

[5] R. Kaas, M.J. Goovaerts, J. Dhaene, M. Denuit, *Modern Actuarial Risk Theory*, Kluwer, 2001.

[6] C. Kleiber, S. Kotz, *Statistical Size Distributions in Economics and Actuarial Sciences*, Wiley, New Jersey, 2003.

[7] P.M. Lee, *Bayesian statistics. An introduction*, Hodder Arnold, London, 2004.

[8] D. Lovelock, M. Mendel, A.L. Wright, *An Introduction to the Mathematics of Money. Saving and Investing*, Springer, New York, 2007.

[9] Y.D. Lyuu, *Financial engineering and computation. Principles, mathematics, algorithms*, Cambridge Univ. Press, 2004.

[10] E. Marceau, *Modelisation et evaluation des risques en Actuariat*, Springer Verlag France, 2013.

[11] I. Mircea, *Matematici financiare şi actuariale*, Ed. Corint, Bucureşti, 2006.

[12] C. Necşulescu, *Econometrie*, Ed. Universitară, Bucureşti, 2009.

[13] I. Negoiţă, *Aplicaţii practice în asigurări şi reasigurări*, Ed. Etape, Bucureşti, 2001.

[14] C. Niculescu, *Probabilităţi şi statistică*, Ed. Universităţii din Bucureşti, 2015.

[15] V. Preda, C. Bălcău, *Entropy optimization with applications*, Ed. Academiei Române, Bucureşti, 2010.

[16] I. Purcaru, *Matematici financiare: Teorie şi practică în operaţiuni bancare. Tranzacţii bursiere. Asigurări*, Ed. Economică, Bucureşti, 1998.

[17] I. Purcaru, I. Mircea, Gh. Lazăr, *Asigurări de persoane şi de bunuri : Aplicaţii. Cazuri. Soluţii*, Ed. Economică, Bucureşti, 1998.

[18] B. Sundt, R. Vernic, *Recursions for Convolutions and Compound Distributions with Insurance Applications*, Springer, 2009.

[19] A. Ullah, D. Giles, *Handbook of applied economic statistics*, Marcel Dekker, New York, 1998.

[20] R. Vernic, *Matematici actuariale*, Ed. Adco, Constanţa, 2004.

[21] J. Wooldridge, *Introductory Econometrics. A Modern Approach*, South-Western Cengage Learning, 2009.

[22] Gh. Zbăganu, *Metode matematice în teoria riscului şi actuariat*, Ed. Univ. Bucureşti, 2004.

[23] Gh. Zbăganu, *Elemente de teoria ruinei*, Geometry Balkan Press, Bucureşti, 2007.

Theme 1

Preparation from Probability Theory

We group here the principal notions and results from probability theory that were used in this course.

1.1 The axiomatic definition of probability

Definition 1.1.1. *Let Ω be any set. We denote by $\mathcal{P}(\Omega)$ the* ***set of all subsets*** *of Ω, i.e. $\mathcal{P}(\Omega) = \{A \;/\; A \subseteq \Omega\}$.*

Definition 1.1.2. *A* ***topology*** *on the set Ω is a family $\mathcal{T}$ of subsets of Ω s.t.*

$$\begin{cases} \varnothing, \Omega \in \mathcal{T}, \\ A, B \in \mathcal{T} \Rightarrow A \cap B \in \mathcal{T}, \\ (A_i)_{i \in I} \subseteq \mathcal{T} \Rightarrow \bigcup_{i \in I} A_i \in \mathcal{T}, \end{cases}$$

for each non-empty index set I.

A ***topological space*** *is a pair $(\Omega, \mathcal{T})$, where Ω is a set and $\mathcal{T}$ is a topology on Ω. Each set $A \in \mathcal{T}$ is called an* ***open set*** *of the topological space $(\Omega, \mathcal{T})$.*

Definition 1.1.3. *A* ***Borel field*** *(**σ-field**, **σ-algebra**) on the set Ω is a non-empty family $\mathcal{B}$ of subsets of Ω s.t.*

$$\begin{cases} A \in \mathcal{B} \Rightarrow \Omega \setminus A \in \mathcal{B}, \\ (A_i)_{i \in \mathbb{N}^*} \subseteq \mathcal{B} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{B}. \end{cases}$$

A ***measurable space*** *is a pair $(\Omega, \mathcal{B})$, where Ω is a set and $\mathcal{B}$ is a Borel field on Ω. Each set $A \in \mathcal{B}$ is called a* ***Borel set*** *(**measurable set**) of the measurable space $(\Omega, \mathcal{B})$.*

Proposition 1.1.1. *If $\mathcal{M}$ is a family of subsets of Ω, then*

$$\mathcal{B}(\mathcal{M}) = \bigcap\{\mathcal{B} \;/\; \mathcal{B} \text{ is a Borel field on } \Omega,\ \mathcal{B} \supseteq \mathcal{M}\}$$

is a Borel field on Ω.

Definition 1.1.4. *In the context of the above proposition, $\mathcal{B}(\mathcal{M})$ is called the **Borel field generated** by $\mathcal{M}$.*

For any $d \in \mathbb{N}^*$, we denote by $\mathcal{B}^d$ the Borel field generated by the intervals of $\mathbb{R}^d$.

Definition 1.1.5. *Let $(\Omega_1, \mathcal{B}_1)$ and $(\Omega_2, \mathcal{B}_2)$ be two measurable spaces. A function $f : \Omega_1 \to \Omega_2$ is called **measurable** (with respect to the Borel fields $\mathcal{B}_1$ and $\mathcal{B}_2$) if $f^{-1}(\mathcal{B}_2) \subseteq \mathcal{B}_1$.*

Definition 1.1.6. *Let Ω be a set and I be a non-empty index set. For every $i \in I$, let $(\Omega_i, \mathcal{B}_i)$ be a measurable space and $f_i : \Omega \to \Omega_i$ be a function. We denote by $\mathcal{B}(f_i \;/\; i \in I)$ the smallest Borel field on Ω with respect to which all the functions f_i, $i \in I$ are measurable, i.e. $\mathcal{B}(f_i \;/\; i \in I) = \mathcal{B}\left(\bigcup_{i\in I} f_i^{-1}(\mathcal{B}_i)\right)$.*

*The **product Borel field** of the Borel fields $\mathcal{B}_i$, $i \in I$ is*

$$\bigotimes_{i\in I} \mathcal{B}_i = \mathcal{B}(pr_i \;/\; i \in I),$$

*where, for every $j \in I$, $pr_j : \prod_{i\in I} \Omega_i \to \Omega_j$, $pr_j((\omega_i)_{i\in I}) = \omega_j$, $\forall (\omega_i)_{i\in I} \in \prod_{i\in I} \Omega_i$ is the **projection function** on j-th component.*

*The measurable space $(\prod_{i\in I} \Omega_i, \bigotimes_{i\in I} \mathcal{B}_i)$ is called the **product of the measurable spaces** $(\Omega_i, \mathcal{B}_i)$, $i \in I$.*

Proposition 1.1.2. *For every $d \in \mathbb{N}^*$ we have $\mathcal{B}^d = \bigotimes_{i=1}^{d} \mathcal{B}^1$.*

Definition 1.1.7. *A **measure** on the measurable space $(\Omega, \mathcal{B})$ is a function $\mu : \mathcal{B} \to [0, \infty]$ s.t.*

$$\begin{cases} \mu(\varnothing) = 0, \\ (A_i)_{i\in\mathbb{N}^*} \subseteq \mathcal{B} \text{ mutually disjoint } (A_i \cap A_j = \varnothing,\ \forall i \neq j) \Rightarrow \mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i). \end{cases}$$

*A measure μ is called a **finite measure** if $\mu(\Omega) < \infty$.*

*A measure μ is called a **σ-finite measure** if there exists a sequence $(A_i)_{i\in\mathbb{N}^*} \subseteq \mathcal{B}$ s.t. $\mu(A_i) < \infty$ $\forall i \in \mathbb{N}^*$, $A_i \subseteq A_{i+1}$ $\forall i \in \mathbb{N}^*$ and $\bigcup\limits_{i=1}^{\infty} A_i = \Omega$.*

*A **measure space** is a triple $(\Omega, \mathcal{B}, \mu)$, where $(\Omega, \mathcal{B})$ is a measurable space and μ is a measure on this space.*

Definition 1.1.8. *Let $(\Omega, \mathcal{B}, \mu)$ be a measure space and let $\mathcal{R}$ be a property on Ω, i.e. $\mathcal{R} : \Omega \to \{0, 1\}$,*

$$\mathcal{R}(\omega) = \begin{cases} 1, & \text{if } \omega \text{ satisfies } \mathcal{R} \\ 0, & \text{otherwise} \end{cases}, \ \forall \omega \in \Omega.$$

*We say that the property $\mathcal{R}$ holds **μ-almost everywhere (μ-a.e.)** if*

$$\mu(\mathcal{R}^{-1}(\{0\})) = 0, \ \text{i.e. } \mu(\{\omega \in \Omega \ / \ \omega \text{ does not satisfy } \mathcal{R}\}) = 0.$$

Definition 1.1.9. *Let μ and ν be two measures on a measurable space $(\Omega, \mathcal{B})$. We say that μ is **absolutely continuous with respect to** ν, and we write $\mu \prec \nu$, if $\mu(A) = 0$ for every $A \in \mathcal{B}$ such that $\nu(A) = 0$.*

Definition 1.1.10. *A **probability (probability measure)** on the measurable space $(\Omega, \mathcal{B})$ is a measure μ on this space with the property that $\mu(\Omega) = 1$.*

*A **probability space** is a triple $(\Omega, \mathcal{B}, \mu)$, where $(\Omega, \mathcal{B})$ is a measurable space and μ is a probability on this space. The elements of $\mathcal{B}$ are called the **events** of the probability space. For every $A \in \mathcal{B}$, $\mu(A)$ is called the **probability of the event** A. For every $\omega \in \Omega$ such that $\{\omega\} \in \mathcal{B}$, the event $\{\omega\}$ is called an **elementary event**.*

Proposition 1.1.3. *Let $(\Omega, \mathcal{B}, \mu)$ be a probability space. We have:*
a) $\mu(\varnothing) = 0$;
b) $\mu(A) \in [0, 1]$, $\forall A \in \mathcal{B}$, i.e. $\mu : \mathcal{B} \to [0, 1]$;
c) If $A, B \in \mathcal{B}$ and $A \subseteq B$, then $\mu(A) \leq \mu(B)$ and $\mu(B \setminus A) = \mu(B) - \mu(A)$;
d) If $A_1, \ldots, A_n \in \mathcal{B}$ are mutually disjoint (i.e. $A_i \cap A_j = \varnothing$, $\forall i \neq j$), then $\mu(\bigcup\limits_{i=1}^{n} A_i) = \sum\limits_{i=1}^{n} \mu(A_i)$;
*e) (**Inclusion-exclusion formula**) If $A_1, \ldots, A_n \in \mathcal{B}$, then*

$$\mu(\bigcup_{i=1}^{n} A_i) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \ldots < i_k \leq n} \mu(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k});$$

f) If $(A_n)_{n\in\mathbb{N}^*} \subseteq \mathcal{B}$ s.t. $A_n \subseteq A_{n+1}$ $\forall n \in \mathbf{N}^*$, then $\mu(\bigcup\limits_{n=1}^{\infty} A_n) = \lim\limits_{n\to\infty} \mu(A_n)$;

g) If $(A_n)_{n\in\mathbb{N}^*} \subseteq \mathcal{B}$ *s.t.* $A_n \supseteq A_{n+1}\ \forall n \in \mathbf{N}^*$, *then* $\mu(\bigcap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mu(A_n)$;

h) If $(A_n)_{n\in\mathbb{N}^*} \subseteq \mathcal{B}$, *then* $\mu(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mu(A_n)$.

Proposition 1.1.4. *Let* Ω *be a countable set and let* μ *be a probability on the measurable space* $(\Omega, \mathcal{P}(\Omega))$. *Then* $\mu(A) = \sum_{\omega\in A} \mu(\{\omega\})\ \forall A \in \mathcal{P}(\Omega)$, *and* $\mu \to (\mu(\{\omega\}))_{\omega\in\Omega}$ *is a bijective correspondence between the set of all the probabilities on* $(\Omega, \mathcal{P}(\Omega))$ *and the set* $\Big\{(p_\omega)_{\omega\in\Omega}\ /\ p_\omega \geq 0\ \forall \omega \in \Omega, \sum_{\omega\in\Omega} p_\omega = 1\Big\}$.

Definition 1.1.11. *In the context of the above proposition, we say that* $(p_\omega)_{\omega\in\Omega}$ *defined by* $p_\omega = \mu(\{\omega\})\ \forall \omega \in \Omega$ *is the **discrete** (or **countable**) **probability distribution** of the discrete (or countable) probability* μ.

Remark 1.1.1. *In the setting of the above definition,* $(p_\omega)_{\omega\in\Omega}$ *is a vector if* Ω *is finite and a sequence if* Ω *is infinite.*

Remark 1.1.2. *If* $(\Omega, \mathcal{B}, P)$ *is a probability space,* $X : \Omega \to \Omega_1$ *is a function and* $x \in \Omega_1$ *s.t.* $\{\omega \in \Omega\ /\ X(\omega) = x\} \in \mathcal{B}$, *then we denote*

$$P(X = x) = P(\{\omega \in \Omega\ /\ X(\omega) = x\}).$$

Similarly one uses the notation $P(X < x)$, $P(X > x)$, $P(X \leq x)$, $P(X \geq x)$, $P(X \neq x)$, $P(X \in A)$, *where* $A \subseteq \Omega_1$.

Also, if $Y : \Omega \to \Omega_1$ *is another function s.t.* $\{\omega \in \Omega\ /\ X(\omega) = Y(\omega)\} \in \mathcal{B}$, *then we denote*

$$P(X = Y) = P(\{\omega \in \Omega\ /\ X(\omega) = Y(\omega)\}).$$

Similarly one uses the notation $P(X < Y)$, $P(X > Y)$, $P(X \leq Y)$, $P(X \geq Y)$, $P(X \neq Y)$.

Also, if $Z : \Omega \to \Omega_2$ *is another function and* $z \in \Omega_2$ *s.t.* $\{\omega \in \Omega\ /\ Z(\omega) = z\} \in \mathcal{B}$, *then we denote*

$$P(X = x, Z = z) = P(\{\omega \in \Omega\ /\ X(\omega) = x \text{ and } Z(\omega) = z\}).$$

Similarly one uses the notation $P(X < x, Z < z)$, $P(X \in A, Y \in B)$, $P(X = x, Y = y, Z = z)$, *etc.*

Definition 1.1.12. *Let* $(\Omega, \mathcal{B}, P)$ *be a probability space and let* $(A_i)_{i\in I} \subseteq \mathcal{B}$ *be a family of events, where* I *is a non-empty index set. The events* A_i, $i \in I$ *are called* ***independent*** *if*

$$P(\bigcap_{i\in J} A_i) = \prod_{i\in J} P(A_i)$$

for every finite non-empty subset $J \subseteq I$ *of indices.*

Proposition 1.1.5. *Let* $(\Omega, \mathcal{B}, P)$ *be a probability space and let* $A \in \mathcal{B}$ *be an event such that* $P(A) > 0$. *Then the function*

$$P_A : \mathcal{B} \to [0, 1],\ P_A(B) = \frac{P(B \cap A)}{P(A)},\ \forall B \in \mathcal{B}$$

is a probability on the measurable space $(\Omega, \mathcal{B})$.

Definition 1.1.13. *In the context of the above proposition,* P_A *is called the* ***conditional probability*** *induced by the event* A. *For every event* $B \in \mathcal{B}$, $P_A(B)$ *is called the conditional probability of the event* B *given the event* A. *We denote* $P(B/A) = P_A(B)$.

Proposition 1.1.1 (Total probability formula). *Let* $(\Omega, \mathcal{B}, P)$ *be a probability space and let* $A_1, \ldots, A_n \in \mathcal{B}$ *be events such that*

$$\Omega = A_1 \cup \cdots \cup A_n,\ A_i \cap A_j = \varnothing\ \forall i \neq j$$

and $P(A_i) > 0\ \forall i \in \{1, \ldots, n\}$. *Then, for every event* $B \in \mathcal{B}$ *we have*

$$P(B) = \sum_{i=1}^{n} P(A_i) P_{A_i}(B).$$

Proposition 1.1.2 (Bayes's formula). *Let* $(\Omega, \mathcal{B}, P)$ *be a probability space and let* $A_1, \ldots, A_n \in \mathcal{B}$ *be events such that*

$$\Omega = A_1 \cup \cdots \cup A_n,\ A_i \cap A_j = \varnothing\ \forall i \neq j$$

and $P(A_i) > 0\ \forall i \in \{1, \ldots, n\}$. *Then, for every event* $B \in \mathcal{B}$ *such that* $P(B) > 0$, *we have*

$$P_B(A_i) = \frac{P(A_i) P_{A_i}(B)}{P(B)} = \frac{P(A_i) P_{A_i}(B)}{\sum\limits_{k=1}^{n} P(A_k) P_{A_k}(B)},\ \forall i \in \{1, \ldots, n\}.$$

Definition 1.1.14. *Let $d \in \mathbb{N}^*$. A* ***distribution (probability distribution)*** *on $\mathbb{R}^d$ is a probability on the measurable space $(\mathbb{R}^d, \mathcal{B}^d)$.*

Let μ be a distribution on $\mathbb{R}^d$. The distribution μ is called ***discrete (countable)*** *if there exists a countable set $A \subset \mathbb{R}^d$ such that $\mu(\mathbb{R}^d \setminus A) = 0$. The distribution μ is called* ***continuous*** *if $\mu(\{\mathrm{x}\}) = 0$ for every $\mathrm{x} \in \mathbb{R}^d$.*

Remark 1.1.3. *A distribution μ on $\mathbb{R}^d$ is discrete if and only if it has the form $\mu = \sum_{\mathrm{x} \in A} p_{\mathrm{x}} \varepsilon_{\mathrm{x}}$, where $A \subset \mathbb{R}^d$ is a countable set, $p_{\mathrm{x}} = \mu(\{\mathrm{x}\})$ $\forall \mathrm{x} \in \mathbb{R}^d$ and ε_{x} is the* ***Dirac measure****, defined by*

$$\varepsilon_{\mathrm{x}}(B) = \begin{cases} 1, & \text{if } \mathrm{x} \in B \\ 0, & \text{if } \mathrm{x} \notin B \end{cases}, \ \forall B \in \mathcal{B}^d.$$

Definition 1.1.15. *Let μ be a distribution on $\mathbb{R}^d$. The* ***distribution function (probability distribution function, cumulative probability distribution function) of*** *μ is the function $F_\mu : \mathbb{R}^d \to [0,1]$ defined by*

$$F_\mu(\mathrm{x}) = \mu((-\infty, \mathrm{x}]), \ \forall \mathrm{x} \in \mathbb{R}^d,$$

where $(-\infty, \mathrm{x}] = (-\infty, x_1] \times \ldots \times (-\infty, x_d]$ for every $\mathrm{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$.

Definition 1.1.16. *Let $d \in \mathbb{N}^*$. For every function $F : \mathbb{R}^d \to \mathbb{R}$ and every vectors $\mathrm{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$, $\mathrm{a} = (a_1, \ldots, a_d) \in \mathbb{R}^d$ and $\mathrm{b} = (b_1, \ldots, b_d) \in \mathbb{R}^d$ we define*

$$\Delta^{(d)}(F; \mathrm{a}; \mathrm{b}) = \sum_{i_1, \ldots, i_d \in \{0,1\}} (-1)^{i_1 + \ldots + i_d + d} F(a_1 + i_1(b_1 - a_1), \ldots, a_d + i_d(b_d - a_d)).$$

Proposition 1.1.6. *Let μ be a distribution on $\mathbb{R}^d$. Then its distribution function F_μ verifies the following properties:*
1) $\Delta^{(d)}(F_\mu; \mathrm{a}; \mathrm{b}) \geq 0$, $\forall \mathrm{a}, \mathrm{b} \in \mathbb{R}^d$ s.t. $\mathrm{a} \leq \mathrm{b}$ (i.e. $a_i \leq b_i$ $\forall i \in \{1, \ldots, d\}$);
2) F_μ is right continuous, i.e. $\lim_{\mathrm{x} \searrow \mathrm{a}} F_\mu(\mathrm{x}) = F_\mu(\mathrm{a})$, $\forall \mathrm{a} \in \mathbb{R}^d$;
3) $\lim_{\mathrm{x} \to \infty} F_\mu(\mathrm{x}) = 1$; $\lim_{x_i \to -\infty} F_\mu(\mathrm{x}) = 0$ for an $i \in \{1, \ldots, d\}$.

Definition 1.1.17. *A function $F : \mathbb{R}^d \to [0,1]$ which verifies all the properties from the above proposition is called a* ***distribution function*** *on $\mathbb{R}^d$. A function $F : \mathbb{R}^d \to \mathbb{R}$ which verifies only the properties 1 and 2 from the above proposition is called a* ***generalized distribution function (Lebesgue-Stieltjes measure function)*** *on $\mathbb{R}^d$.*

Proposition 1.1.7. *The correspondence $\mu \to F_\mu$ is a bijection between the set of all the distributions on $\mathbb{R}^d$ and the set of all the distribution functions on $\mathbb{R}^d$.*

Proposition 1.1.8. *Let $F : \mathbb{R}^d \to \mathbb{R}$ be a generalized distribution function. Then there exists a unique measure μ_F on the measurable space $(\mathbb{R}^d, \mathcal{B}^d)$ with the property that*

$$\mu_F((\mathrm{a}, \mathrm{b}]) = \Delta^{(d)}(F; \mathrm{a}; \mathrm{b}),\ \forall \mathrm{a}, \mathrm{b} \in \mathbb{R}^d \ s.t.\ \mathrm{a} \le \mathrm{b},$$

where $(\mathrm{a}, \mathrm{b}] = (a_1, b_1] \times \ldots \times (a_d, b_d]\ \forall \mathrm{a} = (a_1, \ldots, a_d), \mathrm{b} = (b_1, \ldots, b_d) \in \mathbb{R}^d$.

Definition 1.1.18. *In the context of the above proposition, the measure μ_F is called the* ***Lebesgue-Stieltjes measure generated*** *by F.*

We denote by χ_d the Lebesgue-Stieltjes measure on the space $(\mathbb{R}^d, \mathcal{B}^d)$ generated by the generalized distribution function

$$F(\mathrm{x}) = x_1 \ldots x_d,\ \forall \mathrm{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d.$$

χ_d is called the ***Lebesgue measure*** *on $\mathbb{R}^d$. For $d = 1$ we denote by m_L the Lebesgue measure on $\mathbb{R}$, i.e. $m_L = \chi_1$.*

Proposition 1.1.9. *For every* $\mathrm{a} = (a_1, \ldots, a_d), \mathrm{b} = (b_1, \ldots, b_d) \in \mathbb{R}^d$ *s.t.* $\mathrm{a} \le \mathrm{b}$ *we have* $\chi_d((\mathrm{a}, \mathrm{b}]) = (b_1 - a_1) \ldots (b_d - a_d)$.

In particular, $m_L((a, b]) = b - a,\ \forall a, b \in \mathbb{R}$ s.t. $a \le b$.

Definition 1.1.19. *Let $(\Omega, \mathcal{B})$ be a measurable space and let $f : \Omega \to \mathbb{R}$ be a measurable function (with respect to the Borel fields $\mathcal{B}$ and $\mathcal{B}^1$). We say that the function f is* ***simple*** *if the set $f(\Omega)$ is finite. We denote*

$$\mathcal{S}(\Omega, \mathcal{B}) = \{f : \Omega \to \mathbb{R} \ / \ f \textit{ is measurable (with respect to } \mathcal{B} \textit{ and } \mathcal{B}^1\textit{) and simple}\},$$
$$\bar{\mathcal{S}}_+(\Omega, \mathcal{B}) = \{f : \Omega \to [0, \infty] \ / \ f \textit{ is measurable (with respect to } \mathcal{B} \textit{ and } \bar{\mathcal{B}}^1\textit{)}\},$$

where $\bar{\mathcal{B}}^1$ is the Borel field generated by the open subsets of $\bar{\mathbb{R}}$.

Definition 1.1.20. *Let $A \subseteq \Omega$. The* ***characteristic function*** *of the subset A (with respect to the set Ω) is the function*

$$\mathbf{1}_A : \Omega \to \{0, 1\},\ \mathbf{1}_A(x) = \begin{cases} 1, & \textit{if } \mathrm{x} \in A \\ 0, & \textit{if } \mathrm{x} \notin A \end{cases},\ \forall x \in \Omega.$$

Proposition 1.1.10. *Let $(\Omega, \mathcal{B})$ be a measurable space.*
a) If $f \in \mathcal{S}(\Omega, \mathcal{B})$, then $f = \sum\limits_{a \in f(\Omega)} a \cdot \mathbf{1}_{f^{-1}(\{a\})}$.
b) $f \in \mathcal{S}(\Omega, \mathcal{B})$ if and only if there exist $n \in \mathbb{N}^*$, $a_1, \ldots, a_n \in \mathbb{R}$ and $A_1, \ldots, A_n \in \mathcal{B}$ mutually disjoint s.t. $f = \sum\limits_{i=1}^{n} a_i \mathbf{1}_{A_i}$.
c) If $f \in \bar{\mathcal{S}}_+(\Omega, \mathcal{B})$, then there exists a sequence $(f_n)_{n \in \mathbb{N}^*} \subseteq \mathcal{S}(\Omega, \mathcal{B})$ s.t. $0 \le f_n \le f_{n+1}\ \forall n \in \mathbb{N}^*$ and $\lim\limits_{n \to \infty} f_n = f$.

Definition 1.1.21. *Let $(\Omega, \mathcal{B}, \mu)$ be a measure space and let $f : \Omega \to \bar{\mathbb{R}}$ be a measurable function (with respect to the Borel fields $\mathcal{B}$ and $\bar{\mathcal{B}}^1$).*
*a) If $f \in \mathcal{S}(\Omega, \mathcal{B})$, then the **Lebesgue integral** of the function f with respect to the measure μ is defined by*

$$\int f d\mu = \sum_{a \in f(\Omega)} a\mu(f^{-1}(\{a\}))$$

(with the convention $0 \cdot \infty = 0$, $\infty - \infty = \infty$).
*b) If $f \in \bar{\mathcal{S}}_+(\Omega, \mathcal{B})$ s.t. $f = \lim\limits_{n\to\infty} f_n$, where $(f_n)_{n\in\mathbb{N}^*} \subseteq \mathcal{S}(\Omega, \mathcal{B})$, $0 \leq f_n \leq f_{n+1}$ $\forall n \in \mathbb{N}^*$, then the **Lebesgue integral** of the function f with respect to the measure μ is defined by*

$$\int f d\mu = \lim_{n\to\infty} \int f_n d\mu \ (\in [0, \infty]).$$

*c) f is called **Lebesgue integrable** with respect to the measure μ (μ**-Lebesgue integrable**) if $\displaystyle\int |f| d\mu < \infty$. In this case the **Lebesgue integral** of the function f with respect to the measure μ is defined by*

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu \ (\in \mathbb{R}),$$

where $f^+ = \max\{f, 0\}$ and $f^+ = \max\{-f, 0\}$.
*d) If $A \in \mathcal{B}$ and the function $\mathbf{1}_A \cdot f$ is μ-Lebesgue integrable, then the **Lebesgue integral** of the function f with respect to the measure μ over the set A is defined by*

$$\int_A f d\mu = \int \mathbf{1}_A \cdot f d\mu \ (\in \mathbb{R}).$$

Remark 1.1.4. *Sometimes, in order to avoid any possible confusion, we might choose to emphasize the argument of the function that we are integrating and we write*

$$\int f d\mu = \int f(x) d\mu(x), \ \int_A f d\mu = \int_A f(x) d\mu(x).$$

Proposition 1.1.11. *(**Correctness of Definition 1.1.21.b**) Let $(\Omega, \mathcal{B}, \mu)$ be a measure space. If $f \in \bar{\mathcal{S}}_+(\Omega, \mathcal{B})$ s.t. $f = \lim\limits_{n\to\infty} f_n = \lim\limits_{n\to\infty} g_n$, where $(f_n)_{n\in\mathbb{N}^*}, (g_n)_{n\in\mathbb{N}^*} \subseteq \mathcal{S}(\Omega, \mathcal{B})$, $0 \leq f_n \leq f_{n+1}$, $0 \leq g_n \leq g_{n+1}$ $\forall n \in \mathbb{N}^*$, then*

$$\lim_{n\to\infty} \int f_n d\mu = \lim_{n\to\infty} \int g_n d\mu.$$

Proposition 1.1.12. ***(Properties of Lebesgue integral)*** *Let* $(\Omega, \mathcal{B}, \mu)$ *be a measure space,* $f, f_n : \Omega \to \bar{\mathbb{R}}$ *be* μ*-Lebesgue integrable functions, for every* $n \in \mathbb{N}^*$*, and let* $g : \Omega \to \bar{\mathbb{R}}$ *be a measurable function (with respect to the Borel fields* $\mathcal{B}$ *and* $\bar{\mathcal{B}}^1$*).*
a) ***(Linearity)*** *For every* $\alpha_1, \alpha_2 \in \mathbb{R}$ *the function* $\alpha_1 f_1 + \alpha_2 f_2$ *is* μ*-Lebesgue integrable and* $\displaystyle\int (\alpha_1 f_1 + \alpha_2 f_2) d\mu = \alpha_1 \int f_1 d\mu + \alpha_2 \int f_2 d\mu$.
b) ***(Monotonicity)*** *If* $f_1 \leq f_2$*, then* $\displaystyle\int f_1 d\mu \leq \int f_2 d\mu$.
If $f_1 \leq f_2$ *and* $\mu(\{x \in \Omega \ / \ f_1(x) < f_2(x)\}) > 0$*, then* $\displaystyle\int f_1 d\mu < \int f_2 d\mu$.
c) $\displaystyle\left| \int f d\mu \right| \leq \int |f| d\mu$.
d) f *is finite* μ*-a.e., i.e.* $\mu(\{x \in \Omega \ / \ f(x) = \pm\infty\}) = 0$.
e) If $g = f$ μ*-a.e., then* g *is* μ*-Lebesgue integrable and* $\displaystyle\int g d\mu = \int f d\mu$.
f) If $|g| \leq f$ μ*-a.e., then* g *is* μ*-Lebesgue integrable and* $\displaystyle\left| \int g d\mu \right| \leq \int f d\mu$.

Theorem 1.1.1. ***(Lebesgue's dominated convergence Theorem)*** *Let* $(\Omega, \mathcal{B}, \mu)$ *be a measure space and let the functions* $f, f_n, g : \Omega \to \bar{\mathbb{R}}$*, for every* $n \in \mathbb{N}^*$*. If the functions* f_n*,* $n \geq 1$ *are measurable (with respect to the Borel fields* $\mathcal{B}$ *and* $\bar{\mathcal{B}}^1$*),* $\lim\limits_{n\to\infty} f_n = f$ μ*-a.e.,* $|f_n| \leq g$ μ*-a.e. for every* $n \in \mathbb{N}^*$ *and the function* g *is* μ*-Lebesgue integrable, then* f *is also* μ*-Lebesgue integrable and*

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu.$$

Definition 1.1.22. *Let* $U, V \subseteq \mathbb{R}^d$ *be two non-empty open sets. A function* $\varphi : U \to V$ *is called a* $\mathcal{C}^1$ ***diffeomorphism*** *if it is a bijection and all the components of* φ *and* φ^{-1} *have continuous first partial derivatives.*

Theorem 1.1.2. *i)* ***(Substitution formula)*** *Let* $(\Omega, \mathcal{B}, \mu)$ *be a measure space,* $(\Omega_1, \mathcal{B}_1, \mu)$ *be a measurable space,* $f : \Omega \to \Omega_1$ *be a measurable function (with respect to the Borel fields* $\mathcal{B}$ *and* $\mathcal{B}_1$*) and* $g : \Omega_1 \to \bar{\mathbb{R}}$ *be a measurable function (with respect to the Borel fields* $\mathcal{B}_1$ *and* $\bar{\mathcal{B}}^1$*). Then*

$$\int_\Omega g \circ f d\mu = \int_{\Omega_1} g d(\mu \circ f^{-1}),$$

that is, if either integral exists so does the other and they are equal.
ii) ***(Change of variable formula)*** *Let* $U, V \subseteq \mathbb{R}^d$ *be two non-empty open*

sets, $\varphi : U \to V$ be a $\mathcal{C}^1$ diffeomorphism and $f : V \to \mathbb{R}$ be a measurable function (with respect to the Borel fields $\mathcal{B}^d$ and $\mathcal{B}^1$). Then

$$\int_V f(\mathrm{x}) d\chi_d(\mathrm{x}) = \int_U (f \circ \varphi)(\mathrm{y}) |\mathcal{J}\varphi(\mathrm{y})| d\chi_d(\mathrm{y}),$$

where $\mathcal{J}\varphi(\mathrm{y}) = \det \left(\frac{\partial \varphi_i}{\partial y_j}(\mathrm{y}) \right)_{i,j \in \{1,\ldots,d\}}$ *($\mathcal{J}\varphi$ is called the* ***Jacobian*** *of φ).*

Theorem 1.1.3. *(**Jensen's Inequality**) Let $(\Omega, \mathcal{B}, P)$ be a probability space, $I \subseteq \mathbb{R}$ be an open interval, $f : \Omega \to I$ be a P-Lebesgue integrable function and $F : I \to \mathbb{R}$ be a convex function with the property that the function $F \circ f : \Omega \to \mathbb{R}$ is P-Lebesgue integrable. Then*

$$F\left(\int f dP \right) \leq \int (F \circ f) dP.$$

Moreover, if F is strictly convex then the equality holds if and only if $f = \int f dP$ P-a.e.

Theorem 1.1.4. *(**Radon-Nikodym**) Let μ be a measure and ν be a σ-finite measure on the measurable space $(\Omega, \mathcal{B})$. Then $\mu \prec \nu$ if and only if there exists a ν-Lebesgue integrable function $f : \Omega \to [0, \infty]$ such that*

$$\mu(A) = \int_A f d\nu, \ \forall A \in \mathcal{B}.$$

Moreover, the function f is unique ν-a.e.

Definition 1.1.23. *In the context of the above theorem, the function f is called the **Radon-Nikodym derivative** of μ with respect to ν and is written $f = \frac{d\mu}{d\nu}$.*

Theorem 1.1.5. *(**Fubini**) Let $(\Omega_1, \mathcal{B}_1, \mu_1)$ and $(\Omega_2, \mathcal{B}_2, \mu_2)$ be two measure spaces, where the measures μ_1 and μ_2 are σ-finite. Then there exists a unique measure μ on the measurable space $(\Omega_1 \times \Omega_2, \mathcal{B}_1 \otimes \mathcal{B}_2)$ s.t.*

$$\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2), \ \forall A_1 \in \mathcal{B}_1, \ \forall A_2 \in \mathcal{B}_2.$$

Moreover, for every $A \in \mathcal{B}_1 \otimes \mathcal{B}_2$ we have

$$\mu(A) = \int \mu_2(A_{x_1}) d\mu_1(x_1) = \int \mu_1(A_{x_2}) d\mu_2(x_2),$$

where $A_{x_1} = \{x_2 \in \Omega_2 \ / \ (x_1, x_2) \in A\}$ *and* $A_{x_2} = \{x_1 \in \Omega_1 \ / \ (x_1, x_2) \in A\}$. *Also, for every* μ*-Lebesgue integrable function* $f : \Omega_1 \times \Omega_2 \to \bar{\mathbb{R}}$ *we have*

$$\int f d\mu = \int \left[\int f(x_1, x_2) d\mu_1(x_1) \right] d\mu_2(x_2) = \int \left[\int f(x_1, x_2) d\mu_2(x_2) \right] d\mu_1(x_1)$$

(and all the integrals exist and they are finite).

Definition 1.1.24. *In the context of the above theorem, the measure* μ *is called the* ***product of the measures*** μ_1 *and* μ_2*, and is written* $\mu = \mu_1 \otimes \mu_2$.

The measure space $(\Omega_1 \times \Omega_2, \mathcal{B}_1 \otimes \mathcal{B}_2, \mu_1 \otimes \mu_2)$ *is called the* ***product of the measure spaces*** $(\Omega_1, \mathcal{B}_1, \mu_1)$ *and* $(\Omega_2, \mathcal{B}_2, \mu_2)$.

Theorem 1.1.6. *(**Comparison of the Lebesgue and the Riemann integrals**) a) Let* $\mathrm{a} = (a_1, \ldots, a_d), \mathrm{b} = (b_1, \ldots, b_d) \in \mathbb{R}^d$ *s.t.* $\mathrm{a} \leq \mathrm{b}$. *If the function* $f : [\mathrm{a}, \mathrm{b}] \to \mathbb{R}$ *is measurable (with respect to the Borel fields* $\mathcal{B}^d$ *and* $\mathcal{B}^1$*) and Riemann integrable, then* f *is also* χ_d*-Lebesgue integrable and*

$$\int_{[\mathrm{a},\mathrm{b}]} f(\mathrm{x}) d\chi_d(\mathrm{x}) = \int_{a_1}^{b_1} \cdots \int_{a_d}^{b_d} f(x_1, \ldots, x_d) dx_1 \ldots dx_d.$$

b) Let $f : \mathbb{R}^d \to [0, \infty)$ *be a measurable function (with respect to the Borel fields* $\mathcal{B}^d$ *and* $\mathcal{B}^1$*) such that* f *is Riemann integrable on every compact interval* $[\mathrm{a}, \mathrm{b}] \subset \mathbb{R}^d$. *Then* f *is* χ_d*-Lebesgue integrable if and only if* f *is (improperly) Riemann integrable and*

$$\int_{\mathbb{R}^d} f(\mathrm{x}) d\chi_d(\mathrm{x}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_d) dx_1 \ldots dx_d$$

(where the Riemann integrals from the right side are improper).

Proposition 1.1.13. *Let* $(\Omega_1, \mathcal{B}_1, P_1), \ldots, (\Omega_n, \mathcal{B}_n, P_n)$ *be probability spaces,* $n \in \mathbb{N}^*$. *Then there exists a unique probability* P *on the measurable space* $(\prod_{i=1}^{n} \Omega_i, \bigotimes_{i=1}^{n} \mathcal{B}_i)$ *s.t.*

$$P(\prod_{i=1}^{n} A_i) = \prod_{i=1}^{n} P_i(A_i), \ \forall A_i \in \mathcal{B}_i, \ \forall i \in \{1, \ldots, n\}.$$

Moreover, $P = (\ldots((P_1 \otimes P_2) \otimes P_3) \ldots) \otimes P_n$ *and the operation* $\otimes$ *is associative.*

Definition 1.1.25. *In the context of the above proposition, the probability P is called the* ***product of the probabilities*** *$P_1, \dots, P_n$ and is written $P = \bigotimes_{i=1}^{n} P_i$.*

The probability space $(\prod_{i=1}^{n} \Omega_i, \bigotimes_{i=1}^{n} \mathcal{B}_i, \bigotimes_{i=1}^{n} P_i)$ is called the ***product of the probability spaces*** *$(\Omega_1, \mathcal{B}_1, P_1), \dots, (\Omega_n, \mathcal{B}_n, P_n)$.*

Proposition 1.1.14. *Let I be an infinite index set and, for every $i \in I$, let $(\Omega_i, \mathcal{B}_i, P_i)$ be a probability space. Then there exists a unique probability P on the measurable space $(\prod_{i \in I} \Omega_i, \bigotimes_{i \in I} \mathcal{B}_i)$ s.t.*

$$P \circ pr_J^{-1} = \bigotimes_{j \in J} P_j$$

for every finite non-empty subset $J \subset I$ of indices, where $pr_J : \prod_{i \in I} \Omega_i \to \prod_{j \in J} \Omega_j$, $pr_J((\omega_i)_{i \in I}) = (\omega_j)_{j \in J}$, $\forall (\omega_i)_{i \in I} \in \prod_{i \in I} \Omega_i$.

Definition 1.1.26. *In the context of the above proposition, the probability P is called the* ***product of the probabilities*** *$(P_i)_{i \in I}$, and is written $P = \bigotimes_{i \in I} P_i$.*

The probability space $(\prod_{i \in I} \Omega_i, \bigotimes_{i \in I} \mathcal{B}_i, \bigotimes_{i \in I} P_i)$ is called the ***product of the probability spaces*** *$(\Omega_i, \mathcal{B}_i, P_i)$, $i \in I$.*

1.2 Random variables

Definition 1.2.1. *Let $(\Omega, \mathcal{B}, P)$ be a probability space and $(\Omega_1, \mathcal{B}_1)$ be a measurable space. A function $X : \Omega \to \Omega_1$ which is measurable (with respect to the Borel fields $\mathcal{B}$ and $\mathcal{B}_1$) is called a* ***random variable (r.v., random element)****.*

If $(\Omega_1, \mathcal{B}_1) = (\mathbb{R}^d, \mathcal{B}^d)$, then X is called a ***d-dimensional r.v. (random vector)****. In particular, if $d = 1$ then X is called a* ***real-valued r.v.***

Proposition 1.2.1. *Let $(\Omega, \mathcal{B}, P)$ be a probability space, $(\Omega_1, \mathcal{B}_1)$ be measurable space and $X : \Omega \to \Omega_1$ be a random variable. Then $\mu = P \circ X^{-1}$ is a probability on the space $(\Omega_1, \mathcal{B}_1)$.*

Definition 1.2.2. *In the context of the above proposition, the probability $\mu = P \circ X^{-1}$ is called the* ***distribution (probability distribution) of the r.v.*** *X (with respect to the probability P).*

The ***distribution function (probability distribution function, cumulative probability distribution function) of the r.v.*** *X (with respect to the probability P) is the distribution function of its distribution $P \circ X^{-1}$, i.e. the function*

$$F_X : \mathbb{R}^d \to [0,1],\ F_X(\mathrm{x}) = P(X < \mathrm{x}),\ \forall \mathrm{x} \in \mathbb{R}^d.$$

A d-dimensional r.v. X is called ***discrete*** *if the image $X(\Omega)$ is a countable set. A d-dimensional r.v. X is called* ***continuous*** *(with respect to the probability P) if its distribution function F_X is continuous.*

Remark 1.2.1. *Let X be a d-dimensional discrete r.v. Then its distribution function F_X (with respect to any probability P) is also discrete, i.e. the image $F_X(\mathbb{R}^d)$ is a countable set.*

Proposition 1.2.2. *Let $(\Omega, \mathcal{B}, P)$ be a probability space and $X : \Omega \to \mathbb{R}^d$ be a d-dimensional r.v.*
a) If X is discrete, then its distribution $\mu = P \circ X^{-1}$ is also discrete.
b) If X is continuous (with respect to P), then its distribution $\mu = P \circ X^{-1}$ is also continuous.

Definition 1.2.3. *A function $p : \mathbb{R}^d \to [0, \infty)$ that is measurable (with respect to the Borel fields $\mathcal{B}^d$ and $\bar{\mathcal{B}}^1$) is called a* ***probability density function (probability function, density function)*** *if is χ_d-Lebesgue integrable and $\displaystyle\int_{\mathbb{R}^d} p(\mathrm{x}) d\chi_d(\mathrm{x}) = 1$.*

Proposition 1.2.3. *Let μ be a distribution on $\mathbb{R}^d$ and X be a d-dimensional r.v. with the distribution μ (with respect to a probability P). If $\mu \prec \chi_d$, then the Radon-Nikodym derivative $\dfrac{d\mu}{d\chi_d}$ is a probability density function and*

$$\frac{d\mu}{d\chi_d} = F'_\mu \quad \chi_d\text{-a.e.},$$

where F_μ is the distribution function of μ (and of X, with respect to P), and F'_μ is its derivative.

Definition 1.2.4. *In the context of the above proposition, the function $p = \dfrac{d\mu}{d\chi_d}$ is called the* ***probability density function (probability function, density function) of the distribution*** *μ (and* ***of the r.v.*** *X, with respect to the probability P).*

Definition 1.2.5. *Let* $\mathrm{X}_1, \ldots, \mathrm{X}_n$ *be random variables defined on the probability space* $(\Omega, \mathcal{B}, P)$*, where* X_i *is a* d_i*-dimensional r.v., for every* $i \in \{1, \ldots, n\}$*. Let the* $d_! + \ldots + d_n$*-dimensional r.v.* $\mathrm{X} = (\mathrm{X}_1, \ldots, \mathrm{X}_n) : \Omega \to \mathbb{R}^{d_1} \times \ldots \times \mathbb{R}^{d_n}$.

The distribution of X *(with respect to* P*) is called the* ***joint distribution*** *of the r.v.* $\mathrm{X}_1, \ldots, \mathrm{X}_n$ *(with respect to* P*).*

For any $i \in \{1, \ldots, n\}$*, the distribution of* X_i *(with respect to* P*) is called a* ***marginal distribution*** *of the r.v.* X *(with respect to* P*).*

Definition 1.2.6. *Let* $(\Omega, \mathcal{B}, P)$ *be a probability space,* I *be a non-empty index set and* $(X_i)_{i \in I}$ *be a family of random variables defined on this space with values in a measurable space* $(\Omega_i, \mathcal{B}_i)$*, for every* $i \in I$*. The r.v.* X_i*,* $i \in I$ *are called* ***independent*** *(with respect to the probability* P*) if*

$$P(X_{i_1} \in A_{i_1}, \ldots, X_{i_k} \in A_{i_k}) = P(X_{i_1} \in A_{i_1}) \cdot \ldots \cdot P(X_{i_k} \in A_{i_k}),$$

for every finite non-empty subset $\{i_1, \ldots, i_k\} \subseteq I$ *of indices,* $i_1 < \ldots < i_k$*,* $k \in \mathbb{N}^*$*, and for every events* $A_{i_1} \in \mathcal{B}_{i_1}, \ldots, A_{i_k} \in \mathcal{B}_{i_k}$*.*

Proposition 1.2.4. *Let* $\mathrm{X}_1, \ldots, \mathrm{X}_n$ *be random variables defined on the probability space* $(\Omega, \mathcal{B}, P)$*, where* X_i *is a* d_i*-dimensional r.v., for every* $i \in \{1, \ldots, n\}$*,* $n \in \mathbb{N}^*$*. Let* $\mu_1, \ldots, \mu_n$ *be the distributions of* $\mathrm{X}_1, \ldots, \mathrm{X}_n$*, respectively (with respect to* P*), and let* $F_{\mathrm{X}_1}, \ldots, F_{\mathrm{X}_n}$ *be the distribution functions of* $\mathrm{X}_1, \ldots, \mathrm{X}_n$*, respectively (with respect to* P*).*
a) The following assertions are equivalent:

a1) The r.v. $\mathrm{X}_1, \ldots, \mathrm{X}_n$ *are independent (with respect to* P*);*

a2) For every events $A_1 \in \mathcal{B}^{d_1}, \ldots, A_n \in \mathcal{B}^{d_n}$ *we have*

$$P(\mathrm{X}_1 \in A_1, \ldots, \mathrm{X}_n \in A_n) = P(\mathrm{X}_1 \in A_1) \cdot \ldots \cdot P(\mathrm{X}_n \in A_n);$$

a3) The distribution μ *of* $\mathrm{X} = (\mathrm{X}_1, \ldots, \mathrm{X}_n)$ *(with respect to* P*) verifies the equality* $\quad \mu = \mu_1 \otimes \ldots \otimes \mu_n$*;*

a4) The distribution function F_{X} *of* $\mathrm{X} = (\mathrm{X}_1, \ldots, \mathrm{X}_n)$ *(with respect to* P*) verifies the equality*

$$F_{\mathrm{X}}(\mathrm{x}) = F_{\mathrm{X}_1}(\mathrm{x}_1) \cdot \ldots \cdot F_{\mathrm{X}_n}(\mathrm{x}_n), \ \forall \mathrm{x} = (\mathrm{x}_1, \ldots, \mathrm{x}_n) \in \mathbb{R}^{d_1} \times \ldots \times \mathbb{R}^{d_n}.$$

b) We assume moreover that the r.v. $\mathrm{X}_1, \ldots, \mathrm{X}_n$ *are discrete. Then* $\mathrm{X}_1, \ldots, \mathrm{X}_n$ *are independent (with respect to* P*) if and only if*

$$P(\mathrm{X}_1 = \mathrm{x}_1, \ldots, \mathrm{X}_n = \mathrm{x}_n) = P(\mathrm{X}_1 = \mathrm{x}_1) \cdot \ldots \cdot P(\mathrm{X}_n = \mathrm{x}_n),$$

for every $\mathrm{x}_1 \in \mathbb{R}^{d_1}, \ldots, \mathrm{x}_n \in \mathbb{R}^{d_n}$*.*
c) We assume moreover that the r.v. $\mathrm{X}_1, \ldots, \mathrm{X}_n$ *are continuous, with the*

probability density functions $p_1, \ldots, p_n$, *respectively (with respect to* P*). Then* $\mathrm{X}_1, \ldots, \mathrm{X}_n$ *are independent (with respect to* P*) if and only if the probability density function* p *of the r.v.* $\mathrm{X} = (\mathrm{X}_1, \ldots, \mathrm{X}_n)$ *(with respect to* P*) verifies the equality*

$$p(\mathrm{x}) = p_1(\mathrm{x}_1) \cdot \ldots \cdot p_n(\mathrm{x}_n), \ \forall \mathrm{x} = (\mathrm{x}_1, \ldots, \mathrm{x}_n) \in \mathbb{R}^{d_1} \times \ldots \times \mathbb{R}^{d_n}$$

$(\chi_{d_1+\ldots+d_n}$*-a.e.).*

Definition 1.2.7. *Let* μ *and* ν *be two distributions on* $\mathbb{R}$, *and let* X *and* Y *be two random variables with the distributions* μ *and* ν, *respectively (with respect to a probability* P*). Let* $r \in \mathbb{R},\ r > 0$.
a) The r***th absolute moment*** *of the distribution* μ *(and of the r.v.* X, *with respect to the probability* P*) is defined by*

$$\bar{E}_r(\mu) \equiv \bar{E}_r(X) = \int_{\mathbb{R}} |x|^r d\mu(x).$$

b) If $\bar{E}_r(\mu) < \infty$, *then the* r***th moment*** *of the distribution* μ *(and of the r.v.* X, *with respect to the probability* P*) is defined by*

$$E_r(\mu) \equiv E_r(X) = \int_{\mathbb{R}} x^r d\mu(x).$$

In particular, the ***mean (expected value****,* ***expectation*** *or* ***average****) of the distribution* μ *(and of the r.v.* X, *with respect to the probability* P*) is defined by*

$$E(\mu) \equiv E(X) = E_1(\mu) = \int_{\mathbb{R}} x d\mu(x).$$

c) If $\bar{E}_r(\mu) < \infty$, *then the* r***th central moment*** *of the distribution* μ *(and of the r.v.* X, *with respect to the probability* P*) is defined by*

$$E_r^c(\mu) \equiv E_r^c(X) = \int_{\mathbb{R}} [x - E(\mu)]^r d\mu(x).$$

In particular, the ***variance*** *of the distribution* μ *(and of the r.v.* X, *with respect to the probability* P*) is defined by*

$$\mathrm{var}\,(\mu) \equiv \mathrm{var}\,(X) = E_2^c(\mu) = \int_{\mathbb{R}} [x - E(\mu)]^2 d\mu(x)$$

d) If $\mathrm{var}\,(X) < \infty$ *and* $\mathrm{var}\,(Y) < \infty$, *then the* ***covariance*** *of the r.v.* X *and* Y *(with respect to the probability* P*) is defined by*

$$\mathrm{cov}\,(X, Y) = E\big([X - E(X)][Y - E(Y)]\big).$$

Proposition 1.2.5. ***(Properties of mean, variance and covariance for real-valued random variables)*** *In the context of the above definition, we have:*

$$\begin{gathered}\operatorname{var}(X) = E_2(X) - [E(X)]^2 = \operatorname{cov}(X, X);\\ \operatorname{cov}(Y, X) = \operatorname{cov}(X, Y);\ \operatorname{cov}(X, Y) = E(XY) - E(X)E(Y);\\ E(aX) = aE(X),\ \operatorname{var}(aX) = a^2 \operatorname{var}(X),\ \forall a \in \mathbb{R};\\ E(X+Y) = E(X) + E(Y);\ \operatorname{var}(X+Y) = \operatorname{var}(X) + \operatorname{var}(Y) + 2\operatorname{cov}(X, Y).\end{gathered}$$

If the r.v. X and Y are independent and their means are finite, then

$$E(XY) = E(X)E(Y),\ \operatorname{cov}(X, Y) = 0,\ \operatorname{var}(X+Y) = \operatorname{var}(X) + \operatorname{var}(Y).$$

If the r.v. X are constant, i.e. $X(\omega) = c\ \forall \omega \in \Omega$, where $c \in \mathbb{R}$, then $E(X) = c$ and $\operatorname{var}(X) = 0$.

Proposition 1.2.6. ***(Moments of discrete r.v.)*** *In the context of the above definition, if the r.v. X is discrete then we have:*

$$\begin{gathered}\bar{E}_r(X) \equiv \bar{E}_r(\mu) = \sum_{x \in A} |x|^r \mu(\{x\});\ E_r(X) \equiv E_r(\mu) = \sum_{x \in A} x^r \mu(\{x\});\\ E(X) \equiv E(\mu) = \sum_{x \in A} x\mu(\{x\});\ E_r^c(X) \equiv E_r^c(\mu) = \sum_{x \in A} [x - E(\mu)]^r \mu(\{x\});\\ \operatorname{var}(X) \equiv \operatorname{var}(\mu) = \sum_{x \in A} [x - E(\mu)]^2 \mu(\{x\}) = \sum_{x \in A} x^2 \mu(\{x\}) - \Big[\sum_{x \in A} x\mu(\{x\})\Big]^2,\end{gathered}$$

where $\mu(\{x\}) = P(X = x)$ and $A = \{x \in \mathbb{R} \ /\ \mu(\{x\}) > 0\}$.

Remark 1.2.2. *In the setting of the above proposition, we have $A \subseteq X(\Omega)$, and hence the set A is countable. Obviously, all the formulas for the above proposition remain valid if we replace the set A with the set $X(\Omega)$.*

Proposition 1.2.7. ***(Moments of continuous r.v.)*** *In the context of the above definition, if the r.v. X is continuous, with the probability density function p (with respect to the probability P) then we have:*

$$\begin{gathered}\bar{E}_r(X) \equiv \bar{E}_r(\mu) = \int_{\mathbb{R}} |x|^r p(x) dm_L(x);\\ E_r(X) \equiv E_r(\mu) = \int_{\mathbb{R}} x^r p(x) dm_L(x);\\ E(X) \equiv E(\mu) = \int_{\mathbb{R}} x p(x) dm_L(x);\end{gathered}$$

$$E_r^c(X) \equiv E_r^c(\mu) = \int_{\mathbb{R}} [x - E(\mu)]^r p(x) dm_L(x);$$
$$\mathrm{var}\,(X) \equiv \mathrm{var}\,(\mu) = \int_{\mathbb{R}} [x - E(\mu)]^2 p(x) dm_L(x)$$
$$= \int_{\mathbb{R}} x^2 p(x) dm_L(x) - \Big[\int_{\mathbb{R}} x p(x) dm_L(x) \Big]^2.$$

Remark 1.2.3. *In the context of the above proposition, if the probability density function p is continuous or, more generally, Riemann integrable on every compact interval, then from Theorem 1.1.6 we have:*

$$\bar{E}_r(X) = \int_{-\infty}^{\infty} |x|^r p(x) dx;\ E_r(X) = \int_{-\infty}^{\infty} x^r p(x) dx;$$
$$E(X) = \int_{-\infty}^{\infty} x p(x) dx;\ E_r^c(X) = \int_{-\infty}^{\infty} [x - E(\mu)]^r p(x) dx;$$
$$\mathrm{var}\,(X) = \int_{-\infty}^{\infty} [x - E(\mu)]^2 p(x) dx = \int_{-\infty}^{\infty} x^2 p(x) dx - \Big[\int_{-\infty}^{\infty} x p(x) dx \Big]^2.$$

Definition 1.2.8. *Let μ be a distribution on $\mathbb{R}^d$ and $\mathrm{X} = (X_1, \dots, X_d)$ be a d-dimensional random variable with the distribution μ (with respect to a probability P).*

*If $\displaystyle\int_{\mathbb{R}^d} |x_i| d\mu(\mathrm{x}) < \infty$, $\forall i \in \{1, \dots, d\}$, then the **mean (expected value, expectation** or **average**) of the distribution μ (and of the r.v. X, with respect to the probability P) is defined by*

$$E(\mu) \equiv E(\mathrm{X}) = (E_1(\mathrm{X}), \dots, E_d(\mathrm{X})),\ \textit{where}$$
$$E_i(\mathrm{X}) \equiv E_i(\mu) = \int_{\mathbb{R}^d} x_i d\mu(\mathrm{x}),\ \forall i \in \{1, \dots, d\}.$$

*If $\displaystyle\int_{\mathbb{R}^d} (x_1^2 + \ldots + x_d^2) d\mu(\mathrm{x}) < \infty$, then the **covariance matrix** of the distribution μ (and of the r.v. X, with respect to the probability P) is defined by*

$$\mathrm{cov}\,(\mu) \equiv \mathrm{cov}\,(\mathrm{X}) = \left(E_{ij}(\mathrm{X}) - E_i(\mathrm{X}) E_j(\mathrm{X})\right)_{i,j \in \{1,\dots,d\}},\ \textit{where}$$
$$E_{ij}(\mathrm{X}) \equiv E_{ij}(\mu) = \int_{\mathbb{R}^d} x_i x_j d\mu(\mathrm{x}),\ \forall i, j \in \{1, \dots, d\}.$$

Proposition 1.2.8. ***(Properties of mean and covariance for random vectors)*** *In the context of the above definition, we have:*

$$E(\mathrm{X}) = (E(X_1), \dots, E(X_d));\ \mathrm{cov}\,(\mathrm{X}) = (\mathrm{cov}\,(X_i, X_j))_{i,j \in \{1,\dots,d\}};$$

$$E(\mathrm{AX}^\top) = \mathrm{A}E(\mathrm{X})^\top,\ \mathrm{cov}\,(\mathrm{AX}^\top) = \mathrm{A}\,\mathrm{cov}\,(\mathrm{X})\,\mathrm{A}^\top,\ \forall \mathrm{A} \in \mathbb{R}^{d\times d}.$$

If Y *is another d-dimensional r.v. with the distribution* ν *(with respect to the same probability* P*), then*

$$E(\mathrm{X}+\mathrm{Y}) = E(\mathrm{X}) + E(\mathrm{Y}),$$

and if X *and* Y *are independent, then*

$$\mathrm{cov}\,(\mathrm{X}+\mathrm{Y}) = \mathrm{cov}\,(\mathrm{X}) + \mathrm{cov}\,(\mathrm{Y}).$$

Remark 1.2.4. *Similarly to Proposition 1.2.6, Proposition 1.2.7 and Remark 1.2.3, the formulas of mean and covariance matrix for random vectors can be rewritten in the particular cases of d-dimensional discrete r.v., d-dimensional continuous r.v. with probability density function, and d-dimensional continuous r.v. with a Riemann integrable (on every compact interval) probability density function. The formulas obtained in this way are expressed in terms of sums, Lebesgue integrals (with respect to the Lebesgue measure* χ_d*) and Riemann integrals, respectively.*

Proposition 1.2.9. *Let* $(\Omega, \mathcal{B}, P)$ *be a probability space,* $(\Omega_1, \mathcal{B}_1)$ *be a measurable space,* $X : \Omega \to \Omega_1$ *be a random variable and* $A \in \mathcal{B}$ *be an event such that* $P(A) > 0$*. Then the restriction of* X *to the subset* $A \subseteq \Omega$*, i.e. the function*

$$X/A : A \to \Omega_1,\ (X/A)(\omega) = X(\omega),\ \forall \omega \in A$$

is a random variable on the probability space $(A, \mathcal{B}_A, P_A)$*, where* $\mathcal{B}_A = \{B \cap A \ /\ B \in \mathcal{B}\}$ *and* P_A *is the conditional probability induced by the event* A*.*

Definition 1.2.9. *In the context of the above proposition,* X/A *is called a* ***conditional random variable*** *induced by the event* A*. Its distribution is called the* ***conditional distribution*** *of the r.v.* X *given the event* A*, and its mean* $E(X/A)$ *is called the* ***conditional mean (conditional expectation)*** *of the r.v.* X *given the event* A*.*

Definition 1.2.10. *Let* μ *be a probability on the measurable space* $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ *and let* X *be a real-valued r.v. with the distribution* μ *(with respect to a probability* P*). The* ***probability generating function*** *of* μ *(and of* X*, with respect to* P*) is the function* $G_\mu \equiv G_X$ *defined by*

$$G_\mu(t) \equiv G_X(t) = \sum_{i=0}^{\infty} \mu(\{i\}) t^i,\ \forall t \in [-1, 1],$$

where $\mu(\{i\}) = P(X = i)$*.*

Proposition 1.2.10. *Let $X_1, \ldots, X_n$ be r.v. taking values in $\mathbb{N}$. If $X_1, \ldots, X_n$ are independent, then*

$$G_{X_1+\ldots+X_n} = G_{X_1} \cdot \ldots \cdot G_{X_n}$$

(all the probability generating functions being defined with respect to the same probability P).

Definition 1.2.11. *Let μ be a distribution on $\mathbb{R}^d$ and X be a d-dimensional r.v. with the distribution μ (with respect to a probability P). The **characteristic function** of μ (and of X, with respect to P) is the function $\varphi_\mu \equiv \varphi_{\mathrm{X}}$ defined by*

$$\varphi_\mu(\mathrm{t}) \equiv \varphi_{\mathrm{X}}(\mathrm{t}) = \int_{\mathbb{R}^d} e^{i\langle \mathrm{t},\, \mathrm{x}\rangle} d\mu(\mathrm{x}),\ \forall \mathrm{t} \in \mathbb{R}^d \quad (i^2 = -1).$$

*The **moment generating function** of μ (and of X, with respect to P) is the function $\psi_\mu \equiv \psi_{\mathrm{X}}$ defined by*

$$\psi_\mu(\mathrm{t}) \equiv \psi_{\mathrm{X}}(\mathrm{t}) = \int_{\mathbb{R}^d} e^{\langle \mathrm{t},\, \mathrm{x}\rangle} d\mu(\mathrm{x}),\ \forall \mathrm{t} \in \mathbb{R}^d.$$

Remark 1.2.5. *In the above definition, $\langle \mathrm{t}, \mathrm{x}\rangle$ denotes the **inner product (scalar product)** of vectors $\mathrm{t} = (t_1, \ldots, t_d)$ and $\mathrm{x} = (x_1, \ldots, x_d)$, i.e. $\langle \mathrm{t}, \mathrm{x}\rangle = \sum\limits_{i=1}^{d} t_i x_i$. For $d = 1$ we have:*

$$\varphi_{\mathrm{X}}(t) \equiv \varphi_\mu(t) = \int_{\mathbb{R}} e^{itx} d\mu(x),\ \psi_{\mathrm{X}}(t) \equiv \psi_\mu(t) = \int_{\mathbb{R}} e^{tx} d\mu(x),\ t \in \mathbb{R},$$

Proposition 1.2.11. *a) In the context of the above definition, if the r.v. X is discrete, then*

$$\varphi_{\mathrm{X}}(\mathrm{t}) \equiv \varphi_\mu(\mathrm{t}) = \sum_{\mathrm{x} \in A} e^{i\langle \mathrm{t},\, \mathrm{x}\rangle} \mu(\{\mathrm{x}\}),\ \psi_{\mathrm{X}}(\mathrm{t}) \equiv \psi_\mu(\mathrm{t}) = \sum_{\mathrm{x} \in A} e^{\langle \mathrm{t},\, \mathrm{x}\rangle} \mu(\{\mathrm{x}\}),$$

where $\mu(\{\mathrm{x}\}) = P(\mathrm{X} = \mathrm{x})$ and $A = \{\mathrm{x} \in \mathbb{R}^d \ /\ \mu(\{\mathrm{x}\}) > 0\}$ (or $A = \mathrm{X}(\Omega)$).
If the r.v. X is continuous, with the probability density function p (with respect to the probability P), then

$$\varphi_{\mathrm{X}}(\mathrm{t}) \equiv \varphi_\mu(\mathrm{t}) = \int_{\mathbb{R}^d} e^{i\langle \mathrm{t},\, \mathrm{x}\rangle} p(\mathrm{x}) d\chi_d(\mathrm{x}),\ \psi_{\mathrm{X}}(\mathrm{t}) \equiv \psi_\mu(\mathrm{t}) = \int_{\mathbb{R}^d} e^{\langle \mathrm{t},\, \mathrm{x}\rangle} p(\mathrm{x}) d\chi_d(\mathrm{x}).$$

b) Let $X_1, \ldots, X_n$ *be d-dimensional r.v. If* $X_1, \ldots, X_n$ *are independent (with respect to a probability* P*), then*

$$\varphi_{X_1+\ldots+X_n} = \varphi_{X_1} \cdot \ldots \cdot \varphi_{X_n},\ \psi_{X_1+\ldots+X_n} = \psi_{X_1} \cdot \ldots \cdot \psi_{X_n}$$

(all the characteristic functions and the moment generating functions being defined with respect to the same probability P*).*

Proposition 1.2.12. *Let* X *be a real-valued r.v. with the distribution* μ *(with respect to a probability* P*) such that* $\bar{E}_n(X) < \infty$*, where* $n \in \mathbb{N}^*$*. Then*

$$E_r(X) = \frac{\partial^r \psi_X}{\partial t^r}(0), \forall r \le n,\ r \in \mathbb{N}^*.$$

Proposition 1.2.13. *Let* $\mu_1, \ldots, \mu_n$ *be distributions on* $\mathbb{R}^d$ *and consider the sum function*

$$s_n : \underbrace{\mathbb{R}^d \times \ldots \times \mathbb{R}^d}_{n} \to \mathbb{R}^d, s_n(x_1, \ldots, x_n) = x_1 + \ldots + x_n,\ \forall x_1, \ldots, x_n \in \mathbb{R}^d.$$

Then the function $\mu_1 * \ldots * \mu_n : \mathcal{B}^d \to [0,1]$ *defined by*

$$\mu_1 * \ldots * \mu_n = (\mu_1 \otimes \ldots \otimes \mu_n) \circ s_n^{-1}$$

is a distribution on $\mathbb{R}^d$*.*

Definition 1.2.12. *In the context of the above proposition, the distribution* $\mu_1 * \ldots * \mu_n$ *is called the* ***convolution*** *of the distributions* $\mu_1, \ldots, \mu_n$*.*

Proposition 1.2.14. *Let* $X_1, \ldots, X_n$ *be d-dimensional r.v. with the distributions* $\mu_1, \ldots, \mu_n$*, respectively (with respect to a probability* P*). If* $X_1, \ldots, X_n$ *are independent, then* $X = X_1 + \ldots + X_n$ *is a random variable with the distribution* $\mu = \mu_1 * \ldots * \mu_n$ *(with respect to the same probability* P*).*

Definition 1.2.13. *Let* $(\Omega, \mathcal{B}, P)$ *be a probability space, and let* $(\Omega_1, \mathcal{B}_1)$ *be a measurable space, where* Ω_1 *is a metric space and* $\mathcal{B}_1$ *is the Borel field generated by the open subsets of* Ω_1*.*

Let μ *and* $(\mu_n)_{n\in\mathbb{N}^*}$ *be finite measures on the measurable space* $(\Omega_1, \mathcal{B}_1)$*. We say that the sequence* $(\mu_n)_{n\in\mathbb{N}^*}$ ***converges weakly*** *to* μ*, and we write* $\mu_n \xrightarrow{w} \mu$ *(or* $\mu_n \Rightarrow \mu$*), if*

$$\lim_{n\to\infty} \int f d\mu_n = \int f d\mu$$

for every bounded, continuous function $f : \Omega_1 \to \mathbb{R}$*.*

Let X and $(X_n)_{n\in\mathbb{N}^*}$ be random variables defined on the probability space $(\Omega, \mathcal{B}, P)$ with values in the measurable space $(\Omega_1, \mathcal{B}_1)$. We say that the sequence $(X_n)_{n\in\mathbb{N}^*}$ ***converges in distribution*** *to X (with respect to the probability P), and we write $X_n \xrightarrow{d} X$, if*

$$P \circ X_n^{-1} \xrightarrow{w} P \circ X^{-1}.$$

Proposition 1.2.15. *In the context of the above definition, we have the following equivalences:*
a) $\mu_n \xrightarrow{w} \mu$ if and only if $\lim\limits_{n\to\infty} \displaystyle\int f d\mu_n = \int f d\mu$ for every bounded, uniformly continuous function $f : \Omega_1 \to \mathbb{R}$.
b) $\mu_n \xrightarrow{w} \mu$ if and only if $\lim\limits_{n\to\infty} \mu_n(A) = \mu(A)$ for every $A \in \mathcal{B}_1$ s.t. $\mu(\partial A) = 0$, where ∂A is the boundary of the set A.
c) $X_n \xrightarrow{d} X$ if and only if $\lim\limits_{n\to\infty} \displaystyle\int_\Omega f(X_n) dP = \int_\Omega f(X) dP$ for every bounded, uniformly continuous function $f : \Omega_1 \to \mathbb{R}$.

Theme 2

Statistical Indicators

2.1 Introduction to Economic Statistics

Economic Statistics is the science that deals with the collection, classification, analysis and interpretation of numerical facts or data from economics. It means that by the use of probability theory it imposes order and regularity on aggregate of disparate elements of the same population.

Statistical population (statistical collectivity): the total number of elements of the same properties representing the object of the investigation.

Statistical unit: the basic element of the statistical population, which will be observed within the statistical research, and will represent any individual elements of the population.

Statistical characteristic: a common property of all the population units.

Statistical variable: a statistical characteristic which can take different values from a unit to another unit (or from a group of units to another group).

Statistical indicator: a numerical expression of an economic category, obtained using a statistical calculus characterizing a variable.

Statistical sample: a part of statistical population, which will be investigated.

Descriptive Statistics: methods for representing and describing the statistical population (data summarizing, tabulation and presentation; analysis of data uniformity and consistency and symmetry interpretation; construction of indicators, index numbers, time series; correlation and regression,...).

Inferential Statistics: methods for predicting about the whole statistical population by studying the properties of a statistical sample (estimation of population parameters; construction of confidence intervals; testing statistical hypothesis).

The **main steps** of a statistical research:

1. data collection;
2. data analysis;
3. data conclusions and results interpreting.

The **detailed steps** of a statistical research:

1. Establishing the objective of the research.
2. Defining and identifying the population to be studied according to the objective.
3. Establishing the set of characteristics according to the information we need to obtain.
4. Analyzing the already existing data bases about the studied population, that is analyzing the secondary data sources.
5. For insufficient secondary data, organizing a total research or a partial research (by sampling).
6. Organizing the data collection, which means deciding where, when and how to collect the data for each unit, individually or collectively (using a common recording way, like as a list).
7. Data recording, using a data analysis program like as EXCEL, STATISTICA, MATLAB or R-system.
8. Data summarizing and presentation (by tables, series, graphs...).
9. Data analyzing using descriptive statistics and inferential statistics methods.
10. Data conclusions and results interpreting (by research reports).

2.2 Statistical frequency distribution (series)

Statistical frequency distribution (statistical frequency series) of a statistical variable: the correspondence from the values of the variable, also called **variants** or **units**, to the **frequencies** of these values.

Let $X_1, X_2, \ldots, X_n$ be the observed values of a statistical variable X.

- The number n of recorded values represents **the volume (the size) of distribution**.

The distribution of X can be represented by:

- **the simple distribution (series) with an ungrouped set of values (data)**
$$(X_1, X_2, \ldots, X_n).$$

- **the absolute frequency distribution (series)**
$$\begin{pmatrix} x_1 & x_2 & \ldots & x_k \\ n_1^* & n_2^* & \ldots & n_k^* \end{pmatrix},$$
where $x_1, \ldots, x_k$ represent the distinct values (variants, units) from the sequence $X_1, X_2, \ldots, X_n$, and $n_1^*, \ldots, n_k^*$ represent **the absolute frequencies** of these values, defined by
$$n_j^* = \operatorname{card}\{X_i \,| X_i = x_j,\ i \in \{1, \ldots, n\}\},$$
for any $j \in \{1, \ldots, k\}$.
Obviously, $n_j^* \in \mathbb{N},\ \forall j \in \{1, \ldots, k\}$ and $\displaystyle\sum_{j=1}^{k} n_j^* = n.$

- **the relative frequency distribution (series)**
$$\begin{pmatrix} x_1 & x_2 & \ldots & x_k \\ f_1^* & f_2^* & \ldots & f_k^* \end{pmatrix},$$
where $x_1, \ldots, x_k$ represent the distinct values (variants, units) from the sequence $X_1, X_2, \ldots, X_n$, and $f_1^*, \ldots, f_k^*$ represent **the relative frequencies** of these values, defined by
$$f_j^* = \frac{n_j^*}{n},\ \forall j \in \{1, \ldots, k\}.$$
Obviously, $f_j^* \in [0, 1],\ \forall j \in \{1, \ldots, k\}$ and $\displaystyle\sum_{j=1}^{k} f_j^* = 1.$

- **the cumulated frequencies**: can be obtained from absolute frequencies or from relative frequencies, and represent the number of units with the variable value lower or equal than the upper limit of the current class.

We distinguish two cases:

- **Simple frequency distribution** (single variation frequency distribution or univariated data): the statistical variable is one-dimensional.
- **Multidimensional frequency distribution**: the statistical variable is multidimensional.

2.3 A classification algorithm

A **class (interval) of variation** of the values (data) of a statistical distribution is defined between two boundaries: its lower and upper limit. The **class size** (the **interval size**) represents the difference between the upper limit and the lower limit.

Data grouping assumes solving the following main issues:

- the purpose of the classification is to obtain synthetic data;
- the grouped results should be homogeneous groups;
- their frequency distribution should be as close as possible to the normal distribution (*Gauss bell*).

Let the simple distribution (series) with an ungrouped set of values (data)

$$(X_1, X_2, \ldots, X_n).$$

The **classification algorithm** consists in the following steps:

1. Compute the **amplitude** of distribution:

$$\overline{W} = \max_{i \in \{1,\ldots,n\}} X_i - \min_{i \in \{1,\ldots,n\}} X_i.$$

2. Choose the number k of classes (intervals). For example, according to the **rule of H. Sturges**:

$$k = 1 + \lceil \log_2 n \rceil,$$

where $\lceil x \rceil$ denotes the *upper integer part* (*the ceiling*) of the real number x.

3. Compute the class sizes (the interval sizes).

   For example, for classes with the same size, the size is

$$d = \left\lceil \frac{\overline{W}}{k} \right\rceil .$$

4. Construct the classes, by starting with the minimum value and adding the class sizes step by step.

   For example, for intervals with the same size d, the intervals are

$$[l_0, l_1),\ [l_1, l_2),\ \ldots,\ [l_{k-1}, l_k],$$

   where

$$l_0 = \min_{i\in\{1,\ldots,n\}} X_i \ \text{ and } \ l_i = l_{i-1} + d,\ \forall i \in \{1, \ldots, k\}.$$

5. Compute the absolute frequencies $n_1^*, \ldots, n_k^*$ of the classes (intervals), defined by

$$n_j^* = \text{card}\left\{X_i \,|X_i \text{ belongs to the } j\text{-class },\ i \in \{1, \ldots, n\}\right\}$$

   for any $j \in \{1, \ldots, k\}$.

   Obviously, $n_j^* \in \mathbb{N},\ \forall j \in \{1, \ldots, k\}$ and $\sum_{j=1}^{k} n_j^* = n$.

   Compute also the relative frequencies $f_1^*, \ldots, f_k^*$ of the classes (intervals), defined by

$$f_j^* = \frac{n_j^*}{n},\ \forall j \in \{1, \ldots, k\}.$$

   Obviously, $f_j^* \in [0, 1],\ \forall j \in \{1, \ldots, k\}$ and $\sum_{j=1}^{k} f_j^* = 1$.

2.4 Classification of statistical indicators

Statistical indicators (statistical measures) are numerical expression of a statistical distribution, according to a certain characteristic.

Classification of statistical indicators:

- **Central tendency indicators**: describe in a synthetic manner the typical feature of a statistical distribution and summarize the essential information comprised into it.

  The main central tendency indicators:

  - **Average measures**: the arithmetic mean, the geometric mean, the quadratic mean, the harmonic mean, the absolute moments;
  - **Position measures**: the mode, the median, the quintiles.

  For a central tendency measure to be representative, the set of values of a given distribution need to be homogeneous. This property is evaluate by variation measures.

- **Variation indicators**: evaluate the variability of a statistical distribution from its central tendency measures.

  The main variation indicators:

  - **Simple measures of dispersion**: the amplitude (the absolute range), the relative range, the inter-quintile range, the individual deviation;
  - **Average deviation measures**: the mean absolute deviation, the variance, the standard deviation, the central moments, the covariance.
  - **Shape measures**: the Pearson's skewness coefficients, the Yule's skewness coefficient, the excess coefficient.

2.5 Average measures

2.5.1 The arithmetic mean

- For a **simple distribution** with an ungrouped set of values (data)

  $$(x_1, x_2, \ldots, x_n)$$

  the arithmetic mean is

  $$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i,$$

  where $x_1, \ldots, x_n$ are the values of the distribution, n being the volume (the size) of distribution.

- For a **absolute frequency distribution** or a **relative frequency distribution** obtained from a classification by **variants**

$$\begin{pmatrix} x_1 & x_2 & \dots & x_k \\ n_1^* & n_2^* & \dots & n_k^* \end{pmatrix} \text{ or } \begin{pmatrix} x_1 & x_2 & \dots & x_k \\ f_1^* & f_2^* & \dots & f_k^* \end{pmatrix},$$

respectively, **the arithmetic mean** (or **the weighted arithmetic mean**) is

$$\overline{x} = \frac{\sum\limits_{i=1}^{k} n_i^* x_i}{\sum\limits_{i=1}^{k} n_i^*} = \sum_{i=1}^{k} f_i^* x_i,$$

where $x_1, \dots, x_k$ are the variants (the distinct values of the distribution), $n_1^*, \dots, n_k^*$ are the corresponding absolute frequencies and $f_1^*, \dots, f_k^*$ are the corresponding relative frequencies, k being the number of variants.

- For a **frequency distribution** obtained from a classification by **classes (intervals)**

$$[l_0, l_1),\ [l_1, l_2),\ \dots,\ [l_{k-1}, l_k]$$

the arithmetic mean (or **the weighted arithmetic mean**) is

$$\overline{x} = \frac{\sum\limits_{i=1}^{k} n_i^* x_i}{\sum\limits_{i=1}^{k} n_i^*} = \sum_{i=1}^{k} f_i^* x_i,$$

where $x_1, \dots, x_k$ are **the classes middles (the intervals middles)** given by

$$x_i = \frac{l_{i-1} + l_i}{2},\ \forall i \in \{1, \dots, k\},$$

$n_1^*, \dots, n_k^*$ are the corresponding absolute frequencies of the given classes and $f_1^*, \dots, f_k^*$ are the corresponding relative frequencies of the given classes, k being the number of classes (intervals).

2.5.2 The harmonic mean

We will use the same notations as above.

- For a **simple distribution** with an ungrouped set of values (data)

$$(x_1, x_2, \dots, x_n)$$

the harmonic mean is

$$\overline{x}_h = \frac{n}{\sum\limits_{i=1}^{n} \frac{1}{x_i}}.$$

- For a **frequency distribution** obtained from a classification by **variants** or by **classes (intervals)**, **the harmonic mean** (or **the weighted harmonic mean**) is

$$\overline{x}_h = \frac{\sum\limits_{i=1}^{k} n_i^*}{\sum\limits_{i=1}^{k} \frac{n_i^*}{x_i}} = \frac{1}{\sum\limits_{i=1}^{k} \frac{f_i^*}{x_i}}.$$

2.5.3 The geometric mean

We will use the same notations as above.

- For a **simple distribution** with an ungrouped set of values (data)

$$(x_1, x_2, \ldots, x_n)$$

the geometric mean is

$$\overline{x}_g = \sqrt[n]{\prod_{i=1}^{n} x_i}.$$

- For a **frequency distribution** obtained from a classification by **variants** or by **classes (intervals)**, **the geometric mean** (or **the weighted geometric mean**) is

$$\overline{x}_g = \sqrt[n]{\prod_{i=1}^{k} x_i^{n_i^*}} = \prod_{i=1}^{k} x_i^{f_i^*},$$

where $n = \sum\limits_{i=1}^{k} n_i^*$.

2.5.4 The quadratic mean

We will use the same notations as above.

- For a **simple distribution** with an ungrouped set of values (data)

$$(x_1, x_2, \ldots, x_n)$$

the quadratic mean (the square mean) is

$$\overline{x}_q = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}.$$

- For a **frequency distribution** obtained from a classification by **variants** or by **classes (intervals)**, **the quadratic mean** (**the square mean**, **the weighted quadratic mean** or **the weighted square mean**) is

$$\overline{x}_q = \sqrt{\frac{\sum\limits_{i=1}^{k} n_i^* x_i^2}{\sum\limits_{i=1}^{k} n_i^*}} = \sqrt{\sum_{i=1}^{k} f_i^* x_i^2}.$$

2.5.5 Absolute moments

We will use the same notations as above.

- For a **simple distribution** with an ungrouped set of values (data)

$$(x_1, x_2, \ldots, x_n)$$

the j-th absolute moment is

$$m_j = \frac{1}{n}\sum_{i=1}^{n} |x_i|^j,$$

and **the j-th moment** is

$$\overline{m}_j = \frac{1}{n}\sum_{i=1}^{n} x_i^j.$$

- For a **frequency distribution** obtained from a classification by **variants** or by **classes (intervals)**, **the j-th absolute moment** is

$$m_j = \frac{\sum\limits_{i=1}^{k} n_i^* |x_i|^j}{\sum\limits_{i=1}^{k} n_i^*} = \sum_{i=1}^{k} f_i^* |x_i|^j,$$

and **the j-th moment** is

$$\overline{m}_j = \frac{\sum\limits_{i=1}^{k} n_i^* x_i^j}{\sum\limits_{i=1}^{k} n_i^*} = \sum_{i=1}^{k} f_i^* x_i^j.$$

We remark that $\overline{m}_1 = \overline{x}$.

2.5.6 Properties of the means

1. **The means inequality**:

$$x_{min} \leq \overline{x}_h \leq \overline{x}_g \leq \overline{x} \leq \overline{x}_q \leq x_{max},$$

where x_{min} and x_{max} are the minimum value and the maximum value of the given distribution, respectively.

2. All the above means are affected by extreme values and cannot be used for heterogeneous data.

3. The arithmetic mean is a normal value meaning that the deviations sum of the individual values from the mean will be equal to zero, i.e. $\sum\limits_{i=1}^{n}(x_i - \overline{x}) = 0$.

4. Compared to the arithmetic mean, which is influenced by large values of the given distribution, the harmonic mean value is more influenced by the small values of the distribution. For example, the harmonic mean is used to compute the price index in order to measure the inflation.

5. The geometric mean is used, for example, to compute the average price index for a year: $\overline{I}^p = \sqrt[11]{\overline{I}^p_{feb/jan} \overline{I}^p_{mar/feb} \cdots \overline{I}^p_{dec/nov}}$.

6. The quadratic mean is more influenced by the large values of the variable. This mean is used to compute the standard deviation.

2.6 Position measures

2.6.1 The mode

The mode (the modal value, the dominant value) of a statistical distribution is the most frequent value of this distribution.

- For a **frequency distribution** obtained from a classification by **variants**, **the mode** M_o is the variant with the highest frequency.

- For a **frequency distribution** obtained from a classification by **classes (intervals)**

$$[l_0, l_1),\ [l_1, l_2),\ \ldots,\ [l_{k-1}, l_k]$$

  the mode is

$$M_o = \frac{l_{i-1} + l_i}{2} - \frac{d}{2} \cdot \frac{f^*_{i+1} - f^*_{i-1}}{f^*_{i-1} - 2f^*_i + f^*_{i+1}},$$

  where $[l_{i-1}, l_i)$ is the interval with the maximum frequency, called **the modal interval**, $d = l_i - l_{i-1}$ is the size of the modal interval, f^*_i is the relative frequency of the modal interval, and f^*_{i-1}, f^*_{i+1} are the relative frequencies for the previous interval and for the next interval, respectively.

2.6.2 The median

The median (the median value) of a statistical distribution is the value of distribution that splits the set of values in two equal subsets. Hence half of the population has the characteristic smaller that the median value, and the other half has the characteristic larger than the median value.

- For a **simple distribution** with an ungrouped set of values (data)

$$(x_1, x_2, \ldots, x_n),$$

  let

$$x'_1 \leq x'_2 \cdots \leq x'_n$$

  be the ordered sequence of its values. **The median** of this distribution is

$$M_e = \begin{cases} x'_{\frac{n+1}{2}} \text{ if } n \text{ is odd,} \\ \dfrac{x'_{\frac{n}{2}} + x'_{\frac{n}{2}+1}}{2}, \text{ if } n \text{ is even.} \end{cases}$$

- For a **frequency distribution** obtained from a classification by **variants**

$$\begin{pmatrix} x_1 & x_2 & \dots & x_k \\ n_1^* & n_2^* & \dots & n_k^* \end{pmatrix},$$

the median estimation procedure consists in the following steps:

1. Compute **the median location**:

$$M_{e_{loc}} = \begin{cases} \frac{n}{2} \text{ if } n \geq 100, \\ \frac{n+1}{2}, \text{ if } n < 100, \end{cases}$$

where $n = \sum_{i=1}^{k} n_i^*$ is the volume of distribution.

2. For any variant x_i, $i \in \{1, \dots, k\}$, on compute **the cumulated absolute frequency** $\sum_{j=1}^{i} n_j^*$.

3. Compute **the median** M_e as the variant corresponding to the minimum (or first) cumulated frequency which is gather or equal to the median location, that is,

$$M_e = x_{i^*}, \text{ where } i^* = \min \left\{ i \;\middle|\; i \in \{1, \dots, k\}, \sum_{j=1}^{i} n_j^* \geq M_{e_{loc}} \right\}.$$

- For a **frequency distribution** obtained from a classification by **intervals (classes)**

$$[l_0, l_1), [l_1, l_2), \dots, [l_{r-1}, l_r],$$

let $n_1^*, \dots, n_r^*$ be the corresponding absolute frequencies of these intervals. The median estimation procedure consists in the following steps:

1. Compute **the median location**:

$$M_{e_{loc}} = \begin{cases} \frac{n}{2} \text{ if } n \geq 100, \\ \frac{n+1}{2}, \text{ if } n < 100, \end{cases}$$

where $n = \sum_{i=1}^{k} n_i^*$ is the volume of distribution.

2. For any interval $[l_{i-1}, l_i)$, $i \in \{1, \dots, k\}$, on compute **the cumulated absolute frequency** $\sum_{j=1}^{i} n_j^*$.

3. Compute **the median interval** $[l_{i^*-1}, l_{i^*})$ as the interval corresponding to the minimum (or first) cumulated frequency which is gather or equal to the median location, that is,

$$i^* = \min\left\{ i \,\middle|\, i \in \{1, \ldots, k\},\ \sum_{j=1}^{i} n_j^* \geq M_{e_{loc}} \right\}.$$

4. Compute **the median** M_e as

$$M_e = l_{i^*-1} + d \cdot \frac{M_{e_{loc}} - \sum\limits_{j=1}^{i^*-1} n_j^*}{n_{i^*}^*},$$

where $d = l_{i^*} - l_{i^*-1}$ is the size of the median interval.

2.6.3 Quintiles

The quintiles (the fractiles) of a statistical distribution are the values of distribution that splits the set of values in k equal subsets. They are defined and computed in a similar manner as the median.

The main categories of fractiles:

- **Quartiles**: 3 measures Q_1, Q_2, Q_3 that split the set of values in 4 equal subsets: We remark that the second quartile equals the median: $Q_2 = M_e$;
- **Deciles**: 9 measures that split the set of values in 10 equal subsets;
- **Percentiles**: 99 measures that split the set of values in 100 equal subsets.

2.6.4 Properties of the position measures

1. The mode is a measure of the central tendency very used in sales analysis. Its main advantage is the possibility to be computed also for qualitative variables, and its main disadvantage is the possibility to have multi-modal distribution (distribution with more than one modal value).
2. The main advantage of the median is the fact that the extreme values do not affect it as strong as they are affecting the mean. Also the median is easy to compute and can be used also for ordinal qualitative data. The main disadvantage of the median is that it does not take into account all the observation.

3. For a symmetrical distribution, the mean, the median and the mode are identical. For a skewed distribution the mean, the median and the mode are located in different places.

2.7 Variation measures

The importance of the variation measures:

- provide additional information to analyze the reliability of the central tendency measure;
- characterize in depth the variation and the spread of the value set;
- compare two or many samples selected from the same population.

2.7.1 Simple measures of dispersion

We will use the same notations as above.

1. **The amplitude (the absolute range)**:

$$\overline{W}_x = x_{max} - x_{min},$$

where x_{min} and x_{max} are the minimum value and the maximum value of the given distribution, respectively.

2. **The relative range**:

   - as a coefficient: $\dfrac{\overline{W}_x}{\overline{x}}$;
   - in percentages: $\dfrac{\overline{W}_x}{\overline{x}} \cdot 100$.

3. **The inter-quintile range**:

$$Q = \frac{(M_e - Q_1) + (Q_3 - M_e)}{2} = \frac{Q_3 - Q_1}{2}.$$

It measures how far from the median we should go on either side before including 50% of the observations.

4. **The individual deviations**:

   - the absolute deviation: $d_i = x_i - \overline{x}$;

- the relative deviation: $d_i' = \dfrac{x_i - \overline{x}}{\overline{x}} \cdot 100$.

They provide information only for each recorded value and they are not expressing the overall variation.

2.7.2 Average deviation measures

We will use the same notations as above.

1. **The mean absolute deviation**:

   - For a **simple distribution** with an ungrouped set of values (data)

     $$(x_1, x_2, \dots, x_n)$$

     the mean absolute deviation is

     $$MAD = \frac{1}{n}\sum_{i=1}^{n} |x_i - \overline{x}|.$$

   - For a **frequency distribution** obtained from a classification by **variants** or by **classes (intervals)**, **the mean absolute deviation** is

     $$MAD = \frac{\sum\limits_{i=1}^{k} n_i^* |x_i - \overline{x}|}{\sum\limits_{i=1}^{k} n_i^*} = \sum_{i=1}^{k} f_i^* |x_i - \overline{x}|.$$

2. **The variance**:

   - For a **simple distribution** with an ungrouped set of values (data)

     $$(x_1, x_2, \dots, x_n)$$

     the variance is

     $$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2,$$

     and **the rectified variance** is

     $$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$

- For a **frequency distribution** obtained from a classification by **variants** or by **classes (intervals)**, **the variance** is

$$\sigma^2 = \frac{\sum\limits_{i=1}^{k} n_i^*(x_i - \overline{x})^2}{\sum\limits_{i=1}^{k} n_i^*} = \sum_{i=1}^{k} f_i^*(x_i - \overline{x})^2,$$

and **the rectified variance** is

$$s^2 = \frac{\sum\limits_{i=1}^{k} n_i^*(x_i - \overline{x})^2}{\sum\limits_{i=1}^{k} n_i^* - 1}.$$

For the rectified variance, the average variance computed for many samples extracted from the same population tends to the population variance.

The variance has no measurement unit, being an abstract measure. It is used to compute the standard deviation and other variation and correlation measures.

3. **The standard deviation**: $\sigma = \sqrt{\sigma^2}$.

   The rectified standard deviation: $s = \sqrt{s^2}$.

   Standard deviation allows to determine how the values of a frequency distribution are located in relation to the mean. For example, according to **the Chebyshev's Inequality**

$$P(|X - \overline{x}| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2},\ \forall \varepsilon > 0,$$

   we have:

   - at least 75% of the values will fall within ± 2 standard deviations from the mean of the distribution ($\varepsilon = 2\sigma$);
   - at least 88.89% of the values will fall within ± 3 standard deviations from the mean ($\varepsilon = 3\sigma$).

4. **The coefficient of variation**: $v = \dfrac{\sigma}{\overline{x}}$.

   The rectified coefficient of variation: $v' = \dfrac{s}{\overline{x}}$.

Some average measures of dispersion are expressed in concrete measurements units as the variable. When comparing two or many distribution we cannot use these measures due to possible different measurement units. This inconvenience is over passed using the relative dispersion measures. The coefficient of variation is the main relative dispersion measure.

It takes values between 0 and 1.

- If $0 \leq v \leq 0.17$ then the mean is strictly representative and we have a high level of homogeneity;
- If $0.17 < v \leq 0.35$ then the mean is moderately representative and we have a medium level of homogeneity;
- If $0.35 < v \leq 0.5$ then the mean has a low representativeness;
- If $v > 0.5$ then the mean is not representative for the data set and the population is heterogeneous.

5. **The central moments**:

- For a **simple distribution** with an ungrouped set of values (data)

$$(x_1, x_2, \ldots, x_n)$$

the j-th central moment is

$$m_j^c = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^j.$$

- For a **frequency distribution** obtained from a classification by **variants** or by **classes (intervals)**, **the j-th central moment** is

$$m_j^c = \frac{\sum\limits_{i=1}^{k} n_i^* (x_i - \overline{x})^j}{\sum\limits_{i=1}^{k} n_i^*} = \sum_{i=1}^{k} f_i^* (x_i - \overline{x})^j.$$

We remark that $m_2^c = \sigma^2$.

Remark 2.7.1. *For* ***the frequency distributions*** *obtained from a classification by* ***classes (intervals)****, on use the following* ***Sheepard's corrections*** *for firstly four moments and central moments:*

$$\overline{M}_1 = \overline{m}_1; \qquad M_1^c = m_1^c = 0;$$

$$\overline{M}_2 = \overline{m}_2 + \frac{d^2}{12}; \qquad M_2^c = m_2^c - \frac{d^2}{12};$$
$$\overline{M}_3 = \overline{m}_3 + \frac{d^2}{4} \cdot \overline{m}_1; \qquad M_3^c = m_3^c;$$
$$\overline{M}_4 = \overline{m}_4 + \frac{d^2}{2} \cdot \overline{m}_2 + \frac{d^4}{80}; \qquad M_4^c = m_4^c - \frac{d^2}{2} \cdot m_2^c + \frac{7d^4}{240}.$$

2.7.3 Shape measures

For a perfectly symmetric distribution the mean, the median and the mode are equals. This distribution corresponds to the Gauss Bell shape (the normal distribution). In this case the influence of the random factors is characterized by certain regularity, so the influences are distributed in both directions, compared to the arithmetic mean.

For analyzing the shape of an arbitrary distribution on needs to compare the mean the median and the mode. An arbitrary distributions can be symmetric, slightly skewed or highly skewed. For a skewed distribution the mean, the median and the mode are located in different places. More precisely:

- If the frequencies are concentrated around the small values we have
$$M_o < M_e < \overline{x}$$
and the symmetric distribution was modified by prolonging to $+\infty$ and it is becoming skewed to the right, called **positive skewness**.
- If the frequencies are concentrated around the large values we have
$$\overline{x} < M_e < M_o$$
and the symmetric distribution was modified by prolonging to $-\infty$ and it is becoming skewed to the left, called **negative skewness**.

The main methods for interpreting the frequency distributions shapes:

- The graphical method, by analyzing **the frequency polygon**.
- The analytic method, by computing the skewness coefficients: the Pearson's coefficients, the Yule's coefficients, the excess coefficient, the interquintile coefficient.

We will use the same notations as above.

1. **The Pearson's skewness coefficient based on the mean deviation from the mode**:
$$A_s = \frac{\overline{x} - M_o}{\sigma}.$$
It takes values between -1 and $+1$.
   - If A_s is close to zero, then the distribution is symmetric.
   - If A_s is close to -1, then the distribution is skewed to the left (negative skewness).
   - If A_s is close to $+1$, then the distribution is skewed to the right (positive skewness).

2. **The Pearson's skewness coefficient based on the mean deviation from the median**:
$$A'_s = \frac{3(\overline{x} - M_e)}{\sigma}.$$
It takes values between -3 and +3.
   - If A'_s is close to zero, then the distribution is symmetric.
   - If A'_s is close to -3, then the distribution is skewed to the left (negative skewness).
   - If A'_s is close to +3, then the distribution is skewed to the right (positive skewness).

   This coefficient is mainly used for **slightly skewed distributions** for which we can have the relation
$$\overline{x} - M_o = 3(\overline{x} - M_e), \text{ so } A'_s = A_s.$$

3. **The Yule's skewness coefficient**, based on the quartiles (**the interquintile asymmetry coefficient**):
$$A''_s = \frac{(Q_3 - M_e) - (M_e - Q_1)}{(Q_3 - M_e) + (M_e - Q_1)} = \frac{Q_1 + Q_3 - 2M_e}{Q_3 - Q_1}.$$
It takes values between -1 and 1 and is also close to zero for a symmetrical distribution.

4. **The excess coefficient**:
$$E_s = \frac{m_4^c}{\sigma^4} - 3.$$
If A''_s (A'_s or A_s) and E_s are close to zero, then the distribution is symmetric.

5. **The inter-quintile coefficient**: $q = \dfrac{Q}{M_e} = \dfrac{Q_3 - Q_1}{2M_e}$.

   It takes also values between -1 and 1 and is close to zero for a symmetrical distribution.

2.8 Problems

Exercise 2.8.1. The number of failures produced by an equipment and recorded for the last 24 hours are as follows: 12, 15, 29, 23, 17, 7, 10, 14, 14, 27, 22, 8, 5, 19, 6, 15, 20, 17, 16, 17, 23, 19, 9, 28.

a. Construct a frequency distribution and a relative frequency distribution for these data.

b. Construct a line chart for these data.

c. Group these data using the above classification algorithm.

d. Construct a frequency distribution and a relative frequency distribution for the obtained classes (grouped data).

e. Calculate the above statistical indicators in each of the following cases:

(i) simple distribution with an ungrouped set of values.

(ii) frequency distribution obtained from a classification by variants.

(iii) frequency distribution obtained from a classification by intervals.

Theme 3

Two-dimensional statistical distributions

3.1 Least Squares Method

This method is used to approximate a function when only a partial set of its values is known. Hence we will obtain the trend of the given function.

Let $f : A \subseteq \mathbb{R} \to \mathbb{R}$ be a function and let

$$f(x_i),\ i \in \{1, \dots, n\}$$

be the given values, where $x_1, x_2, \dots, x_n \in A$.

We will approximate the function f by a **trend function** $g : A \to \mathbb{R}$.

Usually, g is **a polynomial function**

$$g(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k,\ \text{where } k \leq n$$

(in particular, g can be **linear** $g(x) = a_0 + a_1 x$ or **quadratic** $g(x) = a_0 + a_1 x + a_2 x^2$), **a hyperbolic function**

$$g(x) = a_0 + \frac{a_1}{x},$$

an exponential function

$$g(x) = a_0 + a_1 \mathrm{e}^x,$$

or **a logarithmic function**

$$g(x) = a_0 + a_1 \ln x.$$

The Least Squares Method consists in the following steps:

1. Select the type of trend function g, according to the graph of the set of given points $(x_i, f(x_i)),\ i \in \{1, \dots, n\}$.
2. Calculate the parameters $a_0, a_1, \dots$ of trend function g by **minimizing the error sum of squares**:
$$\min_{a_0,a_1,\dots} \sum_{i=1}^{n} [f(x_i) - g(x_i)]^2$$
(by **the method of critical points**, for example).

In the case of **polynomial trend function**
$$g(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_k x^k,\ k \le n,$$
using the method of critical points it follows that the parameters $a_0, a_1, a_2, \dots, a_k$ are the solution of the following **equation system**
$$\frac{\partial F}{\partial a_j}(a_0, a_1, \dots, a_k) = 0,\ j \in \{0, \dots, k\},$$
where
$$\begin{aligned} F(a_0, a_1, \dots, a_k) &= \sum_{i=1}^{n} [f(x_i) - g(x_i)]^2 \\ &= \sum_{i=1}^{n} [f(x_i) - a_0 - a_1 x_i - a_2 x_i^2 - \dots - a_k x_i^k]^2 \end{aligned}$$
(since $(a_0, a_1, a_2, \dots, a_k)$ is the unique critical point of the function F).

This system can be derived as
$$-2 \sum_{i=1}^{n} x_i^j [f(x_i) - a_0 - a_1 x_i - a_2 x_i^2 - \dots - a_k x_i^k] = 0,\ j \in \{0, \dots, k\},$$
that is
$$\begin{cases} n a_0 + a_1 \sum_{i=1}^{n} x_i + a_2 \sum_{i=1}^{n} x_i^2 + \dots + a_k \sum_{i=1}^{n} x_i^k = \sum_{i=1}^{n} f(x_i) \\ a_0 \sum_{i=1}^{n} x_i + a_1 \sum_{i=1}^{n} x_i^2 + a_2 \sum_{i=1}^{n} x_i^3 + \dots + a_k \sum_{i=1}^{n} x_i^{k+1} = \sum_{i=1}^{n} x_i f(x_i) \\ \dots \\ a_0 \sum_{i=1}^{n} x_i^k + a_1 \sum_{i=1}^{n} x_i^{k+1} + a_2 \sum_{i=1}^{n} x_i^{k+2} + \dots + a_k \sum_{i=1}^{n} x_i^{2k} = \sum_{i=1}^{n} x_i^k f(x_i) \end{cases} \tag{3.1}$$

(**a linear system** of $k+1$ equations with $k+1$ variables, with a nonzero determinant).

In the particular case of **linear trend function**

$$g(x) = a_0 + a_1 x$$

($k = 1$), this system has the following form

$$\begin{cases} na_0 + a_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} f(x_i) \\ \\ a_0 \sum_{i=1}^{n} x_i + a_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i f(x_i). \end{cases} \tag{3.2}$$

Remark 3.1.1. *If we select two or more trend functions of different types, the best approximation is given by the minimum error sum of squares* $\sum_{i=1}^{n} [f(x_i) - g(x_i)]^2$.

Example 3.1.1. The sales of a company for the last five months are as follows:

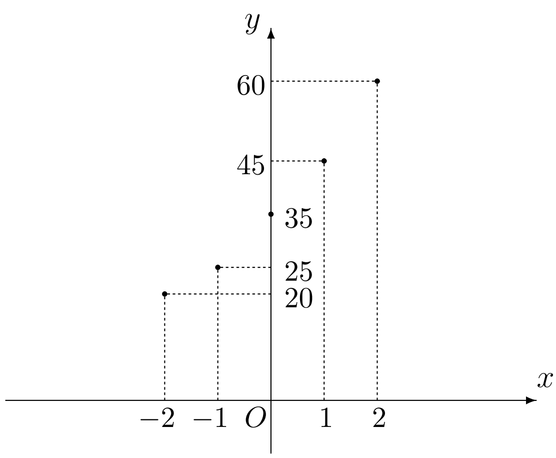
Month	Jan	Feb	March	April	May
Sales	20	25	35	45	60

Determine a trend function of sales and a forecast for June.

Solution. We know the values

$$f(-2) = 20,\ f(-1) = 25,\ f(0) = 35,\ f(1) = 45,\ f(2) = 60.$$

The graph of given points $(x_i, f(x_i))$ is



therefore we can use a linear or a quadratic trend function.

Case 1. For a linear trend function $g(x) = a_0 + a_1 x$, the parameters a_0 and a_1 are the solution of the linear system (3.2), i.e.

$$\begin{cases} 5a_0 + a_1 \sum_{i=1}^{5} x_i = \sum_{i=1}^{5} f(x_i) \\ \\ a_0 \sum_{i=1}^{5} x_i + a_1 \sum_{i=1}^{5} x_i^2 = \sum_{i=1}^{5} x_i f(x_i). \end{cases}$$

The coefficients of this system are calculated in the following table (see the columns corresponding to $x_i,\ f(x_i),\ x_i^2,\ x_i f(x_i)$):

i	x_i	$f(x_i)$	x_i^2	$x_i f(x_i)$	$g(x_i)$	$f(x_i) - g(x_i)$	$[f(x_i) - g(x_i)]^2$
1	-2	20	4	-40	17	3	9
2	-1	25	1	-25	27	-2	4
3	0	35	0	0	37	-2	4
4	1	45	1	45	47	-2	4
5	2	60	4	120	57	3	9
$\sum$	0	185	10	100	—	—	**30**

Therefore

$$\begin{cases} 5a_0 = 185 \\ 10a_1 = 100 \end{cases} \text{ and hence } \begin{cases} a_0 = 37 \\ a_1 = 10. \end{cases}$$

Then the linear trend function is

$$g(x) = 37 + 10x.$$

Hence we estimate that the sales for June will be

$$g(3) = 37 + 10 \cdot 3 = 67.$$

The error sum of squares, calculated in the final column of the above table, has the value

$$\sum_{i=1}^{5} [f(x_i) - g(x_i)]^2 = 30.$$

Case 2. For a quadratic trend function $g(x) = a_0 + a_1 x + a_2 x^2$, the parameters a_0, a_1 and a_2 are the solution of the linear system (3.1) for $k = 2$, i.e.

$$\begin{cases} 5a_0 + a_1 \sum_{i=1}^{5} x_i + a_2 \sum_{i=1}^{5} x_i^2 = \sum_{i=1}^{5} f(x_i) \\ a_0 \sum_{i=1}^{5} x_i + a_1 \sum_{i=1}^{5} x_i^2 + a_2 \sum_{i=1}^{5} x_i^3 = \sum_{i=1}^{5} x_i f(x_i) \\ a_0 \sum_{i=1}^{5} x_i^2 + a_1 \sum_{i=1}^{5} x_i^3 + a_2 \sum_{i=1}^{5} x_i^4 = \sum_{i=1}^{5} x_i^2 f(x_i). \end{cases}$$

The coefficients of this system are calculated in the following table:

i	x_i	$f(x_i)$	x_i^2	x_i^3	x_i^4	$x_i f(x_i)$	$x_i^2 f(x_i)$
1	-2	20	4	-8	16	-40	80
2	-1	25	1	-1	1	-25	25
3	0	35	0	0	0	0	0
4	1	45	1	1	1	45	45
5	2	60	4	8	16	120	240
$\sum$	0	185	10	0	34	100	390

Therefore

$$\begin{cases} 5a_0 + 10a_2 = 185, \\ 10a_1 = 100, \\ 10a_0 + 34a_2 = 390, \end{cases} \quad \text{and hence} \quad \begin{cases} a_0 = \dfrac{239}{7}, \\ a_1 = 10, \\ a_2 = \dfrac{10}{7}. \end{cases}$$

Then the quadratic trend function is

$$g(x) = \frac{239}{7} + 10x + \frac{10}{7} \cdot x^2.$$

Hence in this case we estimate that the sales for June will be

$$g(3) = \frac{239}{7} + 10 \cdot 3 + \frac{10}{7} \cdot 3^2 = 77.$$

The error sum of squares is calculated in the following table:

i	x_i	$f(x_i)$	$g(x_i)$	$f(x_i) - g(x_i)$	$[f(x_i) - g(x_i)]^2$
1	-2	20	19.857	0.143	0.02
2	-1	25	25.571	-0.571	0.33
3	0	35	34.143	0.857	0.73
4	1	45	45.571	-0.571	0.33
5	2	60	59.857	0.143	0.02
$\sum$	0	185	—	—	**1.43**

Then the error sum of squares has now the value

$$\sum_{i=1}^{5}[f(x_i) - g(x_i)]^2 \simeq 143.$$

Obviously, the quadratic approximation is better than the linear approximation. □

3.2 Average measures for two-dimensional statistical distributions

Two-dimensional statistical distribution: the statistical variable is $Z = (X, Y)$, where X and Y are two simple statistical variables, called **the components** of Z.

The distribution of $Z = (X, Y)$ is called **the joint distribution** of X and Y.

The distributions of X and Y are called **the marginal distributions** of $Z = (X, Y)$.

Let $x_1, \ldots, x_n$ be the values of distribution of X, $f_{1\cdot}, \ldots, f_{n\cdot}$ be their corresponding absolute frequencies and $f^*_{1\cdot}, \ldots, f^*_{n\cdot}$ be their corresponding relative frequencies.

Let $y_1, \ldots, y_m$ be the values of distribution of Y, $f_{\cdot 1}, \ldots, f_{\cdot m}$ be their corresponding absolute frequencies and $f^*_{\cdot 1}, \ldots, f^*_{\cdot m}$ be their corresponding relative frequencies.

Then the values of distribution of $Z = (X, Y)$ are the pairs (x_i, y_j), $i \in \{1, \ldots, n\}$, $j \in \{1, \ldots, m\}$.

Let f_{ij} be the absolute frequency and f^*_{ij} be the relative frequency of (x_i, y_j), for any $i \in \{1, \ldots, n\}$, $j \in \{1, \ldots, m\}$.

Obviously,

$$f_{i\cdot} = \sum_{j=1}^{m} f_{ij},\ f^*_{i\cdot} = \sum_{j=1}^{m} f^*_{ij},\ \forall i \in \{1, \ldots, n\},$$
$$f_{\cdot j} = \sum_{i=1}^{n} f_{ij},\ f^*_{\cdot j} = \sum_{i=1}^{n} f^*_{ij},\ \forall j \in \{1, \ldots, m\}.$$

If X and Y are **independent**, then $f_{ij} = f_{i\cdot} f_{\cdot j}$ and $f^*_{ij} = f^*_{i\cdot} f^*_{\cdot j}$, $\forall i, j$.

A two-dimensional statistical distribution and its marginal distributions are represented in **a cross-table** of one of the following forms:

$X \diagdown Y$	y_1	$\dots$	y_j	$\dots$	y_m	Total
x_1	f_{11}	$\dots$	f_{1j}	$\dots$	f_{1m}	$f_{1\cdot}$
$\vdots$						
x_i	f_{i1}	$\dots$	f_{ij}	$\dots$	f_{im}	$f_{i\cdot}$
$\vdots$						
x_n	f_{n1}	$\dots$	f_{nj}	$\dots$	f_{nm}	$f_{n\cdot}$
Total	$f_{\cdot 1}$	$\dots$	$f_{\cdot j}$	$\dots$	$f_{\cdot m}$	—

(absolute frequencies);

$X \diagdown Y$	y_1	$\dots$	y_j	$\dots$	y_m	Total
x_1	f^*_{11}	$\dots$	f^*_{1j}	$\dots$	f^*_{1m}	$f^*_{1\cdot}$
$\vdots$						
x_i	f^*_{i1}	$\dots$	f^*_{ij}	$\dots$	f^*_{im}	$f^*_{i\cdot}$
$\vdots$						
x_n	f^*_{n1}	$\dots$	f^*_{nj}	$\dots$	f^*_{nm}	$f^*_{n\cdot}$
Total	$f^*_{\cdot 1}$	$\dots$	$f^*_{\cdot j}$	$\dots$	$f^*_{\cdot m}$	—

(relative frequencies).

Let the two-dimensional distribution of $Z = (X, Y)$ as above.

- **The conditional distribution** of Y given the event $X = x_i$ has the values $y_1, \dots, y_m$, the corresponding absolute frequencies $f_{i1}, \dots, f_{im}$, and the corresponding relative frequencies

$$\frac{f_{i1}}{f_{i\cdot}} = \frac{f^*_{i1}}{f^*_{i\cdot}}, \dots, \frac{f_{im}}{f_{i\cdot}} = \frac{f^*_{im}}{f^*_{i\cdot}}\ , \text{ suppose that } f_{i\cdot} > 0.$$

  The conditional distribution of X given the event $Y = y_j$ has the values $x_1, \dots, x_n$, the corresponding absolute frequencies $f_{1j}, \dots, f_{nj}$, and the corresponding relative frequencies

$$\frac{f_{1j}}{f_{\cdot j}} = \frac{f^*_{1j}}{f^*_{\cdot j}}, \dots, \frac{f_{nj}}{f_{\cdot j}} = \frac{f^*_{nj}}{f^*_{\cdot j}}\ , \text{ suppose that } f_{\cdot j} > 0.$$

- **The** (u, v)**-th moment** of $Z = (X, Y)$ (or of the two-dimensional distribution of $Z = (X, Y)$) is

$$\overline{m}_{uv} = \frac{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} f_{ij} x_i^u y_j^v}{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} f_{ij}} = \sum_{i=1}^{n}\sum_{j=1}^{m} f^*_{ij} x_i^u y_j^v.$$

We remark that

$$\overline{m}_{10} = \overline{x} \text{ and } \overline{m}_{01} = \overline{y},$$

where

$$\overline{x} = \frac{\sum\limits_{i=1}^{n} f_{i\cdot} x_i}{\sum\limits_{i=1}^{n} f_{i\cdot}} = \sum_{i=1}^{n} f_{i\cdot}^{*} x_i$$

and

$$\overline{y} = \frac{\sum\limits_{j=1}^{m} f_{\cdot j} y_j}{\sum\limits_{j=1}^{m} f_{\cdot j}} = \sum_{j=1}^{m} f_{\cdot j}^{*} y_j$$

are **the means of the marginal distributions** (of X and Y, respectively).

- **The mean** of $Z = (X, Y)$ (or of the two-dimensional distribution of $Z = (X, Y)$) is the pair
$$\overline{z} = (\overline{x}, \overline{y}).$$

- **The conditional mean** of Y given the event $X = x_i$ (**the mean inside x_i-group**) is the mean of the conditional distribution of Y given $X = x_i$, i.e.
$$\beta_i = m_{Y/X=x_i} = \frac{\sum\limits_{j=1}^{m} f_{ij} y_j}{f_{i\cdot}} = \frac{\sum\limits_{j=1}^{m} f_{ij}^{*} y_j}{f_{i\cdot}^{*}}.$$

  The function
$$B(x_i) = \beta_i, \ \forall i \in \{1, \ldots, n\}$$
  is called **the regression function of the mean** of Y with respect to X.

  The conditional mean of X given the event $Y = y_j$ (**the mean inside y_j-group**) is the mean of the conditional distribution of X given $Y = y_j$, i.e.
$$\alpha_j = m_{X/Y=y_j} = \frac{\sum\limits_{i=1}^{n} f_{ij} x_i}{f_{\cdot j}} = \frac{\sum\limits_{i=1}^{n} f_{ij}^{*} x_i}{f_{\cdot j}^{*}}.$$

  The function
$$A(y_j) = \alpha_j, \ \forall j \in \{1, \ldots, m\}$$

is called **the regression function of the mean** of X with respect to Y.

We remark that the regression functions $B(x)$ and $A(y)$ can be approximated by Least Squares Method.

3.3 Variation measures for two-dimensional statistical distributions

Let the two-dimensional distribution of $Z = (X, Y)$ as above.

- **The (u, v)-th central moment** of $Z = (X, Y)$ (or of the two-dimensional distribution of $Z = (X, Y)$) is

$$m^c_{uv} = \frac{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} f_{ij}(x_i - \overline{x})^u (y_j - \overline{y})^v}{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} f_{ij}} = \sum_{i=1}^{n}\sum_{j=1}^{m} f^*_{ij}(x_i - \overline{x})^u (y_j - \overline{y})^v.$$

  We remark that

$$m^c_{20} = \sigma^2_X \text{ and } m^c_{02} = \sigma^2_Y,$$

  where

$$\sigma^2_X = \frac{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} f_{ij}(x_i - \overline{x})^2}{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} f_{ij}} = \sum_{i=1}^{n} f^*_{i\cdot}(x_i - \overline{x})^2$$

  and

$$\sigma^2_Y = \frac{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} f_{ij}(y_j - \overline{y})^2}{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} f_{ij}} = \sum_{j=1}^{m} f^*_{\cdot j}(y_j - \overline{y})^2$$

  are **the variances of the marginal distributions** (of X and Y, respectively).

- **The covariance** of X and Y (or of the two-dimensional distribution of $Z = (X, Y)$) is

$$\text{cov}\,(X, Y) = m^c_{11} = \frac{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} f_{ij}(x_i - \overline{x})(y_j - \overline{y})}{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} f_{ij}} = \sum_{i=1}^{n}\sum_{j=1}^{m} f^*_{ij}(x_i - \overline{x})(y_j - \overline{y}).$$

- **The conditional variance** of Y given the event $X = x_i$ (**the variance inside x_i-group**) is the variance of the conditional distribution of Y given $X = x_i$, i.e.

$$\sigma^2_{Y/X=x_i} = \frac{\sum\limits_{j=1}^{m} f_{ij}(y_j - \beta_i)^2}{f_{i\cdot}} = \frac{\sum\limits_{j=1}^{m} f^*_{ij}(y_j - \beta_i)^2}{f^*_{i\cdot}}.$$

  The conditional variance of X given the event $Y = y_j$ (**the variance inside y_j-group**) is the variance of the conditional distribution of X given $Y = y_j$, i.e.

$$\sigma^2_{X/Y=y_j} = \frac{\sum\limits_{i=1}^{n} f_{ij}(x_i - \alpha_j)^2}{f_{\cdot j}} = \frac{\sum\limits_{i=1}^{n} f^*_{ij}(x_i - \alpha_j)^2}{f^*_{\cdot j}}.$$

- **The average of the variances within x-groups** is

$$\overline{\sigma}^2_Y = \frac{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} f_{ij}(y_j - \beta_i)^2}{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} f_{ij}} = \sum_{i=1}^{n} f^*_{i\cdot}\sigma^2_{Y/X=x_i}.$$

  The average of the variances within y-groups is

$$\overline{\sigma}^2_X = \frac{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} f_{ij}(x_i - \alpha_j)^2}{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} f_{ij}} = \sum_{j=1}^{m} f^*_{\cdot j}\sigma^2_{X/Y=y_j}.$$

- **The conditional variance** of X given Y (**the variance between x-groups**) is

$$\sigma^2_{Y/X} = \frac{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} f_{ij}(\beta_i - \overline{y})^2}{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} f_{ij}} = \sum_{i=1}^{n} f^*_{i\cdot}(\beta_i - \overline{y})^2.$$

  The conditional variance of Y given X (**the variance between y-groups**) is

$$\sigma^2_{X/Y} = \frac{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} f_{ij}(\alpha_j - \overline{x})^2}{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} f_{ij}} = \sum_{j=1}^{m} f^*_{\cdot j}(\alpha_j - \overline{x})^2.$$

The rule of variances:

$$\sigma_Y^2 = \overline{\sigma}_Y^2 + \sigma_{Y/X}^2;$$

$$\sigma_X^2 = \overline{\sigma}_X^2 + \sigma_{X/Y}^2$$

(The overall variation is the combination result between the random factors within each group and the essential factors determining the variation from a group to another.)

3.4 Correlation between variables

The correlations (the dependence) that can be found between two variables X and Y are classified as follows:

- According to **the way of change** we can have:
  - **positive correlation** (**direct dependence**): if X is increasing then Y will also increase and if X is decreasing then Y will also decrease.
  - **negative correlation** (**opposite dependence**): if X is increasing then Y will decrease and if X is decreasing then Y will increase.
- According to **the intensity of the correlation** we can have:
  - **high intensity** (strong or tight);
  - **medium intensity**;
  - **low intensity**.
- According to **the shape** of the correlation we can have:
  - **linear correlation**;
  - **nonlinear correlation**, as exponential growth or logarithmic decrease, for example.

Let the two-dimensional distribution of $Z = (X, Y)$ as above. The degree of correlation between the variables X and Y can be measured by using the following indicators.

1. **The covariance** of X and Y (or of the two-dimensional distribution of $Z = (X, Y)$), i.e.

$$\operatorname{cov}(X,Y) = m_{11}^c = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} f_{ij}(x_i-\overline{x})(y_j-\overline{y})}{\sum_{i=1}^{n}\sum_{j=1}^{m} f_{ij}} = \sum_{i=1}^{n}\sum_{j=1}^{m} f_{ij}^*(x_i-\overline{x})(y_j-\overline{y}).$$

It takes values between $-\sigma_X\sigma_Y$ and $\sigma_X\sigma_Y$.

   - If X and Y are independent, then $\operatorname{cov}(X,Y) = 0$.
   - If $\operatorname{cov}(X,Y)$ is close to zero, then there is no linear dependence between the variables X and Y.
   - If $\operatorname{cov}(X,Y)$ is positive, then we have **a positive correlation** between the variables X and Y.
   - $\operatorname{cov}(X,Y) = \sigma_X\sigma_Y$ in the case of the perfect positive correlation (the linear increasing dependence).
   - If $\operatorname{cov}(X,Y)$ is negative, then we have **a negative correlation** between the variables X and Y.
   - $\operatorname{cov}(X,Y) = -\sigma_X\sigma_Y$ in the case of the perfect negative correlation (the linear decreasing dependence).

2. **The coefficient of correlation** of X and Y (or of the two-dimensional distribution of $Z = (X, Y)$) is

$$\rho = \rho(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sigma_X\sigma_Y}.$$

It takes values between -1 and 1.

The **regression line** (of Y with respect to X) is

$$y - \overline{y} = \rho \cdot \frac{\sigma_Y}{\sigma_X}(x - \overline{x}).$$

   - If X and Y are independent, then $\rho(X,Y) = 0$.
   - If $\rho(X,Y) = 0$, then there is no linear dependence between the variables X and Y (the variables are independent or there is a nonlinear dependence!).
   - If $\rho(X,Y) = 1$, then we have a direct linear dependence between the variables X and Y, given by the regression line

$$y - \overline{y} = \frac{\sigma_Y}{\sigma_X}(x - \overline{x}).$$

- If $\rho(X,Y) = -1$, then we have an opposite linear dependence between the variables X and Y, given by the regression line
$$y - \overline{y} = -\frac{\sigma_Y}{\sigma_X}(x - \overline{x}).$$
- If $0 < \rho(X,Y) < 0.2$, then we have a low positive correlation between the variables X and Y.
- If $-0.2 < \rho(X,Y) < 0$, then we have a low negative correlation between the variables X and Y.
- If $0.2 \leq \rho(X,Y) \leq 0.5$, then we have a weak positive correlation between the variables X and Y, case needing a significance test to be applied (like as the Student test).
- If $-0.5 \leq \rho(X,Y) \leq -0.2$, then we have a weak negative correlation between the variables X and Y, case needing a significance test to be also applied.
- If $0.5 < \rho(X,Y) \leq 0.75$, then we have a medium positive correlation between the variables X and Y.
- If $-0.75 \leq \rho(X,Y) < -0.5$, then we have a medium negative correlation between the variables X and Y.
- If $0.75 < \rho(X,Y) \leq 0.95$, then we have a high positive correlation between the variables X and Y.
- If $-0.95 \leq \rho(X,Y) < -0.75$, then we have a high negative correlation between the variables X and Y.
- If $0.95 < \rho(X,Y) < 1$, then we have an extremely strong positive correlation between the variables X and Y, almost a direct linear dependence.
- If $-1 < \rho(X,Y) < -0.95$, then we have an extremely strong negative correlation between the variables X and Y. almost an opposite linear dependence.

3. **The coefficient of determination** of Y with respect to X is
$$R^2 = \frac{\sigma^2_{Y/X}}{\sigma^2_Y},$$
and **the coefficient of non-determination** of Y with respect to X is
$$K^2 = \frac{\overline{\sigma}^2_Y}{\sigma^2_Y}.$$

By the rule of variances $\sigma_Y^2 = \overline{\sigma}_Y^2 + \sigma_{Y/X}^2$ it follows that

$$R^2 + K^2 = 1.$$

The coefficient R^2 shows the share of the variance between groups in the overall variance, expressing the influence of the classification factors.

- If $R^2 = 1$, then there is a strong functional relation between Y and X.
- If $0.7 < R^2 < 1$, then the classification of the population according to X has a meaning, X variation influencing Y variation.
- If $0.5 < R^2 \leq 0.7$, then the differences between the group means are significant.
- If $R^2 = 0.5$, then we cannot decide whether X variation has a significant influence over Y variation.
- If $0 < R^2 < 0.5$, then X variation has no a significant influence over Y variation.
- If $R^2 = 0$, then X variation has no influence over Y variation.

3.5 Nonparametric measures of correlation

If we do not have sufficient elements to identify the rule of distributions, then we can use nonparametric methods like as the coefficients of ranks correlation proposed by Kendall and Spearman.

Let X and Y be two simple statistical variables for a statistical population or for a statistical sample. Let $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ be the ungrouped values (variants) of the distributions of X and Y, respectively. The distributions of X and Y are represented in a table of the following form:

Units	u_1	u_2	...	u_n
X values	x_1	x_2	...	x_n
Y values	y_1	y_2	...	y_n

where n is **the volume of population** (the number of statistical units u_i).

Let a_i be **the rank of the variant** x_i inside the distribution of X, namely the rank of x_i in the increasing order of $x_1, \ldots, x_n$. Let also b_i be **the rank of the variant** y_i inside the distribution of Y, namely the rank of y_i in the increasing order of $y_1, \ldots, y_n$.

- **The Spearman's coefficient of correlation of the ranks** is

$$\rho_S = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n^3 - n},$$

where

$$d_i = a_i - b_i,\ \forall i \in \{1, \ldots, n\}$$

(the rank differences between variables).

We remark that the Spearman's coefficient of correlation of the ranks ρ_S is even the coefficient of correlation $\rho(A, B)$ of A and B, where A and B are the statistical variables that represent the ranks of X and Y, respectively. The distributions of A and B are represented in the following table:

Units	u_1	u_2	...	u_n
A values	a_1	a_2	...	a_n
B values	b_1	b_2	...	b_n

- If some ranks of X or Y are equal, then on can use **the corrected Spearman's coefficient of correlation of the ranks**, given by

$$\widetilde{\rho}_S = 1 - \frac{6\left(\sum_{i=1}^{n} d_i^2 + \frac{t^3 - t}{12}\right)}{n^3 - n},$$

where t is the number of equal ranks.

- **The Kendall's coefficient of correlation of the ranks** is

$$\rho_K = \frac{2(P - Q)}{n^2 - n},$$

where

$$P = \sum_{i=1}^{n} P_i,\ Q = \sum_{i=1}^{n} Q_i,$$

$$P_i = \text{card}\,\{j = \overline{1, n}\ |a_j > a_i \text{ and } b_j > b_i\},$$

$$Q_i = \text{card}\,\{j = \overline{1, n}\ |a_j > a_i \text{ and } b_j < b_i\},$$

for all $i \in \{1, \ldots, n\}$.

We remark that the numbers P_i are indicators of **the concordance**, and the numbers Q_i are indicators of **the discordance** between the ranks.

- The coefficients of correlation of the ranks take values between -1 and $+1$. They can interpreted similarly with the coefficient ρ of correlation.

These coefficients have the advantage that they can be used in the case of skewed distributions or a small number of units. Also, these coefficients are applicable for studying the relation between qualitative variables that cannot be expressed numerically, but can be classified by their ranks.

3.6 Problems

Exercise 3.6.1. The sales of a product for the last eight years are as follows:

Year	2009	2010	2011	2012	2013	2014	2015	2016
Sales	34	30	23	17	14	13	11	10

Determine a trend function of sales and a forecast for 2017.

Exercise 3.6.2. Two hydrological stations make each a hundred measurements of the level of a river during a year. The recorded data are given in the following table.

$X \diagdown Y$	3.0	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9	4.0
3.2	1	1									
3.3		1	1		2						
3.4			1	2	2	1					
3.5			3	5	2	1	2				
3.6			1	3	5	4	1	1	1		
3.7			1	1	10	3	1	1			
3.8				1	1	11	2		1		
3.9			1		1	12	1		1		1
4.0							2		1	1	
4.1								2		2	1

a. Represent the data into a scatter diagram.

b. Compute the means and the variances of X and Y and the covariance of X and Y.

c. Compute the linear regression function of the mean of Y with respect to X.

d. Compute and interpret the coefficient of correlation of X and Y, the regression line of Y with respect to X, and the coefficient of determination of Y with respect to X.

Exercise 3.6.3. A group of students obtained the following marks over two tests:

Students	1	2	3	4	5	6	7	8	9	10
Test A marks	10	25	13	14	28	16	6	8	24	17
Test B marks	17	23	15	12	26	18	8	13	20	22

Students	11	12	13	14	15	16	17	18	19	20
Test A marks	30	15	23	4	26	12	21	19	29	18
Test B marks	28	13	25	10	27	5	19	14	29	24

Compute and interpret the coefficients of correlation of the ranks between the results of these tests.

Theme 4

Time series and forecasting

Usually, a time series $Y = (y_i)_i$ (i being the time) is influenced by the following **factors (components)**:

- **the trend (the tendency)**;
- **the cyclical factor**;
- **the seasonal factor**;
- **the random factor (the irregular factor)**.

The main **decomposition models** for a time series $Y = (y_i)_i$:

- **The additive model**:

  $$y_i = T_i + C_i + S_i + R_i,$$

  where T_i, C_i, S_i, R_i represent the trend, the cyclical, the seasonal and the random components, respectively.

  This model assumes that the components are independent and they have the same measurement unit.

- **The multiplicative model**:

  $$y_i = T_i \cdot C_i \cdot S_i \cdot R_i.$$

  This model assumes that the components depend each other or they have different measurement units.

4.1 The trend component

This component can be determined by Least Squares Method.

4.2 The cyclical component

The cyclical variation of a time series is the component that tends to oscillate above and below the trend line for periods longer than 1 year (if the time series is composed by annual dates). This component explains most of the variation of evolution that remains unexplained by the trend component.

The cyclical component can be expressed as:

- **The cyclical variation**:

$$C_i = y_i - \widetilde{y}_i,$$

  where

  - y_i is the value of time series Y at time i;
  - $\widetilde{y}_i = T_i$ is the estimated trend value of time series Y at the same time i.

- **The cycle**:

$$\frac{y_i}{\widetilde{y}_i} \cdot 100.$$

- **The relative cyclical residual**:

$$\frac{y_i - \widetilde{y}_i}{\widetilde{y}_i} \cdot 100.$$

Example 4.2.1. The sales of a company for the last nine years are as follows:

Year	2008	2009	2010	2011	2012	2013	2014	2015	2016
Sales	5.7	5.9	6	6.2	6.3	6.3	6.4	6.4	6.6

Determine the trend function of sales and evaluate the cyclical variation.

Solution. Using the Least Squares Method, we obtain that the trend line is

$$\widetilde{y} = 5.7 + 0.1y$$

(where $y_1 = 1, \ldots, y_9 = 9$).

The estimated sales, the cyclical variations, the cycles and the relative cyclical residuals are calculated in the following table:

Year	Sales (y_i)	Estimated sales ($\widetilde{y}_i$)	Cyclical variation	Cycle	Rel. cycl. residual
2008	5.7	5.8	-0.1	98.28	-1.72
2009	5.9	5.9	0	100.00	0.00
2010	6	6	0	100.00	0.00
2011	6.2	6.1	0.1	101.64	1.64
2012	6.3	6.2	0.1	101.61	1.61
2013	6.3	6.3	0	100.00	0.00
2014	6.4	6.4	0	100.00	0.00
2015	6.4	6.5	-0.1	98.46	-1.54
2016	6.6	6.6	0	100.00	0.00

□

4.3 The seasonal component

The seasonal variation of a time series is the repetitive and predictable movement around the trend line in 1 year or less. For detecting the seasonal variation, the time intervals need to be measured in small periods such as quarters, months, weeks, ...

Let $Y = (y_i)_{i=\overline{1,n}}$ be a time series, and let k be the number of equal periods per each year.

The seasonal component can be expressed as:

- **The moving average value for each time interval**:

  If k is odd, the moving average value corresponding to y_i is

$$\overline{y}_i = \frac{1}{k}\left(y_{i-\frac{k-1}{2}} + \cdots + y_i + \cdots + y_{i+\frac{k-1}{2}}\right),$$

  for all $i \in \{1 + \frac{k-1}{2}, \ldots, n - \frac{k-1}{2}\}$.

  If k is even, the moving average value corresponding to y_i is

$$\overline{y}_i = \frac{1}{k}\left(\frac{1}{2}y_{i-\frac{k}{2}} + y_{i-\frac{k}{2}+1} + \cdots + y_i + \cdots + y_{i+\frac{k}{2}-1} + \frac{1}{2}y_{i+\frac{k}{2}}\right),$$

  for all $i \in \{1 + \frac{k}{2}, \ldots, n - \frac{k}{2}\}$.

- **The percentage of actual value to the moving average value**:

$$\frac{y_i}{\overline{y}_i} \cdot 100.$$

- **The seasonal index** for each period is obtained by eliminating the extreme values of the above percentages (of actual value to the moving average value) corresponding to period (that is one minimum value and one maximum value for period) and computing the mean of the remaining values.

The seasonal indexes are used to **deseasonalizing the time series**, in order to remove the effects of the seasonality from the recorded dates. For that, each actual recorded data is dividing by the correspondent seasonal index, before the computation of the trend and the cyclical components of the time series.

After the computation of these components, each of the forecasted values of the deseasonalized time series is multiplied by the correspondent seasonal index to obtain forecasts for the original time series (**reseasonalizing the time series**).

Example 4.3.1. The quarterly sales of a company for the last five years are as follows:

Year	Quarter I	Quarter II	Quarter III	Quarter IV
2012	1022	656	1756	1300
2013	1134	698	1784	1346
2014	1164	712	1864	1380
2015	1208	724	1922	1434
2016	1240	740	1890	1418

Compute the seasonal indexes of the four quarters and deseasonalize the time series.

Solution. We compute the 4-quarter moving averages. We have $n = 20$ and $k = 4$. The moving average value corresponding to y_i is

$$\overline{y}_i = \frac{1}{4}\left(\frac{1}{2}y_{i-2} + y_{i-1} + y_i + y_{i+1} + \frac{1}{2}y_{i+2}\right),$$

for all $i \in \{3, \ldots, 18\}$. For example,

$$\overline{y}_3 = \frac{1}{4}\left(\frac{1}{2} \cdot 1022 + 656 + 1756 + 1300 + \frac{1}{2} \cdot 1134\right) = 1197.5\,.$$

The 4-quarter moving average values are calculated in the following table:

Year	Quarter I	Quarter II	Quarter III	Quarter IV
2012	–	–	1197.50	1216.75
2013	1225.50	1234.75	1244.25	1249.75
2014	1261.50	1275.75	1285.50	1292.50
2015	1301.25	1315.25	1326.00	1332.00
2016	1330.00	1324.00	–	–

The percentages of actual values to the moving average values $\frac{y_i}{\overline{y}_i} \cdot 100\,\%$ are calculated in the following table:

Year	Quarter I	Quarter II	Quarter III	Quarter IV
2012	–	–	146.6388 %	106.8420 %
2013	92.5337 %	56.5297 %	143.3795 %	107.7015 %
2014	92.2711 %	55.8103 %	145.0019 %	106.7698 %
2015	92.8338 %	55.0466 %	144.9472 %	107.6577 %
2016	93.2331 %	55.8912 %	–	–

The seasonal indexes are calculated in the following table:

Quarter I	Quarter II	Quarter III	Quarter IV
92.6837 %	55.8508 %	144.9746 %	107.2498 %

For example, the seasonal index for Quarter I is obtained as

$$\frac{92.5337\,\% + 92.8338\,\%}{2} = 92.6837\,\%.$$

For deseasonalizing the time series, we dividing each actual recorded data by the correspondent seasonal index, as in the following table:

Year	Quarter I	Quarter II	Quarter III	Quarter IV
2012	1102.67	1174.56	1211.25	1212.12
2013	1223.52	1249.76	1230.56	1255.01
2014	1255.88	1274.83	1285.74	1286.72
2015	1303.36	1296.31	1325.75	1337.07
2016	1337.88	1324.96	1303.68	1322.15

□

4.4 Problems

Exercise 4.4.1. The quarterly sales of a company for the last eight years are as follows:

Year	Quarter I	Quarter II	Quarter III	Quarter IV
2009	112	130	103	140
2010	118	128	104	144
2011	116	132	106	147
2012	120	130	110	150
2013	124	133	112	156
2014	126	137	115	160
2015	125	136	119	162
2016	128	141	118	167

a. Compute the four-quarter moving averages.

b. Represent the time series and the moving averages into a scatter diagram.

c. Compute the seasonal indexes of the four quarters.

d. Deseasonalize the time series.

e. Determine the trend function of sales and evaluate the cyclical variation.

f. Calculate the corresponding forecast for the next year.

Exercise 4.4.2. The number of clients of a market for the last six weeks are as follows:

Week	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
1	586	774	758	653	997	1234	354
2	579	770	754	648	987	1226	348
3	570	772	747	644	991	1208	342
4	572	762	742	641	988	1212	337
5	565	758	736	637	982	1195	332
6	560	743	733	640	976	1174	327

a. Compute the 7-day moving averages.

b. Compute the seasonal indexes of the seven days.

c. Deseasonalize the time series.

d. Determine the trend function of clients number and evaluate the cyclical variation.

e. Calculate the corresponding forecast for the next week.

Theme 5

The interest

5.1 A general model of interest

Definition 5.1.1. **The interest** *corresponding to* **the initial value (the present value, the principal)** S_0 *(expressed in units of currency (u.c.)) over* **the time (the period of investment)** t *(usually expressed in years) is a function* $D : [0, \infty) \times [0, \infty) \to [0, \infty)$ *that verifies the following two conditions:*

1. $D(S_0, 0) = 0, \forall S_0 \geq 0;\ D(0, t) = 0, \forall t \geq 0;$

2. *The function* $D(S_0, t)$ *increases in each of the two variables* S_0 *and* t. *Assuming that the function* $D(S_0, t)$ *has partial derivatives, this condition can be expressed as:*

$$\frac{\partial D}{\partial S_0}(S_0, t) > 0,\ \frac{\partial D}{\partial t}(S_0, t) > 0, \forall S_0 > 0, \forall t > 0.$$

Definition 5.1.2. *The sum*

$$S(S_0, t) = S_0 + D(S_0, t)$$

is called **the final value (the future value, the amount)***, and is also denoted by* S_t.

Remark 5.1.1. *The final value is a function* $S : [0, \infty) \times [0, \infty) \to [0, \infty)$ *that verifies the following two conditions:*

$$S(S_0, 0) = S_0, \forall S_0 \geq 0;\ S(0, t) = 0, \forall t \geq 0;$$

$$\frac{\partial S}{\partial S_0}(S_0, t) > 1,\ \frac{\partial S}{\partial t}(S_0, t) > 0, \forall S_0 > 0, \forall t > 0$$

Definition 5.1.3. The annual interest rate, *denoted by* i, *is the interest for* 1 *u.c. over* 1 *year, that is*

$$i = D(1, 1).$$

The annual interest percentage, *denoted by* p, *is the interest for* 100 *u.c. over* 1 *year, that is*

$$p = D(100, 1).$$

Remark 5.1.2. *Usually,* $p = 100i$.

Definition 5.1.4. *The function* $F : [0, \infty) \times [0, \infty) \to [0, \infty)$ *given by*

$$F(S_0, t) = \frac{\partial D}{\partial t}(S_0, t), \forall S_0 \geq 0, \forall t \geq 0$$

is called **the proportionality factor of the interest**.

Remark 5.1.3.

$$F(S_0, t) = \frac{\partial S}{\partial t}(S_0, t), \forall S_0 \geq 0, \forall t \geq 0.$$

Proposition 5.1.1. *We have*

$$D(S_0, t) = \int_0^t F(S_0, x)dx, \forall S_0 \geq 0, \forall t \geq 0;$$

$$S_t \equiv S(S_0, t) = S_0 + \int_0^t F(S_0, x)dx, \forall S_0 \geq 0, \forall t \geq 0$$

Corollary 5.1.1. *We have*

$$i = \int_0^1 F(1, x)dx; \ p = \int_0^1 F(100, x)dx.$$

Definition 5.1.5. *The function* $\delta : [0, \infty) \to [0, \infty)$ *given by*

$$\delta(t) = \frac{\frac{\partial S}{\partial t}(S_0, t)}{S(S_0, t)}, \forall t \geq 0$$

is called **the instantaneous interest rate**.

Remark 5.1.4.

$$\delta(t) = \frac{\partial \ln S}{\partial t}(S_0, t) = \frac{F(S_0, t)}{S(S_0, t)}, \forall t \geq 0.$$

Proposition 5.1.2. *We have*

$$S_t \equiv S(S_0, t) = S_0 e^{\int_0^t \delta(x)dx}, \forall S_0 \geq 0, \forall t \geq 0;$$

$$D(S_0, t) = S_0 \left[e^{\int_0^t \delta(x)dx} - 1 \right], \forall S_0 \geq 0, \forall t \geq 0.$$

Corollary 5.1.2. *We have*

$$i = e^{\int_0^1 \delta(x)dx} - 1;\ p = 100 \left[e^{\int_0^1 \delta(x)dx} - 1 \right].$$

5.2 Equivalence of investments

Definition 5.2.1. ***A multiple (financial) investment*** *consists in n initial values $S_{01},\ S_{02}, \ldots,\ S_{0n}$ invested over the times $t_1,\ t_2,\ \ldots,\ t_n$, with the annual interest rates $i_1,\ i_2,\ \ldots,\ i_n$ (or with the annual interest percentages $p_1,\ p_2,\ \ldots, p_n$). Let $D(S_{01}, t_1),\ D(S_{02}, t_2),\ \ldots,\ D(S_{0n}, t_n)$ be the corresponding interests, and let $S_1 = S(S_{01}, t_1),\ S_2 = S(S_{02}, t_2),\ \ldots,\ S_n = S(S_{0n}, t_n)$ be the corresponding final values. The sums $\sum\limits_{k=1}^{n} S_{0k}$, $\sum\limits_{k=1}^{n} D(S_{0k}, t_k)$ and $\sum\limits_{k=1}^{n} S_k$ are called the* ***total initial value****, **the total interest** and **the total final value** of the given multiple investment, respectively. This multiple investment can be expressed by a matrix of one of the following two forms $\begin{pmatrix} S_{01} & t_1 & i_1 \\ S_{02} & t_2 & i_2 \\ \vdots & \vdots & \vdots \\ S_{0n} & t_n & i_n \end{pmatrix}$ (if the initial values are known), or $\begin{pmatrix} t_1 & i_1 & S_1 \\ t_2 & i_2 & S_2 \\ \vdots & \vdots & \vdots \\ t_n & i_n & S_n \end{pmatrix}$ (if the final values are known).*

Definition 5.2.2. *We say that two multiple investments are* **equivalent by interest** *and we denote $\begin{pmatrix} S_{01} & t_1 & i_1 \\ S_{02} & t_2 & i_2 \\ \vdots & \vdots & \vdots \\ S_{0n} & t_n & i_n \end{pmatrix} \overset{I}{\sim} \begin{pmatrix} S'_{01} & t'_1 & i'_1 \\ S'_{02} & t'_2 & i'_2 \\ \vdots & \vdots & \vdots \\ S'_{0m} & t'_m & i'_m \end{pmatrix}$ if the corresponding total interest are equal, i.e. $\sum\limits_{k=1}^{n} D(S_{0k}, t_k) = \sum\limits_{k=1}^{m} D(S'_{0k}, t'_k)$.*

We say that two multiple investments are **equivalent by present value** *and we denote* $\begin{pmatrix} t_1 & i_1 & S_1 \\ t_2 & i_2 & S_2 \\ \vdots & \vdots & \vdots \\ t_n & i_n & S_n \end{pmatrix} \overset{P}{\sim} \begin{pmatrix} t'_1 & i'_1 & S'_1 \\ t'_2 & i'_2 & S'_2 \\ \vdots & \vdots & \vdots \\ t'_m & i'_m & S'_m \end{pmatrix}$ *if the corresponding total initial values are equal, i.e.* $\sum\limits_{k=1}^{n} S_{0k} = \sum\limits_{k=1}^{m} S'_{0k}$.

Definition 5.2.3. *If*

$$\begin{pmatrix} S_{01} & t_1 & i_1 \\ S_{02} & t_2 & i_2 \\ \vdots & \vdots & \vdots \\ S_{0n} & t_n & i_n \end{pmatrix} \overset{I}{\sim} \left(S_0^{(CI)}, t, i\right) \overset{I}{\sim} \left(S_0, t^{(CI)}, i\right) \overset{I}{\sim} \left(S_0, t, i^{(CI)}\right),$$

then the initial value $S_0^{(CI)}$*, the time of investment* $t^{(CI)}$ *and the annual interest rate* $i^{(CI)}$ *are called* ***commonly replacements by interest****.*

Definition 5.2.4. *If*

$$\begin{pmatrix} S_{01} & t_1 & i_1 \\ S_{02} & t_2 & i_2 \\ \vdots & \vdots & \vdots \\ S_{0n} & t_n & i_n \end{pmatrix} \overset{I}{\sim} \begin{pmatrix} S_0^{(MI)} & t_1 & i_1 \\ S_0^{(MI)} & t_2 & i_2 \\ \vdots & \vdots & \vdots \\ S_0^{(MI)} & t_n & i_n \end{pmatrix} \overset{I}{\sim} \begin{pmatrix} S_{01} & t^{(MI)} & i_1 \\ S_{02} & t^{(MI)} & i_2 \\ \vdots & \vdots & \vdots \\ S_{0n} & t^{(MI)} & i_n \end{pmatrix} \overset{I}{\sim} \begin{pmatrix} S_{01} & t_1 & i^{(MI)} \\ S_{02} & t_2 & i^{(MI)} \\ \vdots & \vdots & \vdots \\ S_{0n} & t_n & i^{(MI)} \end{pmatrix},$$

then the initial value $S_0^{(MI)}$*, the time of investment* $t^{(MI)}$ *and the annual interest rate* $i^{(MI)}$ *are called* ***meanly replacements by interest****.*

Definition 5.2.5. *If*

$$\begin{pmatrix} t_1 & i_1 & S_1 \\ t_2 & i_2 & S_2 \\ \vdots & \vdots & \vdots \\ t_n & i_n & S_n \end{pmatrix} \overset{P}{\sim} \left(t, i, S^{(CP)}\right) \overset{P}{\sim} \left(t^{(CP)}, i, S\right) \overset{P}{\sim} \left(t, i^{(CP)}, S\right).$$

then the final value $S^{(CP)}$*, the time of investment* $t^{(CP)}$ *and the annual interest rate* $i^{(CP)}$ *are called* ***commonly replacements by present value****.*

Definition 5.2.6. *If*

$$\begin{pmatrix} t_1 & i_1 & S_1 \\ t_2 & i_2 & S_2 \\ \vdots & \vdots & \vdots \\ t_n & i_n & S_n \end{pmatrix} \overset{P}{\sim} \begin{pmatrix} t_1 & i_1 & S^{(MP)} \\ t_2 & i_2 & S^{(MP)} \\ \vdots & \vdots & \vdots \\ t_n & i_n & S^{(MP)} \end{pmatrix} \overset{P}{\sim} \begin{pmatrix} t^{(MP)} & i_1 & S_1 \\ t^{(MP)} & i_2 & S_2 \\ \vdots & \vdots & \vdots \\ t^{(MP)} & i_n & S_n \end{pmatrix} \overset{P}{\sim} \begin{pmatrix} t_1 & i^{(MP)} & S_1 \\ t_2 & i^{(MP)} & S_2 \\ \vdots & \vdots & \vdots \\ t_n & i^{(MP)} & S_n \end{pmatrix},$$

*then the final value $S^{(MP)}$, the time of investment $t^{(MP)}$ and the annual interest rate $i^{(MP)}$ are called **meanly replacements by present value**.*

5.3 Simple interest

5.3.1 Basic formulas

Definition 5.3.1. *If the principal is not actualized over the time of investment, then we say that we obtain **a simple interest**.*

Proposition 5.3.1. *For simple interest we have:*

$$D \equiv D(S_0, t) = S_0 it = \frac{S_0 pt}{100}$$

*(**the simple interest formula**),*

$$S_t \equiv S(S_0, t) = S_0 + D = S_0(1 + it)$$

*(**the compounding formula, the rule of interest**),*

$$S_0 = \frac{S_t}{1 + it}$$

*(**the discounting formula**),*

$$i = \frac{D}{S_0 t} = \frac{S_t - S_0}{S_0 t},\ t = \frac{D}{S_0 i} = \frac{S_t - S_0}{S_0 i}.$$

Remark 5.3.1. *According to the above formulas, $1 + it$ is called **the compounding factor**, and $\frac{1}{1+it}$ is called **the discounting factor** for simple interest.*

Corollary 5.3.1. *If $t = \dfrac{h}{k}$ (i.e. k is the number of periods per year and h is the number of such periods), then the simple interest is:*

$$D\left(S, \frac{h}{k}\right) = \frac{S_0 ih}{k} = \frac{S_0 pt}{100k}.$$

Remark 5.3.2. *If the time of investment t is given as the period from the initial date (d_1, m_1, y_1) to the final date (d_2, m_2, y_2), (where d_i, m_i, y_i represents the day, the number of month and the year of the date), then we have three conventions (procedures) to calculate the simple interest:*

1. ***The exact interest ("actual/actual")***:

$$D = \frac{S_0 ih}{365} \quad \text{or} \quad D = \frac{S_0 ih}{366} \quad \text{(for leap years)},$$

where h is the number of calendar days from (d_1, m_1, y_1) to (d_2, m_2, y_2) (excluding either the first or last day);

2. ***The banker's rule ("actual/360")***:

$$D = \frac{S_0 ih}{360},$$

where h is the number of calendar days from (d_1, m_1, y_1) to (d_2, m_2, y_2) (excluding either the first or last day);

3. ***The ordinary interest ("30/360")***:

$$D = \frac{S_0 ih}{360},$$

where

$$h = 360(y_2 - y_1) + 30(m_2 - m_1) + d_2 - d_1$$

*(assumes that all months have 30 days, called **the 30-day month convention**).*

5.3.2 Simple interest with variable rate

Proposition 5.3.2. *If the time of investment is $t = t_1 + t_2 + ... + t_m$ and the annual interest rate is i_1 for the first period t_1, i_2 for the second period t_2,..., i_m for the last period t_m, then we have:*

$$D = S_0 \sum_{k=1}^{m} i_k t_k \quad \text{(the simple interest formula)};$$

$$S_t = S_0 \left(1 + \sum_{k=1}^{m} i_k t_k\right) \quad \text{(the compounding formula)};$$

$$S_0 = \frac{S_t}{1 + \sum\limits_{k=1}^{m} i_k t_k} \quad \text{(the discounting formula)}.$$

5.3.3 Equivalence by simple interest

Proposition 5.3.3. *For simple interest we have:*

$$S_0^{(CI)} = \frac{\sum_{k=1}^{n} S_{0k} i_k t_k}{it}, \ t^{(CI)} = \frac{\sum_{k=1}^{n} S_{0k} i_k t_k}{S_0 i}, \ i^{(CI)} = \frac{\sum_{k=1}^{n} S_{0k} i_k t_k}{S_0 t},$$

$$S_0^{(MI)} = \frac{\sum_{k=1}^{n} S_{0k} i_k t_k}{\sum_{k=1}^{n} i_k t_k}, \ t^{(MI)} = \frac{\sum_{k=1}^{n} S_{0k} i_k t_k}{\sum_{k=1}^{n} S_{0k} i_k}, \ i^{(MI)} = \frac{\sum_{k=1}^{n} S_{0k} i_k t_k}{\sum_{k=1}^{n} S_{0k} t_k}.$$

5.4 Compound interest

5.4.1 Basic formulas

Definition 5.4.1. *If the principal is actualized over each year of investment time (by adding the interest of the previous year), then we say that we obtain* ***a compound interest****.*

Proposition 5.4.1. *For compound interest we have:*

$$D \equiv D(S_0, t) = S_0 \left[(1+i)^t - 1 \right]$$

*(**the compound interest formula**),*

$$S_t \equiv S(S_0, t) = S_0 + D = S_0 (1+i)^t$$

*(**the compounding formula, the rule of interest**),*

$$S_0 = \frac{S_t}{(1+i)^t}$$

*(**the discounting formula**).*

Remark 5.4.1. *According to the above formulas, $(1+i)^t$ is called* ***the compounding factor****, and $\frac{1}{(1+i)^t}$ is called* ***the discounting factor*** *for compound interest. Denoting* ***the annual compounding factor*** *by*

$$u = 1 + i$$

and ***the annual discounting factor*** *by*

$$v = \frac{1}{u} = \frac{1}{1+i},$$

the above formulas can be written as

$$S_t = S_0 u^t, \quad S_0 = S_t v^t.$$

Remark 5.4.2. *If*

$$t = n + \frac{h}{k}$$

*(n being **the integer part** and $\frac{h}{k}$ being **the fractional part**, i.e. the time of investment cover only h periods from a total of k equal periods per the last year), then we have two conventions (procedures) to calculate the compound interest:*

1. ***The rational procedure**: we apply a compound interest for the integer part and a simple interest for the fractional part, and hence*

$$S_t \equiv S_{n+\frac{h}{k}} = S_0(1+i)^n\left(1 + i\cdot\frac{h}{k}\right) \quad \textit{(the compounding formula);}$$

$$D = S_0\left[(1+i)^n\left(1+i\cdot\frac{h}{k}\right) - 1\right] \quad \textit{(the interest formula).}$$

2. ***The commercial procedure**: we extend the compound interest to the fractional part, and hence*

$$S_t \equiv S_{n+\frac{h}{k}} = S_0(1+i)^{n+\frac{h}{k}} = S_0(1+i)^n\sqrt[k]{(1+i)^h} \quad \textit{(the compounding formula);}$$

$$D = S_0\left[(1+i)^{n+\frac{h}{k}} - 1\right] = S_0\left[(1+i)^n\sqrt[k]{(1+i)^h} - 1\right] \quad \textit{(the interest formula).}$$

5.4.2 Nominal rate and effective rate

Definition 5.4.2. *For an initial value S_0 over n years at an annual rate j_k compounded k times per each year, the final value is*

$$S_t = S_0\left(1 + \frac{j_k}{k}\right)^{kn} = S_0\left(1+i\right)^n,$$

where

$$1 + i = \left(1 + \frac{j_k}{k}\right)^k.$$

- *k is called **the number of interest periods per year**;*
- *j_k is called **the nominal rate** (annual interest rate);*

- $i_k = \dfrac{j_k}{k}$ *is called* ***the interest rate per interest period*** *(period interest rate);*
- i *is called* ***the effective rate*** *or* ***the real rate*** *(annual interest rate).*

5.4.3 Compound interest with variable rate

Proposition 5.4.2. *If the time of investment is* $t = t_1 + t_2 + ... + t_m$ *and the annual interest rate is* i_1 *for the first period* $t_1 = n_1 + \dfrac{h_1}{k_1}$, i_2 *for the second period* $t_2 = n_2 + \dfrac{h_2}{k_2}$,..., i_m *for the last period* $t_m = n_m + \dfrac{h_m}{k_m}$, *then we have:*

1. *For the rational procedure:*

$$D = S_0 \left\{ \prod_{l=1}^{m} \left[(1+i_l)^{n_l} \left(1 + i_l \cdot \frac{h_l}{k_l}\right) \right] - 1 \right\} \quad \text{(the compound interest formula);}$$

$$S_t = S_0 \prod_{l=1}^{m} \left[(1+i_l)^{n_l} \left(1 + i_l \cdot \frac{h_l}{k_l}\right) \right] \quad \text{(the compounding formula);}$$

2. *For the commercial procedure:*

$$D = S_0 \left[\prod_{l=1}^{m} (1+i_l)^{t_l} - 1 \right] \quad \text{(the compound interest formula);}$$

$$S_t = S_0 \prod_{l=1}^{m} (1+i_l)^{t_l} \quad \text{(the compounding formula).}$$

5.5 Loans

The amortization table for a loan of size (**original balance**) V_0 u.c. per n years at an annual interest rate i has the following form:

Year	Year's start Remaining principal	Interest part	Principal part	Payment (Rate)	Year's end Remaining principal
1	V_0	$d_1 = V_0 \cdot i$	Q_1	$T_1 = d_1 + Q_1$	$V_1 = V_0 - Q_1$
2	V_1	$d_2 = V_1 \cdot i$	Q_2	$T_2 = d_2 + Q_2$	$V_2 = V_1 - Q_2$
...					
k	V_{k-1}	$d_k = V_{k-1} \cdot i$	Q_k	$T_k = d_k + Q_k$	$V_k = V_{k-1} - Q_k$
...					
n	V_{n-1}	$d_n = V_{n-1} \cdot i$	Q_n	$T_n = d_n + Q_n$	$V_n = V_{n-1} - Q_n = 0$

Obviously, we have:

$$V_0 = Q_1 + Q_2 + \cdots + Q_n,\ V_{n-1} = Q_n,$$

$$T_n = Q_n \cdot u,\ T_{k+1} - T_k = Q_{k+1} - Q_k \cdot u,$$

where $u = 1 + i$ is the the annual compounding factor.

We have two mainly procedures to calculate the payments of a loan:

1. **The fixed-principal amortization**: $Q_1 = Q_2 = \cdots = Q_n = Q$.

   In this case, we have:
   $$Q = \frac{V_0}{n};$$
   $$T_{k+1} - T_k = -Q \cdot i$$
   (arithmetic progression);
   $$T_k = Q[1 + (n - k + 1)i].$$

2. **The fixed-rate amortization**: $T_1 = T_2 = \cdots = T_n = T$.

   In this case, we have:
   $$T = V_0 \cdot \frac{i}{1 - v^n}$$
   (**the fixed-rate formula**);
   $$Q_{k+1} = Q_k \cdot u$$
   (geometric progression);
   $$Q_k = V_0 \cdot \frac{i}{u^n - 1} \cdot u^{k-1},$$
   where $u = 1 + i$ is the the annual compounding factor, and $v = \dfrac{1}{u} = \dfrac{1}{1+i}$ is the annual discounting factor.

Remark 5.5.1. ***The inflation*** *changes the purchasing power of money. After n years,* ***the purchasing power*** *of S_n u.c. is reduced to*

$$S_0 = \frac{S_n}{(1 + a_1)(1 + a_2)\ldots(1 + a_n)},$$

where $a_1, a_2, \ldots, a_n$ are the annual inflation rates. S_n is measured in future units of currency, and S_0 is measured in today's units of currency.

5.6 Problems

Exercise 5.6.1. A person deposits 1000 u.c. on 20 February 2017 at an annual interest percent of 8%. Calculate the amount of this investment on 10 November 2017 in each of the following cases:
a) exact interest;
b) banker's rule;
c) ordinary interest.

Exercise 5.6.2. A person deposits 1000 u.c. for 3 years and seven months at an annual interest percent of 8%. Calculate the final value of this investment in each of the following cases:
a) simple interest;
b) compound interest, the rational procedure;
c) compound interest, the commercial procedure;
d) compounded monthly interest.

Exercise 5.6.3. Consider the following investments: 1000 u.c. for one year at 8% per year, 800 u.c. for 9 months at 10% per year, and 1200 u.c. for 10 months at 9% per year. Calculate the initial value, the time of investment and the annual interest rate meanly replacements by simple interest.

Exercise 5.6.4. A person deposits 100 u.c. at the end of every month for 5 years, at successive annual interest percents of 10%, 10%, 7%, 8%, 8%. Calculate the amount of this investment at the end of 5 years.

Exercise 5.6.5. Construct the amortization table for a loan of 2400 u.c. per 4 years at an annual interest percent of 9%, in each of the following cases:
a) fixed-principal annually amortization;
b) fixed-rate annually amortization;
c) fixed-principal monthly amortization;
d) fixed-rate monthly amortization.
Compare the obtained results when the annual successive inflation rates are 4%, 6%, 5%, 6%.

Theme 6

Introduction to Actuarial Math

6.1 A general model of insurance

In an insurance model **the insurer** agrees to pay **the insured** one or more amounts called **claims (claim payments)**, at fixed times or when **the insured event** occurs. In return of these claims, the insured pays one or more amounts called **premiums**.

Usually the insure events are **random events**.

For a mutually advantageous insurance, the present values (at the initial moment of the insurance) of the premiums need to be equal to the present values of the claims. These values are also called **actuarial present values**.

Definition 6.1.1. *For a given insurance, **the single premium** payable at the initial moment of the insurance is*

$$P = E(X),$$

where $E(X)$ denotes the mean of the random variable X that represents the present value of the claim.

Theorem 6.1.1. *Let A be an insurance consisting in the partial insurances $A_1, A_2, \ldots, A_n$ ($n \in \mathbb{N}^*$), and let $P_1, P_2, \ldots, P_n$ be the single premiums corresponding of these partial insurances. Then the single premium of the total insurance A is*

$$P = P_1 + P_2 + \cdots + P_n.$$

Proof. Let X be the random variable that represents the present value of the total insurance A and let $X_1, X_2, \ldots, X_n$ be the random variables representing the present values of partial insurances $A_1, A_2, \ldots, A_n$, respectively. We have

$$X = X_1 + X_2 + \cdots + X_n,$$

and hence

$$
\begin{aligned}
P &= E(X) = E(X_1 + X_2 + \cdots + X_n) = E(X_1) + e(X_2) + \cdots + E(X_n) \\
&= P_1 + P_2 + \cdots + P_n.
\end{aligned}
$$

□

6.2 Biometric functions

The mortality is the most important factor in the insurances of persons. The frequency of mortality for a population is measured by some statistical function of age called **biometric functions**. We assume that the age is measured in years.

Definition 6.2.1. *We denote by l_0 the total number of persons of the analyzed population (the number of newborns).*

Remark 6.2.1. *Usually, $l_0 = 100000$.*

Remark 6.2.2. *l_0 represents the number of survivors to age 0 (from the analyzed population).*

6.2.1 Probabilities of life and death

Definition 6.2.2. *Let $x, n, m \in \mathbb{N}$. We denote*

- *p_x = the probability that a person of age x will live at least one year;*
- *q_x = the probability that a person of age x will die within one year,*
- *${}_np_x$ = the probability that a person of age x will attain age $x+n$;*
- *${}_nq_x$ = the probability that a person of age x will die until the age $x+n$;*
- *${}_{m|n}q_x$ = the probability that a person of age x will attain age $x+m$ but die until the age $x+m+n$.*

*p_x is called **probability of life** for age x, and q_x is called **probability of death** for age x.*

Proposition 6.2.1. *Let $x, y, z \in \mathbb{N}$, $x \le y \le z$. We have:*

$$q_x = 1 - p_x; \tag{6.1}$$

$$ {}_nq_x = 1 - {}_np_x; \tag{6.2}$$

$$ {}_0p_x = 1;\ {}_0q_x = 0; \tag{6.3} $$

$$ {}_1p_x = p_x;\ {}_1q_x = q_x; \tag{6.4} $$

$$ {}_{0|n}q_x = {}_nq_x; \tag{6.5} $$

$$ {}_{n+m}p_x = {}_np_x \cdot {}_mp_{x+n}; \tag{6.6} $$

$$ {}_np_x = p_x \cdot p_{x+1} \cdot \ldots \cdot p_{x+n-1}; \tag{6.7} $$

$$ {}_nq_x = q_x + p_x \cdot q_{x+1} + p_x \cdot p_{x+1} \cdot q_{x+2} + \ldots + p_x \cdot p_{x+1} \cdot \ldots \cdot p_{x+n-2} \cdot q_{x+n-1}; \tag{6.8} $$

$$ {}_{m|n}q_x = {}_mp_x \cdot {}_nq_{x+m}; \tag{6.9} $$

$$ {}_{m|n}q_x = {}_{m+n}q_x - {}_mq_x = {}_mp_x - {}_{m+n}p_x. \tag{6.10} $$

Proof. Equalities (6.1), (6.2), (6.3), (6.4) and (6.5) are obvious.

Denote by $A(x, y)$ the event that a person of age x will attains age y. Then

$$ A(x, x+n+m) = A(x, x+n) \cap A(x+n, x+n+m), $$

and the events $A(x, x+n)$ and $A(x+n, x+n+m)$ are independent. Hence

$$ {}_{n+m}p_x = P(A(x, x+n+m)) = P(A(x, x+n))P(A(x+n, x+n+m)) = {}_np_x \cdot {}_mp_{x+n} $$

(where $P(A)$ represents the probability of event A). Using (6.6) and (6.4) we have

$$ {}_np_x = {}_1p_x \cdot {}_1p_{x+1} \cdot \ldots \cdot {}_1p_{x+n-1} = p_x \cdot p_{x+1} \cdot \ldots \cdot p_{x+n-1}. $$

Also, we have

$$ \begin{aligned} {}_nq_x &= P(\overline{A(x, x+n)}) \\ &= P\Big(\overline{A(x, x+1)} \cup A(x, x+1) \cap \overline{A(x+1, x+2)} \cup A(x, x+2) \cap \overline{A(x+2, x+3)} \cup \ldots \\ &\quad \cup A(x, x+n-1) \cap \overline{A(x+n-1, x+n)}\Big) \\ &= q_x + p_x \cdot q_{x+1} + p_x \cdot p_{x+1} \cdot q_{x+2} + \ldots + p_x \cdot p_{x+1} \cdot \ldots \cdot p_{x+n-2} \cdot q_{x+n-1} \end{aligned} $$

(where $\overline{A}$ represents the complementary event of A). We have

$$ \begin{aligned} {}_{m|n}q_x &= P(A(x, x+m) \cap \overline{A(x+m, x+m+n)}) \\ &= P(A(x, x+m))P(\overline{A(x+m, x+m+n)}) \\ &= {}_mp_x \cdot {}_nq_{x+m}; \end{aligned} $$

$$ {}_{m|n}q_x = P(\overline{A(x, x+m+n)} \cap A(x, x+m)) $$

$$= P(\overline{A(x, x+m+n)} \setminus \overline{A(x, x+m)})$$
$$= P(\overline{A(x, x+m+n)}) - P(\overline{A(x, x+m)}) = {}_{m+n}q_x - {}_m q_x$$
$$= (1 - {}_{m+n}p_x) - (1 - {}_m p_x) = {}_m p_x - {}_{m+n}p_x.$$

□

6.2.2 The survival function

Definition 6.2.3. *Let $x \in \mathbb{N}$. We denote*

- $l_x =$ ***the expected number of survivors at age*** x *(from the analyzed population).*

Proposition 6.2.2. *For any $x \in \mathbb{N}$ we have*

$$l_x = l_0 p(0, x). \tag{6.11}$$

Proof. Obviously,

$$l_x = E(X),$$

where X is the random variable that represents the number of survivors at age x. Let

$$X : \begin{pmatrix} 0 & \dots & n & \dots & l_0 \\ \alpha_x(0) & \dots & \alpha_x(n) & \dots & \alpha_x(l_0) \end{pmatrix}$$

be the distribution of X, where, for any $n \in \{0, \dots, l_0\}$, $\alpha_x(n)$ denotes the probability that the number of survivors at age x is equal to n. We have

$$\alpha_x(n) = C_{l_0}^n ({}_x p_0)^n ({}_x q_0)^{l_0 - n}, \ \forall n \in \{0, \dots, l_0\}.$$

Then X has a binomial distribution of parameters l_0 and $p(0, x)$. Therefore

$$l_x = E(X) = l_0 \cdot {}_x p_0.$$

□

Definition 6.2.4. *Let $x \in \mathbb{N}$. We denote*

- $s(x) = {}_x p_0 =$ *the probability that a newborn will live to at least x.*

$s(x)$ is called ***the survival function*** *for age x.*

Remark 6.2.3. *${}_x q_0 = 1 - {}_x p_0 = 1 - s(x)$ represents the probability that a newborn will die until the age x.*

Proposition 6.2.3. *Let $x, n, m \in \mathbb{N}$. We have:*

$$l_x = l_0 \cdot p_0 \cdot p_1 \cdot \ldots \cdot p_{x-1}; \tag{6.12}$$

$$_np_x = \frac{l_{x+n}}{l_x};\ _nq_x = \frac{l_x - l_{x+n}}{l_x}; \tag{6.13}$$

$$_{m|n}q_x = \frac{l_{x+m} - l_{x+m+n}}{l_x}; \tag{6.14}$$

$$p_x = \frac{l_{x+1}}{l_x};\ q_x = \frac{d_x}{l_x}, \tag{6.15}$$

where

$$d_x = l_x - l_{x+1}. \tag{6.16}$$

Proof. Equation (6.12) is an immediate consequence of (6.11) and (6.7). Using (6.6) and (6.11) we have

$$_np_x = \frac{_{x+n}p_0}{_xp_0} = \frac{l_{x+n}}{l_0} \cdot \frac{l_0}{l_x} = \frac{l_{x+n}}{l_x}, \text{ and } _nq_x = 1 - {}_np_x = \frac{l_x - l_{x+n}}{l_x}.$$

Using (6.9) and (6.12) we have

$$_{m|n}q_x = {}_mp_x \cdot {}_nq_{x+m} = \frac{l_{x+m}}{l_x} \cdot \frac{l_{x+m} - l_{x+m+n}}{l_{x+m}} = \frac{l_{x+m} - l_{x+m+n}}{l_x}.$$

Taking $n = 1$ into (6.13) we obtain equalities (6.15). □

Remark 6.2.4. *d_x represents* ***the expected number of deaths at age*** *x (i.e. at exactly age x or between ages x and $x + 1$).*

Remark 6.2.5. *There exists an age $\omega \in \mathbb{N}$ such that*

$$l_\omega > 0 \text{ and } l_x = 0\ \forall x > \omega.$$

Definition 6.2.5. *The value ω from the above remark is called* ***the limiting age****.*

Remark 6.2.6. *Usually, $\omega = 100$.*

6.2.3 The life expectancy

Definition 6.2.6. *Let $x \in \mathbb{N}$, $x \leq \omega$. We denote*

- $\overset{\circ}{e}_x$ = *the expected future lifetime for a person of age x (prior to death).*

*$\overset{\circ}{e}_x$ is called **the average remaining lifetime** for age x, and $x+\overset{\circ}{e}_x$ is called **the life expectancy** for age x.*

Remark 6.2.7. *We assume that the deaths are uniform distributed throughout the year.*

Proposition 6.2.4. *For any $x \in \mathbb{N}$, $x \leq \omega$, we have*

$$\overset{\circ}{e}_x = \frac{1}{2} + \frac{1}{l_x} \sum_{n=1}^{\omega - x} l_{x+n}. \tag{6.17}$$

Proof. Obviously,

$$\overset{\circ}{e}_x = E(Y),$$

where Y is the random variable that represents the future lifetime for a person of age x. Let

$$Y : \begin{pmatrix} \frac{1}{2} & \dots & n+\frac{1}{2} & \dots & \omega - x + \frac{1}{2} \\ \beta_x(0) & \dots & \beta_x(n) & \dots & \beta_x(\omega - x) \end{pmatrix}$$

be the distribution of Y, where, for any $n \in \{0, \dots, \omega - x\}$, $\beta_x(n)$ represents the probability that a person of age x will live only n (i.e. will die at age $x+n$). We have

$$\beta_x(n) = {}_{n|1}q_x = {}_n p_x \cdot q_{x+n}, \ \forall n \in \{0, \dots, \omega - x\},$$

Using (6.13) and (6.16) it follows that

$$\beta_x(n) = \frac{l_{x+n}}{l_x} \cdot \frac{d_{x+n}}{l_{x+n}} = \frac{d_{x+n}}{l_x} = \frac{l_{x+n} - l_{x+n+1}}{l_x}, \ \forall n \in \{0, \dots, \omega - x\}. \tag{6.18}$$

Hence

$$\begin{aligned}
\overset{\circ}{e}_x = E(Y) &= \sum_{n=0}^{\omega-x} \left(n + \frac{1}{2}\right) \frac{l_{x+n} - l_{x+n+1}}{l_x} \\
&= \frac{1}{l_x} \sum_{n=0}^{\omega-x} \left[\left(n + \frac{1}{2}\right) l_{x+n} - \left(n + 1 + \frac{1}{2}\right) l_{x+n+1} + l_{x+n+1} \right] \\
&= \frac{1}{l_x} \left[\frac{1}{2} l_x - \left(\omega - x + 1 + \frac{1}{2}\right) l_{\omega+1} + \sum_{n=0}^{\omega-x} l_{x+n+1} \right] = \frac{1}{2} + \frac{1}{l_x} \sum_{n=1}^{\omega-x} l_{x+n},
\end{aligned}$$

since $l_{\omega+1} = 0$. $\square$

Remark 6.2.8. *According to (6.17) and (6.13) we obtain:*

$$\overset{\circ}{e}_x = \frac{1}{2} + \sum_{n=1}^{\omega-x} {}_n p_x, \ \forall x \in \mathbb{N}, \ x \leq \omega. \tag{6.19}$$

6.2.4 Life tables

The value of biometric functions are tabulated in **life table** (**mortality table** or **actuarial table**) of the following form:

Age x	Nr. of survivors l_x	Nr. of deaths d_x	Probab. of death q_x	Average rem. lifetime $\mathring{e}_x$
0	$l_0 = 100000$			
1				
$\vdots$				
$\omega = 100$				

Usually, the values l_x are derived by a census. The values d_x, q_x and $\mathring{e}_x$ are calculated according to (6.16), (6.15) and (6.17), respectively.

The following actuarial table shows the life expectancy for the Romanian population (**www.pensiileprivate.ro**).

x	l_x MALE	l_x FEMALE	q_x MALE	q_x FEMALE	$x + \mathring{e}_x$ MALE	$x + \mathring{e}_x$ FEMALE
18	100000	100000	0.0007	0.0004	66.7	73
19	99930	99960	0.0007	0.0004	66.7	73
20	99860	99920	0.001	0.0004	66.8	73
21	99760	99880	0.0011	0.0004	66.8	73
22	99654	99838	0.0011	0.0004	66.9	73
23	99543	99794	0.0012	0.0005	66.9	73.1
24	99425	99748	0.0012	0.0005	67	73.1
25	99302	99700	0.0013	0.0005	67	73.1
26	99173	99651	0.0014	0.0005	67.1	73.1
27	99032	99597	0.0015	0.0006	67.1	73.2
28	98880	99539	0.0017	0.0006	67.2	73.2
29	98716	99477	0.0018	0.0007	67.3	73.2
30	98540	99412	0.0019	0.0007	67.3	73.2
31	98353	99342	0.0021	0.0008	67.4	73.3
32	98144	99261	0.0023	0.0009	67.5	73.3
33	97914	99167	0.0026	0.0011	67.6	73.3
34	97664	99062	0.0028	0.0012	67.7	73.4
35	97392	98945	0.003	0.0013	67.7	73.4
36	97100	98817	0.0036	0.0015	67.8	73.5
37	96754	98670	0.0041	0.0017	68	73.5
38	96356	98507	0.0047	0.0018	68.1	73.6
39	95905	98325	0.0052	0.002	68.2	73.7

x	l_x MALE	l_x FEMALE	q_x MALE	q_x FEMALE	$x + \overset{\circ}{e}_x$ MALE	$x + \overset{\circ}{e}_x$ FEMALE
40	95402	98127	0.0058	0.0022	68.4	73.7
41	94849	97911	0.0065	0.0025	68.5	73.8
42	94231	97670	0.0072	0.0027	68.7	73.9
43	93548	97404	0.008	0.003	68.9	74
44	92804	97114	0.0087	0.0032	69.1	74.1
45	91998	96799	0.0094	0.0035	69.3	74.2
46	91133	96461	0.0101	0.0039	69.5	74.3
47	90209	96088	0.0109	0.0042	69.8	74.4
48	89228	95683	0.0116	0.0046	70	74.5
49	88191	95245	0.0124	0.0049	70.3	74.6
50	87101	94774	0.0131	0.0053	70.5	74.7
51	85960	94272	0.0143	0.0059	70.8	74.8
52	84731	93717	0.0155	0.0065	71.1	75
53	83417	93112	0.0167	0.007	71.4	75.1
54	82024	92456	0.0179	0.0076	71.7	75.3
55	80556	91752	0.0191	0.0082	72	75.4
56	79017	91000	0.0208	0.009	72.3	75.6
57	77374	90182	0.0225	0.0098	72.7	75.8
58	75633	89302	0.0242	0.0105	73	76
59	73803	88361	0.0259	0.0113	73.4	76.2
60	71891	87361	0.0276	0.0121	73.7	76.3
61	69907	86304	0.0299	0.0137	74.1	76.5
62	67814	85125	0.0323	0.0152	74.5	76.7
63	65625	83829	0.0346	0.0168	74.9	77
64	63353	82423	0.037	0.0183	75.3	77.2
65	61011	80911	0.0393	0.0199	75.7	77.4
66	58614	79301	0.0427	0.0228	76.1	77.7
67	56111	77493	0.0461	0.0257	76.6	77.9
68	53524	75501	0.0495	0.0286	77	78.2
69	50875	73342	0.0529	0.0315	77.4	78.5
70	48183	71032	0.0563	0.0344	77.9	78.8
71	45471	68588	0.0682	0.0474	78.3	79.1
72	42371	65336	0.08	0.0604	78.8	79.5
73	38981	61387	0.0919	0.0735	79.4	79.9
74	35399	56877	0.1037	0.0865	80	80.4
75	31727	51959	0.1156	0.0995	80.6	81
76	28059	46789	0.1275	0.1125	81.3	81.6
77	24483	41524	0.1393	0.1255	82	82.2
78	21072	36311	0.1512	0.1386	82.7	82.9
79	17886	31280	0.163	0.1516	83.4	83.6

x	l_x MALE	l_x FEMALE	q_x MALE	q_x FEMALE	$x+\mathring{e}_x$ MALE	$x+\mathring{e}_x$ FEMALE
80	14970	26538	0.1749	0.1646	84.2	84.3
81	12352	22170	0.1868	0.1776	85	85.1
82	10045	18232	0.1986	0.1906	85.8	85.9
83	8050	14757	0.2105	0.2037	86.6	86.7
84	6356	11751	0.2223	0.2167	87.4	87.5
85	4942	9205	0.2342	0.2297	88.3	88.3
86	3785	7091	0.2461	0.2427	89.1	89.2
87	2854	5370	0.2579	0.2557	90	90
88	2118	3996	0.2698	0.2688	90.9	90.9
89	1546	2922	0.2816	0.2818	91.7	91.7
90	1111	2099	0.2935	0.2948	92.6	92.6
91	785	1480	0.3054	0.3078	93.5	93.5
92	545	1024	0.3172	0.3208	94.4	94.4
93	372	696	0.3291	0.3339	95.3	95.2
94	250	463	0.3409	0.3469	96.2	96.1
95	165	303	0.3528	0.3599	97	97
96	107	194	0.3647	0.3729	97.8	97.8
97	68	122	0.3765	0.3859	98.6	98.6
98	42	75	0.3884	0.399	99.3	99.3
99	26	45	0.4002	0.412	99.8	99.8
100	15	26	1	1	101.8	101.9

Example 6.2.1. Compute the average remaining lifetime and the life expectancy for a 30 years old male.

Solution. The average remaining lifetime is

$$\begin{aligned}\mathring{e}_{30} &= \frac{1}{2} + \frac{1}{l_{30}}\sum_{n=1}^{70} l_{30+n} \\ &= \frac{1}{2} + \frac{l_{31}+l_{32}+\cdots+l_{100}}{l_{30}} \\ &= \frac{1}{2} + \frac{98353+98144+\cdots+15}{98540} \\ &= \frac{1}{2} + \frac{3629652}{98540} \\ &\simeq 0.5+36.8 = 37.3\,.\end{aligned}$$

The life expectancy is $30+\mathring{e}_{30} \simeq 67.3\,.$ □

6.3 Problems

Exercise 6.3.1. Calculate the probability that a 30 years old person will live
a) at least 40 years;
b) exactly 45 years;
c) at most 50 years;
d) at least 35 years but at most 55 years.

Exercise 6.3.2. Consider a family of a 45 years old husband and a 43 years old wife.
a) Calculate the probability that both spouses will die in the same year.
b) Calculate the probability that both spouses will die at the same age.

Exercise 6.3.3. Calculate the average remaining lifetime and the life expectancy for a 50 years old person.

Exercise 6.3.4. Calculate the probability that a 60 years old person will die before the integer number of years of his average remaining lifetime.

Exercise 6.3.5. For a 35 years old person, calculate the life expectancy and the age of death having the maximum probability.

Theme 7

Life annuities

7.1 A general model. Classifications

In a person insurance, the claims are payments while the insured survives. We have the following **classifications**.

1. By **period**, the claims can be:
   - **annuities**;
   - **semiannual**;
   - **quarterly**;
   - **monthly**.
2. By **amount**, the claims can be:
   - **constants**;
   - **variables**.
3. By **time of payment**, the claims can be:
   - **annuity-due**, when the claims are payed at the beginning of each period;
   - **annuity-immediate**, when the claims are payed at the end of each period.
4. By **time of first payment**, the claims can be:
   - **immediate**;
   - **deferred**.

5. By **number of payments**, the claims can be:
   - **single**, when the claim is payed at a fixed time, only if the insured will live at this time;
   - **temporary (limited)**, when the claim is payed at fixed times, while the insured survives;
   - **unlimited**, when the claim is payed whole life.

7.2 Single claim

Definition 7.2.1. *Let $x, n \in \mathbb{N}$ s.t. $x + n \leq \omega$. We denote*

- ${}_nE_x$ *= the single premium payable by a person of age x for a single claim of 1 u.c. over n years if the person survives.*

Remark 7.2.1. ${}_nE_x$ *is called* ***the unitary premium****.*

Proposition 7.2.1. *For any $x, n \in \mathbb{N}$ s.t. $x + n \leq \omega$, we have*

$$ {}_nE_x = \frac{D_{x+n}}{D_x}, \tag{7.1} $$

where

$$ D_x = v^x \cdot l_x, \tag{7.2} $$

$v = \dfrac{1}{1+i}$ *being the annual discounting factor, i being the annual interest rate.*

Proof. For a mutually advantageous insurance, the single premium ${}_nE_x$ need to be equal to the present value of the single claim, that is

$$ {}_nE_x = E(X), $$

where X is the random variable that represents the present value of the claim.
We have

$$ X = \begin{cases} v^n, & \text{if the insurer survives at least } n \text{ years from the time of insurance issue,} \\ 0, & \text{otherwise.} \end{cases} $$

Hence the distribution of X is

$$ X : \begin{pmatrix} v^n & 0 \\ {}_np_x & {}_nq_x \end{pmatrix}. $$

By (6.13) and (7.2) we have

$$_nE_x = E(X) = v^n \cdot {}_np_x + 0 \cdot {}_nq_x = v^n \cdot \frac{l_{x+n}}{l_x} = \frac{v^{x+n} \cdot l_{x+n}}{v^x \cdot l_x} = \frac{D_{x+n}}{D_x}.$$

□

Corollary 7.2.1. *Let $x, n \in \mathbb{N}$, $x + n \leq \omega$, $T \geq 0$. The single premium payable by a person of age x for a single claim of T u.c. over n years if the person survives is*

$$T \cdot {}_nE_x = T \cdot \frac{D_{x+n}}{D_x}.$$

Definition 7.2.2. *D_x defined by (7.2) is called **the commutation number**. The single premium $_nE_x$ defined by (7.1) is called **the life discounting factor**.*

Remark 7.2.2. *By (7.1) it follows that*

$$_{y+z}E_x = {}_yE_x \cdot {}_zE_{x+y},\ \forall x, y, z \in \mathbb{N} \text{ s.t. } x + y \leq \omega. \tag{7.3}$$

Example 7.2.1. Compute the single premium payable by a 44 years old female for a single claim of 25000 u.c. over 18 years if she survives. The annual interest percent is 10%.

Solution. The premium is

$$25000 \cdot {}_{18}E_{44} = 25000 \cdot \frac{D_{62}}{D_{44}} = 25000 \cdot \frac{v^{62} \cdot l_{62}}{v^{44} \cdot l_{44}} = 25000 \cdot \frac{\left(\frac{1}{1.1}\right)^{62} \cdot 85125}{\left(\frac{1}{1.1}\right)^{44} \cdot 97114}$$

$$\simeq 25000 \cdot \frac{231.0524857111}{1465.5603118440} \simeq 3941.36 \text{ u.c.}$$

□

7.3 Life annuities-immediate

7.3.1 Whole life annuities

Definition 7.3.1. *Let $x \in \mathbb{N}$, $x \leq \omega$. We denote*

- a_x = *the single premium payable by a person of age x for a whole life annuity-immediate of 1 u.c. per year.*

Proposition 7.3.1. *For any $x \in \mathbb{N}$, $x \le \omega$, we have*

$$a_x = \frac{N_{x+1}}{D_x}, \tag{7.4}$$

where

$$N_x = D_x + D_{x+1} + \cdots + D_\omega. \tag{7.5}$$

Proof. By Theorem 6.1.1 we have

$$a_x = {}_1E_x + {}_2E_x + \cdots + {}_{\omega-x}E_x.$$

Using (7.1) and (7.5) we obtain

$$a_x = \frac{D_{x+1}}{D_x} + \frac{D_{x+2}}{D_x} + \cdots + \frac{D_\omega}{D_x} = \frac{N_{x+1}}{D_x}.$$

□

Corollary 7.3.1. *Let $x \in \mathbb{N}$, $x \le \omega$, $T \ge 0$. The single premium payable by a person of age x for a whole life annuity-immediate of T u.c. per year is*

$$T \cdot a_x = T \cdot \frac{N_{x+1}}{D_x}.$$

Definition 7.3.2. *N_x defined by (7.5) is called **the cumulative commutation number**.*

Example 7.3.1. Compute the single premium payable by a 49 years old male for a whole life annuity-immediate of 2000 u.c. per year. The annual interest percent is 10%.

Solution. The premium is

$$\begin{aligned}
2000 \cdot a_{49} &= 2000 \cdot \frac{N_{50}}{D_{49}} = 2000 \cdot \frac{D_{50} + D_{51} + \cdots + D_{100}}{D_{49}} \\
&= 2000 \cdot \frac{v^{50} \cdot l_{50} + v^{51} \cdot l_{51} + \cdots + v^{100} \cdot l_{100}}{v^{49} \cdot l_{49}} \\
&= 2000 \cdot \frac{\left(\frac{1}{1.1}\right)^{50} \cdot 87101 + \left(\frac{1}{1.1}\right)^{51} \cdot 85960 + \cdots + \left(\frac{1}{1.1}\right)^{100} \cdot 15}{\left(\frac{1}{1.1}\right)^{49} \cdot 88191} \\
&\simeq 15670.88 \text{ u.c.}
\end{aligned}$$

□

7.3.2 Deferred whole life annuities

Definition 7.3.3. *Let $x, r \in \mathbb{N}$ s.t. $x + r \le \omega$. We denote*

- ${}_{r|}a_x$ = *the single premium payable by a person of age x for an r-year deferred whole life annuity-immediate of* 1 *u.c. per year (payable at the end of each year while the person survives from age $x + r$ onward).*

Proposition 7.3.2. *For any $x, r \in \mathbb{N}$ s.t. $x + r \le \omega$, we have*

$$ {}_{r|}a_x = \frac{N_{x+r+1}}{D_x}. \tag{7.6} $$

Proof. By Theorem 6.1.1 we have

$$ {}_{r|}a_x = {}_{r+1}E_x + {}_{r+2}E_x + \cdots + {}_{\omega-x}E_x. $$

Using (7.1) and (7.5) we obtain

$$ {}_{r|}a_x = \frac{D_{x+r+1}}{D_x} + \frac{D_{x+r+2}}{D_x} + \cdots + \frac{D_\omega}{D_x} = \frac{N_{x+r+1}}{D_x}. $$

□

Corollary 7.3.2. *Let $x, r \in \mathbb{N}$, $x + r \le \omega$, $T \ge 0$. The single premium payable by a person of age x for an r-year deferred whole life annuity-immediate of T u.c. per year is*

$$ T \cdot {}_{r|}a_x = T \cdot \frac{N_{x+r+1}}{D_x}. $$

Remark 7.3.1. *By (7.6), (7.1) and (7.4) it follows that*

$$ \begin{aligned} {}_{r|}a_x &= {}_rE_x \cdot a_{x+r},\ \forall x, r \in \mathbb{N} \text{ s.t. } x + r \le \omega, \qquad (7.7)\\ {}_{0|}a_x &= a_x,\ \forall x \in \mathbb{N},\ x \le \omega,\\ {}_{\omega-x|}a_x &= 0,\ \forall x \in \mathbb{N},\ x \le \omega. \end{aligned} $$

Example 7.3.2. Compute the single premium payable by a 43 years old female for a 17-year deferred whole life annuity-immediate of 500 u.c. per year. The annual interest percent is 10%.

Solution. The premium is

$$ 500 \cdot {}_{17|}a_{43} = 500 \cdot \frac{N_{61}}{D_{43}} = 500 \cdot \frac{D_{61} + D_{62} + \cdots + D_{100}}{D_{43}} $$

$$= 500 \cdot \frac{v^{61} \cdot l_{61} + v^{62} \cdot l_{62} + \cdots + v^{100} \cdot l_{100}}{v^{43} \cdot l_{43}}$$

$$= 500 \cdot \frac{\left(\frac{1}{1.1}\right)^{61} \cdot 86304 + \left(\frac{1}{1.1}\right)^{62} \cdot 85125 + \cdots + \left(\frac{1}{1.1}\right)^{100} \cdot 26}{\left(\frac{1}{1.1}\right)^{43} \cdot 97404}$$

$$\simeq 641.16 \text{ u.c.}$$

□

7.4 Temporary life annuities

Definition 7.4.1. *Let $x, r \in \mathbb{N}$ s.t. $x + r \leq \omega$. We denote*

- $a_{x:\,\overline{r}|}$ = *the single premium payable by a person of age x for an r-year temporary life annuity-immediate of 1 u.c. per year (payable at the end of each year while the person survives during the next r years).*

Proposition 7.4.1. *For any $x, r \in \mathbb{N}$ s.t. $x + r \leq \omega$, we have*

$$a_{x:\,\overline{r}|} = \frac{N_{x+1} - N_{x+r+1}}{D_x}. \tag{7.8}$$

Proof. By Theorem 6.1.1 we have

$$a_{x:\,\overline{r}|} = {}_1E_x + {}_2E_x + \cdots + {}_rE_x.$$

By (7.1) and (7.5) we obtain

$$a_{x:\,\overline{r}|} = \frac{D_{x+1}}{D_x} + \frac{D_{x+2}}{D_x} + \cdots + \frac{D_{x+r}}{D_x} = \frac{N_{x+1} - N_{x+r+1}}{D_x}.$$

□

Corollary 7.4.1. *Let $x, r \in \mathbb{N}$, $x + r \leq \omega$, $T \geq 0$. The single premium payable by a person of age x for an r-year temporary life annuity-immediate of T u.c. per year is*

$$T \cdot a_{x:\,\overline{r}|} = T \cdot \frac{N_{x+1} - N_{x+r+1}}{D_x}.$$

Remark 7.4.1. *By (7.4), (7.8), (7.6) and (7.7) it follows that*

$$\begin{aligned} a_x &= a_{x:\,\overline{r|}} + {}_{r|}a_x,\ \forall x, r \in \mathbb{N}\ s.t.\ x+r \le \omega, \\ a_{x:\,\overline{r|}} &= a_x - {}_rE_x \cdot a_{x+r},\ \forall x, r \in \mathbb{N}\ s.t.\ x+r \le \omega, \\ a_{x:\,\overline{0|}} &= 0,\ \forall x \in \mathbb{N},\ x \le \omega, \\ a_{x:\,\overline{\omega-x|}} &= a_x,\ \forall x \in \mathbb{N},\ x \le \omega. \end{aligned} \tag{7.9}$$

Example 7.4.1. Compute the single premium payable by a 43 years old female for a 17-year temporary life annuity-immediate of 500 u.c. per year. The annual interest percent is 10%.

Solution. The premium is

$$\begin{aligned} 500 \cdot a_{43:\,\overline{17|}} &= 500 \cdot \frac{N_{44} - N_{61}}{D_{43}} = 500 \cdot \frac{D_{44} + D_{45} + \cdots + D_{60}}{D_{43}} \\ &= 500 \cdot \frac{v^{44} \cdot l_{44} + v^{45} \cdot l_{45} + \cdots + v^{60} \cdot l_{60}}{v^{43} \cdot l_{43}} \\ &= 500 \cdot \frac{\left(\frac{1}{1.1}\right)^{44} \cdot 97114 + \left(\frac{1}{1.1}\right)^{45} \cdot 96799 + \cdots + \left(\frac{1}{1.1}\right)^{60} \cdot 87361}{\left(\frac{1}{1.1}\right)^{43} \cdot 97404} \\ &\simeq 3888.22 \text{ u.c.} \end{aligned}$$

□

7.5 Life annuities-immediate with k-thly payments

In this case the claims are payable at the end of each k-th period of the year.

7.5.1 Whole life annuities with k-thly payments

Definition 7.5.1. *Let $x \in \mathbb{N}$, $x \le \omega$ and $k \in \mathbb{N}^*$. We denote*

- $a_x^{(k)}$ = *the single premium payable by a person of age x for a whole life annuity-immediate of $\frac{1}{k}$ u.c. per each k-th period of the year (i.e. 1 u.c. per each year).*

Definition 7.5.2. *For any $x \in \mathbb{N}$, $x \leq \omega$ and $k \in \mathbb{N}^*$, $k \geq 2$, we define* ***the intermediate commutation numbers*** $D_{x+\frac{1}{k}}, D_{x+\frac{2}{k}}, \ldots, D_{x+\frac{k-1}{k}}$ *such that* $D_x, D_{x+\frac{1}{k}}, D_{x+\frac{2}{k}}, \ldots, D_{x+\frac{k-1}{k}}, D_{x+1}$ *is a arithmetic progression.*

Lemma 7.5.1. *For any $x \in \mathbb{N}$, $x \leq \omega$, $k \in \mathbb{N}^*$ and $h \in \{0, 1, \ldots, k\}$ we have*

$$D_{x+\frac{h}{k}} = \frac{k-h}{k} \cdot D_x + \frac{h}{k} \cdot D_{x+1}. \tag{7.10}$$

Proof. The arithmetic progression $D_x, D_{x+\frac{1}{k}}, D_{x+\frac{2}{k}}, \ldots, D_{x+\frac{k-1}{k}}, D_{x+1}$ has $k+1$ terms, so its ratio is $\dfrac{D_{x+1} - D_x}{k}$ and hence its $(h+1)$-th term is

$$D_{x+\frac{h}{k}} = D_x + h \cdot \frac{D_{x+1} - D_x}{k} = \frac{k-h}{k} \cdot D_x + \frac{h}{k} \cdot D_{x+1}.$$

□

Proposition 7.5.1. *For any $x \in \mathbb{N}$, $x \leq \omega$ and $k \in \mathbb{N}^*$ we have*

$$a_x^{(k)} = \frac{N_{x+1}}{D_x} + \frac{k-1}{2k}. \tag{7.11}$$

Proof. By Theorem 6.1.1 we have

$$a_x^{(k)} = \frac{1}{k} \sum_{n=0}^{\omega-x} \sum_{h=1}^{k} {}_{n+\frac{h}{k}}E_x. \tag{7.12}$$

By (7.1), (7.10) and (7.5) we obtain

$$\begin{aligned}
a_x^{(k)} &= \frac{1}{k} \sum_{n=0}^{\omega-x} \sum_{h=1}^{k} \frac{D_{x+n+\frac{h}{k}}}{D_x} = \frac{1}{k \cdot D_x} \sum_{h=1}^{k} \sum_{n=0}^{\omega-x} \left(\frac{k-h}{k} \cdot D_{x+n} + \frac{h}{k} \cdot D_{x+n+1} \right) \\
&= \frac{1}{k \cdot D_x} \sum_{h=1}^{k} \left(\frac{k-h}{k} \sum_{n=0}^{\omega-x} D_{x+n} + \frac{h}{k} \sum_{n=0}^{\omega-x} D_{x+n+1} \right) \\
&= \frac{1}{k \cdot D_x} \sum_{h=1}^{k} \left[\frac{k-h}{k} (D_x + N_{x+1}) + \frac{h}{k} \cdot N_{x+1} \right] \\
&= \frac{1}{k \cdot D_x} \sum_{h=1}^{k} \left(\frac{k-h}{k} \cdot D_x + N_{x+1} \right) = \frac{1}{k \cdot D_x} \left\{ \left[k - \frac{k(k+1)}{2k} \right] \cdot D_x + kN_{x+1} \right\} \\
&= \frac{N_{x+1}}{D_x} + \frac{k-1}{2k}.
\end{aligned}$$

□

Corollary 7.5.1. *Let $x \in \mathbb{N}$, $x \leq \omega$, $k \in \mathbb{N}^*$ and $T \geq 0$. The single premium payable by a person of age x for a whole life annuity-immediate of T u.c. per each k-th period of the year is*

$$T \cdot k \cdot a_x^{(k)} = T\left(k \cdot \frac{N_{x+1}}{D_x} + \frac{k-1}{2}\right).$$

Remark 7.5.1. *By (7.11) and (7.4) it follows that*

$$a_x^{(k)} = a_x + \frac{k-1}{2k}, \ \forall x \in \mathbb{N},\ x \leq \omega,\ \forall k \in \mathbb{N}^*,$$
$$a_x^{(1)} = a_x, \ \forall x \in \mathbb{N},\ x \leq \omega.$$

Example 7.5.1. Compute the single premium payable by a 49 years old male for a whole life annuity-immediate of 200 u.c. per month. The annual interest percent is 10%.

Solution. The premium is

$$\begin{aligned} 200 \cdot 12 \cdot a_{49}^{(12)} &= 200 \cdot 12 \left(a_{49} + \frac{12-1}{2 \cdot 12}\right) \\ &= 200 \cdot 12 \left(\frac{D_{50} + D_{51} + \cdots + D_{100}}{D_{49}} + \frac{11}{24}\right) \\ &\simeq 19905.06 \text{ u.c.} \end{aligned}$$

□

7.5.2 Deferred whole life annuities with k-thly payments

Definition 7.5.3. *Let $x, r \in \mathbb{N}$ s.t. $x + r \leq \omega$ and let $k \in \mathbb{N}^*$. We denote*

- ${}_{r|}a_x^{(k)}$ *= the single premium payable by a person of age x for an r-year deferred whole life annuity-immediate of $\frac{1}{k}$ u.c. per each k-th period of the year (payable at the end of each k-th period of the year while the person survives from age $x + r$ onward).*

Proposition 7.5.2. *For any $x, r \in \mathbb{N}$ s.t. $x + r \leq \omega$ and any $k \in \mathbb{N}^*$ we have*

$${}_{r|}a_x^{(k)} = \frac{N_{x+r+1}}{D_x} + \frac{k-1}{2k} \cdot \frac{D_{x+r}}{D_x}. \tag{7.13}$$

Proof. By Theorem 6.1.1 we have

$$ {}_{r|}a_x^{(k)} = \frac{1}{k} \sum_{n=0}^{\omega-r-x} \sum_{h=1}^{k} {}_{r+n+\frac{h}{k}}E_x. \tag{7.14} $$

By (7.3) and (7.12) we obtain

$$ {}_{r|}a_x^{(k)} = \frac{1}{k} \sum_{n=0}^{\omega-r-x} \sum_{h=1}^{k} {}_rE_x \cdot {}_{n+\frac{h}{k}}E_{x+r} = {}_rE_x \cdot a_{x+r}^{(k)}, $$

and using (7.1) and (7.11) we obtain

$$ {}_{r|}a_x^{(k)} = \frac{D_{x+r}}{D_x} \left(\frac{N_{x+r+1}}{D_{x+r}} + \frac{k-1}{2k} \right). $$

□

Corollary 7.5.2. *Let $x, r \in \mathbb{N}$, $x + r \leq \omega$, $k \in \mathbb{N}^*$ and $T \geq 0$. The single premium payable by a person of age x for an r-year deferred whole life annuity-immediate of T u.c. per each k-th period of the year is*

$$ T \cdot k \cdot {}_{r|}a_x^{(k)} = T \left(k \cdot \frac{N_{x+r+1}}{D_x} + \frac{k-1}{2} \cdot \frac{D_{x+r}}{D_x} \right). $$

Remark 7.5.2. *By (7.13), (7.1), (7.11) and (7.6) it follows that*

$$ {}_{r|}a_x^{(k)} = {}_rE_x \cdot a_{x+r}^{(k)},\ \forall x, r \in \mathbb{N} \text{ s.t. } x + r \leq \omega,\ \forall k \in \mathbb{N}^*, \tag{7.15} $$

$$ {}_{0|}a_x^{(k)} = a_x^{(k)},\ \forall x \in \mathbb{N},\ x \leq \omega,\ \forall k \in \mathbb{N}^*, $$

$$ {}_{\omega-x|}a_x^{(k)} = 0,\ \forall x \in \mathbb{N},\ x \leq \omega,\ \forall k \in \mathbb{N}^*, $$

$$ {}_{r|}a_x^{(1)} = {}_{r|}a_x,\ \forall x, r \in \mathbb{N},\ x \leq \omega. $$

Example 7.5.2. Compute the single premium payable by a 43 years old female for a 17-year deferred whole life annuity-immediate of 50 u.c. per month. The annual interest percent is 10%.

Solution. The premium is

$$ \begin{aligned} 50 \cdot 12 \cdot {}_{17|}a_{43}^{(12)} &= 50 \cdot 12 \left(\frac{N_{61}}{D_{43}} + \frac{11}{24} \cdot \frac{D_{60}}{D_{43}} \right) \\ &= 50 \cdot 12 \left(\frac{D_{61} + D_{62} + \cdots + D_{100}}{D_{43}} + \frac{11}{24} \cdot \frac{D_{60}}{D_{43}} \right) \\ &\simeq 818.19 \text{ u.c.} \end{aligned} $$

□

7.5.3 Temporary life annuities with k-thly payments

Definition 7.5.4. *Let $x, r \in \mathbb{N}$ s.t. $x + r \leq \omega$ and let $k \in \mathbb{N}^*$. We denote*

- $a^{(k)}_{x:\,\overline{r}|}$ = *the single premium payable by a person of age x for an r-year temporary life annuity-immediate of $\dfrac{1}{k}$ u.c. per each k-th period of the year (payable at the end of each each k-th period of the year while the person survives during the next r years).*

Proposition 7.5.3. *For any $x, r \in \mathbb{N}$ s.t. $x + r \leq \omega$ and any $k \in \mathbb{N}^*$ we have*

$$a^{(k)}_{x:\,\overline{r}|} = \frac{N_{x+1} - N_{x+r+1}}{D_x} + \frac{k-1}{2k}\left(1 - \frac{D_{x+r}}{D_x}\right). \tag{7.16}$$

Proof. By Theorem 6.1.1, (7.12) and (7.14) we have

$$\begin{aligned} a^{(k)}_{x:\,\overline{r}|} &= \frac{1}{k}\sum_{n=0}^{r-1}\sum_{h=1}^{k} {}_{n+\frac{h}{k}}E_x = \frac{1}{k}\sum_{n=0}^{\omega-x}\sum_{h=1}^{k} {}_{n+\frac{h}{k}}E_x - \frac{1}{k}\sum_{n=r}^{\omega-x}\sum_{h=1}^{k} {}_{n+\frac{h}{k}}E_x \\ &= \frac{1}{k}\sum_{n=0}^{\omega-x}\sum_{h=1}^{k} {}_{n+\frac{h}{k}}E_x - \frac{1}{k}\sum_{n=0}^{\omega-r-x}\sum_{h=1}^{k} {}_{r+n+\frac{h}{k}}E_x = a^{(k)}_x - {}_{r|}a^{(k)}_x, \end{aligned}$$

and using (7.11) and (7.13) we obtain the equality from enounce. □

Corollary 7.5.3. *Let $x, r \in \mathbb{N}$, $x + r \leq \omega$, $k \in \mathbb{N}^*$ and $T \geq 0$. The single premium payable by a person of age x for an r-year temporary life annuity-immediate of T u.c. per each k-th period of the year is*

$$T \cdot k \cdot a^{(k)}_{x:\,\overline{r}|} = T\left[k \cdot \frac{N_{x+1} - N_{x+r+1}}{D_x} + \frac{k-1}{2}\left(1 - \frac{D_{x+r}}{D_x}\right)\right].$$

Remark 7.5.3. *By (7.11), (7.16), (7.13), (7.15) and (7.8) it follows that*

$$\begin{aligned} a^{(k)}_x &= a^{(k)}_{x:\,\overline{r}|} + {}_{r|}a^{(k)}_x,\ \forall x, r \in \mathbb{N} \text{ s.t. } x + r \leq \omega,\ \forall k \in \mathbb{N}^*, \\ a^{(k)}_{x:\,\overline{r}|} &= a^{(k)}_x - {}_rE_x \cdot a^{(k)}_{x+r},\ \forall x, r \in \mathbb{N} \text{ s.t. } x + r \leq \omega,\ \forall k \in \mathbb{N}^*, \\ a^{(k)}_{x:\,\overline{0}|} &= 0,\ \forall x \in \mathbb{N},\ x \leq \omega,\ \forall k \in \mathbb{N}^*, \\ a^{(k)}_{x:\,\overline{\omega-x}|} &= a^{(k)}_x,\ \forall x \in \mathbb{N},\ x \leq \omega,\ \forall k \in \mathbb{N}^*, \\ a^{(1)}_{x:\,\overline{r}|} &= a_{x:\,\overline{r}|},\ \forall x, r \in \mathbb{N},\ x \leq \omega. \end{aligned}$$

Example 7.5.3. Compute the single premium payable by a 43 years old female for a 17-year temporary life annuity-immediate of 50 u.c. per month. The annual interest percent is 10%.

Solution. The premium is

$$\begin{aligned} 50 \cdot 12 \cdot {}_{17|}a^{(12)}_{43:\,\overline{17|}} &= 50 \cdot 12 \left[\frac{N_{44} - N_{61}}{D_{43}} + \frac{11}{24} \left(1 - \frac{D_{60}}{D_{43}} \right) \right] \\ &= 50 \cdot 12 \left[\frac{D_{44} + D_{45} + \cdots + D_{60}}{D_{43}} + \frac{11}{24} \left(1 - \frac{D_{60}}{D_{43}} \right) \right] \\ &\simeq 4892.07 \text{ u.c.} \end{aligned}$$

□

7.6 Pension

7.6.1 Annually pension

We denote by r the number of years until the time of retirement.

Definition 7.6.1. *Let $x, r \in \mathbb{N}$ s.t. $x + r \leq \omega$. We denote*

- $P_{x:\,\overline{r|}}({}_{r|}a_x)$ = *the r-year temporary life premium payable by a person of age x (at the end of each year while the person survives during the next r years) for an r-year deferred whole life annually pension of* 1 *u.c. per each year (payable at the end of each year while the person survives from age $x + r$ onward).*

Proposition 7.6.1. *For any $x, r \in \mathbb{N}$ s.t. $x + r \leq \omega$, we have*

$$P_{x:\,\overline{r|}}({}_{r|}a_x) = \frac{{}_{r|}a_x}{a_{x:\,\overline{r|}}} = \frac{N_{x+r+1}}{N_{x+1} - N_{x+r+1}}. \tag{7.17}$$

Proof. For a mutually advantageous insurance, the present values (at the initial moment of the insurance) of the premiums need to be equal to the present value of the pensions. By Definition 7.4.1, Corollary 7.4.1, Definition 7.3.3 and Proposition 7.3.2 we have

$$P_{x:\,\overline{r|}}({}_{r|}a_x) \cdot a_{x:\,\overline{r|}} = {}_{r|}a_x, \text{ so } P_{x:\,\overline{r|}}({}_{r|}a_x) \cdot \frac{N_{x+1} - N_{x+r+1}}{D_x} = \frac{N_{x+r+1}}{D_x},$$

and hence we obtain the equality from enounce. □

Corollary 7.6.1. *Let $x, r \in \mathbb{N}$, $x + r \leq \omega$ and $T \geq 0$. The r-year temporary life premium payable by a person of age x for an r-year deferred whole life annually pension of T u.c. per each year is*

$$T \cdot P_{x:\,\overline{r|}}({}_{r|}a_x) = T \cdot \frac{N_{x+r+1}}{N_{x+1} - N_{x+r+1}}.$$

Example 7.6.1. Compute the annuity-immediate premium payable by a 30 years old male for an annuity-immediate pension of 10000 u.c. per year. The annual interest percent is 10% and the age of retirement is 65 years.

Solution. The premium is

$$\begin{aligned}10000 \cdot P_{30:\,\overline{35}|}({}_{35|}a_{30}) &= 10000 \cdot \frac{N_{66}}{N_{31} - N_{66}}\\ &= 10000 \cdot \frac{D_{66} + D_{67} + \cdots + D_{100}}{D_{31} + D_{32} + \cdots + D_{65}}\\ &\simeq 133.37 \text{ u.c.}\end{aligned}$$

□

7.6.2 Monthly pension

We denote by r the number of years until the time of retirement.

Definition 7.6.2. *Let $x, r \in \mathbb{N}$ s.t. $x + r \le \omega$. We denote*

- $P^{(12)}_{x:\,\overline{r}|}({}_{r|}a^{(12)}_x)$ = *the r-year temporary life premium payable by a person of age x at the end of each month (while the person survives during the next r years) for an r-year deferred whole life monthly pension of* 1 *u.c. per each month (payable at the end of each month while the person survives from age $x + r$ onward).*

Proposition 7.6.2. *For any $x, r \in \mathbb{N}$ s.t. $x + r \le \omega$, we have*

$$P^{(12)}_{x:\,\overline{r}|}({}_{r|}a^{(12)}_x) = \frac{{}_{r|}a^{(12)}_x}{a^{(12)}_{x:\,\overline{r}|}} = \frac{24N_{x+r+1} + 11D_{x+r}}{24(N_{x+1} - N_{x+r+1}) + 11(D_x - D_{x+r})}. \tag{7.18}$$

Proof. For a mutually advantageous insurance, the present values (at the initial moment of the insurance) of the premiums need to be equal to the present value of the pensions. By Definition 7.5.4, Corollary 7.5.3, Definition 7.5.3 and Proposition 7.5.2 we have

$$P^{(12)}_{x:\,\overline{r}|}({}_{r|}a^{(12)}_x) \cdot 12 \cdot a^{(12)}_{x:\,\overline{r}|} = 12 \cdot {}_{r|}a^{(12)}_x,$$

so

$$P^{(12)}_{x:\,\overline{r}|}({}_{r|}a^{(12)}_x) \left[12 \cdot \frac{N_{x+1} - N_{x+r+1}}{D_x} + \frac{11}{2}\left(1 - \frac{D_{x+r}}{D_x}\right)\right] = 12 \cdot \frac{N_{x+r+1}}{D_x} + \frac{11}{2} \cdot \frac{D_{x+r}}{D_x},$$

and hence we obtain the equality from enounce. □

Corollary 7.6.2. *Let $x, r \in \mathbb{N}$, $x + r \le \omega$ and $T \ge 0$. The r-year temporary life premium payable by a person of age x at the end of each month for an r-year deferred whole life monthly pension of T u.c. per each month is*

$$T \cdot P^{(12)}_{x:\,\overline{r}|}({}_{r|}a^{(12)}_x) = T \cdot \frac{24N_{x+r+1} + 11D_{x+r}}{24(N_{x+1} - N_{x+r+1}) + 11(D_x - D_{x+r})}.$$

Example 7.6.2. Compute the monthly-immediate premium payable by a 30 years old male for a monthly-immediate pension of 1000 u.c. per month. The annual interest percent is 10% and the age of retirement is 65 years.

Solution. The premium is

$$\begin{aligned} 1000 \cdot P^{(12)}_{30:\,\overline{35}|}({}_{r|}a^{(12)}_{30}) &= 1000 \cdot \frac{24N_{66} + 11D_{65}}{24(N_{31} - N_{66}) + 11(D_{30} - D_{65})} \\ &= 1000 \cdot \frac{24\,(D_{66} + D_{67} + \cdots + D_{100}) + 11D_{65}}{24\,(D_{31} + D_{32} + \cdots + D_{65}) + 11(D_{30} - D_{65})} \\ &\simeq 13.76 \text{ u.c.} \end{aligned}$$

□

7.7 Problems

Exercise 7.7.1. Calculate the single premium payable by a 30 years old person for a single claim of 10000$ over 35 years if the person survives. The annual interest percent is 6%.

Exercise 7.7.2. Calculate the single premium payable by a 30 years old person for a whole life annuity-immediate of 12000RON per year. The annual interest percent is 8%.

Exercise 7.7.3. Calculate the single premium payable by a 30 years old person for a 35-year deferred whole life annuity-immediate of 12000RON per year. The annual interest percent is 8%.

Exercise 7.7.4. Calculate the single premium payable by a 30 years old person for a 35-year temporary life annuity-immediate of 12000RON per year. The annual interest percent is 8%.

Exercise 7.7.5. Calculate the single premium payable by a 30 years old person for a whole life annuity-immediate of 1000RON per month. The annual interest percent is 8%.

Exercise 7.7.6. Calculate the single premium payable by a 30 years old person for a 35-year deferred whole life annuity-immediate of 1000RON per month. The annual interest percent is 8%.

Exercise 7.7.7. Calculate the single premium payable by a 30 years old person for a 35-year temporary life annuity-immediate of 1000RON per month. The annual interest percent is 8%.

Exercise 7.7.8. Calculate the annuity-immediate premium payable by a 30 years old person for an annuity-immediate pension of 12000RON per year. The annual interest percent is 8% and the age of retirement is 65 years.

Exercise 7.7.9. Calculate the monthly-immediate premium payable by a 30 years old person for a monthly-immediate pension of 1000RON per month. The annual interest percent is 8% and the age of retirement is 65 years.

Theme 8

Life insurances

8.1 A general model. Classification

In a life insurance, the single claim is payable at the moment of death, if the death occurs in the period covered by the insurance. The life insurance can be:

- **immediate and unlimited**, when the claim is payed at the moment of death, whenever this occurs.
- **deferred**, when the claim is payed only if the insured dies after a fixed term from the time of insurance issue.
- **temporary (limited)**, when the claim is payed only if the insured dies within a fixed term from the time of insurance issue.

8.2 Whole life insurance

Definition 8.2.1. *Let $x \in \mathbb{N}$, $x \leq \omega$. We denote*

- $\dot{A}_x$ *= the single premium payable by a person of age x for a whole life insurance of 1 u.c. (payable at the moment of death, whenever this occurs).*

Proposition 8.2.1. *For any $x \in \mathbb{N}$, $x \leq \omega$, we have*

$$\dot{A}_x = \frac{M_x}{D_x}, \tag{8.1}$$

where

$$M_x = C_x + C_{x+1} + \cdots + C_\omega, \text{ with } C_x = d_x \cdot v^{x+\frac{1}{2}} = (l_x - l_{x+1})v^{x+\frac{1}{2}}, \tag{8.2}$$

$v = \dfrac{1}{1+i}$ *being the annual discounting factor, i being the annual interest rate.*

Proof. For a mutually advantageous insurance, the single premium $A(x)$ need to be equal to the present value of the single claim, that is

$$\dot{A}_x = E(X),$$

where X is the random variable that represents the present value of the claim. Assuming that the deaths are uniform distributed throughout the year, we have

$X = v^{n+\frac{1}{2}}$, if n is the number of complete years lived by the insured since issue,

for any $n \in \{0, \dots, \omega - x\}$. Hence the distribution of X is

$$X : \begin{pmatrix} v^{\frac{1}{2}} & \dots & v^{n+\frac{1}{2}} & \dots & v^{\omega-x+\frac{1}{2}} \\ \beta_x(0) & \dots & \beta_x(n) & \dots & \beta_x(\omega-x) \end{pmatrix},$$

where, for any $n \in \{0, \dots, \omega - x\}$, $\beta_x(n)$ represents the probability that a person of age x will live only n (i.e. will die at age $x+n$). By (6.18) we have

$$\beta_x(n) = \frac{d_{x+n}}{l_x}, \ \forall n \in \{0, \dots, \omega - x\},$$

where $d_{x+n} = l_{x+n} - l_{x+n+1}$ represents the number of deaths at age $x+n$. Using (8.2) it follows that

$$\begin{aligned} \dot{A}_x &= E(X) = \sum_{n=0}^{\omega-x} \beta_x(n) v^{n+\frac{1}{2}} = \sum_{n=0}^{\omega-x} \frac{d_{x+n}}{l_x} \cdot v^{n+\frac{1}{2}} = \sum_{n=0}^{\omega-x} \frac{d_{x+n} \cdot v^{x+n+\frac{1}{2}}}{l_x \cdot v^x} \\ &= \sum_{n=0}^{\omega-x} \frac{C_{x+n}}{D_x} = \frac{M_x}{D_x}. \end{aligned}$$

$\square$

Corollary 8.2.1. *Let $x \in \mathbb{N}$, $x \leq \omega$ and let $T \geq 0$. The single premium payable by a person of age x for a whole life insurance of T u.c. is*

$$T \cdot \dot{A}_x = T \cdot \frac{M_x}{D_x}.$$

Corollary 8.2.2. *For any $x \in \mathbb{N}$, $x \leq \omega$, we have*

$$\dot{A}_x = \sqrt{v}\,(1 - i \cdot a_x)\,. \tag{8.3}$$

Example 8.2.1. Compute the single premium payable by a 49 years old male for a whole life insurance of 15000 u.c. The annual interest percent is 10%.

Solution. The premium is

$$\begin{aligned}
15000 \cdot \dot{A}_{49} &= 15000 \cdot \frac{M_{49}}{D_{49}} = 15000 \cdot \frac{C_{49} + C_{50} + \cdots + C_{100}}{D_{49}} \\
&= 15000 \cdot \frac{(l_{49} - l_{50})v^{49+\frac{1}{2}} + (l_{50} - l_{51})v^{50+\frac{1}{2}} + \cdots + (l_{100} - l_{101})v^{100+\frac{1}{2}}}{v^{49} \cdot l_{49}} \\
&= 15000 \cdot \frac{\left(\frac{1}{1.1}\right)^{49.5} \cdot 1090 + \left(\frac{1}{1.1}\right)^{50.5} \cdot 1141 + \cdots + \left(\frac{1}{1.1}\right)^{100.5} \cdot 15}{\left(\frac{1}{1.1}\right)^{49} \cdot 88191} \\
&\simeq 3095.73 \text{ u.c.}
\end{aligned}$$

□

8.3 Deferred life insurance

Definition 8.3.1. *Fie $x, r \in \mathbb{N}$ s.t. $x + r \leq \omega$. We denote*

- ${}_{r|}\dot{A}_x$ = *the single premium payable by a person of age x for a r-year deferred life insurance of 1 u.c. (payable at the moment of death only if the insured die at least r years following insurance issue).*

Proposition 8.3.1. *For any $x, r \in \mathbb{N}$ s.t. $x + r \leq \omega$, we have*

$$ {}_{r|}\dot{A}_x = \frac{M_{x+r}}{D_x}. \tag{8.4}$$

Proof. Similar to Proposition 8.2.1 we have

$$ {}_{r|}\dot{A}_x = E({}_{r|}X), $$

where ${}_rX$ is the random variable having the distribution

$$ {}_{r|}X : \begin{pmatrix} 0 & 0 & \ldots & 0 & v^{r+\frac{1}{2}} & v^{r+1+\frac{1}{2}} & \ldots & v^{\omega-x+\frac{1}{2}} \\ \beta_x(0) & \beta_x(1) & \ldots & \beta_x(r-1) & \beta_x(r) & \beta_x(r+1) & \ldots & \beta_x(\omega - x) \end{pmatrix}. $$

Using (6.18) and (8.2) we obtain that

$$\begin{aligned}{}_{r|}\dot{A}_x &= E({}_{r|}X) = \sum_{n=r}^{\omega-x} \beta_x(n) v^{n+\frac{1}{2}} = \sum_{n=r}^{\omega-x} \frac{d_{x+n}}{l_x} \cdot v^{n+\frac{1}{2}} = \sum_{n=r}^{\omega-x} \frac{d_{x+n} \cdot v^{x+n+\frac{1}{2}}}{l_x \cdot v^x} \\ &= \sum_{n=r}^{\omega-x} \frac{C_{x+n}}{D_x} = \frac{M_{x+r}}{D_x}.\end{aligned}$$

□

Corollary 8.3.1. *Let $x, r \in \mathbb{N}$, $x + r \leq \omega$, $T \geq 0$. The single premium payable by a person of age x for a r-year deferred life insurance of T u.c. is*

$$T \cdot {}_{r|}\dot{A}_x = T \cdot \frac{M_{x+r}}{D_x}.$$

Remark 8.3.1. *By (8.4), (7.1) and (8.1) it follows that*

$$\begin{aligned} {}_{r|}\dot{A}_x &= {}_rE_x \cdot \dot{A}_{x+r}, \ \forall x, r \in \mathbb{N} \ s.t. \ x + r \leq \omega, \qquad (8.5) \\ {}_{0|}\dot{A}_x &= \dot{A}_x, \ \forall x \in \mathbb{N}, \ x \leq \omega. \end{aligned}$$

Remark 8.3.2. *By (8.5), (8.3) and (7.7) it follows that*

$$ {}_{r|}\dot{A}_x = \sqrt{v}\left({}_rE_x - i \cdot {}_{r|}a_x\right), \ \forall x, r \in \mathbb{N} \ s.t. \ x + r \leq \omega. \qquad (8.6)$$

Example 8.3.1. Compute the single premium payable by a 49 years old male for a 16-year deferred life insurance of 15000 u.c. The annual interest percent is 10%.

Solution. The premium is

$$\begin{aligned} 15000 \cdot {}_{16|}\dot{A}_{49} &= 15000 \cdot \frac{M_{65}}{D_{49}} = 15000 \cdot \frac{C_{65} + C_{66} + \cdots + C_{100}}{D_{49}} \\ &= 15000 \cdot \frac{(l_{65} - l_{66})v^{65+\frac{1}{2}} + (l_{66} - l_{67})v^{66+\frac{1}{2}} + \cdots + (l_{100} - l_{101})v^{100+\frac{1}{2}}}{v^{49} \cdot l_{49}} \\ &= 15000 \cdot \frac{\left(\frac{1}{1.1}\right)^{65.5} \cdot 2397 + \left(\frac{1}{1.1}\right)^{66.5} \cdot 2393 + \cdots + \left(\frac{1}{1.1}\right)^{100.5} \cdot 15}{\left(\frac{1}{1.1}\right)^{49} \cdot 88191} \\ &\simeq 957.40 \text{ u.c.} \end{aligned}$$

□

8.4 Temporary life insurance

Definition 8.4.1. *Let $x, r \in \mathbb{N}$ s.t. $x + r \leq \omega$. We denote*

- $\dot{A}_{\overset{1}{x}:\, r\rceil}$ *= the single premium payable by a person of age x for a r-year term life insurance of 1 u.c. (payable at the moment of death only if the insured die within r years following insurance issue).*

Proposition 8.4.1. *For any $x, r \in \mathbb{N}$ s.t. $x + r \leq \omega$, we have*

$$\dot{A}_{\overset{1}{x}:\, r\rceil} = \frac{M_x - M_{x+r}}{D_x}. \tag{8.7}$$

Proof. Similar to Proposition 8.2.1 we have

$$\dot{A}_{\overset{1}{x}:\, r\rceil} = E(X_{r\rceil}),$$

where $X_{r\rceil}$ is the random variable having the distribution

$$X_{r\rceil} : \begin{pmatrix} v^{\frac{1}{2}} & v^{1+\frac{1}{2}} & \dots & v^{r-1+\frac{1}{2}} & 0 & 0 & \dots & 0 \\ \beta_x(0) & \beta_x(1) & \dots & \beta_x(r-1) & \beta_x(r) & \beta_x(r+1) & \dots & \beta_x(\omega - x) \end{pmatrix}.$$

Using(6.18) and (8.2) we obtain

$$\dot{A}_{\overset{1}{x}:\, r\rceil} = E(X_{r\rceil}) = \sum_{n=0}^{r-1} \beta_x(n) v^{n+\frac{1}{2}} = \sum_{n=0}^{r-1} \frac{d_{x+n}}{l_x} \cdot v^{n+\frac{1}{2}} = \sum_{n=0}^{r-1} \frac{d_{x+n} \cdot v^{x+n+\frac{1}{2}}}{l_x \cdot v^x}$$
$$= \sum_{n=0}^{r-1} \frac{C_{x+n}}{D_x} = \frac{M_x - M_{x+r}}{D_x}.$$

□

Corollary 8.4.1. *Let $x, r \in \mathbb{N}$, $x + r \leq \omega$, $T \geq 0$. The single premium payable by a person of age x for a r-year term life insurance of T u.c. is*

$$T \cdot \dot{A}_{\overset{1}{x}:\, r\rceil} = T \cdot \frac{M_x - M_{x+r}}{D_x}.$$

Remark 8.4.1. *By (8.1), (8.7), (8.4) and (8.5) it follows that*

$$\dot{A}_x = \dot{A}_{\overset{1}{x}:\, r\rceil} + {}_{r|}\dot{A}_x, \ \forall x, r \in \mathbb{N} \text{ s.t. } x + r \leq \omega, \tag{8.8}$$

$$\dot{A}_{\overset{1}{x}:\, r\rceil} = \dot{A}_x - {}_rE_x \cdot \dot{A}_{x+r}, \ \forall x, r \in \mathbb{N} \text{ s.t. } x + r \leq \omega,$$

$$\dot{A}_{\overset{1}{x}:\, 0\rceil} = 0, \ \forall x \in \mathbb{N}, \ x \leq \omega.$$

Remark 8.4.2. *By (8.8), (8.3), (8.6) and (7.9) it follows that*

$$\dot{A}_{\overset{1}{x:\,\overline{r|}}} = \sqrt{v}\left(1 - {}_rE_x - i \cdot a_{x:\,\overline{r|}}\right),\ \forall x, r \in \mathbb{N}\ s.t.\ x + r \leq \omega. \tag{8.9}$$

Example 8.4.1. Compute the single premium payable by a 49 years old male for a 16-year term life insurance of 15000 u.c. The annual interest percent is 10%.

Solution. The premium is

$$\begin{aligned}
15000 \cdot \dot{A}_{\overset{1}{49:\,\overline{16|}}} &= 15000 \cdot \frac{M_{49} - M_{65}}{D_{49}} = 15000 \cdot \frac{C_{49} + C_{50} + \cdots + C_{64}}{D_{49}} \\
&= 15000 \cdot \frac{(l_{49} - l_{50})v^{49+\frac{1}{2}} + (l_{50} - l_{51})v^{50+\frac{1}{2}} + \cdots + (l_{64} - l_{65})v^{64+\frac{1}{2}}}{v^{49} \cdot l_{49}} \\
&= 15000 \cdot \frac{\left(\frac{1}{1.1}\right)^{49.5} \cdot 1090 + \left(\frac{1}{1.1}\right)^{50.5} \cdot 1141 + \cdots + \left(\frac{1}{1.1}\right)^{64.5} \cdot 2342}{\left(\frac{1}{1.1}\right)^{49} \cdot 88191} \\
&\simeq 2138.33 \text{ u.c.}
\end{aligned}$$

□

8.5 Problems

Exercise 8.5.1. Calculate the single premium payable by a 30 years old person for a whole life insurance of 1000RON. The annual interest percent is 8%.

Exercise 8.5.2. Calculate the single premium payable by a 30 years old person for a 35-year deferred life insurance of 1000RON. The annual interest percent is 8%.

Exercise 8.5.3. Calculate the single premium payable by a 30 years old person for a 35-year term life insurance of 1000RON. The annual interest percent is 8%.

Theme 9

Collective annuities and insurances

Next, we consider an insured group of m persons having the ages $x_1, x_2, \ldots, x_m$ ($m \in \mathbb{N}^*$, $x_j \in \mathbb{N}$, $x_j \leq \omega\ \forall j \in \{1, \ldots, m\}$).

9.1 Multiple life probabilities

Definition 9.1.1. *Let a group of m persons having the ages $x_1, x_2, \ldots, x_m$, where $m \in \mathbb{N}^*$, $x_j \in \mathbb{N}$, $x_j \leq \omega\ \forall j \in \{1, \ldots, m\}$. Let $n, k \in \mathbb{N}$ s.t. $k \leq m$. We denote*

- ${}_np_{x_1x_2\ldots x_m}$ *= the probability that all members of the group will survive n years;*
- ${}_np_{\overline{x_1x_2\ldots x_m}}^{[k]}$ *= the probability that exactly k of the group members will survive n years;*
- ${}_np_{\overline{x_1x_2\ldots x_m}}^{k}$ *= the probability that at least k of the group members will survive n years;*

${}_np_{x_1x_2\ldots x_m}$ *is called **the probability of joint survival** (**probability of the joint-life**) for the group, and ${}_np_{\overline{x_1x_2\ldots x_m}}^{[k]}$ and ${}_np_{\overline{x_1x_2\ldots x_m}}^{k}$ are called **probabilities of partial survival** for the group.*

Remark 9.1.1. *We assume that the deaths of the group members are independent.*

Definition 9.1.2. *We denote by $\widetilde{x}$ **the maximum age of the group**, i.e.*

$$\widetilde{x} = \max\{x_1, x_2, \ldots, x_m\}.$$

Also, we denote by $\widehat{x}$ ***the minimum age of the group****, i.e.*

$$\widehat{x} = \min\{x_1, x_2, \ldots, x_m\}.$$

Remark 9.1.2. *Obviously, if* $n > \omega - \widetilde{x}$ *then* ${}_np_{x_1x_2\ldots x_m} = 0$.

Proposition 9.1.1. *Let* $n, k \in \mathbb{N}$ *s.t.* $k \leq m$. *We have:*

$$ {}_np_{x_1x_2\ldots x_m} = {}_np_{x_1} \cdot {}_np_{x_2} \cdot \ldots \cdot {}_np_{x_m} = \frac{l_{x_1+n}}{l_{x_1}} \cdot \frac{l_{x_2+n}}{l_{x_2}} \cdot \ldots \cdot \frac{l_{x_m+n}}{l_{x_m}}; \qquad (9.1)$$

$$ {}_np_{\overset{[k]}{\overline{x_1x_2\ldots x_m}}} = \sum_{s=0}^{m-k} (-1)^s \cdot C_{k+s}^{s} \sum_{1\leq i_1<\ldots<i_{k+s}\leq m} {}_np_{x_{i_1}x_{i_2}\ldots x_{i_{k+s}}}; \qquad (9.2)$$

$$ {}_np_{\overset{k}{\overline{x_1x_2\ldots x_m}}} = \sum_{s=0}^{m-k} (-1)^s \cdot C_{k+s-1}^{s} \sum_{1\leq i_1<\ldots<i_{k+s}\leq m} {}_np_{x_{i_1}x_{i_2}\ldots x_{i_{k+s}}}. \qquad (9.3)$$

Example 9.1.1. Consider a family composed of the father (56 years old), mother (52 years old), daughter (32 years old) and sun (30 years old).
a) Compute the probability that all of the members survive 20 years.
b) Compute the probability that exactly 2 of the members survive 20 years.
c) Compute the probability that at least 2 of the members survive 20 years.

Solution. a) The probability is

$$\begin{aligned} {}_{20}p_{56,52,32,30} &= {}_{20}p_{56} \cdot {}_{20}p_{52} \cdot {}_{20}p_{32} \cdot {}_{20}p_{30} \\ &= \frac{l_{76}}{l_{56}} \cdot \frac{l_{72}}{l_{52}} \cdot \frac{l_{52}}{l_{32}} \cdot \frac{l_{50}}{l_{30}} \\ &= \frac{28059}{79017} \cdot \frac{65336}{93717} \cdot \frac{93717}{99261} \cdot \frac{87101}{98540} \\ &\simeq 0{,}206603. \end{aligned}$$

b) The probability is

$$\begin{aligned} {}_{20}p_{\overset{[2]}{\overline{56,52,32,30}}} &= \sum_{s=0}^{2} (-1)^s \cdot C_{2+s}^{s} \sum_{1\leq i_1<\ldots<i_{2+s}\leq 4} {}_{20}p_{x_{i_1}x_{i_2}\ldots x_{i_{2+s}}} \\ &= C_2^0 \sum_{1\leq i_1<i_2\leq 4} {}_{20}p_{x_{i_1}x_{i_2}} - C_3^1 \sum_{1\leq i_1<i_2<i_3\leq 4} {}_{20}p_{x_{i_1}x_{i_2}x_{i_3}} + C_4^2 \sum_{1\leq i_1<i_2<i_3<i_4\leq 4} {}_{20}p_{x_{i_1}x_{i_2}x_{i_3}x_{i_4}} \\ &= C_2^0 \left({}_{20}p_{x_1x_2} + {}_{20}p_{x_1x_3} + {}_{20}p_{x_1x_4} + {}_{20}p_{x_2x_3} + {}_{20}p_{x_2x_4} + {}_{20}p_{x_3x_4}\right) - C_3^1 \big({}_{20}p_{x_1x_2x_3} + {}_{20}p_{x_1x_2x_4} \\ &\quad + {}_{20}p_{x_1x_3x_4} + {}_{20}p_{x_2x_3x_4}\big) + C_4^2 \cdot {}_{20}p_{x_1x_2x_3x_4} \\ &= {}_{20}p_{56,52} + {}_{20}p_{56,32} + {}_{20}p_{56,30} + {}_{20}p_{52,32} + {}_{20}p_{52,30} + {}_{20}p_{32,30} - 3\big({}_{20}p_{56,52,32} + {}_{20}p_{56,52,30} \end{aligned}$$

$$
\begin{aligned}
&\quad + {}_{20}p_{56,32,30} + {}_{20}p_{52,32,30}\big) + 6 \cdot {}_{20}p_{56,52,32,30} \\
&= {}_{20}p_{56} \cdot {}_{20}p_{52} + {}_{20}p_{56} \cdot {}_{20}p_{32} + {}_{20}p_{56} \cdot {}_{20}p_{30} + {}_{20}p_{52} \cdot {}_{20}p_{32} + {}_{20}p_{52} \cdot {}_{20}p_{30} + {}_{20}p_{32} \cdot {}_{20}p_{30} \\
&\quad - 3\big({}_{20}p_{56} \cdot {}_{20}p_{52} \cdot {}_{20}p_{32} + {}_{20}p_{56} \cdot {}_{20}p_{52} \cdot {}_{20}p_{30} + {}_{20}p_{56} \cdot {}_{20}p_{32} \cdot {}_{20}p_{30} + {}_{20}p_{52} \cdot {}_{20}p_{32} \cdot {}_{20}p_{30}\big) \\
&\quad + 6 \cdot {}_{20}p_{56} \cdot {}_{20}p_{52} \cdot {}_{20}p_{32} \cdot {}_{20}p_{30} \\
&= \frac{l_{76}}{l_{56}} \cdot \frac{l_{72}}{l_{52}} + \frac{l_{76}}{l_{56}} \cdot \frac{l_{52}}{l_{32}} + \frac{l_{76}}{l_{56}} \cdot \frac{l_{50}}{l_{30}} + \frac{l_{72}}{l_{52}} \cdot \frac{l_{52}}{l_{32}} + \frac{l_{72}}{l_{52}} \cdot \frac{l_{50}}{l_{30}} + \frac{l_{52}}{l_{32}} \cdot \frac{l_{50}}{l_{30}} \\
&\quad - 3\left(\frac{l_{76}}{l_{56}} \cdot \frac{l_{72}}{l_{52}} \cdot \frac{l_{52}}{l_{32}} + \frac{l_{76}}{l_{56}} \cdot \frac{l_{72}}{l_{52}} \cdot \frac{l_{50}}{l_{30}} + \frac{l_{76}}{l_{56}} \cdot \frac{l_{52}}{l_{32}} \cdot \frac{l_{50}}{l_{30}} + \frac{l_{72}}{l_{52}} \cdot \frac{l_{52}}{l_{32}} \cdot \frac{l_{50}}{l_{30}}\right) \\
&\quad + 6 \cdot \frac{l_{76}}{l_{56}} \cdot \frac{l_{72}}{l_{52}} \cdot \frac{l_{52}}{l_{32}} \cdot \frac{l_{50}}{l_{30}} \\
&\simeq 0.253160.
\end{aligned}
$$

c) The probability is

$$
\begin{aligned}
&{}_{20}p_{\overline{\overset{2}{56,52,32,30}}} = \sum_{s=0}^{2} (-1)^s \cdot C_{1+s}^{s} \sum_{1 \le i_1 < \ldots < i_{2+s} \le 4} {}_{20}p_{x_{i_1} x_{i_2} \ldots x_{i_{2+s}}} \\
&= C_1^0 \sum_{1 \le i_1 < i_2 \le 4} {}_{20}p_{x_{i_1} x_{i_2}} - C_2^1 \sum_{1 \le i_1 < i_2 < i_3 \le 4} {}_{20}p_{x_{i_1} x_{i_2} x_{i_3}} + C_3^2 \sum_{1 \le i_1 < i_2 < i_3 < i_4 \le 4} {}_{20}p_{x_{i_1} x_{i_2} x_{i_3} x_{i_4}} \\
&= C_1^0 \left({}_{20}p_{x_1x_2} + {}_{20}p_{x_1x_3} + {}_{20}p_{x_1x_4} + {}_{20}p_{x_2x_3} + {}_{20}p_{x_2x_4} + {}_{20}p_{x_3x_4}\right) - C_2^1\big({}_{20}p_{x_1x_2x_3} + {}_{20}p_{x_1x_2x_4} \\
&\quad + {}_{20}p_{x_1x_3x_4} + {}_{20}p_{x_2x_3x_4}\big) + C_3^2 \cdot {}_{20}p_{x_1x_2x_3x_4} \\
&= {}_{20}p_{56,52} + {}_{20}p_{56,32} + {}_{20}p_{56,30} + {}_{20}p_{52,32} + {}_{20}p_{52,30} + {}_{20}p_{32,30} - 2\big({}_{20}p_{56,52,32} + {}_{20}p_{56,52,30} \\
&\quad + {}_{20}p_{56,32,30} + {}_{20}p_{52,32,30}\big) + 3 \cdot {}_{20}p_{56,52,32,30} \\
&= {}_{20}p_{56} \cdot {}_{20}p_{52} + {}_{20}p_{56} \cdot {}_{20}p_{32} + {}_{20}p_{56} \cdot {}_{20}p_{30} + {}_{20}p_{52} \cdot {}_{20}p_{32} + {}_{20}p_{52} \cdot {}_{20}p_{30} + {}_{20}p_{32} \cdot {}_{20}p_{30} \\
&\quad - 2\big({}_{20}p_{56} \cdot {}_{20}p_{52} \cdot {}_{20}p_{32} + {}_{20}p_{56} \cdot {}_{20}p_{52} \cdot {}_{20}p_{30} + {}_{20}p_{56} \cdot {}_{20}p_{32} \cdot {}_{20}p_{30} + {}_{20}p_{52} \cdot {}_{20}p_{32} \cdot {}_{20}p_{30}\big) \\
&\quad + 3 \cdot {}_{20}p_{56} \cdot {}_{20}p_{52} \cdot {}_{20}p_{32} \cdot {}_{20}p_{30} \\
&= \frac{l_{76}}{l_{56}} \cdot \frac{l_{72}}{l_{52}} + \frac{l_{76}}{l_{56}} \cdot \frac{l_{52}}{l_{32}} + \frac{l_{76}}{l_{56}} \cdot \frac{l_{50}}{l_{30}} + \frac{l_{72}}{l_{52}} \cdot \frac{l_{52}}{l_{32}} + \frac{l_{72}}{l_{52}} \cdot \frac{l_{50}}{l_{30}} + \frac{l_{52}}{l_{32}} \cdot \frac{l_{50}}{l_{30}} \\
&\quad - 2\left(\frac{l_{76}}{l_{56}} \cdot \frac{l_{72}}{l_{52}} \cdot \frac{l_{52}}{l_{32}} + \frac{l_{76}}{l_{56}} \cdot \frac{l_{72}}{l_{52}} \cdot \frac{l_{50}}{l_{30}} + \frac{l_{76}}{l_{56}} \cdot \frac{l_{52}}{l_{32}} \cdot \frac{l_{50}}{l_{30}} + \frac{l_{72}}{l_{52}} \cdot \frac{l_{52}}{l_{32}} \cdot \frac{l_{50}}{l_{30}}\right) \\
&\quad + 3 \cdot \frac{l_{76}}{l_{56}} \cdot \frac{l_{72}}{l_{52}} \cdot \frac{l_{52}}{l_{32}} \cdot \frac{l_{50}}{l_{30}} \\
&\simeq 0.964075.
\end{aligned}
$$

$\square$

9.2 Single claim for joint survival

Definition 9.2.1. *Let a group of m persons having the ages $x_1, x_2, \ldots, x_m$, where $m \in \mathbb{N}^*$, $x_j \in \mathbb{N}$, $x_j \leq \omega\ \forall j \in \{1, \ldots, m\}$. Let $n \in \mathbb{N}$. We denote*

- ${}_nE_{x_1,x_2,\ldots,x_m}$ = *the single premium payable by the group for a single claim of* 1 *u.c. over n years if all of the members survive.*

Remark 9.2.1. ${}_nE_{x_1,x_2,\ldots,x_m}$ *is called* ***the unitary premium****.*

Proposition 9.2.1. *We have*

$$ {}_nE_{x_1,x_2,\ldots,x_m} = \frac{D_{x_1+n,x_2+n,\ldots,x_m+n}}{D_{x_1,x_2,\ldots,x_m}}, \tag{9.4}$$

where

$$D_{x_1,x_2,\ldots,x_m} = l_{x_1} \cdot l_{x_2} \cdot \ldots \cdot l_{x_m} \cdot v^{\frac{x_1+x_2+\ldots+x_m}{m}}, \tag{9.5}$$

$v = \dfrac{1}{1+i}$ *being the annual discounting factor, i being the annual interest rate.*

Corollary 9.2.1. *Let a group of m persons having the ages $x_1, x_2, \ldots, x_m$, where $m \in \mathbb{N}^*$, $x_j \in \mathbb{N}$, $x_j \leq \omega\ \forall j \in \{1, \ldots, m\}$. Let $n \in \mathbb{N}$ and $T \geq 0$. The single premium payable by the group for a single claim of T u.c. over n years if all of the members survive is*

$$T \cdot {}_nE_{x_1,x_2,\ldots,x_m} = T \cdot \frac{D_{x_1+n,x_2+n,\ldots,x_m+n}}{D_{x_1,x_2,\ldots,x_m}}.$$

Example 9.2.1. Consider a family composed of the father (55 years old), mother (53 years old) and sun (28 years old). Compute the single premium payable by the family for a single claim of 25000 u.c. over 30 years if all the members will be alive. The annual interest percent is 10%.

Solution. The premium is

$$\begin{aligned}
25000 \cdot {}_{30}E_{55,53,28} &= 25000 \cdot \frac{D_{85,83,58}}{D_{55,53,28}} = 25000 \cdot \frac{l_{85} \cdot l_{83} \cdot l_{58} \cdot v^{\frac{85+83+58}{3}}}{l_{55} \cdot l_{53} \cdot l_{28} \cdot v^{\frac{55+53+28}{3}}} \\
&= 25000 \cdot \frac{4942 \cdot 14757 \cdot 75633 \cdot \left(\dfrac{1}{1.1}\right)^{\frac{85+83+58}{3}}}{80556 \cdot 93112 \cdot 98880 \cdot \left(\dfrac{1}{1.1}\right)^{\frac{55+53+28}{3}}} \\
&\simeq 10.65 \text{ u.c.}
\end{aligned}$$

□

9.3 Single claims for partial survival

Definition 9.3.1. *Let a group of m persons having the ages $x_1, x_2, \ldots, x_m$, where $m \in \mathbb{N}^*$, $x_j \in \mathbb{N}$, $x_j \leq \omega\ \forall j \in \{1, \ldots, m\}$. Let $n, k \in \mathbb{N}$ s.t. $k \leq m$. We denote*

- ${}_nE_{\overline{x_1,x_2,\ldots,x_m}^{[k]}}$ = *the single premium payable by the group for a single claim of 1 u.c. over n years if exactly k of the members survive;*
- ${}_nE_{\overline{x_1,x_2,\ldots,x_m}^{k}}$ = *the single premium payable by the group for a single claim of 1 u.c. over n years if at least k of the members survive.*

Proposition 9.3.1. *We have*

$$ {}_nE_{\overline{x_1,x_2,\ldots,x_m}^{[k]}} = \sum_{s=0}^{m-k} (-1)^s \cdot C_{k+s}^{s} \sum_{1 \leq i_1 < \ldots < i_{k+s} \leq m} {}_nE_{x_{i_1},x_{i_2},\ldots,x_{i_{k+s}}}; \tag{9.6} $$

$$ {}_nE_{\overline{x_1,x_2,\ldots,x_m}^{k}} = \sum_{s=0}^{m-k} (-1)^s \cdot C_{k+s-1}^{s} \sum_{1 \leq i_1 < \ldots < i_{k+s} \leq m} {}_nE_{x_{i_1},x_{i_2},\ldots,x_{i_{k+s}}}. \tag{9.7} $$

Corollary 9.3.1. *Let a group of m persons having the ages $x_1, x_2, \ldots, x_m$, where $m \in \mathbb{N}^*$, $x_j \in \mathbb{N}$, $x_j \leq \omega\ \forall j \in \{1, \ldots, m\}$. Let $n, k \in \mathbb{N}$ s.t. $k \leq m$ and let $T \geq 0$.*

1. *The single premium payable by the group for a single claim of T u.c. over n years if exactly k of the members survive is*

$$ T \cdot {}_nE_{\overline{x_1,x_2,\ldots,x_m}^{[k]}} = T \cdot \sum_{s=0}^{m-k} (-1)^s \cdot C_{k+s}^{s} \sum_{1 \leq i_1 < \ldots < i_{k+s} \leq m} {}_nE_{x_{i_1},x_{i_2},\ldots,x_{i_{k+s}}}. $$

2. *The single premium payable by the group for a single claim of T u.c. over n years if at least k of the members survive is*

$$ T \cdot {}_nE_{\overline{x_1,x_2,\ldots,x_m}^{k}} = T \cdot \sum_{s=0}^{m-k} (-1)^s \cdot C_{k+s-1}^{s} \sum_{1 \leq i_1 < \ldots < i_{k+s} \leq m} {}_nE_{x_{i_1},x_{i_2},\ldots,x_{i_{k+s}}}. $$

Example 9.3.1. Consider a family composed of the father (55 years old), mother (53 years old) and sun (28 years old). Compute the single premium payable by the family for a single claim of 25000 u.c. over 30 years if:
a) just one member will be alive;
b) exactly two of the members will be alive;

c) at least one member will be alive;
d) at least two of the members will be alive.
The annual interest percent is 10%.

Solution. a) The premium is

$$25000 \cdot {}_{30}E_{\overline{55,53,28}}^{[1]} = 25000 \cdot \sum_{s=0}^{2}(-1)^s \cdot C_{1+s}^{s} \sum_{1\le i_1<\ldots<i_{1+s}\le 3} {}_{30}E_{x_{i_1},x_{i_2},\ldots,x_{i_{1+s}}}$$

$$= 25000\left(C_1^0 \sum_{1\le i_1\le 3} {}_{30}E_{x_{i_1}} - C_2^1 \sum_{1\le i_1<i_2\le 3} {}_{30}E_{x_{i_1},x_{i_2}} + C_3^2 \sum_{1\le i_1<i_2<i_3\le 3} {}_{30}E_{x_{i_1},x_{i_2},x_{i_3}}\right)$$

$$= 25000\Big[C_1^0({}_{30}E_{x_1} + {}_{30}E_{x_2} + {}_{30}E_{x_3}) - C_2^1({}_{30}E_{x_1,x_2} + {}_{30}E_{x_1,x_3} + {}_{30}E_{x_2,x_3}) + C_3^2 \cdot {}_{30}E_{x_1,x_2,x_3}\Big]$$

$$= 25000\Big[{}_{30}E_{55} + {}_{30}E_{53} + {}_{30}E_{28} - 2({}_{30}E_{55,53} + {}_{30}E_{55,28} + {}_{30}E_{53,28}) + 3 \cdot {}_{30}E_{55,53,28}\Big]$$

$$= 25000\left[\frac{D_{85}}{D_{55}} + \frac{D_{83}}{D_{53}} + \frac{D_{58}}{D_{28}} - 2\left(\frac{D_{85,83}}{D_{55,53}} + \frac{D_{85,58}}{D_{55,28}} + \frac{D_{83,58}}{D_{53,28}}\right) + 3 \cdot \frac{D_{85,83,58}}{D_{55,53,28}}\right]$$

$$= 25000\Bigg[\frac{l_{85} \cdot v^{85}}{l_{55} \cdot v^{55}} + \frac{l_{83} \cdot v^{83}}{l_{53} \cdot v^{53}} + \frac{l_{58} \cdot v^{58}}{l_{28} \cdot v^{28}} - 2\left(\frac{l_{85} \cdot l_{83} \cdot v^{\frac{85+83}{2}}}{l_{55} \cdot l_{53} \cdot v^{\frac{55+53}{2}}} + \frac{l_{85} \cdot l_{58} \cdot v^{\frac{85+58}{2}}}{l_{55} \cdot l_{28} \cdot v^{\frac{55+28}{2}}} + \frac{l_{83} \cdot l_{58} \cdot v^{\frac{83+58}{2}}}{l_{53} \cdot l_{28} \cdot v^{\frac{53+28}{2}}}\right)$$

$$+ 3 \cdot \frac{l_{85} \cdot l_{83} \cdot l_{58} \cdot v^{\frac{85+83+58}{3}}}{l_{55} \cdot l_{53} \cdot l_{28} \cdot v^{\frac{55+53+28}{3}}}\Bigg].$$

By replacing

$$v = \frac{1}{1.1},$$
$$l_{85} = 4942,\ l_{55} = 80556 \text{ (MALE)},$$
$$l_{83} = 14757,\ l_{53} = 93112 \text{ (FEMALE)},$$
$$l_{58} = 75633,\ l_{28} = 98880 \text{ (MALE)},$$

we obtain that

$$25000 \cdot {}_{30}E_{\overline{55,53,28}}^{[1]} \simeq 933.11 \text{ u.c.}$$

b) The premium is

$$25000 \cdot {}_{30}E_{\overline{55,53,28}}^{[2]} = 25000 \cdot \sum_{s=0}^{1}(-1)^s \cdot C_{2+s}^{s} \sum_{1\le i_1<\ldots<i_{2+s}\le 3} {}_{30}E_{x_{i_1},x_{i_2},\ldots,x_{i_{2+s}}}$$

$$= 25000\left(C_2^0 \sum_{1\le i_1<i_2\le 3} {}_{30}E_{x_{i_1},x_{i_2}} - C_3^1 \sum_{1\le i_1<i_2<i_3\le 3} {}_{30}E_{x_{i_1},x_{i_2},x_{i_3}}\right)$$

$$= 25000\Big[C_2^0({}_{30}E_{x_1,x_2} + {}_{30}E_{x_1,x_3} + {}_{30}E_{x_2,x_3}) - C_3^1 \cdot {}_{30}E_{x_1,x_2,x_3}\Big]$$
$$= 25000\Big({}_{30}E_{55,53} + {}_{30}E_{55,28} + {}_{30}E_{53,28} - 3 \cdot {}_{30}E_{55,53,28}\Big)$$
$$= 25000\left(\frac{D_{85,83}}{D_{55,53}} + \frac{D_{85,58}}{D_{55,28}} + \frac{D_{83,58}}{D_{53,28}} - 3 \cdot \frac{D_{85,83,58}}{D_{55,53,28}}\right)$$
$$= 25000\left(\frac{l_{85} \cdot l_{83} \cdot v^{\frac{85+83}{2}}}{l_{55} \cdot l_{53} \cdot v^{\frac{55+53}{2}}} + \frac{l_{85} \cdot l_{58} \cdot v^{\frac{85+58}{2}}}{l_{55} \cdot l_{28} \cdot v^{\frac{55+28}{2}}} + \frac{l_{83} \cdot l_{58} \cdot v^{\frac{83+58}{2}}}{l_{53} \cdot l_{28} \cdot v^{\frac{53+28}{2}}} - 3 \cdot \frac{l_{85} \cdot l_{83} \cdot l_{58} \cdot v^{\frac{85+83+58}{3}}}{l_{55} \cdot l_{53} \cdot l_{28} \cdot v^{\frac{55+53+28}{3}}}\right)$$
$$\simeq 222.87 \text{ u.c.}$$

c) The premium is

$$25000 \cdot {}_{30}E_{\overline{55,53,28}^{\,1}} = 25000 \cdot \sum_{s=0}^{2}(-1)^s \cdot C_s^s \sum_{1\le i_1<\ldots<i_{1+s}\le 3} {}_{30}E_{x_{i_1},x_{i_2},\ldots,x_{i_{1+s}}}$$
$$= 25000\left(C_0^0 \sum_{1\le i_1\le 3} {}_{30}E_{x_{i_1}} - C_1^1 \sum_{1\le i_1<i_2\le 3} {}_{30}E_{x_{i_1},x_{i_2}} + C_2^2 \sum_{1\le i_1<i_2<i_3\le 3} {}_{30}E_{x_{i_1},x_{i_2},x_{i_3}}\right)$$
$$= 25000\Big[C_0^0({}_{30}E_{x_1} + {}_{30}E_{x_2} + {}_{30}E_{x_3}) - C_1^1({}_{30}E_{x_1,x_2} + {}_{30}E_{x_1,x_3} + {}_{30}E_{x_2,x_3}) + C_2^2 \cdot {}_{30}E_{x_1,x_2,x_3}\Big]$$
$$= 25000\Big[{}_{30}E_{55} + {}_{30}E_{53} + {}_{30}E_{28} - ({}_{30}E_{55,53} + {}_{30}E_{55,28} + {}_{30}E_{53,28}) + {}_{30}E_{55,53,28}\Big]$$
$$= 25000\left[\frac{D_{85}}{D_{55}} + \frac{D_{83}}{D_{53}} + \frac{D_{58}}{D_{28}} - \left(\frac{D_{85,83}}{D_{55,53}} + \frac{D_{85,58}}{D_{55,28}} + \frac{D_{83,58}}{D_{53,28}}\right) + \frac{D_{85,83,58}}{D_{55,53,28}}\right]$$
$$= 25000\Bigg[\frac{l_{85} \cdot v^{85}}{l_{55} \cdot v^{55}} + \frac{l_{83} \cdot v^{83}}{l_{53} \cdot v^{53}} + \frac{l_{58} \cdot v^{58}}{l_{28} \cdot v^{28}} - \left(\frac{l_{85} \cdot l_{83} \cdot v^{\frac{85+83}{2}}}{l_{55} \cdot l_{53} \cdot v^{\frac{55+53}{2}}} + \frac{l_{85} \cdot l_{58} \cdot v^{\frac{85+58}{2}}}{l_{55} \cdot l_{28} \cdot v^{\frac{55+28}{2}}} + \frac{l_{83} \cdot l_{58} \cdot v^{\frac{83+58}{2}}}{l_{53} \cdot l_{28} \cdot v^{\frac{53+28}{2}}}\right)$$
$$+ \frac{l_{85} \cdot l_{83} \cdot l_{58} \cdot v^{\frac{85+83+58}{3}}}{l_{55} \cdot l_{53} \cdot l_{28} \cdot v^{\frac{55+53+28}{3}}}\Bigg]$$
$$\simeq 1166.65 \text{ u.c.}$$

d) The premium is

$$25000 \cdot {}_{30}E_{\overline{55,53,28}^{\,2}} = 25000 \cdot \sum_{s=0}^{1}(-1)^s \cdot C_{1+s}^s \sum_{1\le i_1<\ldots<i_{2+s}\le 3} {}_{30}E_{x_{i_1},x_{i_2},\ldots,x_{i_{2+s}}}$$
$$= 25000\left(C_1^0 \sum_{1\le i_1<i_2\le 3} {}_{30}E_{x_{i_1},x_{i_2}} - C_2^1 \sum_{1\le i_1<i_2<i_3\le 3} {}_{30}E_{x_{i_1},x_{i_2},x_{i_3}}\right)$$
$$= 25000\Big[C_1^0({}_{30}E_{x_1,x_2} + {}_{30}E_{x_1,x_3} + {}_{30}E_{x_2,x_3}) - C_2^1 \cdot {}_{30}E_{x_1,x_2,x_3}\Big]$$
$$= 25000\Big({}_{30}E_{55,53} + {}_{30}E_{55,28} + {}_{30}E_{53,28} - 2 \cdot {}_{30}E_{55,53,28}\Big)$$

$$= 25000\left(\frac{D_{85,83}}{D_{55,53}}+\frac{D_{85,58}}{D_{55,28}}+\frac{D_{83,58}}{D_{53,28}}-2\cdot\frac{D_{85,83,58}}{D_{55,53,28}}\right)$$
$$= 25000\left(\frac{l_{85}\cdot l_{83}\cdot v^{\frac{85+83}{2}}}{l_{55}\cdot l_{53}\cdot v^{\frac{55+53}{2}}}+\frac{l_{85}\cdot l_{58}\cdot v^{\frac{85+58}{2}}}{l_{55}\cdot l_{28}\cdot v^{\frac{55+28}{2}}}+\frac{l_{83}\cdot l_{58}\cdot v^{\frac{83+58}{2}}}{l_{53}\cdot l_{28}\cdot v^{\frac{53+28}{2}}}-2\cdot\frac{l_{85}\cdot l_{83}\cdot l_{58}\cdot v^{\frac{85+83+58}{3}}}{l_{55}\cdot l_{53}\cdot l_{28}\cdot v^{\frac{55+53+28}{3}}}\right)$$
$$\simeq 233.53 \text{ u.c.}$$

□

9.4 Whole life annuities for joint survival

Definition 9.4.1. *Let a group of m persons having the ages $x_1, x_2, \ldots, x_m$, where $m \in \mathbb{N}^*$, $x_j \in \mathbb{N}$, $x_j \leq \omega\ \forall j \in \{1, \ldots, m\}$. We denote*

- $a_{x_1,x_2,\ldots,x_m}$ = *the single premium payable by the group for a whole joint-life annuity-immediate of* 1 *u.c. per year (payable at the end of each year while all of the members survive).*

Proposition 9.4.1. *We have*

$$a_{x_1,x_2,\ldots,x_m} = \frac{N_{x_1+1,x_2+1,\ldots,x_m+1}}{D_{x_1,x_2,\ldots,x_m}}, \tag{9.8}$$

where

$$N_{x_1,x_2,\ldots,x_m} = \sum_{n=0}^{\omega-\widetilde{x}} D_{x_1+n,x_2+n,\ldots,x_m+n}, \tag{9.9}$$

$\widetilde{x} = \max\{x_1, x_2, \ldots, x_m\}$ *being the maximum age of the group.*

Corollary 9.4.1. *Let a group of m persons having the ages $x_1, x_2, \ldots, x_m$, where $m \in \mathbb{N}^*$, $x_j \in \mathbb{N}$, $x_j \leq \omega\ \forall j \in \{1, \ldots, m\}$. Let $T \geq 0$. The single premium payable by the group for a whole joint-life annuity-immediate of T u.c. per year is*

$$T \cdot a_{x_1,x_2,\ldots,x_m} = T \cdot \frac{N_{x_1+1,x_2+1,\ldots,x_m+1}}{D_{x_1,x_2,\ldots,x_m}}.$$

Example 9.4.1. Consider a family composed of the father (55 years old), mother (53 years old) and sun (28 years old). Compute the single premium payable by the family for a life annuity-immediate of 250 u.c. per year while all of the members survive. The annual interest percent is 10%.

Solution. The premium is

$$\begin{aligned}
&250 \cdot a_{55,53,28} = 250 \cdot \frac{N_{56,54,29}}{D_{55,53,28}} \\
&= 250 \cdot \frac{D_{56,54,29} + D_{57,55,30} + D_{58,56,31} + \cdots + D_{100,98,73}}{D_{55,53,28}} \\
&= 250 \cdot \frac{l_{56} \cdot l_{54} \cdot l_{29} \cdot v^{\frac{56+54+29}{3}} + l_{57} \cdot l_{55} \cdot l_{30} \cdot v^{\frac{57+55+30}{3}} + \cdots + l_{100} \cdot l_{98} \cdot l_{73} \cdot v^{\frac{100+98+73}{3}}}{l_{55} \cdot l_{53} \cdot l_{28} \cdot v^{\frac{55+53+28}{3}}} \\
&\simeq 1606.06 \text{ u.c.}
\end{aligned}$$

□

9.5 Whole life annuities for partial survival

Definition 9.5.1. *Let a group of m persons having the ages $x_1, x_2, \ldots, x_m$, where $m \in \mathbb{N}^*$, $x_j \in \mathbb{N}$, $x_j \leq \omega$ $\forall j \in \{1, \ldots, m\}$. Let $k \in \mathbb{N}$ s.t. $k \leq m$. We denote*

- $a_{\overline{x_1, x_2, \ldots, x_m}^{[k]}}$ = *the single premium payable by the group for a whole life annuity-immediate of 1 u.c. per year payable (at the end of each year) while exactly k of the members survive;*
- $a_{\overline{x_1, x_2, \ldots, x_m}^{k}}$ = *the single premium payable by the group for a whole life annuity-immediate of 1 u.c. per year payable (at the end of each year) while at least k of the members survive.*

Proposition 9.5.1. *We have*

$$a_{\overline{x_1, x_2, \ldots, x_m}^{[k]}} = \sum_{s=0}^{m-k} (-1)^s \cdot C_{k+s}^{s} \sum_{1 \leq i_1 < \ldots < i_{k+s} \leq m} a_{x_{i_1}, x_{i_2}, \ldots, x_{i_{k+s}}}; \tag{9.10}$$

$$a_{\overline{x_1, x_2, \ldots, x_m}^{k}} = \sum_{s=0}^{m-k} (-1)^s \cdot C_{k+s-1}^{s} \sum_{1 \leq i_1 < \ldots < i_{k+s} \leq m} a_{x_{i_1}, x_{i_2}, \ldots, x_{i_{k+s}}}. \tag{9.11}$$

Corollary 9.5.1. *Let a group of m persons having the ages $x_1, x_2, \ldots, x_m$, where $m \in \mathbb{N}^*$, $x_j \in \mathbb{N}$, $x_j \leq \omega$ $\forall j \in \{1, \ldots, m\}$. Let $k \in \mathbb{N}$ s.t. $k \leq m$ and let $T \geq 0$.*

1. *The single premium payable by the group for a whole life annuity-immediate of T u.c. per year payable while exactly k of the members survive is*

$$T \cdot a_{\overline{x_1,x_2,\ldots,x_m}}^{[k]} = T \cdot \sum_{s=0}^{m-k} (-1)^s \cdot C_{k+s}^s \sum_{1 \le i_1 < \ldots < i_{k+s} \le m} a_{x_{i_1},x_{i_2},\ldots,x_{i_{k+s}}}.$$

2. *The single premium payable by the group for a whole life annuity-immediate of T u.c. per year payable while at least k of the members survive is*

$$T \cdot a_{\overline{x_1,x_2,\ldots,x_m}}^{k} = T \cdot \sum_{s=0}^{m-k} (-1)^s \cdot C_{k+s-1}^s \sum_{1 \le i_1 < \ldots < i_{k+s} \le m} a_{x_{i_1},x_{i_2},\ldots,x_{i_{k+s}}}.$$

Example 9.5.1. Consider a family composed of the father (55 years old), mother (53 years old) and sun (28 years old). Compute the single premium payable by the family for a life annuity-immediate of 250 u.c. per year while:
a) just one member survives;
b) exactly two of the members survive;
c) at least one member survives;
d) at least two of the members survive.
The annual interest percent is 10%.

Solution. a) The premium is

$$\begin{aligned}
250 \cdot a_{\overline{55,53,28}}^{[1]} &= 250 \cdot \sum_{s=0}^{2} (-1)^s \cdot C_{1+s}^s \sum_{1 \le i_1 < \ldots < i_{1+s} \le 3} a_{x_{i_1},x_{i_2},\ldots,x_{i_{1+s}}} \\
&= 250 \left(C_1^0 \sum_{1 \le i_1 \le 3} a_{x_{i_1}} - C_2^1 \sum_{1 \le i_1 < i_2 \le 3} a_{x_{i_1},x_{i_2}} + C_3^2 \sum_{1 \le i_1 < i_2 < i_3 \le 3} a_{x_{i_1},x_{i_2},x_{i_3}} \right) \\
&= 250 \Big[C_1^0 (a_{x_1} + a_{x_2} + a_{x_3}) - C_2^1 (a_{x_1,x_2} + a_{x_1,x_3} + a_{x_2,x_3}) + C_3^2 \cdot a_{x_1,x_2,x_3} \Big] \\
&= 250 \Big[a_{55} + a_{53} + a_{28} - 2(a_{55,53} + a_{55,28} + a_{53,28}) + 3 \cdot a_{55,53,28} \Big] \\
&\simeq 248.29 \text{ u.c.}
\end{aligned}$$

b) The premium is

$$250 \cdot a_{\overline{55,53,28}}^{[2]} = 250 \cdot \sum_{s=0}^{1} (-1)^s \cdot C_{2+s}^s \sum_{1 \le i_1 < \ldots < i_{2+s} \le 3} a_{x_{i_1},x_{i_2},\ldots,x_{i_{2+s}}}$$

$$
\begin{aligned}
&= 250\left(C_2^0 \sum_{1\le i_1<i_2\le 3} a_{x_{i_1},x_{i_2}} - C_3^1 \sum_{1\le i_1<i_2<i_3\le 3} a_{x_{i_1},x_{i_2},x_{i_3}}\right)\\
&= 250\Big[C_2^0(a_{x_1,x_2}+a_{x_1,x_3}+a_{x_2,x_3}) - C_3^1\cdot a_{x_1,x_2,x_3}\Big]\\
&= 250\Big(a_{55,53}+a_{55,28}+a_{53,28}-3\cdot a_{55,53,28}\Big)\\
&\simeq 555.85 \text{ u.c.}
\end{aligned}
$$

c) The premium is

$$
\begin{aligned}
250\cdot a_{\overline{55,53,28}}^{\;1} &= 250\cdot\sum_{s=0}^{2}(-1)^s\cdot C_s^s \sum_{1\le i_1<\ldots<i_{1+s}\le 3} a_{x_{i_1},x_{i_2},\ldots,x_{i_{1+s}}}\\
&= 250\left(C_0^0\sum_{1\le i_1\le 3} a_{x_{i_1}} - C_1^1 \sum_{1\le i_1<i_2\le 3} a_{x_{i_1},x_{i_2}} + C_2^2 \sum_{1\le i_1<i_2<i_3\le 3} a_{x_{i_1},x_{i_2},x_{i_3}}\right)\\
&= 250\Big[C_0^0(a_{x_1}+a_{x_2}+a_{x_3}) - C_1^1(a_{x_1,x_2}+a_{x_1,x_3}+a_{x_2,x_3}) + C_2^2\cdot a_{x_1,x_2,x_3}\Big]\\
&= 250\Big[a_{55}+a_{53}+a_{28}-(a_{55,53}+a_{55,28}+a_{53,28})+a_{55,53,28}\Big]\\
&\simeq 2410.21 \text{ u.c.}
\end{aligned}
$$

d) The premium is

$$
\begin{aligned}
250\cdot a_{\overline{55,53,28}}^{\;2} &= 250\cdot\sum_{s=0}^{1}(-1)^s\cdot C_{1+s}^s \sum_{1\le i_1<\ldots<i_{2+s}\le 3} a_{x_{i_1},x_{i_2},\ldots,x_{i_{2+s}}}\\
&= 250\left(C_1^0 \sum_{1\le i_1<i_2\le 3} a_{x_{i_1},x_{i_2}} - C_2^1 \sum_{1\le i_1<i_2<i_3\le 3} a_{x_{i_1},x_{i_2},x_{i_3}}\right)\\
&= 250\Big[C_1^0(a_{x_1,x_2}+a_{x_1,x_3}+a_{x_2,x_3}) - C_2^1\cdot a_{x_1,x_2,x_3}\Big]\\
&= 250\Big(a_{55,53}+a_{55,28}+a_{53,28}-2\cdot a_{55,53,28}\Big)\\
&\simeq 2161.91 \text{ u.c.}
\end{aligned}
$$

□

9.6 Deferred whole life annuities for joint survival

Definition 9.6.1. *Let a group of m persons having the ages $x_1, x_2, \ldots, x_m$, where $m \in \mathbb{N}^*$, $x_j \in \mathbb{N}$, $x_j \le \omega\ \forall j \in \{1, \ldots, m\}$. Let $r \in \mathbb{N}$ s.t. $\widetilde{x} + r \le \omega$.*

We denote

- ${}_{r|}a_{x_1,x_2,\ldots,x_m}$ = *the single premium payable by the group for an r-year deferred whole joint-life annuity-immediate of* 1 *u.c. per year (payable after r-years, at the end of each year while all of the members survive).*

Proposition 9.6.1. *We have*

$$ {}_{r|}a_{x_1,x_2,\ldots,x_m} = \frac{N_{x_1+r+1,x_2+r+1,\ldots,x_m+r+1}}{D_{x_1,x_2,\ldots,x_m}}. \tag{9.12}$$

Corollary 9.6.1. *Let a group of m persons having the ages $x_1, x_2, \ldots, x_m$, where $m \in \mathbb{N}^*$, $x_j \in \mathbb{N}$, $x_j \le \omega\ \forall j \in \{1, \ldots, m\}$. Let $r \in \mathbb{N}$ s.t. $\widetilde{x}+r \le \omega$ and let $T \ge 0$. The single premium payable by the group for an r-year deferred whole joint-life annuity-immediate of T u.c. per year is*

$$T \cdot {}_{r|}a_{x_1,x_2,\ldots,x_m} = T \cdot \frac{N_{x_1+r+1,x_2+r+1,\ldots,x_m+r+1}}{D_{x_1,x_2,\ldots,x_m}}.$$

Example 9.6.1. Consider a family composed of the father (55 years old), mother (53 years old) and sun (28 years old). Compute the single premium payable by the family for a 20-year deferred life annuity-immediate of 250 u.c. per year while all of the members survive. The annual interest percent is 10%.

Solution. The premium is

$$\begin{aligned}
&250 \cdot {}_{20|}a_{55,53,28} = 250 \cdot \frac{N_{76,74,49}}{D_{55,53,28}} \\
&= 250 \cdot \frac{D_{76,74,49} + D_{77,75,50} + D_{78,76,51} + \cdots + D_{100,98,73}}{D_{55,53,28}} \\
&= 250 \cdot \frac{l_{76} \cdot l_{74} \cdot l_{49} \cdot v^{\frac{76+74+49}{3}} + l_{77} \cdot l_{75} \cdot l_{50} \cdot v^{\frac{77+75+50}{3}} + \cdots + l_{100} \cdot l_{98} \cdot l_{73} \cdot v^{\frac{100+98+73}{3}}}{l_{55} \cdot l_{53} \cdot l_{28} \cdot v^{\frac{55+53+28}{3}}} \\
&\simeq 19.94 \text{ u.c.}
\end{aligned}$$

□

9.7 Deferred whole life annuities for partial survival

Definition 9.7.1. *Let a group of m persons having the ages $x_1, x_2, \ldots, x_m$, where $m \in \mathbb{N}^*$, $x_j \in \mathbb{N}$, $x_j \le \omega\ \forall j \in \{1, \ldots, m\}$. Let $r \in \mathbb{N}$ s.t. $\widehat{x} + r \le \omega$ and let $k \in \mathbb{N}$ s.t. $k \le m$. We denote*

- ${}_{r|}a_{\overline{x_1,x_2,\ldots,x_m}}^{[k]}$ = *the single premium payable by the group for an r-year deferred whole life annuity-immediate of* 1 *u.c. per year payable (after r-years, at the end of each year) while exactly k of the members survive;*
- ${}_{r|}a_{\overline{x_1,x_2,\ldots,x_m}}^{k}$ = *the single premium payable by the group for an r-year deferred whole life annuity-immediate of* 1 *u.c. per year payable (after r-years, at the end of each year) while at least k of the members survive.*

Proposition 9.7.1. *We have*

$$ {}_{r|}a_{\overline{x_1,x_2,\ldots,x_m}}^{[k]} = \sum_{s=0}^{m-k}(-1)^s \cdot C_{k+s}^{s} \sum_{1\le i_1<\ldots<i_{k+s}\le m} {}_{r|}a_{x_{i_1},x_{i_2},\ldots,x_{i_{k+s}}}; \tag{9.13}$$

$$ {}_{r|}a_{\overline{x_1,x_2,\ldots,x_m}}^{k} = \sum_{s=0}^{m-k}(-1)^s \cdot C_{k+s-1}^{s} \sum_{1\le i_1<\ldots<i_{k+s}\le m} {}_{r|}a_{x_{i_1},x_{i_2},\ldots,x_{i_{k+s}}}. \tag{9.14}$$

Corollary 9.7.1. *Let a group of m persons having the ages $x_1, x_2, \ldots, x_m$, where $m \in \mathbb{N}^*$, $x_j \in \mathbb{N}$, $x_j \le \omega$ $\forall j \in \{1, \ldots, m\}$. Let $r \in \mathbb{N}$ s.t. $\widehat{x} + r \le \omega$ and let $k \in \mathbb{N}$ s.t. $k \le m$. Let $T \ge 0$.*

1. *The single premium payable by the group for an r-year deferred whole life annuity-immediate of T u.c. per year payable while exactly k of the members survive is*
$$T \cdot {}_{r|}a_{\overline{x_1,x_2,\ldots,x_m}}^{[k]} = T \cdot \sum_{s=0}^{m-k}(-1)^s \cdot C_{k+s}^{s} \sum_{1\le i_1<\ldots<i_{k+s}\le m} {}_{r|}a_{x_{i_1},x_{i_2},\ldots,x_{i_{k+s}}}.$$
2. *The single premium payable by the group for an r-year deferred whole life annuity-immediate of T u.c. per year payable while at least k of the members survive is*
$$T \cdot {}_{r|}a_{\overline{x_1,x_2,\ldots,x_m}}^{k} = T \cdot \sum_{s=0}^{m-k}(-1)^s \cdot C_{k+s-1}^{s} \sum_{1\le i_1<\ldots<i_{k+s}\le m} {}_{r|}a_{x_{i_1},x_{i_2},\ldots,x_{i_{k+s}}}.$$

Example 9.7.1. Consider a family composed of the father (55 years old), mother (53 years old) and sun (28 years old). Compute the single premium payable by the family for a 20-year deferred life annuity-immediate of 250 u.c. per year while:
a) just one member survives;
b) exactly two of the members survive;
c) at least one member survives;
d) at least two of the members survive.
The annual interest percent is 10%.

Solution. a) The premium is

$$250 \cdot {}_{20|}a_{\overline{55,53,28}}^{[1]} = 250 \cdot \sum_{s=0}^{2} (-1)^s \cdot C_{1+s}^{s} \sum_{1 \le i_1 < \ldots < i_{1+s} \le 3} {}_{20|}a_{x_{i_1}, x_{i_2}, \ldots, x_{i_{1+s}}}$$

$$= 250 \left(C_1^0 \sum_{1 \le i_1 \le 3} {}_{20|}a_{x_{i_1}} - C_2^1 \sum_{1 \le i_1 < i_2 \le 3} {}_{20|}a_{x_{i_1}, x_{i_2}} + C_3^2 \sum_{1 \le i_1 < i_2 < i_3 \le 3} {}_{20|}a_{x_{i_1}, x_{i_2}, x_{i_3}} \right)$$

$$= 250 \Big[C_1^0 ({}_{20|}a_{x_1} + {}_{20|}a_{x_2} + {}_{20|}a_{x_3}) - C_2^1 ({}_{20|}a_{x_1,x_2} + {}_{20|}a_{x_1,x_3} + {}_{20|}a_{x_2,x_3}) + C_3^2 \cdot {}_{20|}a_{x_1,x_2,x_3} \Big]$$

$$= 250 \Big[{}_{20|}a_{55} + {}_{20|}a_{53} + {}_{20|}a_{28} - 2({}_{20|}a_{55,53} + {}_{20|}a_{55,28} + {}_{20|}a_{53,28}) + 3 \cdot {}_{20|}a_{55,53,28} \Big]$$

$$\simeq 175.35 \text{ u.c.}$$

b) The premium is

$$250 \cdot {}_{20|}a_{\overline{55,53,28}}^{[2]} = 250 \cdot \sum_{s=0}^{1} (-1)^s \cdot C_{2+s}^{s} \sum_{1 \le i_1 < \ldots < i_{2+s} \le 3} {}_{20|}a_{x_{i_1}, x_{i_2}, \ldots, x_{i_{2+s}}}$$

$$= 250 \left(C_2^0 \sum_{1 \le i_1 < i_2 \le 3} {}_{20|}a_{x_{i_1}, x_{i_2}} - C_3^1 \sum_{1 \le i_1 < i_2 < i_3 \le 3} {}_{20|}a_{x_{i_1}, x_{i_2}, x_{i_3}} \right)$$

$$= 250 \Big[C_2^0 ({}_{20|}a_{x_1,x_2} + {}_{20|}a_{x_1,x_3} + {}_{20|}a_{x_2,x_3}) - C_3^1 \cdot {}_{20|}a_{x_1,x_2,x_3} \Big]$$

$$= 250 \Big({}_{20|}a_{55,53} + {}_{20|}a_{55,28} + {}_{20|}a_{53,28} - 3 \cdot {}_{20|}a_{55,53,28} \Big)$$

$$\simeq 89.99 \text{ u.c.}$$

c) The premium is

$$250 \cdot {}_{20|}a_{\overline{55,53,28}}^{1} = 250 \cdot \sum_{s=0}^{2} (-1)^s \cdot C_{s}^{s} \sum_{1 \le i_1 < \ldots < i_{1+s} \le 3} {}_{20|}a_{x_{i_1}, x_{i_2}, \ldots, x_{i_{1+s}}}$$

$$= 250 \left(C_0^0 \sum_{1 \le i_1 \le 3} {}_{20|}a_{x_{i_1}} - C_1^1 \sum_{1 \le i_1 < i_2 \le 3} {}_{20|}a_{x_{i_1}, x_{i_2}} + C_2^2 \sum_{1 \le i_1 < i_2 < i_3 \le 3} {}_{20|}a_{x_{i_1}, x_{i_2}, x_{i_3}} \right)$$

$$= 250 \Big[C_0^0 ({}_{20|}a_{x_1} + {}_{20|}a_{x_2} + {}_{20|}a_{x_3}) - C_1^1 ({}_{20|}a_{x_1,x_2} + {}_{20|}a_{x_1,x_3} + {}_{20|}a_{x_2,x_3}) + C_2^2 \cdot {}_{20|}a_{x_1,x_2,x_3} \Big]$$

$$= 250 \Big[{}_{20|}a_{55} + {}_{20|}a_{53} + {}_{20|}a_{28} - ({}_{20|}a_{55,53} + {}_{20|}a_{55,28} + {}_{20|}a_{53,28}) + {}_{20|}a_{55,53,28} \Big]$$

$$\simeq 285.29 \text{ u.c.}$$

d) The premium is

$$250 \cdot {}_{20|}a_{\overline{55,53,28}}^{2} = 250 \cdot \sum_{s=0}^{1} (-1)^s \cdot C_{1+s}^{s} \sum_{1 \le i_1 < \ldots < i_{2+s} \le 3} {}_{20|}a_{x_{i_1}, x_{i_2}, \ldots, x_{i_{2+s}}}$$

$$= 250\left(C_1^0 \sum_{1\le i_1<i_2\le 3} {}_{20|}a_{x_{i_1},x_{i_2}} - C_2^1 \sum_{1\le i_1<i_2<i_3\le 3} {}_{20|}a_{x_{i_1},x_{i_2},x_{i_3}}\right)$$
$$= 250\Big[C_1^0({}_{20|}a_{x_1,x_2} + {}_{20|}a_{x_1,x_3} + {}_{20|}a_{x_2,x_3}) - C_2^1 \cdot {}_{20|}a_{x_1,x_2,x_3}\Big]$$
$$= 250\Big({}_{20|}a_{55,53} + {}_{20|}a_{55,28} + {}_{20|}a_{53,28} - 2\cdot {}_{20|}a_{55,53,28}\Big)$$
$$\simeq 109.94 \text{ u.c.}$$

□

9.8 Temporary life annuities for joint survival

Definition 9.8.1. *Let a group of m persons having the ages $x_1, x_2, \ldots, x_m$, where $m \in \mathbb{N}^*$, $x_j \in \mathbb{N}$, $x_j \le \omega\ \forall j \in \{1, \ldots, m\}$. Let $r \in \mathbb{N}$ s.t. $\widetilde{x} + r \le \omega$. We denote*

- $a_{x_1,x_2,\ldots,x_m:\overline{r}|}$ = *the single premium payable by the group for an r-year temporary joint-life annuity-immediate of 1 u.c. per year (payable at the end of each year while all of the members survive during the next r years).*

Proposition 9.8.1. *We have*

$$a_{x_1,x_2,\ldots,x_m:\overline{r}|} = \frac{N_{x_1+1,x_2+1,\ldots,x_m+1} - N_{x_1+r+1,x_2+r+1,\ldots,x_m+r+1}}{D_{x_1,x_2,\ldots,x_m}}. \tag{9.15}$$

Corollary 9.8.1. *Let a group of m persons having the ages $x_1, x_2, \ldots, x_m$, where $m \in \mathbb{N}^*$, $x_j \in \mathbb{N}$, $x_j \le \omega\ \forall j \in \{1, \ldots, m\}$. Let $r \in \mathbb{N}$ s.t. $\widetilde{x}+r \le \omega$ and let $T \ge 0$. The single premium payable by the group for an r-year temporary joint-life annuity-immediate of T u.c. per year is*

$$T \cdot a_{x_1,x_2,\ldots,x_m:\overline{r}|} = T \cdot \frac{N_{x_1+1,x_2+1,\ldots,x_m+1} - N_{x_1+r+1,x_2+r+1,\ldots,x_m+r+1}}{D_{x_1,x_2,\ldots,x_m}}.$$

Example 9.8.1. Consider a family composed of the father (55 years old), mother (53 years old) and sun (28 years old). Compute the single premium payable by the family for a 20-year temporary life annuity-immediate of 250 u.c. per year while all of the members survive. The annual interest percent is 10%.

Solution. The premium is

$$250 \cdot a_{55,53,28:\overline{20}|} = 250 \cdot \frac{N_{56,54,29} - N_{76,74,49}}{D_{55,53,28}}$$
$$= 250 \cdot \frac{D_{56,54,29} + D_{57,55,30} + D_{58,56,31} + \cdots + D_{75,73,48}}{D_{55,53,28}}$$
$$= 250 \cdot \frac{l_{56} \cdot l_{54} \cdot l_{29} \cdot v^{\frac{56+54+29}{3}} + l_{57} \cdot l_{55} \cdot l_{30} \cdot v^{\frac{57+55+30}{3}} + \cdots + l_{75} \cdot l_{73} \cdot l_{48} \cdot v^{\frac{75+73+48}{3}}}{l_{55} \cdot l_{53} \cdot l_{28} \cdot v^{\frac{55+53+28}{3}}}$$
$$\simeq 1586.11 \text{ u.c.}$$

□

9.9 Temporary life annuities for partial survival

Definition 9.9.1. *Let a group of m persons having the ages $x_1, x_2, \ldots, x_m$, where $m \in \mathbb{N}^*$, $x_j \in \mathbb{N}$, $x_j \leq \omega\ \forall j \in \{1, \ldots, m\}$. Letr $\in \mathbb{N}$ s.t. $\widehat{x} + r \leq \omega$ and let $k \in \mathbb{N}$ s.t. $k \leq m$. We denote*

- $a_{\overset{[k]}{\overline{x_1,x_2,\ldots,x_m}}:\overline{r}|}$ = *the single premium payable by the group for an r-year temporary life annuity-immediate of 1 u.c. per year payable (at the end of each year) while exactly k of the members survive (during the next r years);*

- $a_{\overset{k}{\overline{x_1,x_2,\ldots,x_m}}:\overline{r}|}$ = *the single premium payable by the group for an r-year temporary life annuity-immediate of 1 u.c. per year payable (at the end of each year) while at least k of the members survive (during the next r years).*

Proposition 9.9.1. *We have*

$$a_{\overset{[k]}{\overline{x_1,x_2,\ldots,x_m}}:\overline{r}|} = \sum_{s=0}^{m-k} (-1)^s \cdot C_{k+s}^{s} \sum_{1 \leq i_1 < \ldots < i_{k+s} \leq m} a_{x_{i_1},x_{i_2},\ldots,x_{i_{k+s}}:\overline{r}|}; \tag{9.16}$$

$$a_{\overset{k}{\overline{x_1,x_2,\ldots,x_m}}:\overline{r}|} = \sum_{s=0}^{m-k} (-1)^s \cdot C_{k+s-1}^{s} \sum_{1 \leq i_1 < \ldots < i_{k+s} \leq m} a_{x_{i_1},x_{i_2},\ldots,x_{i_{k+s}}:\overline{r}|}. \tag{9.17}$$

Corollary 9.9.1. *Let a group of m persons having the ages $x_1, x_2, \ldots, x_m$, where $m \in \mathbb{N}^*$, $x_j \in \mathbb{N}$, $x_j \leq \omega\ \forall j \in \{1, \ldots, m\}$. Letr $\in \mathbb{N}$ s.t. $\widehat{x} + r \leq \omega$ and let $k \in \mathbb{N}$ s.t. $k \leq m$. Let $T \geq 0$.*

1. *The single premium payable by the group for an r-year temporary life annuity-immediate of T u.c. per year payable while exactly k of the members survive is*

$$T \cdot a_{\overline{\overset{[k]}{x_1,x_2,\ldots,x_m}}:\overline{r}|} = T \cdot \sum_{s=0}^{m-k} (-1)^s \cdot C_{k+s}^{s} \sum_{1 \le i_1 < \ldots < i_{k+s} \le m} a_{x_{i_1},x_{i_2},\ldots,x_{i_{k+s}}:\overline{r}|}.$$

2. *The single premium payable by the group for an r-year temporary life annuity-immediate of T u.c. per year payable while at least k of the members survive is*

$$T \cdot a_{\overline{\overset{k}{x_1,x_2,\ldots,x_m}}:\overline{r}|} = T \cdot \sum_{s=0}^{m-k} (-1)^s \cdot C_{k+s-1}^{s} \sum_{1 \le i_1 < \ldots < i_{k+s} \le m} a_{x_{i_1},x_{i_2},\ldots,x_{i_{k+s}}:\overline{r}|}.$$

Example 9.9.1. Consider a family composed of the father (55 years old), mother (53 years old) and sun (28 years old). Compute the single premium payable by the family for a 20-year temporary life annuity-immediate of 250 u.c. per year while:
a) just one member survives;
b) exactly two of the members survive;
c) at least one member survives;
d) at least two of the members survive.
The annual interest percent is 10%.

Solution. a) The premium is

$$250 \cdot a_{\overline{\overset{[1]}{55,53,28}}:\overline{20}|} = 250 \cdot \sum_{s=0}^{2} (-1)^s \cdot C_{1+s}^{s} \sum_{1 \le i_1 < \ldots < i_{1+s} \le 3} a_{x_{i_1},x_{i_2},\ldots,x_{i_{1+s}}:\overline{20}|}$$

$$= 250 \left(C_1^0 \sum_{1 \le i_1 \le 3} a_{x_{i_1}:\overline{20}|} - C_2^1 \sum_{1 \le i_1 < i_2 \le 3} a_{x_{i_1},x_{i_2}:\overline{20}|} + C_3^2 \sum_{1 \le i_1 < i_2 < i_3 \le 3} a_{x_{i_1},x_{i_2},x_{i_3}:\overline{20}|} \right)$$

$$= 250 \Big[C_1^0 (a_{x_1:\overline{20}|} + a_{x_2:\overline{20}|} + a_{x_3:\overline{20}|}) - C_2^1 (a_{x_1,x_2:\overline{20}|} + a_{x_1,x_3:\overline{20}|} + a_{x_2,x_3:\overline{20}|}) + C_3^2 \cdot a_{x_1,x_2,x_3:\overline{20}|} \Big]$$

$$= 250 \Big[a_{55:\overline{20}|} + a_{53:\overline{20}|} + a_{28:\overline{20}|} - 2(a_{55,53:\overline{20}|} + a_{55,28:\overline{20}|} + a_{53,28:\overline{20}|}) + 3 \cdot a_{55,53,28:\overline{20}|} \Big]$$

$$\simeq 72.94 \text{ u.c.}$$

b) The premium is

$$250 \cdot a_{\overline{\overset{[2]}{55,53,28}}:\overline{20}|} = 250 \cdot \sum_{s=0}^{1} (-1)^s \cdot C_{2+s}^{s} \sum_{1 \le i_1 < \ldots < i_{2+s} \le 3} a_{x_{i_1},x_{i_2},\ldots,x_{i_{2+s}}:\overline{20}|}$$

$$= 250 \left(C_2^0 \sum_{1 \le i_1 < i_2 \le 3} a_{x_{i_1}, x_{i_2} : \overline{20|}} - C_3^1 \sum_{1 \le i_1 < i_2 < i_3 \le 3} a_{x_{i_1}, x_{i_2}, x_{i_3} : \overline{20|}} \right)$$

$$= 250 \Big[C_2^0 (a_{x_1, x_2 : \overline{20|}} + a_{x_1, x_3 : \overline{20|}} + a_{x_2, x_3 : \overline{20|}}) - C_3^1 \cdot a_{x_1, x_2, x_3 : \overline{20|}} \Big]$$

$$= 250 \Big(a_{55,53 : \overline{20|}} + a_{55,28 : \overline{20|}} + a_{53,28 : \overline{20|}} - 3 \cdot a_{55,53,28 : \overline{20|}} \Big)$$

$$\simeq 465.86 \text{ u.c.}$$

c) The premium is

$$250 \cdot a_{\overline{\overset{1}{55,53,28}} : \overline{20|}} = 250 \cdot \sum_{s=0}^{2} (-1)^s \cdot C_s^s \sum_{1 \le i_1 < \ldots < i_{1+s} \le 3} a_{x_{i_1}, x_{i_2}, \ldots, x_{i_{1+s}} : \overline{20|}}$$

$$= 250 \left(C_0^0 \sum_{1 \le i_1 \le 3} a_{x_{i_1} : \overline{20|}} - C_1^1 \sum_{1 \le i_1 < i_2 \le 3} a_{x_{i_1}, x_{i_2} : \overline{20|}} + C_2^2 \sum_{1 \le i_1 < i_2 < i_3 \le 3} a_{x_{i_1}, x_{i_2}, x_{i_3} : \overline{20|}} \right)$$

$$= 250 \Big[C_0^0 (a_{x_1 : \overline{20|}} + a_{x_2 : \overline{20|}} + a_{x_3 : \overline{20|}}) - C_1^1 (a_{x_1, x_2 : \overline{20|}} + a_{x_1, x_3 : \overline{20|}} + a_{x_2, x_3 : \overline{20|}}) + C_2^2 \cdot a_{x_1, x_2, x_3 : \overline{20|}} \Big]$$

$$= 250 \Big[a_{55 : \overline{20|}} + a_{53 : \overline{20|}} + a_{28 : \overline{20|}} - (a_{55,53 : \overline{20|}} + a_{55,28 : \overline{20|}} + a_{53,28 : \overline{20|}}) + a_{55,53,28 : \overline{20|}} \Big]$$

$$\simeq 2124.92 \text{ u.c.}$$

d) The premium is

$$250 \cdot a_{\overline{\overset{2}{55,53,28}} : \overline{20|}} = 250 \cdot \sum_{s=0}^{1} (-1)^s \cdot C_{1+s}^s \sum_{1 \le i_1 < \ldots < i_{2+s} \le 3} a_{x_{i_1}, x_{i_2}, \ldots, x_{i_{2+s}} : \overline{20|}}$$

$$= 250 \left(C_1^0 \sum_{1 \le i_1 < i_2 \le 3} a_{x_{i_1}, x_{i_2} : \overline{20|}} - C_2^1 \sum_{1 \le i_1 < i_2 < i_3 \le 3} a_{x_{i_1}, x_{i_2}, x_{i_3} : \overline{20|}} \right)$$

$$= 250 \Big[C_1^0 (a_{x_1, x_2 : \overline{20|}} + a_{x_1, x_3 : \overline{20|}} + a_{x_2, x_3 : \overline{20|}}) - C_2^1 \cdot a_{x_1, x_2, x_3 : \overline{20|}} \Big]$$

$$= 250 \Big(a_{55,53 : \overline{20|}} + a_{55,28 : \overline{20|}} + a_{53,28 : \overline{20|}} - 2 \cdot a_{55,53,28 : \overline{20|}} \Big)$$

$$\simeq 2051.97 \text{ u.c.}$$

□

9.10 Group insurance payable at the first death

Definition 9.10.1. *Let a group of m persons having the ages $x_1, x_2, \ldots, x_m$, where $m \in \mathbb{N}^*$, $x_j \in \mathbb{N}$, $x_j \le \omega$ $\forall j \in \{1, \ldots, m\}$. We denote*

- $\dot{A}_{x_1,x_2,\ldots,x_m}$ = *the single premium payable by the group for an insurance of* 1 *u.c. payable at the moment of the first death, whenever this occurs.*

Proposition 9.10.1. *We have*

$$\dot{A}_{x_1,x_2,\ldots,x_m} = \frac{M_{x_1,x_2,\ldots,x_m}}{D_{x_1,x_2,\ldots,x_m}}, \tag{9.18}$$

where

$$M_{x_1,x_2,\ldots,x_m} = \sum_{n=0}^{\omega-\widetilde{x}} C_{x_1+n,x_2+n,\ldots,x_m+n}, \tag{9.19}$$

$\widetilde{x} = \max\{x_1, x_2, \ldots, x_m\}$ *being the maximum age of the group, with*

$$C_{x_1,x_2,\ldots,x_m} = (l_{x_1} \cdot l_{x_2} \cdot \ldots \cdot l_{x_m} - l_{x_1+1} \cdot l_{x_2+1} \cdot \ldots \cdot l_{x_m+1}) \cdot v^{\frac{x_1+x_2+\ldots+x_m}{m}+\frac{1}{2}}, \tag{9.20}$$

$v = \dfrac{1}{1+i}$ *being the annual discounting factor, i being the annual interest rate.*

Corollary 9.10.1. *Let a group of m persons having the ages $x_1, x_2, \ldots, x_m$, where $m \in \mathbb{N}^*$, $x_j \in \mathbb{N}$, $x_j \leq \omega\ \forall j \in \{1, \ldots, m\}$. Let $T \geq 0$. The single premium payable by the group for an insurance of T u.c. payable at the moment of the first death, whenever this occurs, is*

$$T \cdot \dot{A}_{x_1,x_2,\ldots,x_m} = T \cdot \frac{M_{x_1,x_2,\ldots,x_m}}{D_{x_1,x_2,\ldots,x_m}}.$$

Example 9.10.1. Consider a family composed of the father (55 years old), mother (53 years old) and sun (28 years old). Compute the single premium payable by the family for an insurance of 25000 u.c. payable at the moment of the first death. The annual interest percent is 10%.

Solution. The premium is

$$\begin{aligned}
25000 \cdot \dot{A}_{55,53,28} &= 25000 \cdot \frac{M_{55,53,28}}{D_{55,53,28}} \\
&= 25000 \cdot \frac{C_{55,53,28} + C_{56,54,29} + C_{57,55,30} + \cdots + C_{100,98,73}}{D_{55,53,28}} \\
&= 25000 \cdot \frac{1}{l_{55} \cdot l_{53} \cdot l_{28} \cdot v^{\frac{55+53+28}{3}}} \cdot \Big[(l_{55} \cdot l_{53} \cdot l_{28} - l_{56} \cdot l_{54} \cdot l_{29}) \cdot v^{\frac{55+53+28}{3}+\frac{1}{2}} + \cdots + \\
&\quad + (l_{100} \cdot l_{98} \cdot l_{73} - l_{101} \cdot l_{99} \cdot l_{74}) \cdot v^{\frac{100+98+73}{3}+\frac{1}{2}} \Big] \\
&\simeq 8523.42 \text{ u.c.}
\end{aligned}$$

We recall that $l_{101} = 0$. □

9.11 Group insurance payable at the k-th death

Definition 9.11.1. *Let a group of m persons having the ages $x_1, x_2, \ldots, x_m$, where $m \in \mathbb{N}^*$, $x_j \in \mathbb{N}$, $x_j \leq \omega\ \forall j \in \{1, \ldots, m\}$. Let $k \in \mathbb{N}$ s.t. $k \leq m$. We denote*

- $\dot{A}_{\overline{x_1, x_2, \ldots, x_m}}^{\ [k]}$ *= the single premium payable by the group for an insurance of 1 u.c. payable at the moment of the k-th death, whenever this occurs.*

Proposition 9.11.1. *We have*

$$\dot{A}_{\overline{x_1, x_2, \ldots, x_m}}^{\ [k]} = \sum_{s=0}^{k-1} (-1)^s \cdot C_{m-k+s}^{s} \sum_{1 \leq i_1 < \ldots < i_{m-k+s+1} \leq m} \dot{A}_{x_{i_1}, x_{i_2}, \ldots, x_{i_{m-k+s+1}}}. \quad (9.21)$$

Corollary 9.11.1. *Let a group of m persons having the ages $x_1, x_2, \ldots, x_m$, where $m \in \mathbb{N}^*$, $x_j \in \mathbb{N}$, $x_j \leq \omega\ \forall j \in \{1, \ldots, m\}$. Let $k \in \mathbb{N}$ s.t. $k \leq m$ and let $T \geq 0$. The single premium payable by the group for an insurance of T u.c. payable at the moment of the k-th death, whenever this occurs, is*

$$T \cdot \dot{A}_{\overline{x_1, x_2, \ldots, x_m}}^{\ [k]} = T \cdot \sum_{s=0}^{k-1} (-1)^s \cdot C_{m-k+s}^{s} \sum_{1 \leq i_1 < \ldots < i_{m-k+s+1} \leq m} \dot{A}_{x_{i_1}, x_{i_2}, \ldots, x_{i_{m-k+s+1}}}.$$

9.12 Problems

Exercise 9.12.1. Consider a group of 4 members of 55, 53, 30, and 28 years old.
a) Calculate the probability that all of the members survive 15 years.
b) Calculate the probability that exactly 2 of the members survive 25 years.
c) Calculate the probability that at least 3 of the members survive 20 years.
d) Calculate the probability that at most 3 of the members survive 10 years.

Exercise 9.12.2. Calculate the single premium payable by a family of two persons of 32 and 30 years old for a single claim of 20000$ over 35 years if both members will be alive. The annual interest percent is 8%.

Exercise 9.12.3. Calculate the single premium payable by a family of two persons of 32 and 30 years old for a single claim of 20000$ over 35 years if just one member will be alive. The annual interest percent is 8%.

Exercise 9.12.4. Calculate the single premium payable by a family of two persons of 32 and 30 years old for a single claim of 20000$ over 35 years if at least one member will be alive. The annual interest percent is 8%.

Exercise 9.12.5. Calculate the single premium payable by a family of three persons of 46, 44 and 22 years old for a life annuity-immediate of 10000$ per year while all of the members survive. The annual interest percent is 8%.

Exercise 9.12.6. Calculate the single premium payable by a family of three persons of 46, 44 and 22 years old for a life annuity-immediate of 10000$ per year while exactly two of the members survive. The annual interest percent is 8%.

Exercise 9.12.7. Calculate the single premium payable by a family of three persons of 46, 44 and 22 years old for a life annuity-immediate of 10000$ per year while at least two of the members survive. The annual interest percent is 8%.

Exercise 9.12.8. Calculate the single premium payable by a family of two persons of 42 and 37 years old for a 10-year deferred life annuity-immediate of 10000$ per year while all of the members survive. The annual interest percent is 8%.

Exercise 9.12.9. Calculate the single premium payable by a family of two persons of 42 and 37 years old for a 10-year deferred life annuity-immediate of 10000$ per year while just one member survives. The annual interest percent is 8%.

Exercise 9.12.10. Calculate the single premium payable by a family of two persons of 42 and 37 years old for a 10-year deferred life annuity-immediate of 10000$ per year while at least one member survives. The annual interest percent is 8%.

Exercise 9.12.11. Calculate the single premium payable by a family of three persons of 60, 54 and 35 years old for a 30-year temporary life annuity-immediate of 10000$ per year while all of the members survive. The annual interest percent is 8%.

Exercise 9.12.12. Calculate the single premium payable by a family of three persons of 60, 54 and 35 years old for a 30-year temporary life annuity-immediate of 10000$ per year while just one member survives. The annual interest percent is 8%.

Exercise 9.12.13. Calculate the single premium payable by a family of three persons of 60, 54 and 35 years old for a 30-year temporary life annuity-immediate of 10000$ per year while at least one member survives. The annual interest percent is 8%.

Exercise 9.12.14. Calculate the single premium payable by a family of two persons of 28 and 25 years old for an insurance of 50000$ payable at the moment of the first death. The annual interest percent is 8%.

Exercise 9.12.15. Calculate the single premium payable by a family of two persons of 28 and 25 years old for an insurance of 50000$ payable at the moment of the last death. The annual interest percent is 8%.

Exercise 9.12.16. Calculate the single premium payable by a family of three persons of 50, 49 and 25 years old for an insurance of 50000$ payable at the moment of the first death. The annual interest percent is 8%.

Exercise 9.12.17. Calculate the single premium payable by a family of three persons of 50, 49 and 25 years old for an insurance of 50000$ payable at the moment of the second death. The annual interest percent is 8%.

Exercise 9.12.18. Calculate the single premium payable by a family of three persons of 50, 49 and 25 years old for an insurance of 50000$ payable at the moment of the last death. The annual interest percent is 8%.

Theme 10

Bonus-Malus system in automobile insurance

10.1 A general model

The Bonus-Malus system is the most well known system of goods insurance, especially car insurance. In this type of insurance, policies are categorized based on characteristics of the insured vehicle (the insured good), and on **Bonus-Malus level**, given by the previous number of claims. The insurance period for goods is usually one year. In this case, one policy remains in a certain payment class for one year and then it can be transferred to another payment class, based on the number of accidents from the previous year. If the insured vehicle didn't have any accident, then the new payment class will be "better", so the premium will be reduced (**bonus**). As the number of accidents grows, the new class will be "worst", so the premium will be increased (**malus**).

Definition 10.1.1. *A **Bonus-Malus insurance system** can be represented as $S = (C, D, T, \Pi)$, where:*

- $C = \{1, \ldots, c\}$ *represents the set of **payment classes** ($c \in \mathbb{N}^*$). If $i > j$, $i, j \in C$, we say thai i is a **better class** than j.*

- $D = \{0, \ldots, r\}$ *represents the set of **annual number of accidents** possible for an insurance policy ($r \in \mathbb{N}^*$).*

- $T : C \times D \to C$ *is a function called **the rule of passing** of the system; for any $i \in C$ and $j \in D$, $T(i, j)$ represents the payment class in which it will be transferred the next year every insurance policy from class i that had j accidents during the current year. The function $T(i, j)$ increases in i (for any fixed j) and decreases in j (for any fixed i).*

- $\Pi : C \to (0, \infty)$ *is a decreasing function; for any* $i \in C$, $\Pi(i)$ *represents the* ***premium insurance*** *for an insurance policy from class* i.

Remark 10.1.1. *The premium insurances* $\Pi(i)$ *calculated using the previous system are also called* ***mathematical premiums****. In practice, to this premiums are also added*

- *values for reducing the probability of ruin for the insurer;*
- *charges spent on employees;*
- *taxes.*

10.2 Bayes model based on a mixed Poisson distribution

Definition 10.2.1. *We denote by* X *the random variable that represents the* ***number of accidents during one year*** *(for a random insurance policy).*

In order to place a policy in a payment class and calculate the corresponding insurance premium, it is necessary to estimate the value of the number of accidents X_{n+1} for the next year based on the recorded values of the number of accidents $X_1, \ldots, X_n$ of the ensured vehicle in n previous years (years $1, \ldots, n$), where $n \in \mathbb{N}^*$. In this context, the known distributions of the r.v. $X_1, \ldots, X_n$ are also called the **prior distributions**, and the estimated distribution of r.v. X_{n+1} is also called the **posterior distribution**.

For estimating the posterior distribution of number of accidents and for calculating the premium for year $n+1$, we consider a **Bayes model based on a mixed Poisson distribution for the annual number of accidents**, in which we assume:

- The annual number of accidents X (for a random policy) has a mixed Poisson-H distribution, where H is the distribution of the random variable $\Lambda > 0$ which represents the **average number of annual accidents** (for a random policy), so $X|(\Lambda = \lambda) \sim \text{Po}(\lambda)$ for any $\lambda > 0$.
- For any $\lambda > 0$ the conditioned random variables $X_1|(\Lambda = \lambda), \ldots, X_n|(\Lambda = \lambda), X_{n+1}|(\Lambda = \lambda)$ are independent and identically distributed with $X|(\Lambda = \lambda)$ (i.e. $X_i|(\Lambda = \lambda) \sim \text{Po}(\lambda)$ for any $i \in \{1, \ldots, n+1\}$).

Definition 10.2.2. *For any* $n \in \mathbb{N}^*$, *the number*

$$I_{n+1}(x_1, \ldots, x_n) = \frac{E(X_{n+1}|X_1 = x_1, \ldots, X_n = x_n)}{E(X_{n+1})}, \ \forall x_1, \ldots, x_n \in D \tag{10.1}$$

is called the ***frequency index for year*** $n+1$ *when* $X_1 = x_1, \ldots, X_n = x_n$ *are known.*

Remark 10.2.1. *The frequency index represents the rapport between the posterior mean and the prior mean of* X_{n+1}.

Proposition 10.2.1. *For any* $n \in \mathbb{N}^*$ *and* $x_1, \ldots, x_n \in D$ *we have*

$$I_{n+1}(x_1, \ldots, x_n) = \frac{E(\Lambda|X_1 = x_1, \ldots, X_n = x_n)}{E(\Lambda)} \tag{10.2}$$

$$= \frac{\textit{premium for year } n+1 \textit{ given } (x_1, \ldots, x_n)}{\textit{initial premium, from year 1}}. \tag{10.3}$$

Remark 10.2.2. *The premium for year* $n+1$ *given* $(x_1, \ldots, x_n)$ *is called a* ***posterior premium****, and the initial premium, from year 1, is called a* ***prior premium****.*

Algorithm 10.2.1 (Bayes model for calculating the premiums in Bonus- Malus system).

Step 0. *Estimate the distribution of r.v.* Λ *representing the average number of annual accidents (for a random insurance policy). This distribution is also called the* ***prior distribution of the r.v.*** Λ. *It can use the maximum likelihood estimation, based on the frequency of the accidents from the initial year.*

To calculate the premium for year $n+1$ *knowing the number of accidents in the previous years* $x_1, \ldots, x_n$ *the following four steps will be proceeded:*

Step 1. *Calculate the distribution of the conditioned random variable* $\Lambda|(X_1 = x_1, \ldots, X_n = x_n)$, *also called the* ***posterior distribution of r.v.*** Λ. *For this it can use the* ***Bayes's formula****.*

Step 2. *Calculate the* ***posterior distribution of r.v.*** X_{n+1}, *i.e. the distribution of conditioned r.v.* $X_{n+1}|(X_1 = x_1, \ldots, X_n = x_n)$. *For this it can use the* ***total probability formula****.*

Step 3. *Calculate the* ***frequency index*** $I_{n+1}(x_1, \ldots, x_n)$, *using the formulas (10.1) or (10.2).*

*Step 4. Calculate the **premium for year** $n+1$ by formula (10.3).*

Remark 10.2.3. *In the next section will be prove that the posterior distributions of r.v. Λ and X_{n+1} and the frequency indexes $I_{n+1}(x_1,\ldots,x_n)$ depend only on the total number of accidents $\sum\limits_{i=1}^{n} x_i$ observed in the previous n years, and not on the distribution of the accidents during these n years. Therefore the values of the frequency indexes are tabled according to the year n and the total number of accidents $\sum\limits_{i=1}^{n} x_i$.*

10.3 Gamma distribution for the average number of accidents

We will apply the described model in the particular case when the r.v. Λ (which represents the average number of annual accidents for a random insurance policy) has a Gamma prior distribution of parameters a and b, where $a, b > 0$, with the probability density function

$$f(\lambda) = \frac{1}{\Gamma(a)b^a}\lambda^{a-1}\mathrm{e}^{-\frac{\lambda}{b}},\ \forall\lambda > 0.$$

Then the r.v. X (which represents the number of accidents during one year for a random policy) has a mixed Poisson-Gamma distribution of parameters a and b, which is equivalent with a Negative Binomial distribution of parameters a and $\frac{1}{b+1}$. So $X \sim BN(a, \frac{1}{b+1})$.

Step 0 of the previous algorithm requires the estimation of the parameters a and b for the prior distribution of r.v. Λ. In the next example we will apply the **maximum likelihood estimation method** to estimate these parameters, based on the number of accidents during one year.

Step 1 consists of calculating the a posteriori distribution for r.v. Λ, that is the distribution of conditioned r.v. $\Lambda|(X_1 = x_1, \ldots, X_n = x_n)$. According to Bayes' formula, this distribution has the probability density function

$$f(\lambda|x_1,\ldots,x_n) = \frac{f(\lambda)P(X_1 = x_1,\ldots,X_n = x_n|\Lambda = \lambda)}{\int_0^\infty f(t)P(X_1 = x_1,\ldots,X_n = x_n|\Lambda = t)dt}.$$

Using the hypothesis that for any $\lambda > 0$ the conditioned random variables $X_1|(\Lambda = \lambda), \ldots, X_n|(\Lambda = \lambda)$ are independent and identically distributed with $X|(\Lambda = \lambda)$ (i.e. $X_i|(\Lambda = \lambda) \sim \mathrm{Po}(\lambda)$ for any $i \in \{1,\ldots,n\}$), it follows that

$$P(X_1 = x_1,\ldots,X_n = x_n|\Lambda = \lambda) = P(X_1 = x_1|\Lambda = \lambda)\cdot\ldots\cdot P(X_n = x_n|\Lambda = \lambda)$$

$$= P(X = x_1|\Lambda = \lambda) \cdot \ldots \cdot P(X = x_n|\Lambda = \lambda)$$

$$= \mathrm{e}^{-\lambda}\frac{\lambda^{x_1}}{x_1!} \cdot \ldots \cdot \mathrm{e}^{-\lambda}\frac{\lambda^{x_n}}{x_n!} = \mathrm{e}^{-n\lambda}\frac{\lambda^{\sum\limits_{i=1}^{n} x_i}}{x_1! \cdot \ldots \cdot x_n!},$$

so

$$f(\lambda|x_1,\ldots,x_n) = \frac{\dfrac{1}{\Gamma(a)b^a}\lambda^{a-1}\mathrm{e}^{-\frac{\lambda}{b}}\mathrm{e}^{-n\lambda}\dfrac{\lambda^{\sum\limits_{i=1}^{n} x_i}}{x_1! \cdot \ldots \cdot x_n!}}{\dfrac{1}{\Gamma(a)b^a} \cdot \dfrac{1}{x_1! \cdot \ldots \cdot x_n!}\displaystyle\int_0^\infty t^{a-1}\mathrm{e}^{-\frac{t}{b}}\mathrm{e}^{-nt}t^{\sum\limits_{i=1}^{n} x_i}dt}$$

$$= \frac{\lambda^{a+\sum\limits_{i=1}^{n} x_i - 1}\mathrm{e}^{-\frac{\lambda(1+bn)}{b}}}{\int_0^\infty t^{a+\sum\limits_{i=1}^{n} x_i - 1}\mathrm{e}^{-\frac{t(1+bn)}{b}}dt}.$$

Therefore the posterior distribution of the r.v. Λ is a Gamma distribution of parameters $a + \sum\limits_{i=1}^{n} x_i$ and $\frac{b}{1+bn}$.

Step 2 consists of calculating the posterior distribution for the r.v. X_{n+1}, that is the distribution of the conditioned r.v. $X_{n+1}|(X_1 = x_1, \ldots, X_n = x_n)$. According to the total probability formula, this distribution is given by

$$P(X_{n+1} = x|X_1 = x_1, \ldots, X_n = x_n)$$
$$= \int_0^\infty P(X_{n+1} = x|X_1 = x_1, \ldots, X_n = x_n, \Lambda = \lambda)f(\lambda|x_1, \ldots, x_n)d\lambda,\ \forall x \in \mathbb{N}.$$

Using the hypothesis that for any $\lambda > 0$ the conditioned random variables $X_1|(\Lambda = \lambda), \ldots, X_n|(\Lambda = \lambda), X_{n+1}|(\Lambda = \lambda)$ are independent and identically distributed with $X|(\Lambda = \lambda)$, it follows that

$$P(X_{n+1} = x|X_1 = x_1, \ldots, X_n = x_n, \Lambda = \lambda) = P(X_{n+1} = x|\Lambda = \lambda)$$
$$= P(X = x|\Lambda = \lambda),$$

so

$$P(X_{n+1} = x|X_1 = x_1, \ldots, X_n = x_n) = \int_0^\infty P(X = x|\Lambda = \lambda)f(\lambda|x_1, \ldots, x_n)d\lambda,$$

for any $x \in \mathbb{N}$. Since $X|(\Lambda = \lambda) \sim \mathrm{Po}(\lambda)$ and $f(\lambda|x_1, \ldots, x_n)$ is the probability density function of Gamma distribution of parameters $a + \sum\limits_{i=1}^{n} x_i$ and $\frac{b}{1+bn}$, we get that the posterior distribution of the r.v. X_{n+1} is a mixed Poisson-Gamma distribution of parameters $a + \sum_{i=1}^{n} x_i$ and $\frac{b}{1+bn}$, which is equivalent with the Negative Binomial distribution of parameters $a + \sum_{i=1}^{n} x_i$ and $\frac{1+bn}{1+b+bn}$. Hence $X_{n+1}|(X_1 = x_1, \ldots, X_n = x_n) \sim BN(a + \sum_{i=1}^{n} x_i, \frac{1+bn}{1+b+bn})$.

Step 3 consists of calculating the frequency indexes $I_{n+1}(x_1, \ldots, x_n)$. According to formula (10.1) we have

$$I_{n+1}(x_1, \ldots, x_n) = \frac{E(X_{n+1}|X_1 = x_1, \ldots, X_n = x_n)}{E(X_{n+1})},\ \forall x_1, \ldots, x_n \in D.$$

According to the formula of the mean for the Negative Binomial distribution, from $X \sim BN(a, \frac{1}{b+1})$ it follows that

$$E(X_{n+1}) = E(X) = a \cdot \frac{1 - \frac{1}{b+1}}{\frac{1}{b+1}} = ab,$$

and from $X_{n+1}|(X_1 = x_1, \ldots, X_n = x_n) \sim BN(a + \sum_{i=1}^{n} x_i, \frac{1+bn}{1+b+bn})$ it follows that

$$E(X_{n+1}|X_1 = x_1, \ldots, X_n = x_n) = \left(a + \sum_{i=1}^{n} x_i\right) \frac{1 - \frac{1+bn}{1+b+bn}}{\frac{1+bn}{1+b+bn}} = \frac{b\left(a + \sum_{i=1}^{n} x_i\right)}{1+bn},$$

so

$$I_{n+1}(x_1, \ldots, x_n) = \frac{a + \sum_{i=1}^{n} x_i}{a(1+bn)} = \frac{1 + \frac{1}{a}\sum_{i=1}^{n} x_i}{1+bn},\ \forall x_1, \ldots, x_n \in D. \qquad (10.4)$$

After calculating the frequency indexes using formula (10.4), the premiums for year $n+1$ can be obtained (at Step 4) by using the formula (10.3).

Example 10.3.1. We apply the model discussed for the following data set, that one French insurance company had during year 1979. The data set consists of $m = 1044454$ policyholders.

No. of accidents (j)	Absolute frequency (m_j)
0	881705
1	142217
2	18088
3	2118
4	273
≥ 5	53
Total	1044454

According to our model, $X \sim BN(a, \frac{1}{b+1})$.

By using the maximum likelihood estimation method, the estimated values of parameters $a > 0$ and b verify the following equations

$$\frac{1}{b+1} = \frac{a}{a+\overline{X}}, \tag{10.5}$$

$$\sum_{j=1}^{5} m_j \left(\frac{1}{a} + \frac{1}{a+1} + \ldots + \frac{1}{a+j-1}\right) - m \ln\left(1 + \frac{\overline{X}}{a}\right) = 0, \tag{10.6}$$

where $\overline{X}$ is the mean of the sample formed by the recorded data. We have

$$\overline{X} = \sum_{j=0}^{5} j \cdot \frac{m_j}{m} \simeq 0.178183051.$$

Dividing by m, the equation (10.6) can be rewritten as

$$\sum_{j=1}^{5} \frac{m_j}{m} \left(\frac{1}{a} + \frac{1}{a+1} + \ldots + \frac{1}{a+j-1}\right) - \ln\left(1 + \frac{\overline{X}}{a}\right) = 0. \tag{10.7}$$

For the given values of m_j, m and $\overline{X}$, we derive that the equation (10.7) has a unique positive solution, namely

$$a \simeq 1.672974126.$$

By (10.5) we obtain that

$$b = \frac{\overline{X}}{a} \simeq 0.106506758.$$

It follows from (10.4) that

$$I_{n+1}(x_1, \ldots, x_n) = \frac{1 + \frac{1}{1.672974126} \sum\limits_{i=1}^{n} x_i}{1 + 0.106506758 \cdot n}, \ \forall x_1, \ldots, x_n \in D. \tag{10.8}$$

Using the constructed model (**Classical Bonus-Malus System**) we consider the case when the maximum number of consecutive years is $n = 10$ and the maximum number of accidents is 5. According to formula (10.8) we obtain the following table containing the values of frequency indexes $I_{n+1}(x_1, \ldots, x_n)$ based on year n and the total number of accidents $\sum\limits_{i=1}^{n} x_i$.

$n \diagdown \sum_{i=1}^{n} x_i$	0	1	2	3	4	5
0	1.0000	—	—	—	—	—
1	0.9037	1.4439	1.9842	2.5244	3.0646	3.6048
2	0.8244	1.3172	1.8099	2.3027	2.7955	3.2882
3	0.7579	1.2108	1.6638	2.1168	2.5698	3.0228
4	0.7012	1.1204	1.5396	1.9587	2.3779	2.7971
5	0.6525	1.0425	1.4326	1.8226	2.2126	2.6027
6	0.6101	0.9748	1.3395	1.7042	2.0689	2.4336
7	0.5729	0.9153	1.2578	1.6002	1.9426	2.2851
8	0.5399	0.8627	1.1854	1.5082	1.8309	2.1537
9	0.5106	0.8158	1.1210	1.4262	1.7313	2.0365
10	0.4842	0.7737	1.0631	1.3526	1.6421	1.9315

For example, consider a policyholder that had only one accident in the first three years and the initial premium was 200 u.c. According to formula (10.3) and the values from the table of frequency indexes, the premium for the fourth year is obtained as $200 \cdot I_{3+1}(x_1, x_2, x_3)$ for $\sum_{i=1}^{3} x_i = 1$, that is $200 \cdot 1.2108 = 242.16$ u.c.

10.4 Problems

Exercise 10.4.1. Calculate and extend the above table of frequency indexes.

Exercise 10.4.2. The initial premium for a policyholder was 300 u.c. Calculate the premium for the next 12 years, if the policyholder had only one accident in the second year, two accidents in the sixth year and one accident in the seventh year.

Exercise 10.4.3. a) A policyholder had only one accident in the first year. Calculate the number of years over which the premium will be less that the initial premium.
b) The same question for a policyholder who had three accidents in the first year.

Theme 11

Some optimization models

11.1 Portfolio planning

Consider a portfolio model where an investor wishes to invest in n assets. For any $i \in \{1, \ldots, n\}$, let r_i be the return (the rate of profit) of the asset i. Obviously, $\mathrm{r} = (r_1, \ldots, r_n)^\top$ is a random vector.

Let $\mathrm{m} = (m_1, \ldots, m_n)^\top$ and $\mathrm{V} = (v_{ij})_{i,j \in \{1,\ldots,n\}}$ be the mean and the covariance matrix of r, respectively. The matrix V represents the *risk matrix* of the investment.

Assume that the investor disposes of estimated values of m and V. The portfolio planning consists in determining the proportions $p_1, p_2, \ldots, p_n$ of the investment to asset $1, 2, \ldots, n$, respectively. Obviously,

$$\sum_{i=1}^{n} p_i = 1 \text{ and } p_i \geq 0, \ \forall i \in \{1, \ldots, n\}.$$

Let $\mathrm{p} = (p_1, \ldots, p_n)^\top$. Then the value

$$\mathrm{m}^\top \mathrm{p} = \sum_{i=1}^{n} m_i p_i$$

represents the *expected return* of the portfolio p, and the value

$$\mathrm{p}^\top \mathrm{V} \mathrm{p} = \sum_{i=1}^{n} \sum_{j=1}^{n} V_{ij} p_i p_j$$

represents the *expected risk* of the portfolio p.

To optimize the portfolio, on the one hand one needs to minimize the expected risk, and on the other hand one needs to maximize the expected return. So, by choosing the minimization of expected risk as optimization criterion, we obtain that an optimal portfolio p is an optimal solution for the following problem

$$(\text{P11.1.1}) \left| \begin{array}{l} \min \mathrm{p}^\top \mathrm{V} \mathrm{p} \text{ s.t.} \\ \mathrm{m}^\top \mathrm{p} \geq c, \\ \sum\limits_{i=1}^{n} p_i = 1, \\ p_i \geq 0, \ \forall i \in \{1, \ldots, n\}, \end{array} \right.$$

where c represents a given lower bound for the expected return.

On the other hand, by choosing the maximization of expected return as optimization criterion, we obtain that an optimal portfolio p is an optimal solution for the following problem

$$(\text{P11.1.2}) \left| \begin{array}{l} \max \mathrm{m}^\top \mathrm{p} \text{ s.t.} \\ \mathrm{p}^\top \mathrm{V} \mathrm{p} \leq d, \\ \sum\limits_{i=1}^{n} p_i = 1, \\ p_i \geq 0, \ \forall i \in \{1, \ldots, n\}, \end{array} \right.$$

where d represents a given upper bound for the expected risk.

Also, by choosing the MEP (maximum entropy principle) as optimization criterion, we obtain that an optimal portfolio p is an optimal solution for the following problem

$$(\text{P11.1.3}) \left| \begin{array}{l} \max H(\mathrm{p}) = -\sum\limits_{i=1}^{n} p_i \ln p_i \text{ s.t.} \\ \mathrm{m}^\top \mathrm{p} \geq c, \\ \mathrm{p}^\top \mathrm{V} \mathrm{p} \leq d, \\ \sum\limits_{i=1}^{n} p_i = 1, \\ p_i \geq 0, \ \forall i \in \{1, \ldots, n\}. \end{array} \right.$$

By combining the above criteria, we can define an optimal portfolio p as an optimal solution for the following problem

$$(\text{P11.1.4}) \left| \begin{array}{l} \min \left(w_1 \mathrm{p}^\top \mathrm{V} \mathrm{p} - w_2 \mathrm{m}^\top \mathrm{p} + w_3 \sum\limits_{i=1}^{n} p_i \ln p_i \right) \text{ s.t.} \\ \sum\limits_{i=1}^{n} p_i = 1, \\ p_i \geq 0, \ \forall i \in \{1, \ldots, n\}, \end{array} \right.$$

where w_1, w_2 and w_3 are given weights such that $w_1, w_2, w_3 > 0$ and $w_1 + w_2 + w_3 = 1$.

We remark that the problem (P11.1.1) is a quadratic programming problem, and the problem (P11.1.2) is a linear programming problem with quadratic constraints. Also, the problem (P11.1.3) is an entropy optimization problem with quadratic constraints, and the problem (P11.1.4) is a convex quadratic programming problem with entropic perturbation.

11.2 Regional planning

An important problem in regional or urban planning is the allocation of new houses. Let n be the number of zones dividing the city or the region, K be the number of different household types to be located, and let L be the number of different house types to be allocated. For any $i \in \{1, \ldots, n\}$, $k \in \{1, \ldots, K\}$ and $l \in \{1, \ldots, L\}$ we assume that the following elements are known:

- b_{ikl} = the budget that a type k household is willing to allocate for purchasing and living in a type l house from zone i, including house price, the housing costs and the daily transportation costs (the *residential budget*);

- c_{ikl} = the cost that must be allocated by a type k household to living in a type l house from zone i, including the daily transportation (the *necessary budget*);

- s_{ikl} = the area allocated for a type k household with a type l house from zone i;

- S_i = the total area allocated for housing in zone i;

- F_k = the total number of type k households to be located.

The difference $b_{ikl} - c_{ikl}$ represents the *bidding power* of a type k household for purchasing a type l house from zone i.

The urban or regional planning for households locating consists in determining the numbers x_{ikl} of type k households that will be located in a type l house in zone i, for any $i \in \{1, \ldots, n\}$, $k \in \{1, \ldots, K\}$ and $l \in \{1, \ldots, L\}$.

To optimize the households locating, one needs to introduce an optimization criterion. A such criterion is the maximization of total bidding power.

Hence $(x_{ikl})_{i,k,l}$ is an optimal solution for the following linear programming problem

$$\text{(P11.2.1)}\quad \left| \begin{array}{l} \max \sum_{i=1}^{n} \sum_{k=1}^{K} \sum_{l=1}^{L} (b_{ikl} - c_{ikl}) x_{ikl} \text{ s.t.} \\ \sum_{k=1}^{K} \sum_{l=1}^{L} s_{ikl} x_{ikl} \leq S_i,\ \forall i \in \{1, \dots, n\}, \\ \sum_{i=1}^{n} \sum_{l=1}^{L} x_{ikl} = F_k,\ \forall k \in \{1, \dots, K\}, \\ x_{ikl} \geq 0,\ \forall i \in \{1, \dots, n\},\ \forall k \in \{1, \dots, K\},\ \forall l \in \{1, \dots, L\}. \end{array} \right.$$

When the planner disposes of a lower bound M for the total bidding power, then $(x_{ikl})_{i,k,l}$ is an optimal solution for the following problem, according to the MEP:

$$\text{(P11.2.2)}\quad \left| \begin{array}{l} \max - \sum_{i=1}^{n} \sum_{k=1}^{K} \sum_{l=1}^{L} x_{ikl} \ln x_{ikl} \text{ s.t.} \\ \sum_{k=1}^{K} \sum_{l=1}^{L} s_{ikl} x_{ikl} \leq S_i,\ \forall i \in \{1, \dots, n\}, \\ \sum_{i=1}^{n} \sum_{l=1}^{L} x_{ikl} = F_k,\ \forall k \in \{1, \dots, K\}, \\ \sum_{i=1}^{n} \sum_{k=1}^{K} \sum_{l=1}^{L} (b_{ikl} - c_{ikl}) x_{ikl} \geq M, \\ x_{ikl} \geq 0,\ \forall i \in \{1, \dots, n\},\ \forall k \in \{1, \dots, K\},\ \forall l \in \{1, \dots, L\}. \end{array} \right.$$

Adding auxiliary variables to the inequality constraints, this problem becomes a linear programming problem with partial entropic perturbation.

Another model for households locating is obtained by choosing the minimization of the total of journey-to-work transportation costs. By refining the elements of the above model, we assume that the following elements are now known, for any $i, j \in \{1, \dots, n\}$, $k \in \{1, \dots, K\}$ and $l \in \{1, \dots, L\}$:

- b_{ijkl} = the budget that a type k household is willing to allocate for purchasing and living in a type l house in zone i and working in zone j (under the assumption that every household has only one key worker);
- c_{ijkl} = the cost that must be allocated by a type k household to living in a type l house in zone i and working in zone j;
- L_{il} = the number of type l house available in zone i;

- S_{jk} = the (estimated) number of jobs in zone j for key workers of a type k household.

The urban or regional planning for households locating consists now in determining the numbers x_{ijkl} of type k households living in a type l house in zone i and working in zone j, for any $i, j \in \{1, \ldots, n\}$, $k \in \{1, \ldots, K\}$ and $l \in \{1, \ldots, L\}$. Hence $(x_{ijkl})_{i,j,k,l}$ is now an optimal solution for the following linear programming problem

$$\text{(P11.2.3)} \quad \left| \begin{array}{l} \max \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{K} \sum_{l=1}^{L} (b_{ijkl} - c_{ijkl}) x_{ijkl} \text{ s.t.} \\ \sum_{j=1}^{n} \sum_{k=1}^{K} x_{ijkl} \leq L_{il},\ \forall i \in \{1, \ldots, n\},\ \forall l \in \{1, \ldots, L\}, \\ \sum_{i=1}^{n} \sum_{l=1}^{L} x_{ijkl} = S_{jk},\ \forall j \in \{1, \ldots, n\},\ \forall k \in \{1, \ldots, K\}, \\ x_{ijkl} \geq 0,\ \forall i, j \in \{1, \ldots, n\},\ \forall k \in \{1, \ldots, K\},\ \forall l \in \{1, \ldots, L\}. \end{array} \right.$$

Again, when the planner disposes of a lower bound M for the total bidding power, then $(x_{ijkl})_{i,j,k,l}$ is an optimal solution for the following problem, according to the MEP:

$$\text{(P11.2.4)} \quad \left| \begin{array}{l} \max - \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{K} \sum_{l=1}^{L} x_{ijkl} \ln x_{ijkl} \text{ s.t.} \\ \sum_{j=1}^{n} \sum_{k=1}^{K} x_{ijkl} \leq L_{il},\ \forall i \in \{1, \ldots, n\},\ \forall l \in \{1, \ldots, L\}, \\ \sum_{i=1}^{n} \sum_{l=1}^{L} x_{ijkl} = S_{jk},\ \forall j \in \{1, \ldots, n\},\ \forall k \in \{1, \ldots, K\}, \\ \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{K} \sum_{l=1}^{L} (b_{ijkl} - c_{ijkl}) x_{ijkl} \geq M, \\ x_{ijkl} \geq 0,\ \forall i, j \in \{1, \ldots, n\},\ \forall k \in \{1, \ldots, K\},\ \forall l \in \{1, \ldots, L\}. \end{array} \right.$$

Similarly to (P11.2.2), problem (P11.2.4) can be written as a linear programming problem with partial entropic perturbation.

11.3 Industrial production planning

When planning the industrial production of a country or region, one important problem consists of estimating the *technical coefficient matrix*

$\mathrm{A} = (a_{ij})_{i,j\in\{1,\ldots,n\}}$, where n is the number of industry sectors and, for any $i, j \in \{1, \ldots, n\}$, a_{ij} represents the amount of input from sector i to sector j per unit of the output of sector j. Therefore

$$a_{ij} = \frac{z_{ij}}{X_j},\ \forall i, j \in \{1, \ldots, n\},$$

where z_{ij} represents the sales input from sector i to sector j, and X_j represents the total output of sector j. Also, we have

$$X_i = \sum_{j=1}^{n} z_{ij} + Y_i,\ \forall i \in \{1, \ldots, n\},$$

where Y_i represents the amount of output from sector i to beneficiaries outside the analyzed sectors, like as the government or the foreign markets.

Usually, the estimation of the current values a_{ij} of technical coefficients is based on some known estimated values $a_{ij}^{(0)}$ of the previous values of these coefficients and on some known estimated values l_i and c_i of the total inter-industry outputs (sales) and inputs (purchases), respectively, for each sector $i \in \{1, \ldots, n\}$, i.e.

$$\sum_{j=1}^{n} z_{ij} = l_i,\ \forall i \in \{1, \ldots, n\}, \quad \sum_{i=1}^{n} z_{ij} = c_j,\ \forall j \in \{1, \ldots, n\},$$

where

$$\sum_{i=1}^{n} l_i = \sum_{j=1}^{n} c_j.$$

Also, by using some known estimated values Y_i, $i \in \{1, \ldots, n\}$ of the current outputs from industry sectors to beneficiaries outside the analyzed sectors, we obtain the estimated values

$$X_i = l_i + Y_i,\ i \in \{1, \ldots, n\}$$

of total output for each sector i.

Among the measures of deviation of current values a_{ij} of technical coefficients from their previous values $a_{ij}^{(0)}$, we can use the generalized relative entropy

$$H(\mathrm{A}; \mathrm{A}^{(0)}) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \ln \frac{a_{ij}}{a_{ij}^{(0)}},$$

where $\mathrm{A}^{(0)} = (a_{ij}^{(0)})_{i,j\in\{1,\ldots,n\}}$, with the assumption that $a_{ij} = 0$ for any $i, j \in \{1, \ldots, n\}$ for which $a_{ij}^{(0)} = 0$.

Hence, the technical coefficient matrix A is an optimal solution for the following optimization problem

$$
\text{(P11.3.1)}\quad \left|\begin{array}{l}
\min H(\mathrm{A};\mathrm{A}^{(0)}) = \displaystyle\sum_{(i,j)\in\mathcal{K}} f_{ij}a_{ij} + \sum_{(i,j)\in\mathcal{K}} a_{ij}\ln a_{ij} \text{ s.t.}\\
\displaystyle\sum_{j=1}^{n} X_j a_{ij} = l_i,\ \forall i\in\{1,\ldots,n\},\\
\displaystyle\sum_{i=1}^{n} a_{ij} = \frac{c_j}{X_j},\ \forall j\in\{1,\ldots,n\},\\
a_{ij}\geq 0,\ \forall i,j\in\{1,\ldots,n\},\\
a_{ij}=0,\ \forall (i,j)\in\{1,\ldots,n\}\times\{1,\ldots,n\}\setminus\mathcal{K},
\end{array}\right.
$$

where

$$\mathcal{K} = \left\{(i,j)\ \middle|\ a_{ij}^{(0)} > 0,\ i,j\in\{1,\ldots,n\}\right\}$$

and

$$f_{ij} = -\ln a_{ij}^{(0)},\ \forall (i,j)\in\mathcal{K}.$$

We remark that this problem is a linear programming problem with entropic perturbation. We can use the geometric programming method to solve problem (P11.3.1) and we obtain that this problem has a unique optimal solution of the form

$$a_{ij} = r_i a_{ij}^{(0)} s_j,\ \forall i,j\in\{1,\ldots,n\},$$

i.e.

$$\mathrm{A} = \mathrm{R}\mathrm{A}^{(0)}\mathrm{S},$$

where R and S are the $n\times n$ diagonal matrices with the diagonal entries r_i ($i\in\{1,\ldots,n\}$) and s_j ($j\in\{1,\ldots,\text{n}\}$), respectively (and all other entries equal to zero). We can iterate the last equality to estimate the technical coefficient matrix at m consecutive times, namely

$$\mathrm{A}^{(k)} = \mathrm{R}^{(k)}\mathrm{A}^{(k-1)}\mathrm{S}^{(k)},\ k\in\{1,\ldots,m\}.$$

Therefore, we obtain a method for estimating the matrix $\mathrm{A}=\mathrm{A}^{(m)}$, called the *RAS algorithm*.