TOMAS CIPRA

Financial and Insurance Formulas

Physica-Verlag

A Springer Company

Financial and Insurance Formulas

Tomas Cipra

Financial and Insurance Formulas

Physica-Verlag

Tomas Cipra
Charles University of Prague
Dept. of Statistics
Faculty of Mathematics and Physics
Sokolovska 83
186 75 Prague
Czech Republic
cipra@karlin.mff.cuni.cz

ISBN 978-3-7908-2592-3 e-ISBN 978-3-7908-2593-0
DOI 10.1007/978-3-7908-2593-0
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2010925641

Printed on acid-free paper

Physica-Verlag is a brand of Springer-Verlag Berlin Heidelberg
Springer-Verlag is part of Springer Science+Business Media (www.springer.com)

To my wife and parents

Contents

List of Symbols

a.s.	almost surely (i.e. with probability one)
argmin	argument minimizing a function (similarly for argmax)
B	lag operator
cov(X, Y)	covariance of random variables X and Y
E(X)	mean value (mean, expectation) of random variable X
$\mathrm{E}_P(X)$	mean value of X with respect to probability P
$\mathrm{E}(X\,\|\,\mathbf{Y})$	conditional mean value of X under condition $\mathbf{Y}$ (σ-algebra)
$\mathrm{E}(X\,\|\,Y)$	conditional mean value of X under condition Y (random variable)
e	base of natural logarithms
$r\mathrm{e}^{\mathrm{i}x}$	trigonometric form of complex numbers ($\mathrm{e}^{\mathrm{i}x} = \cos x + i \sin x$)
$f(x)$	probability density of random variable X
$f(x_1, ..., x_n)$	joint probability density of random variables $X_1, ..., X_n$
$f(x_1\|x_2)$	conditional probability density
$F(x)$	distribution function of random variable X
$F(x_1, ..., x_n)$	joint distribution function of random variables $X_1, ..., X_n$
$F_\alpha(n_1,n_2)$	α -quantile of F-distribution with n_1 and n_2 degrees of freedom
i	imaginary unit
iid	independent and identically distributed (random variables)
$I_{(a,\ b)}(x)$	indicator function: $I_{(a,\ b)}(x) = 1$ for $x \in (a, b)$, $= 0$ others
I_A	random indicator function: $I_A = 1$, if a random event A occurred, $= 0$ others
$\ln x$	natural logarithm of x
N	set of natural numbers 1, 2, 3, ...
N_0	set of natural numbers including zero 0, 1, 2, ...
$o(h)$	denotation of an arbitrary function f of argument h such that $\lim\limits_{h\to\infty} f(h)/h = 0$
$O(h)$	denotation of an arbitrary function f of argument h such that the ratio $f(h)/h$ is constrained in vicinity of a point a (mostly $a = 0$ or $+\infty$)
$P(A)$	probability of random event A

$P(A\|B)$	conditional probability of random event A given random event B
Q	set of rational numbers p/q (p, $q \in$ Z, $q \neq 0$)
R	set of real numbers ($-\infty$, ∞)
R^n	set of n-tuples of real numbers
$\mathbf{R}_{xx}$	correlation matrix of random variables $X_1, ..., X_n$
$t_\alpha(n)$	α-quantile of t-distribution with n degrees of freedom
$u_\alpha(n)$	α-quantile of standard normal distribution
var(X)	variance of random variable X
$\bar{x}$	arithmetic mean
$\hat{x}$	mode
$\bar{x}_G$	geometric mean
$\bar{x}_H$	harmonic mean
$\bar{x}_Q$	quadratic mean
x_p	p-quantile
$\hat{x}_{t+k}(t)$	k-period-ahead prediction (for time $t + k$ at time t)
Z	set of integers ..., –1, 0, 1, ...

a^+ or a^-	positive or negative part of number a ($a^+ = \max(0, a)$, $a^- = \min(0, a)$)
$[a]$ or $\{a\}$	integral or nonintegral part of number a
$(a_i)_{i=1, ..., n}$	column vector of type $n \times 1$ with components a_i ($i = 1, ..., n$)
$(a_{ij})_{\substack{i=1, ..., m \\ j=1, ..., n}}$	matrix of type $m \times n$ with elements a_{ij} ($i = 1, ..., m$, $j = 1, ..., n$)
$f(x - 0)$ or $f(x + 0)$	limit from left or right of function f at point x
$f'(x)$	first derivative of function f at point x
$f''(x)$	second derivative of function f at point x
$f^{(j)}(x)$	j-th derivative of function f at point x ($j \in N_0$)
$\frac{dP}{dQ}$	Radon-Nikodym derivative of probability P with respect to Q
$\binom{n}{k}$	binomial coefficient (combinatorial number)

$\mathbf{x}'$	transpose of vector $\mathbf{x}$
$\mathbf{x} \geq \mathbf{y}$	inequality between vectors: it means that $x_i \geq y_i$ for all components
$\mathbf{x} > \mathbf{y}$	inequality between vectors: it means that $x_i \geq y_i$ for all components but at least one of these inequalities must be sharp
$\mathbf{0} = (0, ..., 0)'$	column zero vector for suitable dimension
$\mathbf{e} = (1, ..., 1)'$	column vector of unities for suitable dimension
$\mathbf{I}$	identity matrix for suitable dimension

$\mathbf{A}'$	transpose of matrix $\mathbf{A}$
$\mathbf{A}' = \mathbf{A}$	symmetric matrix (square one)
$\mathbf{A}^{-1}$	inverse matrix to regular (i.e. nonsingular) matrix (square one)
$\mathbf{A} > 0$	positively definite matrix (symmetric one): $\mathbf{x}'\mathbf{Ax} > 0$ for arbitrary $\mathbf{x} \neq 0$
$\mathbf{A} \geq 0$	positively semidefinite matrix (symmetric one): $\mathbf{x}'\mathbf{Ax} \geq 0$ for arbitrary $\mathbf{x}$
$\mathbf{A}_{ij}$	(i, j)-th element of matrix $\mathbf{A}$
$\mathbf{A}_{i\cdot}$	i-th row of matrix $\mathbf{A}$ (row vector)
$\mathbf{A}_{\cdot j}$	j-th column of matrix $\mathbf{A}$ (column vector)
$\mathbf{A}{\cdot}\mathbf{A} = \mathbf{A}$	idempotent matrix (square one)
$\det(\mathbf{A}) = \lvert\mathbf{A}\rvert$	determinant of matrix $\mathbf{A}$ (square one)
$\mathbf{diag}\{d_1, ..., d_n\}$	diagonal matrix (square one) with elements d_1, ..., d_n on principal diagonal
$\mathrm{tr}(\mathbf{A})$	trace of matrix $\mathbf{A}$ (square one), i.e. the sum of diagonal elements
$h(\mathbf{A})$	rank of matrix $\mathbf{A}$
$\mathbf{A}{\cdot}\mathbf{a} = \lambda{\cdot}\mathbf{a}$	eigenvalue λ and corresponding eigenvector $\mathbf{a}$ of matrix $\mathbf{A}$ (square one)
$\mathbf{a}_1, ..., \mathbf{a}_n$	orthonormal eigenvectors of matrix $\mathbf{A}$ $(n \times n)$: $\mathbf{a}'_i\mathbf{a}_j = \delta_{ij}$, $(i, j = 1, \ldots, n)$

$\chi^2_\alpha(n)$	α-quantile of χ^2-distribution with n degrees of freedom
δ_{ij}	Kronecker delta: $\delta_{ij} = 1$ for $i = j$, $= 0$ for $i \neq j$
Φ	distribution function of normal distribution $N(0, 1)$
γ_1	kurtosis
γ_2	skewness
μ_k	kth central moment
μ'_k	kth moment
$\rho(X, Y)$	correlation coefficient of random variables X and Y
$\sigma(X)$	standard deviation of random variable X
$\mathbf{\Sigma}_{xx}$	covariance matrix of random variables $X_1, ..., X_n$
$(\Omega, \boldsymbol{B}, P)$	probability space with set of elementary events Ω, σ-algebra $\boldsymbol{B}$ and probability measure P

Chapter 1
Introduction

Financial and insurance calculations become more and more frequent and helpful for many users not only in their profession life but sometimes even in their personal life. Therefore a survey of formulas of financial and insurance mathematics that can be applied to such calculations seems to be a suitable aid. In some cases one should use instead of the term *formula* more suitable terms of the type *method*, *procedure* or *algorithm* since the corresponding calculations cannot be simply summed up to a single expression, and a verbal description without introducing complicated symbols is more appropriate.

The survey has the following ambitions:

- The formulas should be applicable in practice: it has motivated their choice for this survey first and foremost. On the other hand it is obvious that by time one puts to use in practice seemingly very abstract formulas of higher mathematics, e.g. when pricing financial derivatives, evaluating financial risks, applying accounting principles based on fair values, choosing alternative risk transfers ARL in insurance, and the like.
- The formulas should be error-free (though such a goal is not achievable in full) since in the financial and insurance framework one publishes sometimes in a hectic way various untried formulas and methods that may be incorrect. Of course, the formulas are introduced here without proofs because their derivation is not the task of this survey.
- The formulas should be systematically sorted and described including a simple denotation that enables a quick and operative searching. Explanation and references to related parts of the survey are often attached to some formulas so that one can browse and look up in the text in an effective way. The detailed *Index* is also helpful for this purpose.
- The formulas should be presented in the form that is in average the most frequent and the most conventional one in practice.
- The formulas should be sufficiently self-contained. Therefore formulas of related disciplines (e.g. from statistics, theory of probability, demography and others) are also given in final chapters.

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_1,

The mathematical level of the formulas and methods ranges from simple ones exploiting only an arithmetic to very sophisticated matters of higher mathematics (e.g. the stochastic calculus, and the like). The author hopes that users find in this survey their level of acceptability corresponding to the problems they solve. The survey contains also “Mathematical Compendium” to remind some basic mathematical principles, and chapters that are related in direct or indirect way to financial and insurance analysis: “Descriptive and Mathematical Statistics”, “Econometrics”, “Index Numbers”, “Stochastic Processes” and “Statistical Analysis of Time Series”. One attaches also “List of Symbols” for symbols that are frequent in the text (however, special symbols may be explained in the context of particular formulas).

Acknowledgement The author thanks for various forms of help to Dr. K. Janeček, Dr. P. Myška and Dr. H. Zaňková. The work is a part of the research project MSM0021620839 (Czech Republic).

Part I
Financial Formulas

Chapter 2
Simple Interest and Discount

Abstract Chapter 2 contains basic formulas for simple interest and discount: 2.1. Simple Interest, 2.2. Calendar Conventions, 2.3. Simple Interest with Principals Credited *m*thly, 2.4. Simple Discount.

2.1 Simple Interest

- Interest (1) from the point of view of *debtor*: is the price paid for borrowed money; (2) from the point of view of *creditor*: is the reward for postponed consumption and uncertainty of investment; the interests can be of various types:
  - *active*: are interests paid by clients for credits (loans) from banks
  - *effective*: are actual interests respecting various effects that influence returns (due to conversions m times per year, time shifts, charges, and the like)
  - *nominal*: are quoted interests that do not respect the actual conversion periods
  - *passive*: are interests paid by banks for deposits from clients
  - *real*: are interests adjusted for inflation
  - *returns*: are yields (gains, profits) due to investment
- *Simple interest* model: means that a loan has an interest calculated at any period entirely from the original principal so that the amount due increases linearly

  Denotation:

i annual (p.a.) interest rate given as a decimal value
p annual (p.a.) interest rate given as a per cent value: $i = p/100$
t time measured in years (e.g. $t = 0.5$ means half a year)
k time measured in days: $t = k/365$ (see Sect. 2.2)
I_t interest credited at time t
P *principal* (principal capital, *present value*)
S_t *amount due* at time t (*future value*, terminal value): $S_0 = P$

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_2,

$$I_t = Pit = Pi\frac{k}{365} = P\frac{p}{100}\frac{k}{365}$$

(*simple interest* credited at time t: it is applied when the time t does not exceed 1 year, i.e. $0 \leq t \leq 1$ or $0 \leq k \leq 365$)

$$P = \frac{I_t}{it} \quad (\textit{principal } P)$$

$$i = \frac{I_t}{Pt} \quad (\textit{interest rate } i); \qquad p = \frac{100 I_t}{Pt} \quad (\textit{interest rate } p \text{ given as percent})$$

$$t = \frac{I_t}{Pi} \quad (\textit{time } t); \qquad k = \frac{365 I_t}{Pi} \quad (\textit{time } k \text{ measured in days})$$

$$I = \frac{\mathrm{IN}_1 + \cdots + \mathrm{IN}_n}{\mathrm{ID}}$$

where $\mathrm{IN}_j = (P_j k_j)/100$ is *interest number* for principal P_j and time k_j measured in days ($j = 1, \ldots, n$); $\mathrm{ID} = 365/p$ is *interest divisor* for interest rate p given as per cent

(*simple interest for checking account* (*demand deposit*): principals $P_1, \ldots, P_n$ bear interests within $k_1, \ldots, k_n$ days, respectively, due to a fixed interest rate p)

$$S_t = P + I_t = P(1 + it) = P\left(1 + i\frac{k}{365}\right) = P\left(1 + \frac{p}{100}\frac{k}{365}\right)$$

(*amount due* at time t: is the principal P with simple interest I_t accrued up to time t; it is applied when the time t does not exceed 1 year, i.e. $0 \leq t \leq 1$ or $0 \leq k \leq 365$)

$$P = \frac{S_t}{1 + it} \quad (\textit{principal } P)$$

$$i = \frac{S_t - P}{Pt} \quad (\textit{interest rate } i); \qquad p = \frac{100(S_t - P)}{Pt} \quad (\textit{interest rate } p \text{ given as percent})$$

$$t = \frac{S_t - P}{Pi} \quad (\textit{time } t \text{ in years}); \qquad k = \frac{365(S_t - P)}{Pi} \quad (\textit{time } k \text{ measured in days})$$

$$\bar{t} = \frac{S_1 + \cdots + S_n - P}{Pi}, \quad \text{where } P = \frac{S_1}{1 + it_1} + \cdots + \frac{S_n}{1 + it_n}$$

(*mean pay-off time* for simple interest: is an equivalent time at which all amounts $S_1, \ldots, S_n$ corresponding originally to times $t_1, \ldots, t_n$ could be paid off all at once)

- *Discount rate*: is the interest rate charged on discount loans (short-term funds) by the central bank to commercial banks (such loans provide reserves to banks in a time of need and are a tool of monetary policy: e.g. the central bank increases the discount rate when a higher inflation is expected)
- *Repo rate*: is the interest rate charged by the central bank when purchasing bills of exchange discounted by commercial banks (more generally, repo rates are the rates applied in any repo operation)
- *Interbank interest rates*: are the interest rates for short-term loans among commercial banks (their motivation is the same as in the case of discount rate); e.g. LIBOR (London Interbank Offered Rate) and LIBID (London Interbank Bid Rate) are published daily for leading currencies and various maturities as the trimmed average of eight or sixteen leading interest rates on the interbank market in UK; similarly, one applies FIBOR in Germany (Frankfurt) or EURIBOR in EU

2.2 Calendar Conventions

- *Calendar conventions*: are rules how to count the difference between two dates; in particular, one applies them in formulas of simple interest model (see Sect. 2.1) and simple discount model (see Sect. 2.4)

Denotation:

DMY symbol for date (e.g. the date November 6, 2009 corresponds to $D = 6$, $M = 11$, $Y = 2009$); for a time interval one needs dates $T_1 = D_1M_1R_1$ and $T_2 = D_2M_2R_2$

$$t = \frac{k}{360} \quad \text{or} \quad t = \frac{k}{365} \quad \text{or} \quad t = \frac{k}{\text{act}}$$

(*calendar conventions*: express the time t in interest or discount models as a fraction of year (see examples *thereinafter*); they differ according to countries and financial products)

$$t = \frac{360(R_2 - R_1) + 30(M_2 - M_1) + \min\{D_2;\ 30\} - \min\{D_1;\ 30\}}{360}$$

(*calendar Euro-30/360*: all months have 30 days and all years have 360 days)

$$t = \frac{360(R_2 - R_1) + 30(M_2 - M_1) + D_2^* - D_1^*}{360}$$

(*calendar US-30/360*: the asterisks mean that all dates ending on 31st are changed to 30th as for Euro-30/360 with the only exception, namely if $D_1 < 30$ and $D_2 = 31$ then one changes T_2 to the first day of the next month)

$$t = \frac{T_2 - T_1}{360}$$

(*calendar act/360*: uses the actual number of days of the given period (one denotes it as $T_2 - T_1$), but 360 days of the year are considered in denominator of the corresponding fraction; it is used e.g. in Germany for operations with eurocurrencies and for floaters)

$$t = \frac{T_2 - T_1}{365}$$

(*calendar act/365*: uses the actual number of days of the given period (one denotes it as $T_2 - T_1$) and 365 days of the year are considered in denominator (also for the leap year); it is used e.g. in UK or for short-termed securities on German money-market)

$$t = \frac{k_1}{\text{number of days in beginning year } R_1} + R_2 - R_1 - 1 + \frac{k_2}{\text{number of days in ending year } R_2}$$

(*calendar act/act*: uses the actual number of days of the given period and the actual number of days of particular years; k_1 is the actual number of days of the given period in the beginning year R_1 and k_2 is the actual number of days of the given period in the ending year R_2 (if $R_1 = R_2$, then k_1 is the actual number of days from the beginning of the given period till the end of this year and k_2 is the actual number of days from the beginning of this year till the end of the given period)

$$i_2 = i_1 \frac{t_1}{t_2}$$

(conversion of the rate of return i_1 to i_2 due to change from the calendar convention t_1 to t_2)

$$i_2 = i_1 \frac{365}{360}$$

(example of conversion of the rate of return i_1 to i_2 due to change from the calendar convention $t_1 = \text{act/360}$ to $t_2 = \text{act/365}$)

- Conventions that are applied when the maturity date is not the bank day:
  - *following* day: the maturity date is taken as the following bank day
  - *modified following* day: the maturity date is taken as the following bank day, if it still lies in the same month; otherwise one takes the preceding bank day
  - *preceding* day: the maturity date is taken as the preceding bank day
  - *modified preceding* day: the maturity date is taken as the preceding bank day, if it still lies in the same month; otherwise one takes the following bank day
  - *second-day-after*: the maturity date is taken as the following second bank day

2.3 Simple Interest with Principals Credited *m*thly

- *Simple interest with principals credited mthly*: instead of the usual calculation of the simple interest at the end of an annual period, one divides it to m subperiods (e.g. to months for $m = 12$) as if the same principals r were credited from the beginnings or from the ends of particular subperiods to the end of the given annual period; at this date a single *annual compensation* for all subperiods is paid up

Denotation:

i	annual (p.a.) interest rate
r	amounts of principals credited $m$thly (e.g. monthly for $m = 12$)
R	annual compensation for all subperiods

$$R = rm + r\frac{i}{m}\,(m + (m-1) + \cdots + 1) = r\left(m + \frac{m+1}{2}i\right)$$

(simple interest with principals credited $m$thly from the *beginnings* of particular subperiods: e.g. for $m = 12$ one obtains $R = r\cdot(12 + 6.5\cdot i)$)

$$R = rm + r\frac{i}{m}\,((m-1) + (m-2) + \cdots + 1) = r\left(m + \frac{m-1}{2}i\right)$$

(simple interest with principals credited m-thly from the *ends* of particular subperiods: e.g. for $m = 12$ one obtains $R = r\cdot(12 + 5.5\cdot i)$)

2.4 Simple Discount

- *Simple discount*: is an interest transaction common mainly for short-term loan instruments, i.e. with maturity up to 1 year (bills of exchange, certificates of deposits (CD's), Treasury bills (T-bills), and the like), where the price of the corresponding loan is set down by subtracting the so-called *discount* from the amount due; such a loan makes use of the *discount principle*, i.e. the corresponding interest is credited at the beginning of the discount period (*interest-in-advance*), while in the simple interest model the interest is credited *in arrears* at the end of the interest period

Denotation:

d	annual (p.a.) discount rate given as a decimal value
t	time measured in years (e.g. $t = 0.5$ means half a year)
k	time measured in days: $t = k/365$ (see Sect. 2.2)
D_t	discount (interest-in-advance) credited at the beginning of discount period t

P *principal*
S_t *amount due* at time t

$$D_t = S_t dt = S_t d \frac{k}{365} \quad (\textit{discount} \text{ credited at the beginning of discount period } t)$$

$$P = S_t - D_t = S_t(1 - dt) = S_t \left(1 - d\frac{k}{365}\right) \ (\textit{principal } P)$$

$$d = \frac{S_t - P}{S_t t} \ (\textit{discount rate } d)$$

$$t = \frac{S_t - P}{S_t d} \ (\textit{discount period } t \text{ measured in years})$$

$$k = \frac{365(S_t - P)}{S_t d} \ (\textit{discount period } k \text{ measured in days})$$

$$i = \frac{S_t - P}{Pt} \quad > \quad d = \frac{S_t - P}{S_t t} \ (\text{comparison of interest rate } i \text{ and discount rate } d)$$

$$i = \frac{d}{1 - dt}$$

(relation between interest rate i and discount rate d)

Further Reading

Bosch, K.: Finanzmathematik für Banker. Oldenbourg Verlag, München (2001)
Cissell, R., Cissell, H., Flaspohler, D.C.: Mathematics of Finance. Houghton Mifflin, Boston, MA (1982)
Dupacova, J., Hurt, J., Stepan, J.: Stochastic Modeling in Economics and Finance. Kluwer, Dordrecht (2002)
Grundmann, W.: Finanz- und Versicherungsmathematik. Teubner, Leipzig (1996)
Grundmann, W., Luderer, B.: Formelsammlung: Finanzmathematik, Versicherungsmathematik, Wertpapieranalyse. Teubner, Stuttgart (2001)
Ihrig, H., Pfaumer, P.: Finanzmathematik. Intensivkurs. Oldenbourg Verlag, München (1999)
Khoury, S.J., Parsons, T.D.: Mathematical Methods in Finance and Economics. North Holland, New York (1981)
Knox, D.M., Zima, P., Brown, R.L.: Mathematics of Finance. McGraw-Hill, Sydney, NSW (1984)
Luderer, B., Nollau, V., Vetters, K.: Mathematische Formeln für Wirtschaftswissenschaftler. Teubner, Stuttgart (2000)
McCutcheon, J.J., Scott, W.F.: An Introduction to the Mathematics of Finance. Heinemann, London (1986)
Prakash, A.J., Karels, G.V., Fernandez, R.: Financial, Commercial, and Mortgage Mathematics and Their Applications. Praeger, New York (1987)
Thomsett, M.C.: The Mathematics of Investing. Wiley, New York (1989)

Chapter 3
Compound Interest and Discount

Abstract Chapter 2 contains basic formulas for compound interest and discount: 3.1. Compound Interest, 3.2. Compound Discount, 3.3. Compound Interest and Discount Convertible *m*thly, 3.4. Combination of Simple and Compound Interest.

3.1 Compound Interest

- *Compound interest* model: means that interest after each *interest period* (or *conversion period*) is added to the previous principal and the interest for the next period is calculated from this increased value of the principal so that the amount due increases exponentially (i.e. interest itself earns interest)

Denotation:

i	annual (p.a.) interest rate given as a decimal value
p	annual (p.a.) interest rate given as a per cent value: $i = p/100$
q	*interest factor*: $q = 1 + i$
q^n	*accumulation factor*: $q^n = (1 + i)^n$
n	number of annual interest periods
P	*principal* (principal capital, *present value*)
S_n	*amount due* after n annual interest periods (*future value*, terminal value, time value of capital at time n): $S_0 = P$

$$S_n = P(1+i)^n = P\left(1+\frac{p}{100}\right)^n = Pq^n$$

(*amount due* after n interest periods: is calculated by multiplying the principal P by the accumulation factor, i.e. by the appropriate power of the interest factor)

$$i = \left(\frac{S_n}{P}\right)^{1/n} - 1 = \sqrt[n]{\frac{S_n}{P}} - 1 \text{ (\textit{interest rate } } i\text{)}$$

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_3,

$$n = \frac{\ln S_n - \ln P}{\ln(1+i)} = \frac{\ln(S_n/P)}{\ln q} = \frac{\log S_n - \log P}{\log(1+i)} = \frac{\log(S_n/P)}{\log q}$$

(*number of annual interest periods n*)

$$n \approx \frac{69}{p} + 0.35$$

(*rule 69*: approximative number of interest periods to double the principal)

$$n \approx \frac{72}{p}$$

(*rule 72*: approximative number of interest periods to double the principal)

$$n \approx \frac{110}{p} + 0.52$$

(*rule 110*: approximative number of interest periods to triple the principal)

$$\bar{t} = \frac{\ln(S_1 + \cdots + S_n) - \ln P}{\ln q}, \quad \text{where } P = \frac{S_1}{q^{t_1}} + \cdots + \frac{S_n}{q^{t_n}}$$

(*mean pay-off time* for compound interest model: is an equivalent term at which all amounts $S_1, \ldots, S_n$ due at terms $t_1, \ldots, t_n$ could be paid up all at once)

$$K_n = K_0(1+i_1)(1+i_2)\cdots(1+i_n) = K_0 q_1 q_2 \cdots q_n$$

(amount due after n interest periods with *varying* interest rate)

3.2 Compound Discount

- *Compound discount* model: is the inverse one to the compound interest model (see Sect. 3.1)

  Denotation:

i	annual (p.a.) interest rate given as a decimal value
v	*discount factor*: $v = \frac{1}{1+i}$
v^n	*present value of* 1: $v^n = 1/(1+i)^n$
d	annual (p.a.) *discount rate* given as a decimal value: $d = 1 - v = iv$
n	number of annual discount periods
P	*principal* (*present value*)
S_n	*amount due* after n annual interest periods (*future value*)

$$P = \frac{S_n}{(1+i)^n} = S_n v^n = S_n (1-d)^n$$

(*present value* before n interest periods, i.e. the principal that will accrue to the given amount due S_n: is calculated by multiplying the amount due S_n by the appropriate power of the discount factor; the difference between the amount due and its present value is *compound discount* (compare with the simple discount model $P = S_t(1 - dt)$ in Sect. 2.4))

$$P = S_n \frac{1}{1+i_1} \frac{1}{1+i_2} \cdots \frac{1}{1+i_n} = P v_1 v_2 \cdots v_n = P(1-d_1)(1-d_2)\cdots(1-d_n)$$

(present value before n interest periods with *varying* interest rate)

3.3 Compound Interest and Discount Convertible *m*thly

- *Compound interest convertible mthly*: when the interest (conversion) period does not coincide with the basic annual time unit, the interest rate is called *nominal* but convertible *m*thly with interest rate i/m per conversion period, where in particular
  - $m = 1$, i.e. annually (p.a. = per annum)
  - $m = 2$, i.e. semiannually (p.s. = per semestre)
  - $m = 4$, i.e. quarterly (p.q. = per quartale)
  - $m = 12$, i.e. monthly (p.m. = per mensem)
  - $m = 52$, i.e. weekly (p.sept.= per septimanam)
  - $m = 365$, i.e. daily (p.d. = per diem)

  Denotation:

i annual (p.a.) interest rate (*nominal* interest rate convertible *m*thly)
i/m interest rate per conversion period
k number of conversion periods of length $1/m$ (e.g. the number of months for $m = 12$); in particular, it is $k = m\,n$ when the interests are calculated over n years
P principal (present value)
S_k amount due after k conversion periods (future value)

$$S_k = P\left(1 + \frac{i}{m}\right)^k$$

(*amount due after k conversion periods*: is calculated using the interest rate per conversion period i/m and the number k of conversion periods of length $1/m$)

$$S_k \approx P\left(\sqrt[m]{1+i}\right)^k = P\,(1+i)^{k/m}$$

(amount due after k conversion periods with *root approximation* of the interest factor over one conversion period)

$$i_{\mathrm{ef}} = \left(1 + \frac{i}{m}\right)^m - 1$$

(*effective interest rate*: is the interest rate converted annually that will produce the same amount of interest per year as the nominal rate i converted m times per year; more generally, an interest rate is called effective, if its conversion period and the basic time unit, which is used for all interest rates, are identical; it is applied to put different rates and frequencies of conversion on a comparable basis; it holds $i_{\mathrm{ef}} > i$ (excepting the trivial case $m = 1$ with $i_{\mathrm{ef}} = i$); for a fixed nominal rate i, the effective rate i_{ef} increases with increasing frequency m)

$$i^{(m)} = m\left((1+i)^{1/m} - 1\right)$$

(*conformal interest rate* to given (annual) effective interest rate i and frequency of conversion m: is the nominal interest rate, convertible m times per year, which is equivalent (conformal) to i and m; it holds $i^{(m)} < i$ (excepting the trivial case $m = 1$ with $i^{(1)} = i$))

- *Compound discount* model *convertible mthly*: is the inverse one to the compound interest model convertible $m$thly

Denotation:

d	annual (p.a.) discount rate (*nominal* discount rate convertible $m$thly)
d/m	discount rate per conversion period
k	number of conversion periods of length $1/m$ (e.g. the number of months for $m = 12$); in particular, it is $k = m \cdot n$ when the discounts are calculated over n years

$$P = S_k\left(1 - \frac{d}{m}\right)^k$$

(*present value before k conversion periods*, i.e. the principal that will accrue to the given amount due S_k: is calculated using the discount rate per conversion period d/m and the number k of conversion periods of length $1/m$)

$$P \approx S_k\left(\sqrt[m]{1-d}\right)^k = S_k\,(1-d)^{k/m}$$

(present value before k conversion periods with *root approximation* of the discount factor over one conversion period)

$$d_{\mathrm{ef}} = 1 - \left(1 - \frac{d}{m}\right)^m$$

(*effective discount rate*: is the discount rate converted annually that will produce the same discount per year as the nominal discount rate d converted m times per year; it is applied to put different discount rates and frequencies of conversion on a comparable basis; it holds $d_{ef} < d$ (excepting the trivial case $m = 1$ with $d_{ef} = d$); for a fixed nominal discount rate d, the effective discount rate d_{ef} decreases with increasing frequency m)

$$d^{(m)} = m\left(1 - (1-d)^{1/m}\right)$$

(*conformal discount rate* to given (annual) effective discount rate d and frequency of conversion m: is the nominal discount rate, convertible m times per year, which is equivalent (conformal) to d and m; it holds $d^{(m)} > d$ (excepting the trivial case $m = 1$ with $d^{(1)} = d$))

$$\left(1 + \frac{i^{(m)}}{m}\right)^m = 1 + i; \quad \left(1 - \frac{d^{(m)}}{m}\right)^m = 1 - d$$

$$\frac{1}{d^{(m)}} = \frac{1}{i^{(m)}} + \frac{1}{m}; \quad d^{(m)} = \frac{i^{(m)}}{1 + i^{(m)}/m}; \quad i^{(m)} = \frac{d^{(m)}}{1 - d^{(m)}/m}$$

3.4 Combination of Simple and Compound Interest

- *Combination of simple and compound interest*: combines both types of models in such a way that the simple interest model is applied only to the first and last incomplete years of the model

Denotation:

i	annual (p.a.) interest rate given as a decimal value
p	annual (p.a.) interest rate given as a per cent value: $i = p/100$
n	number of annual interest periods
t_1	incomplete part of the first year ($0 \leq t_1 < 1$)
t_2	incomplete part of the last year ($0 \leq t_2 < 1$)
k_1	number of days credited in the first year
k_1	number of days credited in the last year
P	principal (present value)
S_t	amount due after time t

$$S_t = P(1 + it_1)(1 + i)^n(1 + it_2) = P\left(1 + \frac{p}{100}\frac{k_1}{365}(1 + i)^n\right)\left(1 + \frac{p}{100}\frac{k_2}{365}\right)$$

(*amount due* after time t: the reason of such a combination of both types of models is the fact that creditors prefer for $0 < t < 1$ the simple interest model, but for $t > 1$ the compound interest model (this strategy guarantees them higher interests))

$$S_t \approx P(1+i)^t$$

(amount due after time t with *approximation* of the combined models: $t = t_1 + n + t_2$)

$$S_t = P(1+i)^{[t]}(1+i\{t\})$$

(amount due after time t with *approximation* of the combined models: $[t]$ and $\{t\}$ are integral and nonintegral parts of time t, respectively)

Further Reading

Bosch, K.: Finanzmathematik für Banker. Oldenbourg Verlag, München (2001)

Cissell, R., Cissell, H., Flaspohler, D.C.: Mathematics of Finance. Houghton Mifflin, Boston, MA (1982)

Dupacova, J., Hurt, J., Stepan, J.: Stochastic Modeling in Economics and Finance. Kluwer, Dordrecht (2002)

Grundmann, W.: Finanz- und Versicherungsmathematik. Teubner, Leipzig (1996)

Grundmann, W., Luderer, B.: Formelsammlung: Finanzmathematik, Versicherungsmathematik, Wertpapieranalyse. Teubner, Stuttgart (2001)

Ihrig, H., Pfaumer, P.: Finanzmathematik. Intensivkurs. Oldenbourg Verlag, München (1999)

Khoury, S.J., Parsons, T.D.: Mathematical Methods in Finance and Economics. North Holland, New York (1981)

Knox, D.M., Zima, P., Brown, R.L.: Mathematics of Finance. McGraw-Hill, Sydney, NSW (1984)

Luderer, B., Nollau, V., Vetters, K.: Mathematische Formeln für Wirtschaftswissenschaftler. Teubner, Stuttgart (2000)

McCutcheon, J.J., Scott, W.F.: An Introduction to the Mathematics of Finance. Heinemann, London (1986)

Prakash, A.J., Karels, G.V., Fernandez, R.: Financial, Commercial, and Mortgage Mathematics and Their Applications. Praeger, New York (1987)

Thomsett, M.C.: The Mathematics of Investing. Wiley, New York (1989)

Chapter 4
Continuous Interest and Discount

Abstract Chapter 4 deals with basic formulas for continuous interest and discount.

- *Continuous interest* model: comes from the usual compounding interest model for the number of conversion periods approaching to infinity (i.e. for conversion periods of infinitesimal lengths corresponding to the frequency of conversion $m \to \infty$, see Sect. 3.3); the corresponding conformal interest rate $i^{(\infty)}$ is called the *force of interest* δ

Denotation:

i	annual (p.a.) interest rate given as a decimal value
d	annual (p.a.) discount rate given as a decimal value: $1-d=\frac{1}{1+i}$
t	time measured in years (e.g. $t=1.5$ means 1 year and 6 months)
$i^{(m)}$	conformal interest rate (see Sect. 3.3): $i^{(m)}=m\left((1+i)^{1/m}-1\right)$
$d^{(m)}$	conformal discount rate (see Sect. 3.3): $d^{(m)}=m\left(1-(1-d)^{1/m}\right)$ $=m\left(1-(1+i)^{-1/m}\right)$
δ	force of interest
$\delta(\tau)$	force of interest at time τ
P	principal (present value)
S_t	amount due after time t (future value)

$$\delta=\ln(1+i)=-\ln(1-d)$$

(*force of interest* corresponding to (effective) interest rate i)

$$i=\mathrm{e}^{\delta}-1;\quad d=1-\mathrm{e}^{-\delta}$$

$$i^{(\infty)}=\lim_{m\to\infty} i^{(m)}=\delta;\quad d^{(\infty)}=\lim_{m\to\infty} d^{(m)}=\delta$$

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_4,

$$S_t = P \lim_{m \to \infty} \left(1 + \frac{i^{(m)}}{m}\right)^{tm} = P(1+i)^t = P\mathrm{e}^{\delta t}$$

(*amount due after time t*: is calculated by multiplying the principal P by the appropriate power of the exponential value e^{δ})

$$P = S_t \lim_{m \to \infty} \left(1 - \frac{d^{(m)}}{m}\right)^{tm} = S_t(1-d)^t = S_t\mathrm{e}^{-\delta t}$$

(*present value before time t*, i.e. the principal that will accrue to the given amount due S_t: is calculated by multiplying the amount due S_t by the appropriate power of the exponential value $\mathrm{e}^{-\delta}$)

$$\bar{t} = \frac{\ln(S_1 + \cdots + S_n) - \ln P}{\delta}, \text{ where } P = S_1\mathrm{e}^{-\delta t_1} + \cdots + S_n\mathrm{e}^{-\delta t_n}$$

(*mean pay-off time* for continuous interest model: is an equivalent term at which all amounts $S_1, \ldots, S_n$ due at terms $t_1, \ldots, t_n$ could be paid up all at once)

$$S_t = S_s\mathrm{e}^{\int_s^t \delta(\tau)\,\mathrm{d}\tau} = S_s q(s,t)$$

(*future value* at the end of time interval from s to t ($s < t$) with *varying* force of interest $\delta(\tau)$ (*amount due*): $q(s,t) = \exp\left\{\int_s^t \delta(\tau)\,\mathrm{d}\tau\right\}$ is the accumulation factor corresponding to the force of interest $\delta(\tau)$; in particular, for $s = 0$ one denotes $q(t) = q(0,t)$)

$$S_s = S_t\mathrm{e}^{-\int_s^t \delta(\tau)\,\mathrm{d}\tau} = S_t v(s,t)$$

(*present value* at the beginning of time interval from s to t ($s < t$) with *varying* force of interest $\delta(\tau)$ (*principal*): $v(s,t) = \exp\left\{-\int_s^t \delta(\tau)\,\mathrm{d}\tau\right\}$ is the present value of 1 corresponding to the force of interest $\delta(\tau)$; in particular, for $s = 0$ one denotes $v(t) = v(0,t)$)

$$\frac{1}{S_t}\frac{\mathrm{d}S_t}{\mathrm{d}t} = \frac{\mathrm{d}\ \ln S_t}{\mathrm{d}t} = \delta(t)$$

(differential equation for S_t: in addition, one must prescribe a value S_s at time s ($s < t$) as the boundary condition)

$$\mathrm{d}S_t = \delta(t)S_t\,\mathrm{d}t \text{ (equivalent form of differential equation for } S_t\text{)}$$

	i	p	q	v	d	δ
i	i	$100i$	$q-1$	$\frac{1-v}{v}$	$\frac{d}{1-d}$	$e^{\delta}-1$
p	$100i$	p	$100(q-1)$	$100\frac{1-v}{v}$	$100\frac{d}{1-d}$	$100(e^{\delta}-1)$
q	$1+i$	$1+\frac{p}{100}$	q	$\frac{1}{v}$	$\frac{1}{1-d}$	e^{δ}
v	$\frac{1}{1+i}$	$\frac{100}{100+p}$	$\frac{1}{q}$	v	$1-d$	$e^{-\delta}$
d	$\frac{i}{1+i}$	$\frac{p}{100+p}$	$\frac{q-1}{q}$	$1-v$	d	$1-e^{-\delta}$
δ	$\ln(1+i)$	$\ln\left(1+\frac{p}{100}\right)$	$\ln q$	$-\ln v$	$\ln\left(\frac{1}{1-d}\right)$	δ

$$\delta(\tau) = \alpha + \frac{\beta}{1 + \beta\gamma\ e^{\beta\cdot\tau}}$$

(*Stoodley's model* of force of interest: in practice, the parameters α, β, γ must be estimated or chosen properly)

$$v(t) = \frac{1}{1+\gamma} e^{-(\alpha+\beta)t} + \frac{\gamma}{1+\gamma} e^{-\alpha t}$$

(present value of 1 according to Stoodley's model)

Further Reading

Cissell, R., Cissell, H., Flaspohler, D.C.: Mathematics of Finance. Houghton Mifflin, Boston, MA (1982)

Dupacova, J., Hurt, J., Stepan, J.: Stochastic Modeling in Economics and Finance. Kluwer, Dordrecht (2002)

McCutcheon, J.J., Scott, W.F.: An Introduction to the Mathematics of Finance. Heinemann, London (1986)

Chapter 5
Classical Analysis of Interest Rates

Abstract Chapter 5 deals with basic theory of interest rates: 5.1. Risk-Free Interest Rate and Real Interest Rate, 5.2. Term Structure of Interest Rates.

5.1 Risk-Free Interest Rate and Real Interest Rate

- *Decomposition of interest rate* to factor components: separates factors corresponding usually to particular interest premiums that are charged for in practice

  Denotation:

i	*nominal* interest rate (before decomposition to factor components)
i_{f}	*risk-free* interest rate *rfr* (considered in practice to be really "free of risk")
i_{infl}	*inflation premium* (expected *rate of inflation*)
i_{default}	*default risk premium* (charged for the risk that the debtor will not pay the principal or the interest)
i_{liquid}	*liquidity premium* (charged for the risk that particular assets are not readily convertible into cash without considerable costs)
i_{mat}	*maturity risk premium* (charged for the risk produced by possible changes of interest rates during the life of particular assets, e.g. reinvestment risk or volatility risk)

$$1 + i = (1 + i_{\text{f}})(1 + i_{\text{infl}})(1 + i_{\text{deault}})(1 + i_{\text{liquid}})(1 + i_{\text{mat}})$$

(*decomposition of interest rate*)

$$i \approx i_{\text{f}} + i_{\text{infl}} + i_{\text{default}} + i_{\text{liquid}} + i_{\text{mat}}$$

(*approximative* decomposition of interest rate)

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_5,

- *Inflation*: is the currency depreciation in consequence of rising prices (the opposite phenomenon is *deflation*)
- *Rate of inflation* i_{infl}: is the relative increase of the corresponding CPI (*consumer price index*) or the relative increase of the average CPI over a given period; various variants of i_{infl} may be published, e.g.
  - rate of inflation as the increase in average annual CPI (see below table for the Euro area)
  - rate of inflation as the increase in CPI compared with the corresponding month of preceding year
  - rate of inflation as the increase in CPI compared with preceding month
  - rate of inflation as the increase in CPI compared with a base period (e.g. year 2005 = 100)

Year	Annual inflation rate in Euro area i_{infl} (%)
1996	2.2
1997	1.6
1998	1.1
1999	1.1
2000	2.1
2001	2.3
2002	2.2
2003	2.1
2004	2.1
2005	2.2
2006	2.2
2007	2.1
2008	3.3

Source: European Central Bank (Statistical Data Warehouse, sdw.ecb.europa.eu)

- *Real interest rate*: is the nominal interest rate adjusted for inflation

Denotation:

i_{real} real interest rate
i_{tax} tax rate

$$1 + i = (1 + i_{\text{real}})(1 + i_{\text{infl}})$$

(*Fisher's formula*: decomposition of interest rate when adjusting for inflation)

$$i_{\text{real}} = \frac{i - i_{\text{infl}}}{1 + i_{\text{infl}}}$$

(*real interest rate*)

$$i_{\text{real}} \approx i - i_{\text{infl}}$$

(*approximation* for real interest rate: can be applied only when inflation is low)

$$i_{\text{real}}^{\text{net}} = \frac{i \cdot (1 - i_{\text{tax}}) - i_{\text{infl}}}{1 + i_{\text{infl}}}$$

(*net* (i.e. after tax) real interest rate)

5.2 Term Structure of Interest Rates

- *Term structure of interest rates*: concerns the relations among interest rates in time with respect to the current time
- *Spot interest rate*: is the interest rate that is in force during a stipulated period since the current time (i.e. it holds immediately)
- *Forward interest rate*: is the interest rate that is in force during a stipulated period since a future time (i.e. it will start to hold in future)

Denotation:

i_n annual spot interest rate for n years (i.e. for period $(0, n)$, if $t = 0$ denotes the current time)

$i_{t,n}$ annual forward interest rate for n years (i.e. for period $(t, t + n)$)

$$(1 + i_t)^t (1 + i_{t,n})^n = (1 + i_{t+n})^{t+n}$$

$$(1 + i_t)^t = (1 + i_1)(1 + i_{1,1})(1 + i_{2,1}) \cdot \ldots \cdot (1 + i_{t-1,1})$$

$$i_{t,n} = \left(\frac{(1 + i_{t+n})^{t+n}}{(1 + i_t)^t} \right)^{1/n} - 1; \quad i_{t,1} = \frac{(1 + i_{t+1})^{t+1}}{(1 + i_t)^t} - 1$$

(*forward interest rate* expressed by means of spot interest rate)

Further Reading

Brigham, E.F.: Fundamentals of Financial Management. The Dryden Press, Forth Worth, TX (1992)
Cissell, R., Cissell, H., Flaspohler, D.C.: Mathematics of Finance. Houghton Mifflin, Boston, MA (1982)
Dupacova, J., Hurt, J., Stepan, J.: Stochastic Modeling in Economics and Finance. Kluwer, Dordrecht (2002)
McCutcheon, J.J., Scott, W.F.: An Introduction to the Mathematics of Finance. Heinemann, London (1986)
Sears, R.S., Trennepohl, G.L.: Investment Management. The Dryden Press, Forth Worth, TX (1993)
Van Horne, J.C.: Financial Market Rates and Flows. Prentice-Hall, Englewood Cliffs, NJ (1978)

Chapter 6
Systems of Cash Flows

Abstract Chapter 6 deals with analysis of cash flow systems including rules of investment decision: 6.1. Present and Future Value, 6.2. Internal Rate of Return, 6.3. Pay-Back Period, 6.4. Duration, 6.5. Convexity.

- *Cash flows* (CF): are payments (sums) related to various time date in the framework of financial, investment or business transactions and projects. One distinguishes
  - inflows: i.e. amounts received (with positive signs in formulas)
  - outflows: i.e. amounts paid (with negative signs in formulas)
- *Net* cash flows: originate by netting all inflows and outflows that are related to the same time date

6.1 Present and Future Value

- *Valuation interest rate* (*cost of capital*, *opportunity cost rate*): is the rate of return (see Sect. 12.1) that can be earned on alternative investments of the given capital sums; it is used to price cash flow systems
- *Present value* PV of cash flow system: is the price of the given system if we price it by means of the valuation interest rate at the present time $t = 0$

Denotation:

CF_t	cash flow at time t
i	valuation interest rate
PV	present value
NPV	net present value

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_6,

$$\text{PV} = \text{CF}_0 + \frac{\text{CF}_1}{1+i_1} + \cdots + \frac{\text{CF}_n}{(1+i_n)^n} = \sum_{t=0}^{n} \frac{\text{CF}_t}{(1+i_t)^t}$$

(*present value* of cash flows $\text{CF}_0, \text{CF}_1, \ldots, \text{CF}_n$ using *spot* valuation interest rates (i_t is the spot interest rate (over one interest period) for t interest periods), see Sect. 5.2)

$$\text{PV} = \text{CF}_0 + \text{CF}_1 v_1 + \cdots + \text{CF}_n v_n^n = \sum_{t=0}^{n} \text{CF}_t v_t^t, \text{ where } v_t = \frac{1}{1+i_t}$$

(present value calculated by means of *discount factors*, see Sect. 3.2)

$$\text{PV} = \text{CF}_0 + \text{CF}_1 \text{e}^{-\delta_1} + \cdots + \text{CF}_n \text{e}^{-\delta_n n} = \sum_{t=0}^{n} \text{CF}_t\ \text{e}^{-\delta_t t}, \text{ where } \delta_t = \ln(1+i_t)$$

(present value in *continuous interest* model, see Chap. 4)

$$\text{NPV} = \frac{\text{CF}_1}{1+i_1} + \cdots + \frac{\text{CF}_n}{(1+i_n)^n} - C$$

(*net* present value of cash flows $\text{CF}_1, \ldots, \text{CF}_n$ when one separates initial costs C (C can be e.g. the purchase price of a security $P = C$); it is a special case of the general formula for $\text{CF}_0, \text{CF}_1, \ldots, \text{CF}_n$ (see *thereinbefore*) when one puts $\text{CF}_0 = -C$ or $\text{CF}_0 = -P$)

$$\text{PV} = \text{CF}_0 + \frac{\text{CF}_1}{1+i_1} + \frac{\text{CF}_2}{(1+i_1)(1+i_{1,\,1})} + \cdots + \frac{\text{CF}_n}{(1+i_1)(1+i_{1,\,1})\ldots(1+i_{n-1,\,1})}$$
$$= \sum_{t=0}^{n} \frac{\text{CF}_t}{\prod_{k=1}^{t}(1+i_{k-1,\,1})}$$

(present value of cash flows $\text{CF}_0, \text{CF}_1, \ldots, \text{CF}_n$ using *forward* valuation interest rates ($i_{t,\,1}$ is the forward interest rate since time t for the next interest period, i.e. for period $(t, t+1)$), see Sect. 5.2)

$$\text{PV} = \text{CF}_0 + \text{CF}_{t_1} v_{t_1}^{t_1} + \cdots + \text{CF}_{t_n} v_{t_n}^{t_n} = \sum_{k=0}^{n} \text{CF}_{t_k} v_{t_k}^{t_k},$$

where $v_t = \dfrac{1}{1+i_t}$

(present value of cash flows at *irregular* time instants $0 = t_0 < t_1 < \ldots < t_n$, see Sect. 3.2)

$$\mathrm{PV} = \sum_{t:\,\mathrm{CF}_t > 0} \frac{\mathrm{CF}_t}{(1+i_{\mathrm{I}t})^t} + \sum_{t:\,\mathrm{CF}_t < 0} \frac{\mathrm{CF}_t}{(1+i_{\mathrm{O}t})^t}$$

(*present value* of cash flows using different valuation interest rates for inflows (i_I) and for outflows (i_O))

- Cash flows in *continuous time*: can model continuous changes of capital in time due to a varying intensity of payments

  Denotation:

$\mathrm{CF}(t)$	cash flow coming for the period $(0, t)$
$\mathrm{cf}(t)$	force of capital at time t: $\mathrm{cf}(t) = \mathrm{CF}'(t)$
$v(t)$	discount factor corresponding to the force of interest $\delta(\tau)$ over period $(0, t)$: $v(t) = \exp\left\{-\int_0^t \{\delta(\tau)\,\mathrm{d}\tau\}\right\}$ (see Chap. 4)

$$\mathrm{CF}(t_2) - \mathrm{CF}(t_2) = \int_{t_1}^{t_2} \mathrm{cf}(t)\,\mathrm{d}t$$

(cash flows for the period (t_1, t_2))

$$\mathrm{PV} = \int_0^t \mathrm{cf}(t) v(t)\,\mathrm{d}t$$

(present value of cash flows for the period $(0, t)$)

- *Investment decision*: is usually based on (1) *profitability* (see *thereinafter*), (2) *risk* (i.e. the level of uncertainty for expected returns, see Sect. 12.2) and (3) *liquidity* (i.e. the level of convertibility for investments readily into cash); in the case of several investment alternatives, the corresponding projects can be
  - *independent* (an arbitrary number of them can be accepted)
  - *mutually exclusive* (at most one of them can be accepted)
- Profitability of investments can be compared by means of various *rules of investment decision*, mainly
  - *present value rule* (see *thereinafter*)
  - *internal rate of return rule* (see Sect. 6.2)
  - *payback period rule* (see Sect. 6.3)

$$\mathrm{PV} > 0$$

(*present value rule*: if the investment projects are independent (see *thereinbefore*), then all of them with positive present values are accepted; if the investment projects are mutually exclusive, then the one with the highest positive present value is chosen (all calculations are performed with the given valuation interest rate))

- *Future value* FV of system of cash flows: is the price of the given system if we price it by means of the valuation interest rate at a given future time

$$\mathrm{FV} = \mathrm{CF}_0(1+i_n)^n + \mathrm{CF}_1(1+i_{n-1})^{n-1} + \cdots + \mathrm{CF}_n = \sum_{t=0}^{n} \mathrm{CF}_t(1+i_{n-t})^{n-t}$$

(*future value* of cash flows CF_0, CF_1, ..., CF_n using *spot* valuation interest rates (i_t is the spot interest rate (over one interest period) for t interest periods), see Sect. 5.2)

$$\begin{aligned}\mathrm{FV} &= \mathrm{CF}_0(1+i_1)(1+i_{1,\,1})\cdots(1+i_{n-1,\,1}) + \cdots + \mathrm{CF}_{n-1}(1+i_{n-1,\,1}) + \mathrm{CF}_n \\ &= \sum_{t=0}^{n} \mathrm{CF}_t \prod\nolimits_{k=t+1}^{n} (1+i_{k-1,\,1})\end{aligned}$$

(future value of cash flows CF_0, CF_1, ..., CF_n using *forward* valuation interest rates ($i_{t,\,1}$ is the forward interest rate since time t for the next interest period, i.e. for period $(t, t+1)$), see Sect. 5.2)

$$\mathrm{FV} = \mathrm{PV}(1+i)^n,$$

where the valuation interest rate i is constant

6.2 Internal Rate of Return

- *Internal rate of return IRR*: is the valuation interest rate for which the present value of inflows is equal to the present value of outflows; equivalently, it is such an interest rate that equates the present value of the whole cash flow system to zero; sometimes (e.g. for securities) one also uses the term *yield to maturity YTM*; *IRR* does not depend on t (unlike the spot valuation interest rates i_t for calculation of PV, see Sect. 6.1))

  Denotation:

CF_t cash flow at time t

P market price of cash flow system; in the case of securities that are not quoted, P can be also their present value (see Sect. 6.1) estimated by means of *reference* valuation interest rates; therefore the term *price of cash flow system* or simply *price of cash flows* is simply used

y internal rate of return *IRR* (yield to maturity *YTM*)

y^* reference valuation rate of return

$P(y)$ price P of cash flow system considered as a function of internal rate of return y

$$\sum_{t:\,\mathrm{CF}_t>0} \frac{\mathrm{CF}_t}{(1+y)^t} = -\sum_{t:\,\mathrm{CF}_t<0} \frac{\mathrm{CF}_t}{(1+y)^t}$$

(definition relation for *internal rate of return*)

$$P = \frac{\mathrm{CF}_1}{1+y} + \dots + \frac{\mathrm{CF}_n}{(1+y)^n} = \sum_{t=1}^{n} \frac{\mathrm{CF}_t}{(1+y)^t}$$

(equation for calculation of *internal rate of return* of cash flows $\mathrm{CF}_1, \dots, \mathrm{CF}_n$ with market price P at time $t = 0$; in practice, it is the most frequent way how to calculate y (it corresponds to the previous formula for $\mathrm{CF}_0 = -P$); sometimes this relation is looked upon as the definition relation for function $P(y)$ with argument y)

$$P = \sum_{k=1}^{n} \frac{\mathrm{CF}_{t_k}}{(1+y)^{t_k}}$$

(equation for calculation of internal rate of return of cash flows at irregular times $0 < t_1 < \cdots < t_n$)

$$S_j = \sum\nolimits_{t=0}^{j} \mathrm{CF}_t,\ \ j = 0,\ 1,\ \dots,\ n;\ \ S_0 \neq 0\ ;\ \ S_n \neq 0$$

(*existence condition of exactly one positive root* y: the sequence of $S_0, S_1, \dots, S_n$ with zeroes excluded changes the sign just once; $S_0 = \mathrm{CF}_0 = -P$)

$$y \begin{cases} > y^*, & \text{if price of cash flows is decreasing function of internal rate of return} \\ < y^*, & \text{if price of cash flows is increasing function of internal rate of return} \end{cases}$$

(*internal rate of return rule* (see Sect. 6.1): a given investment project is accepted, if its internal rate of return y fulfils the corresponding inequality, where y^* is the reference valuation rate of return)

$$y > y^*, \text{where } \mathrm{CF}_0 = -P < 0, \mathrm{CF}_{t_1} > 0,\ \ldots,\ \mathrm{CF}_{t_n} > 0$$

(internal rate of return rule for an investment project *with exactly one outflow initializing the project* (all remaining cash flows are inflows): the given investment project is accepted, if its internal rate of return y exceeds the reference valuation rate of return y^*)

6.3 Payback Period

- *Payback period* PB: is the number of periods required to recover the initial outflows

Denotation:

CF_t cash flow at time t
i valuation interest rate

$$S_j = \sum\nolimits_{t=0}^{j} \mathrm{CF}_t,\ \ j = 0,\ 1,\ \ldots,\ n$$

$$S_j^{(i)} = \sum\nolimits_{t=0}^{j} \frac{\mathrm{CF}_t}{(1+i_t)^t},\ \ j = 0,\ 1,\ \ldots,\ n$$

$$\mathrm{PB} = k - 1\ - \frac{S_{k-1}}{CF_k},$$

where k is the first index such that $S_k > 0$ ($k - 1$ is the period just preceding the full recovery)
(*payback period* of cash flows $\mathrm{CF}_0 < 0, \ldots, \mathrm{CF}_n > 0$, where only inflows follow several initial outflows)

$$\mathrm{PB} = k - 1\ - \frac{S_{k-1}^{(i)}}{\mathrm{CF}_k/(1+i_k)^k}, \text{where } k \text{ is the first index such that } S_k^{(i)} > 0$$

(*discounted payback period* of cash flows $\mathrm{CF}_0 < 0, \ldots, \mathrm{CF}_n > 0$, where only inflows follow several initial outflows)

$$\min\{PB\}$$

(*payback period rule* (see Sect. 6.1): if the investment projects are mutually exclusive, then the one with the shortest (discounted) payback period is chosen)

6.4 Duration

- *Duration* (expected life of cash flow system, also see Sect. 9.2): is an important instrument of cash flow analysis that enables to investigate
  - sensitivity of a cash flow system to varying interest rates
  - various aspects concerning the length of life of investment portfolios (e.g. in the framework of bond portfolio matching and immunization, see Sect. 9.2)

Denotation:

CF_t	cash flow at time t
y	internal rate of return *IRR* (yield to maturity *YTM*, see Sect. 6.2)
$P(y)$	price P of cash flow system considered as a function of internal rate of return y

$$D = \frac{\sum_{t=1}^{n} t\mathrm{CF}_t(1+y)^{-t}}{\sum_{t=1}^{n} \mathrm{CF}_t(1+y)^{-t}} = \frac{\sum_{t=1}^{n} t\mathrm{CF}_t(1+y)^{-t}}{P} = -\frac{1+y}{P(y)}\frac{\mathrm{d}P(y)}{\mathrm{d}y}$$

((*Macaulay*) *duration* of cash flows $\mathrm{CF}_1, \ldots, \mathrm{CF}_n$ with price P and internal rate of return y; D is a weighted average of times prior to particular cash flows, using the relative present values of the payments as weights; D can be also characterized as the ratio of price and interest elasticity; in particular, for $\mathrm{CF}_1 = \cdots = \mathrm{CF}_{n-1} = 0$ and $\mathrm{CF}_n \neq 0$ (i.e. for deposits or zero-coupon bonds) it holds $D = n$)

$$D \approx -\frac{\Delta P(y)/P(y)}{\Delta y/(1+y)} \approx -\frac{\left(P(y+\Delta y) - P(y-\Delta y)\right)/P(y)}{2\Delta y/(1+y)}$$

(approximation of the duration by means of the relative change of price $\Delta P(y)/P(y)$ and the corresponding relative change of interest factor $\Delta(1+y)/(1+y) = \Delta y/(1+y)$)

$$\Delta P(y)/P(y) \approx -D \cdot \Delta y/(1+y) \quad \text{or} \quad \Delta P(y) \approx -P(y) \cdot D \cdot \frac{\Delta y}{1+y}$$

(approximation of the relative change of price $\Delta P(y)/P(y)$ by means of the duration and the corresponding relative change of interest factor $\Delta(1+y)/(1+y) = \Delta y/(1+y)$; this approximation and its modifications (see *thereinafter*) are used in practice when investigating sensitivity of cash flow systems to varied interest rates (it is based on Taylor expansion))

$$\mathrm{DD} = \sum\nolimits_{t=1}^{n} t\mathrm{CF}_t(1+y)^{-t-1} = \frac{1}{1+y}\sum\nolimits_{t=1}^{n} t\mathrm{CF}_t(1+y)^{-t} = -\frac{\mathrm{d}P(y)}{\mathrm{d}y}$$

(*dollar duration* of cash flows $\mathrm{CF}_1, \ldots, \mathrm{CF}_n$ with price P and internal rate of return y; unlike D, it is a linear function of cash flows, but for $\mathrm{CF}_1 = \ldots = \mathrm{CF}_{n-1} = 0$ and $CF_n \neq 0$ it does not hold $DD = n$)

$$\mathrm{DD} \approx -\frac{\Delta P(y)}{\Delta y} \approx -\frac{P(y+\Delta y) - P(y-\Delta y)}{2\Delta y}$$

(approximation of the dollar duration by means of the change of price $\Delta P(y)$ and the corresponding change of internal rate of return Δy)

$$\Delta P(y) \approx -\mathrm{DD}\Delta y$$

(approximation of the change of price $\Delta P(y)$ by means of the dollar duration and the corresponding change of internal rate of return Δy)

$$y \approx \frac{\mathrm{DD}_m i_m + \mathrm{DD}_n i_n}{\mathrm{DD}_m + \mathrm{DD}_n},$$

where

$$\frac{\mathrm{CF}_m}{(1+y)^m} + \frac{\mathrm{CF}_n}{(1+y)^n} = \frac{\mathrm{CF}_m}{(1+i_m)^m} + \frac{\mathrm{CF}_n}{(1+i_n)^n}$$

(relation among internal rate of return y of the system of two cash flows $\{\mathrm{CF}_m, \mathrm{CF}_n\}$ and the spot interest rates i_m and i_n expressed by means of the dollar durations (e.g. $\mathrm{DD}_m = m \cdot \mathrm{CF}_m \cdot (1+y)^{-m-1}$ is the dollar duration of the cash flow CF_m using the internal rate of return y))

$$\mathrm{MD} = \frac{\sum_{t=1}^{n} t\mathrm{CF}_t(1+y)^{-t-1}}{\sum_{t=1}^{n} \mathrm{CF}_t(1+y)^{-t}} = \frac{1}{1+y}\frac{\sum_{t=1}^{n} t\mathrm{CF}_t(1+y)^{-t}}{P(y)} = -\frac{1}{P(y)}\frac{\mathrm{d}P(y)}{\mathrm{d}y}$$
$$= \frac{D}{1+y} = \frac{\mathrm{DD}}{P(y)}$$

(*modified duration* of cash flows $\mathrm{CF}_1, \ldots, \mathrm{CF}_n$ with price P and internal rate of return y; MD is a ratio of the dollar duration DD and the price P)

$$\mathrm{MD} \approx -\frac{\Delta P(y)/P(y)}{\Delta y} \approx -\frac{\left(P(y+\Delta y) - P(y-\Delta y)\right)/P(y)}{2\Delta y}$$

(approximation of the modified duration by means of the relative change of price $\Delta P(y)/P(y)$ and the corresponding change of internal rate of return Δy)

$$\Delta P(y)/P(y) \approx -\text{MD}\Delta y \quad \text{or} \quad \Delta P(y) \approx -P(y)\text{MD}\Delta y$$

(approximation of the relative change of price $\Delta P(y)/P(y)$ by means of the modified duration and the corresponding change of internal rate of return Δy)

$$\text{MD} = \frac{\sum_{t=1}^{n} t\text{CF}_t \mathrm{e}^{-\delta t}}{\sum_{t=1}^{n} \text{CF}_t \mathrm{e}^{-\delta t}} = \frac{\sum_{t=1}^{n} t\text{CF}_t \mathrm{e}^{-\delta t}}{P(\delta)} = -\frac{1}{P(\delta)} \frac{\mathrm{d}P(\delta)}{\mathrm{d}\delta}$$

(modified duration of cash flows CF_1, ..., CF_n in *continuous interest* model)

6.5 Convexity

$$\text{CX} = \frac{\sum_{t=1}^{n} t(t+1)\text{CF}_t(1+y)^{-t}}{\sum_{t=1}^{n} \text{CF}_t(1+y)^{-t}} = \frac{\sum_{t=1}^{n} t(t+1)\text{CF}_t(1+y)^{-t}}{P} = \frac{(1+y)^2}{P(y)} \frac{\mathrm{d}^2P(y)}{\mathrm{d}y^2}$$

((*Macaulay*) *convexity* of cash flows $\text{CF}_1, \ldots, \text{CF}_n$ with price P and internal rate of return y; CX and its modifications (see *thereinafter*) improve the approximations based only on durations (see Sect. 6.4))

$$\Delta P(y) \approx -P(y)D\frac{\Delta y}{1+y} + \frac{1}{2}P(y)\text{CX}\frac{(\Delta y)^2}{(1+y)^2}$$

(approximation of the change of price $\Delta P(y)$ by means of the duration and the convexity)

$$\text{DCX} = \sum_{t=1}^{n} t(t+1)\text{CF}_t(1+y)^{-t-2} = \frac{1}{(1+y)^2}\sum_{t=1}^{n} t(t+1)\text{CF}_t(1+y)^{-t}$$
$$= \frac{\mathrm{d}^2P(y)}{\mathrm{d}y^2}$$

(*dollar convexity*)

$$\text{DCX} \approx \frac{P(y+\Delta y) + P(y-\Delta y) - 2P(y)}{(\Delta y)^2}$$

(approximation of the dollar convexity)

$$\Delta P(y) \approx -\text{DD}\Delta y + \frac{1}{2}\text{DCX}(\Delta y)^2$$

(approximation of the change of price $\Delta P(y)$ by means of the dollar duration and the dollar convexity)

$$\mathrm{MCX} = \frac{1}{(1+y)^2} \cdot \frac{\sum_{t=1}^{n} t(t+1)\mathrm{CF}_t(1+y)^{-t}}{P(y)} = \frac{1}{P(y)} \frac{\mathrm{d}^2 P(y)}{\mathrm{d}y^2} = \frac{\mathrm{CX}}{(1+y)^2} = \frac{\mathrm{DCX}}{P(y)}$$

(*modified convexity*)

$$\mathrm{MCX} \approx \frac{(P(y+\Delta y) + P(y-\Delta y) - 2P(y))\,/P(y)}{(\Delta y)^2}$$

(approximation of the modified convexity)

$$\Delta P(y) \approx -P(y)\mathrm{MD}\Delta y + \frac{1}{2}P(y)\mathrm{MCX}(\Delta y)^2$$

(approximation of the change of price $\Delta P(y)$ by means of the modified duration and the modified convexity)

Further Reading

Brealey, R.A., Myers, S.C.: Principles of Corporate Finance. McGraw-Hill, New York (1988)
Brigham, E.F.: Fundamentals of Financial Management. The Dryden Press, Forth Worth, TX (1992)
Cissell, R., Cissell, H., Flaspohler, D.C.: Mathematics of Finance. Houghton Mifflin, Boston, MA (1982)
Dupacova, J., Hurt, J., Stepan, J.: Stochastic Modeling in Economics and Finance. Kluwer, Dordrecht (2002)
Khoury, S.J., Parsons, T.D.: Mathematical Methods in Finance and Economics. North Holland, New York (1981)
Knox, D.M., Zima, P., Brown, R.L.: Mathematics of Finance. McGraw-Hill, Sydney, NSW (1984)
McCutcheon, J.J., Scott, W.F.: An Introduction to the Mathematics of Finance. Heinemann, London (1986)
Prakash, A.J., Karels, G.V., Fernandez, R.: Financial, Commercial, and Mortgage Mathematics and Their Applications. Praeger, New York (1987)
Sears, R.S., Trennepohl, G.L.: Investment Management. The Dryden Press, Forth Worth, TX (1993)
Thomsett, M.C.: The Mathematics of Investing. Wiley, New York (1989)
Van Horne, J.C.: Financial Market Rates and Flows. Prentice-Hall, Englewood Cliffs, NJ (1978)

Chapter 7
Annuities

Abstract Chapter 7 contains formulas on financial annuities: 7.1. Annuity Calculus, 7.2. Dynamic Annuities, 7.3. Annuities Payable *m*thly, 7.4. Continuously Payable Annuities, 7.5. Amortization of Debt.

- *Annuity*: is a series of periodic payments (installments) that are repeated regularly in time and are of the same amount, or change according to a given schedule (it is a system of regularly distributed CF's (see Sect. 6)); one can classify the following types of annuities:

  – *annuity-certain*: its payments are guaranteed (e.g. by a contract)
  – *contingent annuity*: its payments are qualified by given conditions which have usually a random character (e.g. a life annuity)
  – *annuity-due*: its periodic payments are made at the beginning of each *payment period* (i.e. the payments are made in advance)
  – *immediate annuity*: its periodic payments are made at the end of each payment period (i.e. the payments are made in arrear)
  – *temporary annuity*: its payments have a limited duration (by a contract, e.g. a 10-year annuity)
  – *perpetuity*: there is no contractual limit of duration (e.g. dividends of a common stock, coupons of the British government bonds known as consols that have no maturity date, awards of Nobel prize)
  – *deferred annuity*: its payments begin not in the first period but after several periods of a fixed *deferment*
  – *annual annuity*: its payment periods are years (e.g. the payments are made once year on the third February)
  – *monthly annuity*: its payment periods are months (e.g. payments are made on the seventh day of each month); similarly for *quarterly annuity*, *daily annuity* etc.

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_7,

  - *unit annuity*: pays out a unit amount in total each year (e.g. the unit annual annuity has annual payments in amount of 1 for a given monetary unit); it is used when constructing annuity formulas

- The formulas in this chapter assume for simplicity that *payment periods coincide with interest periods* (otherwise the formal notation must be more complicated but the formulas are similar)

7.1 Annuity Calculus

- *Present value of annuity*: is PV (see Sect. 6.1) of such a system of CF's that consists of the annuity payments; it can be interpreted as
  - the price of the annuity (*capitalized annuity*)
  - the principal of the debt amortized by the annuity payments
- *Future value of annuity*: is FV (see Sect. 6.1) of such a system of CF's that consists of the annuity payments; it can be interpreted as
  - the amount accumulated by the annuity payments (*accumulation*)
- Classical annuity calculus produces explicit formulas for PV and FV of annuities if one assumes that the interest rate i over one payment (= interest) period remains *the same for all payment periods*, i.e. $i_t = i$ for all t (see Sect. 6.1)

Denotation:

$\mathrm{PV}^{\mathrm{due}}$	present value of annuity-due
$\mathrm{FV}^{\mathrm{due}}$	future value of annuity-due
$\mathrm{PV}^{\mathrm{imm}}$	present value of immediate annuity
$\mathrm{FV}^{\mathrm{imm}}$	future value of immediate annuity
V_t^{due}	value of annuity-due at time t
V_t^{imm}	value of immediate annuity at time t
K	amount of periodic payment
i	interest rate over one payment (= interest) period
q	interest factor (see Sect. 3.1): $q = 1 + i$
v	discount factor (see Sect. 3.2): $v = \frac{1}{1+i}$
d	discount rate over one payment (= interest) period (see Sect. 3.2): $d = 1 - v = i \cdot v$
n	number of payment (= interest) periods (run time of annuity)

$$\ddot{a}_{\overline{n}|} = 1 + v + \cdots + v^{n-1} = \frac{1 - v^n}{1 - v} = \frac{1 - v^n}{d} = \frac{q^n - 1}{q^{n-1}(q - 1)}$$

(present value of unit annuity-due)

$$a_{\overline{n}|} = v + v^2 + \cdots + v^n = \frac{1-v^n}{i} = v\frac{1-v^n}{1-v} = \frac{q^n-1}{q^n(q-1)}$$

(present value of unit immediate annuity)

$$\ddot{s}_{\overline{n}|} = q + q^2 + \cdots + q^n = \frac{q^n-1}{1-v} = \frac{q^n-1}{d} = q\frac{q^n-1}{q-1}$$

(future value of unit annuity-due)

$$s_{\overline{n}|} = 1 + q + \cdots + q^{n-1} = \frac{q^n-1}{i} = \frac{q^n-1}{q-1}$$

(future value of unit immediate annuity)

$$\ddot{a}_{\overline{\infty}|} = 1 + v + \cdots = \frac{1}{1-v} = \frac{1}{d} = \frac{q}{q-1}$$

(present value of unit perpetuity-due)

$$a_{\overline{\infty}|} = v + v^2 + \cdots = \frac{1}{i} = \frac{1}{q-1}$$

(present value of unit immediate perpetuity)

$$\ddot{a}_{\overline{1}|} = 1; \quad a_{\overline{1}|} = v; \quad \ddot{s}_{\overline{1}|} = q; \quad s_{\overline{1}|} = 1$$

$$\ddot{a}_{\overline{0}|} = a_{\overline{0}|} = \ddot{s}_{\overline{0}|} = s_{\overline{0}|} = 0$$

$$\ddot{a}_{\overline{n}|} = qa_{\overline{n}|}; \quad \ddot{s}_{\overline{n}|} = qs_{\overline{n}|}$$

(relations between values of due and immediate annuities)

$$\ddot{a}_{\overline{n}|} = v^n\ddot{s}_{\overline{n}|}; \qquad a_{\overline{n}|} = v^n s_{\overline{n}|}; \qquad \ddot{s}_{\overline{n}|} = q^n\ddot{a}_{\overline{n}|}; \qquad s_{\overline{n}|} = q^n a_{\overline{n}|}; \qquad 1/a_{\overline{n}|} = 1/s_{\overline{n}|} + i$$

(relations between present and future values of annuities)

$$\ddot{a}_{\overline{n+1}|} = v\ddot{a}_{\overline{n}|} + 1; \qquad a_{\overline{n+1}|} = v(a_{\overline{n}|} + 1); \qquad \ddot{s}_{\overline{n+1}|} = q(\ddot{s}_{\overline{n}|} + 1); \qquad s_{\overline{n+1}|} = qs_{\overline{n}|} + 1$$

(recursive relations for present or future values of annuities)

$$\ddot{a}_{\overline{n+1}|} = a_{\overline{n}|} + 1; \quad s_{\overline{n+1}|} = \ddot{s}_{\overline{n}|} + 1$$

(relations between adjacent present or future values of due and immediate annuities)

	$\ddot{a}_{n\rceil}$	$a_{n\rceil}$	$\ddot{s}_{n\rceil}$	$s_{n\rceil}$	q^n	v^n
$\ddot{a}_{n\rceil}$	$\ddot{a}_{n\rceil}$	$qa_{n\rceil}$	$\dfrac{\ddot{s}_{n\rceil}}{1+d\ddot{s}_{n\rceil}}$	$\dfrac{qs_{n\rceil}}{1+is_{n\rceil}}$	$\dfrac{q^n-1}{dq^n}$	$\dfrac{1-v^n}{d}$
$a_{n\rceil}$	$v\ddot{a}_{n\rceil}$	$a_{n\rceil}$	$\dfrac{v\ddot{s}_{n\rceil}}{1+d\ddot{s}_{n\rceil}}$	$\dfrac{s_{n\rceil}}{1+is_{n\rceil}}$	$\dfrac{q^n-1}{iq^n}$	$\dfrac{1-v^n}{i}$
$\ddot{s}_{n\rceil}$	$\dfrac{\ddot{a}_{n\rceil}}{1-d\ddot{a}_{n\rceil}}$	$\dfrac{qa_{n\rceil}}{1-ia_{n\rceil}}$	$\ddot{s}_{n\rceil}$	$qs_{n\rceil}$	$\dfrac{q^n-1}{d}$	$\dfrac{1-v^n}{dv^n}$
$s_{n\rceil}$	$\dfrac{v\ddot{a}_{n\rceil}}{1-d\ddot{a}_{n\rceil}}$	$\dfrac{a_{n\rceil}}{1-ia_{n\rceil}}$	$v\ddot{s}_{n\rceil}$	$s_{n\rceil}$	$\dfrac{q^n-1}{i}$	$\dfrac{1-v^n}{iv^n}$
q^n	$\dfrac{1}{1-d\ddot{a}_{n\rceil}}$	$\dfrac{1}{1-ia_{n\rceil}}$	$1+d\ddot{s}_{n\rceil}$	$1+is_{n\rceil}$	q^n	$\dfrac{1}{v^n}$
v^n	$1-d\ddot{a}_{n\rceil}$	$1-ia_{n\rceil}$	$\dfrac{1}{1+d\ddot{s}_{n\rceil}}$	$\dfrac{1}{1+is_{n\rceil}}$	$\dfrac{1}{q^n}$	v^n

$$\mathrm{PV}_{n\rceil}^{\mathrm{due}} = K\ddot{a}_{n\rceil} = K\frac{1-v^n}{d}; \quad \mathrm{FV}_{n\rceil}^{\mathrm{due}} = K\ddot{s}_{n\rceil} = K\frac{q^n-1}{d};$$

$$\mathrm{PV}_{\infty\rceil}^{\mathrm{due}} = K\ddot{a}_{\infty\rceil} = K\frac{q}{i}$$

(present and future value of annuity-*due*)

$$\mathrm{PV}_{n\rceil}^{\mathrm{imm}} = Ka_{n\rceil} = K\frac{1-v^n}{i}; \quad \mathrm{FV}_{n\rceil}^{\mathrm{imm}} = Ks_{n\rceil} = K\frac{q^n-1}{i};$$

$$\mathrm{PV}_{\infty\rceil}^{\mathrm{imm}} = Ka_{\infty\rceil} = K\frac{1}{i}$$

(present and future value of *immediate* annuity)

$$K = \frac{\mathrm{PV}_{n\rceil}^{\mathrm{due}}}{\ddot{a}_{n\rceil}} = \mathrm{PV}_{n\rceil}^{\mathrm{due}}\frac{d}{1-v^n} = \frac{\mathrm{FV}_{n\rceil}^{\mathrm{due}}}{\ddot{s}_{n\rceil}} = \mathrm{FV}_{n\rceil}^{\mathrm{due}}\frac{d}{q^n-1}$$

(payments-*due* for a given debt or for a given accumulation)

$$K = \frac{\mathrm{PV}_{n\rceil}^{\mathrm{imm}}}{a_{n\rceil}} = \mathrm{PV}_{n\rceil}^{\mathrm{imm}}\frac{i}{1-v^n} = \frac{\mathrm{FV}_{n\rceil}^{\mathrm{imm}}}{s_{n\rceil}} = \mathrm{FV}_{n\rceil}^{\mathrm{imm}}\frac{i}{q^n-1}$$

(*immediate* payments for a given debt or for a given accumulation)

$$n = \frac{1}{\ln q} \ln \frac{Kq}{Kq - \mathrm{PV}^{\mathrm{due}}_{n\rceil} i} = \frac{1}{\ln q} \ln \left(\mathrm{FV}^{\mathrm{due}}_{n\rceil} \frac{i}{Kq} + 1 \right)$$

(number of payments-*due* for a given debt or for a given accumulation)

$$n = \frac{1}{\ln q} \ln \frac{K}{K - \mathrm{PV}^{\mathrm{imm}}_{n\rceil} i} = \frac{1}{\ln q} \ln \left(\mathrm{FV}^{\mathrm{imm}}_{n\rceil} \frac{i}{K} + 1 \right)$$

(number of *immediate* payments for a given debt or for a given accumulation)

$$V_t^{\mathrm{due}} = K(\ddot{s}_t + \ddot{a}_{n-\,t}) = Kq^t \ddot{a}_{n\rceil} \qquad \text{(value of annuity-}\textit{due}\text{ at time } t)$$

$$V_t^{\mathrm{imm}} = K(s_{t\rceil} + a_{n-\,t\rceil}) = Kq^t a_{n\rceil} \qquad \text{(value of }\textit{immediate}\text{ annuity at time } t)$$

$$\bar{t}^{\mathrm{due}} = \frac{1}{\ln q} \ln \frac{nq^n i}{q^n - 1} - 1 \approx \frac{n-1}{2}$$

(mean settlement time of annuity-due: is the run time after which one can settle the annuity-due with n unit payments by a one-off payment of amount n, i.e.: $nv^{\bar{t}^{\mathrm{due}}} = \ddot{a}_{n\rceil}$)

$$\bar{t}^{\mathrm{imm}} = \frac{1}{\ln q} \ln \frac{nq^n i}{q^n - 1} \approx \frac{n+1}{2}$$

(mean settlement time of immediate annuity: is the run time after which one can settle the immediate annuity with n unit payments by a one-off payment of amount n, i.e.:
$nv^{\bar{t}^{\mathrm{imm}}} = a_{n\rceil}$)

$$D^{\mathrm{due}} = \frac{v + 2v^2 + \cdots + (n-1)v^{n-1}}{1 + v + \cdots + v^{n-1}} = \frac{1}{i} - \frac{n}{q^n - 1}$$

(*duration* of annuity-*due* (see Sect. 6.4): is the mean time of payments of the annuity-due weighted by their present values; it can be interpreted as a centre of gravity of present values of payments of the annuity-due or as the time in which the value of the annuity-due is insensitive to changes in the interest rate)

$$D^{\mathrm{imm}} = \frac{v + 2v^2 + \cdots + nv^n}{v + v^2 + \cdots + v^n} = \frac{1 + 2v + \cdots + nv^{n-1}}{1 + v + \cdots + v^{n-1}} = \frac{q}{i} - \frac{n}{q^n - 1} = D^{\mathrm{due}} + 1$$

(*duration* of *immediate* annuity (see Sect. 6.4): is the mean time of payments of the immediate annuity weighted by their present values; it can be interpreted as a centre of gravity of present values of payments of the immediate annuity or as the time in which the value of the immediate annuity is insensitive to changes in the interest rate)

7.2 Dynamic Annuities

$$(I\ddot{a})_{\overline{n}|} = 1 + 2v + \cdots + nv^{n-1} = \frac{1}{d}(\ddot{a}_{\overline{n}|} - nv^n)$$

(present value of unit increasing annuity-due of the type 1, 2, . . . , n)

$$(Ia)_{\overline{n}|} = v + 2v^2 + \cdots + nv^n = \frac{1}{i}(\ddot{a}_{\overline{n}|} - nv^n)$$

(present value of unit increasing immediate annuity of the type 1, 2, . . . , n)

$$(I\ddot{s})_{\overline{n}|} = q^n + 2q^{n-1} + \cdots + nq = \frac{1}{d}(\ddot{s}_{\overline{n}|} - n)$$

(future value of unit increasing annuity-due of the type 1, 2, . . . , n)

$$(Is)_{\overline{n}|} = q^{n-1} + 2q^{n-2} + \cdots + n = \frac{1}{i}(\ddot{s}_{\overline{n}|} - n)$$

(future value of unit increasing immediate annuity of the type 1, 2, . . . , n)

$$(D\ddot{a})_{\overline{n}|} = n + (n-1)v + \cdots + v^{n-1} = \frac{1}{d}(n - a_{\overline{n}|})$$

(present value of unit decreasing annuity-due of the type n, n – 1, . . . , 1)

$$(Da)_{\overline{n}|} = nv + (n-1)v^2 + \cdots + v^n = \frac{1}{i}(n - a_{\overline{n}|})$$

(present value of unit decreasing immediate annuity of the type n, n – 1, . . . , 1)

$$(D\ddot{s})_{\overline{n}|} = nq^n + (n-1)q^{n-1} + \cdots + q = \frac{1}{d}(nq^n - s_{\overline{n}|})$$

(future value of unit decreasing annuity-due of the type n, n – 1, . . . , 1)

$$(Ds)_{\overline{n}|} = nq^{n-1} + (n-1)q^{n-2} + \cdots + 1 = \frac{1}{i}(nq^n - s_{\overline{n}|})$$

(future value of unit decreasing immediate annuity of the type n, n – 1, . . . , 1)

$$(I\ddot{a})_{\overline{\infty}|} = 1 + 2v + \cdots = \frac{1}{d^2}$$

(present value of unit increasing perpetuity-due of the type 1, 2, . . .)

$$(Ia)_{\infty\rceil} = v + 2v^2 + \cdots = \frac{1}{id}$$

(present value of unit increasing immediate perpetuity of the type 1, 2, ...)

$$(I\ddot{a})_{n\rceil} = q(Ia)_{n\rceil}; \qquad (I\ddot{s})_{n\rceil} = q(Is)_{n\rceil};$$
$$(D\ddot{a})_{n\rceil} = q(Da)_{n\rceil}; \qquad (D\ddot{s})_{n\rceil} = q(Ds)_{n\rceil}$$

(relations between values of due and immediate dynamic annuities)

$$(I\ddot{s})_{n\rceil} = q^n(I\ddot{a})_{n\rceil}; \qquad (Is)_{n\rceil} = q^n(Ia)_{n\rceil};$$
$$(D\ddot{s})_{n\rceil} = q^n(D\ddot{a})_{n\rceil}; \qquad (Ds)_{n\rceil} = q^n(Da)_{n\rceil}$$

(relations between present and future values of dynamic annuities)

$$(I\ddot{a})_{n\rceil} + (D\ddot{a})_{n\rceil} = (n+1)\ddot{a}_{n\rceil}; \quad (Ia)_{n\rceil} + (Da)_{n\rceil} = (n+1)a_{n\rceil}$$

$$\mathrm{PV}_{n\rceil}^{\mathrm{due}} = K\left(\ddot{a}_{n\rceil} + \delta v(I\ddot{a})_{n-1\rceil}\right); \quad \mathrm{FV}_{n\rceil}^{\mathrm{due}} = K\left(\ddot{s}_{n\rceil} + \delta(I\ddot{s})_{n-1\rceil}\right);$$
$$\mathrm{PV}_{\infty\rceil}^{\mathrm{due}} = K\frac{q}{i}\left(1 + \frac{\delta}{i}\right)$$

(present and future value of *arithmetically increasing* annuity-*due* of the type K, $K(1+\delta)$, $K(1+2\delta)$, ... with increments *proportional* to the payment K; in fact it means an arithmetic decrease for $\delta < 0$)

$$\mathrm{PV}_{n\rceil}^{\mathrm{imm}} = K\left(a_{n\rceil} + \delta v(Ia)_{n-1\rceil}\right); \quad \mathrm{FV}_{n\rceil}^{\mathrm{imm}} = K\left(s_{n\rceil} + \delta(Is)_{n-1\rceil}\right);$$
$$\mathrm{PV}_{\infty\rceil}^{\mathrm{imm}} = K\frac{1}{i}\left(1 + \frac{\delta}{i}\right)$$

(present and future value of *arithmetically increasing immediate* annuity of the type K, $K(1+\delta)$, $K(1+2\delta)$, ... with increments *proportional* to the payment K)

$$\mathrm{PV}_{n\rceil}^{\mathrm{due}} = K\left(\ddot{a}_{n\rceil} + \delta(D\ddot{a})_{n-1\rceil}\right); \quad \mathrm{FV}_{n\rceil}^{\mathrm{due}} = K\left(\ddot{s}_{n\rceil} + \delta q(D\ddot{s})_{n-1\rceil}\right);$$
$$\mathrm{PV}_{\infty\rceil}^{\mathrm{due}} = K\frac{q}{i}\left(1 + \delta\left(n - \frac{q}{i}\right)\right)$$

(present and future value of *arithmetically decreasing* annuity-*due* of the type ..., $K(1+2\delta)$, $K(1+\delta)$, K with decrements *proportional* to the payment K; in fact it means an arithmetic increase for $\delta < 0$)

$$\mathrm{PV}_{n\rceil}^{\mathrm{imm}} = K\left(a_{n\rceil} + \delta(Da)_{n-1\rceil}\right); \quad \mathrm{FV}_{n\rceil}^{\mathrm{imm}} = K\left(s_{n\rceil} + \delta q(Ds)_{n-1\rceil}\right);$$

$$\mathrm{PV}_{\infty\rceil}^{\mathrm{imm}} = K\frac{1}{i}\left(1 + \delta\left(n - \frac{q}{i}\right)\right)$$

(present and future value of *arithmetically decreasing immediate* annuity of the type $\ldots, K(1+2\delta), K(1+\delta), K$ with decrements *proportional* to the payment K)

$$\mathrm{PV}_{n\rceil}^{\mathrm{due}} = K\ddot{a}_{n\rceil} + \frac{\Delta}{i}\left(\ddot{a}_{n\rceil} - \frac{n}{q^{n-1}}\right); \ \mathrm{FV}_{n\rceil}^{\mathrm{due}} = K\ddot{s}_{n\rceil} + \frac{\Delta}{i}\left(\ddot{s}_{n\rceil} - nq\right);$$

$$\mathrm{PV}_{\infty\rceil}^{\mathrm{due}} = \frac{q}{i}\left(K + \frac{\Delta}{i}\right)$$

(present and future value of *arithmetically increasing* annuity-*due* of the type K, $K+\Delta$, $K+2\Delta, \ldots$ with increments *independent* of the payment K; in fact it means an arithmetic decrease for $\Delta < 0$)

$$\mathrm{PV}_{n\rceil}^{\mathrm{imm}} = Ka_{n\rceil} + \frac{\Delta}{i}\left(a_{n\rceil} - \frac{n}{q^n}\right); \quad \mathrm{FV}_{n\rceil}^{\mathrm{imm}} = Ks_{n\rceil} + \frac{\Delta}{i}\left(s_{n\rceil} - n\right);$$

$$\mathrm{PV}_{\infty\rceil}^{\mathrm{imm}} = \frac{1}{i}\left(K + \frac{\Delta}{i}\right)$$

(present and future value of *arithmetically increasing immediate* annuity of the type $K, K+\Delta, K+2\Delta, \ldots$ with increments *independent* of the payment K)

$$\mathrm{PV}_{n\rceil}^{\mathrm{due}} = K\ddot{a}_{n\rceil} - \frac{\Delta q}{i}\left(\ddot{a}_{n\rceil} - n\right); \ \mathrm{FV}_{n\rceil}^{\mathrm{due}} = K\ddot{s}_{n\rceil} - \frac{\Delta q}{i}\left(\ddot{s}_{n\rceil} - nq^n\right);$$

$$\mathrm{PV}_{\infty\rceil}^{\mathrm{due}} = \frac{q}{i}\left(K - \frac{\Delta q}{i} + n\Delta\right)$$

(present and future value of *arithmetically decreasing* annuity-*due* of the type $\ldots$, $K+2\Delta, K+\Delta, K$ with decrements *independent* of the payment K; in fact it means an arithmetic increase for $\Delta < 0$)

$$\mathrm{PV}_{n\rceil}^{\mathrm{imm}} = Ka_{n\rceil} - \frac{\Delta q}{i}\left(a_{n\rceil} - \frac{n}{q}\right); \quad \mathrm{FV}_{n\rceil}^{\mathrm{imm}} = Ks_{n\rceil} - \frac{\Delta q}{i}\left(s_{n\rceil} - nq^{n-1}\right);$$

$$\mathrm{PV}_{\infty\rceil}^{\mathrm{imm}} = \frac{1}{i}\left(K - \frac{\Delta q}{i} + n\Delta\right)$$

(present and future value of *arithmetically decreasing immediate* annuity of the type $\ldots, K+2\Delta, K+\Delta, K$ with decrements *independent* of the payment K)

$$\mathrm{PV}_{n\rceil}^{\mathrm{due}} = v^n \mathrm{FV}_{n\rceil}^{\mathrm{due}}; \quad \mathrm{FV}_{n\rceil}^{\mathrm{due}} = \frac{q}{i}\left(K(q^n - 1) + \Delta\left(\frac{q^n - q^{n-(k-1)f}}{q^f - 1} - k + 1\right)\right)$$

(present and future value of *periodic-arithmetically increasing* annuity-*due* of the following type: the first f payments in amount K, the second f payments in amount $K + \Delta$, the third f payments in amount $K + 2\Delta$, etc.)

$$\mathrm{PV}_{n\rceil}^{\mathrm{imm}} = v^n \mathrm{FV}_{n\rceil}^{\mathrm{imm}}; \quad \mathrm{FV}_{n\rceil}^{\mathrm{imm}} = \frac{1}{i}\left(K(q^n - 1) + \Delta\left(\frac{q^n - q^{n-(k-1)f}}{q^f - 1} - k + 1\right)\right)$$

(present and future value of *periodic-arithmetically increasing immediate* annuity of the following type: the first f payments in amount K, the second f payments in amount $K + \Delta$, the third f payments in amount $K + 2\Delta$, etc.)

$$\mathrm{PV}_{n\rceil}^{\mathrm{due}} = \begin{cases} \dfrac{K}{q^{n-1}} \dfrac{q^n - b^n}{q - b}, & b \neq q \\ Kn, & b = q \end{cases}; \quad \mathrm{FV}_{n\rceil}^{\mathrm{due}} = q^n \mathrm{PV}_{n\rceil}^{\mathrm{due}};$$

$$\mathrm{PV}_{\infty\rceil}^{\mathrm{due}} = K\tfrac{q}{q-b}, \quad b < q$$

(present and future value of *geometrically increasing* annuity-*due* of the type $K, Kb, Kb^2, \ldots, Kb^{n-1}$; in fact it means a geometric decrease for $0 < b < 1$)

$$\mathrm{PV}_{n\rceil}^{\mathrm{imm}} = \begin{cases} \dfrac{K}{q^n} \dfrac{q^n - b^n}{q - b}, & b \neq q \\ K\dfrac{n}{q}, & b = q \end{cases}; \quad \mathrm{FV}_{n\rceil}^{\mathrm{imm}} = q^n \mathrm{PV}_{n\rceil}^{\mathrm{imm}};$$

$$\mathrm{PV}_{\infty\rceil}^{\mathrm{imm}} = K\frac{1}{q-b}, \quad b < q$$

(present and future value of *geometrically increasing immediate* annuity of the type $K, Kb, Kb^2, \ldots, Kb^{n-1}$)

$$\mathrm{FV}_{n\rceil}^{\mathrm{imm}} = \begin{cases} K\dfrac{q^f - 1}{i} kq^{(k-1)f}, & b = q^f; \quad n = kf \\ K\dfrac{q^f - 1}{i} \dfrac{q^{kf} - b^k}{q^f - b}, & b \neq q^f; \quad n = kf \\ \mathrm{FV}_{(k-1)f\rceil} q^{n-(k-1)f} + K\dfrac{1}{i} b^{k-1}(q^{n-(k-1)f} - 1), & (k-1)f < n \leq kf \end{cases}$$

(future value of *periodic-geometrically increasing immediate* annuity of the following type: the first f payments in amount K, the second f payments in amount Kb, the third f payments in amount Kb^2, etc.)

7.3 Annuities Payable *m*thly

- Annuity payable *m*thly: is an annuity that is payable more frequently than once year (usually on monthly basis for $m = 12$, on quarterly basis for $m = 4$ or on semiannual basis for $m = 2$); one still assumes for simplicity that payment periods coincide with interest periods, i.e. the interest is payable also *m*thly per year (see Sect. 3.3)
- *Unit* annuity payable *m*thly: each year one pays *m*thly payments in amount of $1/m$ (as if the unit amount would be paid off in total each year)

Denotation:

i	annual effective interest rate
q	interest factor (see Sect. 3.1): $q = 1 + i$
v	discount factor (see Sect. 3.2): $v = \frac{1}{1+i}$
d	annual effective discount rate: $d = 1 - v = iv$
n	number of years
m	frequency of payments (e.g. monthly for $m = 12$)
$i^{(m)}$	nominal interest rate conformal to i and m (see Sect. 3.3): $i^{(m)} = m\left((1+i)^{1/m} - 1\right)$
$d^{(m)}$	nominal discount rate conformal to i and m (see Sect. 3.3): $d^{(m)} = m\left(1 - (1-d)^{1/m}\right)$

$$\ddot{a}_{\overline{n|}}^{(m)} = \frac{1}{m}\sum_{k=0}^{mn-1} v^{\frac{k}{m}} = \frac{d}{d^{(m)}}\ddot{a}_{\overline{n|}} = \frac{1-v^n}{d^{(m)}} \approx \ddot{a}_{\overline{n|}} + \frac{m-1}{2m}(1-v^n)$$

(present value of unit annuity-due payable *m*thly)

$$a_{\overline{n|}}^{(m)} = \frac{1}{m}\sum_{k=1}^{mn} v^{\frac{k}{m}} = \frac{i}{i^{(m)}}a_{\overline{n|}} = \frac{1-v^n}{i^{(m)}} \approx a_{\overline{n|}} - \frac{m-1}{2m}(1-v^n)$$

(present value of unit immediate annuity payable *m*thly)

$$\ddot{s}_{\overline{n|}}^{(m)} = \frac{1}{m}\sum_{k=1}^{mn} v^{\frac{k}{m}} = \frac{d}{d^{(m)}}\ddot{s}_{\overline{n|}} = \frac{q^n-1}{d^{(m)}} \approx \ddot{s}_{\overline{n|}} + \frac{m-1}{2m}(q^n-1)$$

(future value of unit annuity-due payable *m*thly)

$$s_{\overline{n|}}^{(m)} = \frac{1}{m}\sum_{k=0}^{mn-1} v^{\frac{k}{m}} = \frac{i}{i^{(m)}}s_{\overline{n|}} = \frac{q^n-1}{i^{(m)}} \approx s_{\overline{n|}} - \frac{m-1}{2m}(q^n-1)$$

(future value of unit immediate annuity payable *m*thly)

$$\ddot{a}_{\overline{\infty}|}^{(m)} = \frac{1}{d^{(m)}}$$

(present value of unit perpetuity-due payable $m$thly)

$$a_{\overline{\infty}|}^{(m)} = \frac{1}{i^{(m)}}$$

(present value of unit immediate perpetuity payable $m$thly)

$$(I^{(k)}\ddot{a})_{\overline{n}|}^{(m)} = \frac{1}{d^{(m)}}(\ddot{a}_{\overline{n}|}^{(k)} - nv^n); \quad (I^{(k)}a)_{\overline{n}|}^{(m)} = \frac{1}{i^{(m)}}(\ddot{a}_{\overline{n}|}^{(k)} - nv^n),$$

where m is divisible by k
(present value of unit due and immediate annuity increasing $k$thly by $1/k$ and payable $m$thly each year during n years with the first payment of amount $1/k$)

$$(D^{(k)}\ddot{a})_{\overline{n}|}^{(m)} = \frac{1}{d^{(m)}}(n - a_{\overline{n}|}^{(k)}); \quad (D^{(k)}a)_{\overline{n}|}^{(m)} = \frac{1}{i^{(m)}}(n - a_{\overline{n}|}^{(k)}),$$

where m is divisible by k
(present value of unit due and immediate annuity decreasing $k$thly by $1/k$ and payable $m$thly each year during n years with the first payment of amount n)

$$(I^{(k)}\ddot{a})_{\overline{\infty}|}^{(m)} = \frac{1}{d^{(m)}d^{(k)}}; \quad (I^{(k)}a)_{\overline{\infty}|}^{(m)} = \frac{1}{i^{(m)}d^{(k)}}, \text{ where } m \text{ is divisible by } k$$

(present value of unit due and immediate perpetuity increasing $k$thly by $1/k$ and payable $m$thly each year with the first payment of amount $1/k$)

7.4 Continuously Payable Annuities

- Payments in continuous time (payment streams, also see Sect. 6.1): are determined by theirs annual *instantaneous rate of payment* $c(t)$ that is a function in continuous time such that the amount $c(t)\mathrm{d}t$ is paid off in total in the infinitesimal interval $(t, t + \mathrm{d}t)$

  Denotation:

i	annual interest rate with interest factor $q = 1 + i$ and discount factor
d	annual discount rate: $1 - d = \frac{1}{1+i}$
δ	force of interest (see Chap. 4): $\delta = \ln(1 + i) = \ln q = -\ln v = -\ln(1 - d)$
$\delta(\tau)$	force of interest at time τ (see Sect. 4)

$$\mathrm{PV} = \int_0^t c(\tau)\mathrm{e}^{-\int_0^\tau \delta(s)\,\mathrm{d}s}\mathrm{d}\tau; \quad \mathrm{FV} = \int_0^t c(\tau)\mathrm{e}^{\int_\tau^t \delta(s)\,\mathrm{d}s}\mathrm{d}\tau$$

(present and future value of payments in continuous time $(0, t)$)

$$\bar{a}_{\overline{n}|} = \int_0^n e^{-\delta t} dt = \frac{1 - e^{-\delta n}}{\delta} = \frac{1 - v^n}{\delta};$$

$$\bar{s}_{\overline{n}|} = \int_0^n e^{\delta t} dt = \frac{e^{\delta n} - 1}{\delta} = \frac{q^n - 1}{\delta}$$

(present and future value of annuity with unit rate of payment $c(t) = 1$)

$$\bar{a}_{\overline{\infty}|} = \int_0^\infty e^{-\delta t} dt = \frac{1}{\delta}$$

(present value of perpetuity with unit rate of payment $c(t) = 1$)

$$(\bar{I}\,\bar{a})_{\overline{\infty}|} = \int_0^\infty t e^{-\delta t} dt = \frac{1}{\delta^2}$$

(present value of continuously increasing perpetuity with rate of payment $c(t) = t$)

$$(I\,\bar{a})_{\overline{\infty}|} = \int_0^\infty [t + 1] e^{-\delta t} dt = \frac{1}{\delta d}$$

(present value of perpetuity increasing in steps with rate of payment $c(t) = [t + 1]$, where [ ] denotes the integer part of a number)

7.5 Amortization of Debt

- *Amortization of debt*: is a gradual repayment of interest-bearing debt (credit, loan) by its debtor (borrower) to its creditor (lender) according to an *amortization schedule* (*repayment plan*); each *repayment* (*installment*) in this schedule comprises of two components:
  - *payment on principal*: it reduces gradually the *principal* (unpaid balance) of the debt
  - *payment on interest*: it settles interest from the unpaid balance of the debt due to the given credit interest rate (it is important for tax declarations)
- *Repayments*: can be
  - unequal: e.g. interest, uniform, accelerated or general amortization of the debt
  - equal (the so-called annuity amortization of the debt): e.g. with a prescribed number of repayments and the corresponding amount of repayments,

with a prescribed amount of repayments and the corresponding number of repayments, and the like)

Denotation:

i	credit interest rate over one repayment period with factors $q = 1 + i$ and $v = \frac{1}{1+i}$
d	discount rate: $d = 1 - v = i \cdot v$
S_0	(initial) principal
S_k	principal outstanding at the end of the kth repayment period (after the kth repayment)
K_k	repayment at the end of the kth repayment period: $K_k = T_k + U_k$
T_k	payment on principal in the kth repayment
U_k	payment on interest in the kth repayment
U	total interest: $U = U_1 + U_2 + \ldots + U_n$
PVU	present value of the total interest
n	number of repayment periods (*time of amortization*)

- *Interest amortization* (one repays only interests, and the principal is amortized as late as in the last repayment):

$$K_k = S_0 i, \quad k = 1,\ 2,\ \ldots,\ n-1; \quad K_n = S_0(1+i)$$

- *Uniform amortization* (repayments amortize always the same part of the principal):

$$K_k = \frac{S_0}{n}\,(1 - (n - k + 1)i)$$

$$T_k = \frac{S_0}{n}$$

$$U_k = S_0\left(1 - \frac{k-1}{n}\right)i$$

$$S_k = S_0\left(1 - \frac{k}{n}\right)$$

$$U = \frac{n+1}{2} S_0 i;$$

$$\text{PVU} = \frac{S_0 i}{n}\left((n+1)a_{\overline{n|}} - \left(\frac{1-v^n}{id} - \frac{nv^n}{i}\right)\right)$$

- *Annuity amortization* (repayments of the same amount K result from the prescribed time of amortization n):

$$K = \frac{S_0}{a_{n\rceil}} = S_0 \frac{i}{1 - v^n}$$

$$T_1 = K - S_0 i = K v^n; \quad T_k = T_1 q^{k-1} = K q^{k-n-1} = K v^{n-k+1}$$

$$U_k = K - T_k = K - T_1 q^{k-1} = K i a_{n-k+1\rceil} = K(1 - v^{n-k+1})$$

$$S_k = S_0 q^k - K s_{k\rceil} = S_0 - T_1 s_{k\rceil} = S_0 \frac{1 - v^{n-k}}{1 - v^n} = K(a_{n-k+1\rceil} - v^{n-k+1}) = K a_{n-k\rceil}$$

	Repayment	T_k	U_k	Principal outstanding S_k
1	K	Kv^n	$Kia_{n\rceil} = K(1 - v^n)$	$K(a_{n\rceil} - v^n) = Ka_{n-1\rceil}$
2	K	Kv^{n-1}	$Kia_{n-1\rceil} = K(1 - v^{n-1})$	$K(a_{n-1\rceil} - v^{n-1}) = Ka_{n-2\rceil}$
3	K	Kv^{n-2}	$Kia_{n-2\rceil} = K(1 - v^{n-2})$	$K(a_{n-2\rceil} - v^{n-2}) = Ka_{n-3\rceil}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$n-1$	K	Kv^2	$Kia_{2\rceil} = K(1 - v^2)$	$K(a_{2\rceil} - v^2) = Ka_{1\rceil}$
n	K	Kv	$Kia_{1\rceil} = K(1 - v)$	$K(a_{1\rceil} - v) = 0$
Σ	Kn	S_0	$Kn - S_0$	

- *Annuity amortization* (time of amortization n results from the prescribed repayments of the same amount K):

$$n = \frac{\ln K - \ln(K - S_0 i)}{\ln q}$$

- *General amortization* (using repayments of general amounts $K_1, \ldots, K_{n-1}$ one calculates the amount K_n of the last repayment to amortize the debt):

$$S_k = S_0 q^k - \sum_{j=1}^{k} K_j q^{k-j}, \quad k = 1, 2, \ldots, n-1$$

$$K_n = S_0 q^n - \sum_{j=1}^{n-1} K_j q^{n-j}$$

Further Reading

Bosch, K.: Finanzmathematik für Banker. Oldenbourg Verlag, München (2001)
Cissell, R., Cissell, H., Flaspohler, D.C.: Mathematics of Finance. Houghton Mifflin, Boston (1982)
Gerber, H.U.: Life Insurance Mathematics. Springer, Berlin (1990)
Grundmann, W.: Finanz- und Versicherungsmathematik. Teubner, Leipzig (1996)

Grundmann, W., Luderer, B.: Formelsammlung: Finanzmathematik, Versicherungsmathematik, Wertpapieranalyse. Teubner, Stuttgart (2001)

Ihrig, H., Pfaumer, P.: Finanzmathematik. Intensivkurs. Oldenbourg Verlag, München (1999)

Khoury, S.J., Parsons, T.D.: Mathematical Methods in Finance and Economics. North Holland, New York (1981)

Knox, D.M., Zima, P., Brown, R.L.: Mathematics of Finance. McGraw-Hill, Sydney, NSW (1984)

Luderer, B., Nollau, V., Vetters, K.: Mathematische Formeln für Wirtschaftswissenschaftler. Teubner, Stuttgart (2000)

McCutcheon, J.J., Scott, W.F.: An Introduction to the Mathematics of Finance. Heinemann, London (1986)

Thomsett, M.C.: The Mathematics of Investing. Wiley, New York (1989)

Chapter 8
Depreciation

Abstract Chapter 8 is an overview of basic accounting depreciation methods.

- *Depreciation*: is the reduction in the value of an asset due to usage, passage of time, wear and tear, technological outdating or obsolescence or other such factors; in accounting, this term describes any method of allocation of purchase cost of a depreciable asset across time when the asset is employed to generate revenues; the depreciation affects the financial statements and in most countries also taxes
- *Methods of depreciation*: are various schedules how to distribute the depreciations across the useful life of the asset; the key factors in the corresponding formulas are:

  – *original cost* of the given depreciable asset
  – *salvage value* (*scrap value*) of the asset: this value is not further depreciated
  – *depreciation time period* (*useful life*) of the asset
  – *depreciation strategy*: the annual *depreciation expenses* can be uniform (*linear* depreciation methods) or gradually decreasing (*declining* or *degressive* depreciation methods, *accelerated depreciation*) or gradually increasing (*progressive* depreciation methods); e.g. when using the accelerated depreciation, one can speed up the investment payback due to lower tax payments in the first years of an investment

  Denotation:

P_0	original cost of the given asset
P_k	*depreciated cost* (*book value*) after k years
n	depreciation time measured in years
o_k	depreciation charges of the kth year
O	accumulated depreciation: $O = P_0 - P_n = o_1 + o_2 + \ldots + o_n$
s	depreciation rate (with respect to P_0)
s_k	depreciation rate (with respect to P_k)

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_8,

- *Straight line depreciation method* (the annual depreciation charges are equal across the useful asset's life):

$$o_k = o = \frac{P_0 - P_n}{n}$$
$$P_k = P_0 - ko$$
$$s = \frac{o}{P_0}$$
$$s_k = \frac{o}{P_k} = \frac{P_0 - P_n}{nP_0 - k(P_0 - P_n)}$$

- *Arithmetical declining balance depreciation method* (the annual depreciation charges decrease arithmetically):

$$\text{condition for}\, o_1: \frac{P_0 - P_n}{n} \le o_1 \le \frac{2(P_0 - P_n)}{n} \quad (\text{then}\, d \ge 0,\ o_n \ge 0)$$
$$o_k = o_{k-1} - d = o_1 - (k-1)d, \ \text{ where}\, d = \frac{2}{n-1}\left(o_1 - \frac{P_0 - P_n}{n}\right)$$
$$P_k = P_0 - ko_1 + \frac{(k-1)k}{(n-1)n}\left(no_1 - (P_0 - P_n)\right)$$

- *Sum of years digit depreciation method*: is a special case of the previous method for $o_n = d$ (i.e. $o_{n+1} = 0$):

$$o_k = (n - k + 1) \cdot d, \ \text{where}\, d = \frac{(P_0 - P_n)}{n(n+1)/2}$$
$$o_1 = \frac{P_0 - P_n}{(n+1)/2} = nd$$
$$P_k = P_0 - \frac{k(P_0 - P_n)}{n(n+1)}(2n + 1 - k)$$

- *Geometrical declining balance depreciation method* (the annual depreciation charges decrease geometrically):

$$o_k = o_{k-1}s = P_0 s(1-s)^{k-1}, \ \text{where}\, s = 1 - \sqrt[n]{\frac{P_n}{P_0}}$$
$$P_k = P_0(1-s)^k$$

- *Arithmetical progressive balance depreciation method* (the annual depreciation charges increase arithmetically):
$$o_k = o_{k-1} + d = o_1 + (k-1)d, \quad \text{where}\, d = \frac{2}{n-1}\left(\frac{P_0 - P_n}{n} - o_1\right)$$

$$o_1 = \frac{P_0 - P_n}{n} - \frac{n-1}{2}d$$

Further Reading

Cissell, R., Cissell, H., Flaspohler, D.C.: Mathematics of Finance. Houghton Mifflin, Boston, MA (1982)
Grundmann, W., Luderer, B.: Formelsammlung: Finanzmathematik, Versicherungsmathematik, Wertpapieranalyse. Teubner, Stuttgart (2001)
IASB Standards (http://www.iasb.org)

Chapter 9
Financial Instruments

Abstract Chapter 9 provides important formulas for basic financial instruments and securities: 9.1. Discount Securities, 9.2. Bonds, 9.3. Stocks, 9.4. Currencies.

9.1 Discount Securities

- *Discount securities*: are quoted (priced) using the discount principle, i.e. their price is set down by subtracting the discount from the *face value* (amount due) using the corresponding discount rate (see Sect. 2.4) and a suitable calendar convention (frequently, for simplicity, calendar Euro-30/360, see Sect. 2.2); discount securities are mostly *money market instruments* (i.e. *short-term* ones meaning 1 year or less) and also *fixed-income instruments* (i.e. promising the investor to receive a specified cash flow at a specified time in the future); examples are as follows:
- *Certificates of deposit* (CD): are securities representing savings deposits at commercial banks or savings institutions
- *Eurodollars*: are CD's denominated in US dollars at banks outside the United States (mainly in Europe)
- *Bills of exchange* (*drafts*): are unconditional orders issued by a first party (*drawer*) directing a second party (*drawee*, mainly a bank) to make a certain payment to a third party (*payee*) at a future date (sometimes they may be not "orders" but "promises"); a common type of bills of exchange is the *cheque*
- *Commercial papers*: are unsecured promissory notes issued by large banks and corporations collecting money to meet their short-term debt obligations, and backed only by an issuing bank or corporation's promise to pay the face amount on the maturity date specified on the note
- *Banker's acceptances*: are drafts used in foreign trade; upon acceptance, which occurs when an authorized bank accepts them, the drafts become an unconditional liability of the bank

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_9,

- *Treasury bills* (*T-bills*): are government securities issued by central banks (by the Fed in the United States) to finance the government in short-terms, i.e. to cover the short-term state or communal deficit; their maturities vary from weeks to 1 year

  Denotation:

P quoted price
F face value
d annual discount rate for pricing the discount security
k *maturity* (i.e. the length of time to maturity measured in days mostly according to the calendar Euro-30/360, see Sect. 2.2)
d_1 annual discount rate for pricing the discount security at the moment of its purchase
k_1 maturity at the moment of purchase
d_2 annual discount rate for pricing the discount security at the moment of its sale
k_2 maturity at the moment of sale
i annual *yield to maturity* (i.e. the interest rate corresponding to the simple discount, see Sect. 2.4)

$$D = Fd\frac{k}{360}$$

(*discount* for pricing the discount security)

$$P = F - D = F\left(1 - d\frac{k}{360}\right)$$

(*quoted price*)

$$d = \frac{F - P}{F}\frac{360}{k}$$

(*discount rate* corresponding to the quoted price P)

$$i = \frac{d}{1 - d\frac{k}{360}}$$

(annual *yield to maturity*)

$$i = \left(\frac{1 - d_2\frac{k_2}{360}}{1 - d_1\frac{k_1}{360}} - 1\right)\frac{360}{k_1 - k_2}$$

(annual yield to maturity at the moment of sale prior to the maturity date)

9.2 Bonds

- *Bonds*: are long-term debt securities in which the authorized *issuer* (central or local government, bank, corporation in the position of borrower) promises to the *bond holder* (in the position of lender) to pay periodically the interest (*coupon*) and to repay at the *maturity date* the *face value* (*par value*, *nominal value*, *principal*); one sometimes uses the following classification of fixed-income securities according to their maturities:

  – *bills* with short-term maturities up to 1 year (see Sect. 9.1)
  – *notes* with medium-term maturities between 1 and 5 years
  – *bonds* with long-term maturities greater than 5 years

- *Coupons*: are paid regularly at the ends of (annual or semiannual) *coupon periods* up to and including the maturity date (the exception are the consol bonds with no maturity); the coupon expressed relatively to the face value is called *coupon rate*; when a new bond holder obtained the bond after the *date ex-coupon*, then the first forthcoming coupon belongs to the previous holder
- Types of bonds:

  – *zero-coupon bonds*: pay only the face value at maturity (no regular interests) but they are issued at a substantial discount to the face value; such "long-term CD's" are common in some countries mainly due to tax reasons
  – *coupon bonds*: pay periodically interests in the form of coupons (see *thereinbefore*)
  – *consol bonds* (*perpetual bonds*): have no maturity date so that their face value is never paid up
  – *fixed rate bonds*: have the coupon rate that remains constant till the maturity date
  – *indexed bonds* (*inflation linked bonds*): have the coupons and sometimes also the face value indexed to inflation or to another business indicator (they are common in pension fund portfolios, e.g. inflation linked UK Gilts)
  – *floating rate notes* (FRN's, *floaters*): have a variable coupon rate that is linked to a reference interest rate (e.g. to LIBOR, see Sect. 2.1); the coupon rate is recalculated periodically (typically every 1 or 3 months) and charged by a margin (spread) that reflects the credit risk of the issuer and the liquidity of the market
  – *callable bonds*: give the issuer the right of option to pay up the face value (usually plus a call premium) prior to the original maturity date ("bonds with embedded call option"); *putable bonds* enable a similar option for the bondholders
  – *convertible bonds*: give the bondholders the right to exchange the bond for another security (typically for the common stock issued by the same company)
  – *bonds with warrant*: unlike the convertible bonds, the warrant can be separated from such a bond and offered as an individual call option (on common stock, see Sect. 10.5)

- *asset-backed securities*: are bonds whose interest and principal payments are backed by underlying cash flows from other assets; examples are *mortgage-backed securities* (MBS's), *collateralized mortgage obligations* (CMO's) and *collateralized debt obligations* (CDO's) that enable banks to finance mortgages
- *subordinate bonds*: have a lower priority than other bonds of the same issuer in the case of default
- *junk bonds*: are high-yield bonds of speculative (non-investment) grade with a higher risk of default
- *structured bonds*: have payments dependent on various market factors (they are linked e.g. to development of FX courses at FOREX, segments of stock market and the like)
- *state bonds* (*Treasury notes*, *Treasury bonds*): are government securities issued by central banks (by the Fed in the United States) to finance the government in long-terms, i.e. to cover the long-term state deficit (see also *T-bills* in Sect. 9.1)
- *municipal bonds*: are issued by local governments, regions, cities and the like
- *corporate bonds*: are issued by corporations (firms or banks)

• Bond pricing:

- *market price*: is given by the current supply and demand on the capital market; if a bond is not regularly quoted (marked-to-market), then it can be priced by its present value using a suitable reference interest rate (as benchmarks, one uses e.g. YTM's of government securities that must be for the priced bond charged by various spreads)
- *present value* (*fair value*): is calculated as the present value of future cash flows corresponding to the bond (i.e. the future coupons and the face value) by means of the valuation interest rate for discounting (see Sect. 6.1)
- *accrued interest* AI: is the part of the next forthcoming coupon to which the seller of the bond is entitled if the bond is sold prior this next coupon payment
- *gross price* (*dirty price*, *full price*, *settlement price*): is equal to the market price if the accrued interest is explicitly involved; in practice, bonds are usually traded in the gross prices
- *net price* (*pure price*): is the gross price without the accrued interest; in practice, bonds are usually quoted in the net prices

Denotation:

F face value
c annual coupon rate
C annual coupon: $C = F \cdot c$
n number of annual coupons remaining to maturity date
τ time period (a part of year) since the last preceding coupon (with respect to the current date of pricing): $0 \leq \tau < 1$
m annual frequency of coupon payments (mainly semiannually for $m = 2$)

PV	present value
i	annual valuation interest rate (see Sect. 6.1)
P	market price of bond
y	yield to maturity (YTM, see Sect. 6.2)
$P(y)$	price P of bond considered as a function of yield to maturity y
AI	accrued interest
P_{gross}	gross price
P_{net}	net price

$$\mathrm{PV} = \frac{C}{1+i_1} + \cdots + \frac{C+F}{(1+i_n)^n} = \sum_{t=1}^{n} \frac{C}{(1+i_t)^t} + \frac{F}{(1+i_n)^n}$$

(*present value of bond* at the end of an annual coupon period, where i_t denotes the annual valuation spot interest rate for t years (see Sect. 5.2))

$$y_c = \frac{C}{P}$$

(*current yield*)

$$P = \frac{C}{1+y} + \cdots + \frac{C}{(1+y)^{n-1}} + \frac{C+F}{(1+y)^n} = \frac{F}{y}\left(c + \frac{y-c}{(1+y)^n}\right) = \frac{C}{y} + \frac{1}{(1+y)^n}\left(F - \frac{C}{y}\right)$$

(equation for calculation of *yield to maturity* y at the end of an annual coupon period; yield to maturity is the internal rate of return *IRR* (see Sect. 6.2) of the cash flow system corresponding to the given bond)

$$y \approx \frac{C + (F-P)/n}{P}$$

(approximate calculation of yield to maturity y (*commercial method*))

$$y \approx \frac{C + (F-P)/n}{0.6P + 0.4F}$$

(approximate calculation of yield to maturity y (*Hawawini-Vora*))

$$\begin{aligned} P(y) &= \frac{C}{1+y} + \cdots + \frac{C}{(1+y)^{n-1}} + \frac{C+F}{(1+y)^n} = C\frac{1-(1+y)^{-n}}{y} + \frac{F}{(1+y)^n} \\ &= \frac{F}{y}\left(c + \frac{y-c}{(1+y)^n}\right) = \frac{C}{y} + \frac{1}{(1+y)^n}\left(F - \frac{C}{y}\right) \end{aligned}$$

(price of bond at the end of an annual coupon period considered as a function of yield to maturity y (on the contrary, if P is given, then one obtains the equation for calculation of y, see *thereinbefore*))

- Properties of the price $P = P(y)$:
  - there is exactly one value y for given P
  - P decreases at a decreasing rate with increasing yield to maturity y (the convex relationship between price and yield)
  - the absolute change of the price P due to a change of y is larger (and the relative change of the price P due to a change of y is smaller), if the coupon rate c is larger
  - if $y = c$, or $y > c$, or $y < c$, then with the decreasing number n of remaining coupons (i.e. with the decreasing maturity) the price P remains equal to F ($P = F$ is called "*sell at par*"), or P increases to F ($P < F$ is called "*sell at discount*"), or P decreases to F ($P > F$ is called "*sell at premium*"), respectively; in addition, such an increase of P, or such a decrease of P, accelerates with decreasing maturity n, respectively
  - empirical rules: (1) "yields of bonds decrease (or increase) with increasing (or decreasing) average yields on the capital market, respectively"; (2) "yields of bonds decrease (or increase) with increasing (or decreasing) inflation, respectively"

$$\begin{cases} P = F, \text{ if and only if } y = c \\ P = F, \text{ if and only if } y > c \\ P = F, \text{ if and only if } y < c \end{cases}$$

$$P(y) = \frac{C/m}{(1+y)^{1/m}} + \frac{C/m}{(1+y)^{2/m}} + \cdots + \frac{C/m + F}{(1+y)^n} = \frac{C}{m}\frac{1-(1+y)^{-n}}{(1+y)^{1/m}-1} + \frac{F}{(1+y)^n}$$

(price of bond at the end of an annual coupon, if the annual coupon C is divided to m regular installments during each year (mainly semiannually for $m = 2$))

$$\mathrm{AI} = \frac{C}{y}\left((1+y)^{\tau} - 1\right)$$

(theoretical calculation of *accrued interest*, when τ is the time period (a part of year) since the last preceding coupon ($0 \le \tau < 1$))

$$\mathrm{AI} \approx \frac{k}{360}C$$

(practical calculation (calendar convention 30E/360, see Sect. 2.2) of *accrued interest*, when k is the time period (measured in days) since the last preceding coupon ($0 \le k < 360$))

$$P_{\text{gross}} = \frac{C}{(1+y)^{1-\tau}} + \cdots + \frac{C}{(1+y)^{n-1-\tau}} + \frac{C+F}{(1+y)^{n-\tau}} = \frac{C}{y}(1+y)^{\tau} + \frac{1}{(1+y)^{n-\tau}}\left(F - \frac{C}{y}\right)$$

(*gross price* of bond, when τ is the time period (a part of year) since the last preceding coupon ($0 \le \tau < 1$) and n is the remaining number of annual coupons)

$$P_{\text{gross}} \approx P_0 + \frac{k}{360}(P_1 - P_0 + C)$$

(approximate calculation of *gross price* by means of interpolation between the price P_0 at the date of the last preceding coupon and the price P_1 at the date of the next forthcoming coupon, when k is the time period (measured in days) since the last preceding coupon ($0 \leq k < 360$))

$$P_{net} = P_{\text{gross}} - AI = \frac{C}{y} + \frac{1}{(1+y)^{n-\tau}}\left(F - \frac{C}{y}\right) \approx P_{\text{gross}} - \frac{k}{360}C \approx P_0 + \frac{k}{360}(P_1 - P_0)$$

(*net price* of bond; in practice, the quoted net prices are usually obtained using YTM's of government securities as benchmarks that must be charged for the priced bonds by various spreads reflecting credibility, liquidity and market psychology)

$$P_0 = \frac{C}{1+y} + \cdots + \frac{C}{(1+y)^{n-1}} + \frac{C+P_n}{(1+y)^n}$$

(equation for calculation of *holding-period return* (*realized return*, *yield to call*) y across n annual coupon periods, where P_0 and P_n are the realized prices at the beginning and at the end of the holding period)

- *Yield curve*: plots interest rates paid on interest bearing securities against the time to maturity (for bonds, it is simply a graph of YTM for various maturities); such a plot makes sense only for classes of comparable securities (e.g. for Treasury bonds with rating AA) and for the current date that must be fixed for all plotted maturities (yield curves constructed today and next month may be different); one distinguishes:

  – *yield curve of zero-coupon bonds*: plots annual yields to maturity y_n of zero-coupon bonds against their maturities n (i.e. $y_n = (F/P)^{1/n} - 1$); in practice, it is usually impossible to collect the corresponding data so that one uses the *yield curve of swap interest rates* (see Sect. 10.4) charged by suitable spreads for this purpose
  – *yield curve of coupon bonds*: plots annual yields to maturity y_n of coupon bonds against their maturities n
  – *forward yield curve*: plots annual forward yields to maturity against maturities n (see forward interest rate in Sect. 5.2); one can construct e.g. the curve of annual forward yields to maturity $y_{1,\,n}$ valid in 1 year for zero-coupon bonds with maturity n

$$(1+y_1)(1+y_{1,\,n})^n = (1+y_{n+1})^{n+1}$$

(construction of the forward yield curve $y_{1,\,n}$ by means of the spot yield curve y_n for zero-coupon bonds)

$$P_n = \frac{C_n}{1+y_1} + \frac{C_n}{(1+y_2)^2} + \cdots + \frac{C_n}{(1+y_{n-1})^{n-1}} + \frac{C_n + F_n}{(1+y_n)^n}, \quad n = 1,\ 2,\ \ldots$$

(*bootstrapping*: is a recursive method for construction of the yield curve y_n of zero-coupon bonds (not available in practice) by means of market data on coupon bonds (available in practice): y_n is constructed by means of previous $y_1, \ldots, y_{n-1}$ and observed values $\{P_n, C_n, F_n\}$ for the coupon bond with maturity n)

$$y(t) = \beta_0 + (\beta_1 + \beta_2)\frac{1 - \exp(-t/\tau)}{t/\tau} - \beta_2 \exp(-t/\tau) \ (\beta_0,\ \beta_1,\ \beta_2,\ \tau \text{ are parameters})$$

(*Nelson-Siegel curve* for modeling yield curves $y(t)$ of zero-coupon bonds against maturities t by means of the nonlinear regression methods, see Sect. 27.11)

- *Shape of yield curve*: is explained by various theories; the typical shape is upward-sloping (i.e. increasing and concave), but it can be also downward-sloping (i.e. decreasing and convex), humped (i.e. increasing first and decreasing for longer maturities), flat, U-shaped, and the like
- The *increasing* yield curves have the following ordering: coupon bonds < zero-coupon bonds < forward bonds
- The *decreasing* yield curves have the following ordering: forward bonds < zero-coupon bonds < coupon bonds

- *Duration*: is a measure of length of financial instruments (in particular, the bonds) that enables (similarly as for general cash flow systems in Sect. 6.4) to investigate
  - a sensitivity of bond prices to yields
  - a length of bond lives (e.g. for constructing dedicated or immunized bond portfolios, for matching bond assets to corresponding liabilities, and the like)

Denotation:

y	yield to maturity of bond
$P = P(y)$	price P of bond at the end of an annual coupon period considered as a function of its yield to maturity y
D	duration of bond
CX	convexity of bond

$$\begin{aligned} D &= \frac{\sum_{t=1}^{n-1} tC(1+y)^{-t} + n(C+F)(1+y)^{-n}}{\sum_{t=1}^{n} C(1+y)^{-t} + F(1+y)^{-n}} \\ &= \frac{\sum_{t=1}^{n-1} tC(1+y)^{-t} + n(C+F)(1+y)^{-n}}{P} \\ &= -\frac{1+y}{P(y)}\frac{\mathrm{d}P(y)}{\mathrm{d}y} = \frac{1+y}{y} - \frac{n(c-y)+1+y}{c(1+y)^n - (c-y)} \end{aligned}$$

((*Macaulay*) *duration of coupon bond*: is the expected bond life calculated as the weighted average of times prior to particular coupons, using the relative present values of particular payments (i.e. the discounted coupons and the discounted last coupon plus face value, all related to the bond price) as weights; the 1-year bond has $D = 1$; the zero-coupon bond has $D = n$; the consol bond has $D = (1 + y)/y$)

- Properties of the duration D:
  - D decreases with increasing coupon rate c
  - D decreases at a decreasing rate with increasing yield to maturity y
  - D increases at a decreasing rate with increasing maturity n (only for bonds with long maturities and deep discounts, the duration can turn to decline after some fairly long maturities)

$$\Delta P(y)/P(y) \approx -D\Delta y/(1+y), \text{resp. } \Delta P(y) \approx -P(y)D\frac{\Delta y}{1+y}$$

(approximation of the relative change of price $\Delta P(y)/P(y)$ by means of the duration and the corresponding relative change of interest factor $\Delta(1 + y)/(1 + y) = \Delta y/(1 + y)$)

$$\mathrm{DD} = \frac{1}{1+y}\left(\sum_{t=1}^{n-1} tC(1+y)^{-t} + n(C+F)(1+y)^{-n}\right) = -\frac{\mathrm{d}P(y)}{\mathrm{d}y} = \frac{P(y)D}{1+y}$$

(*dollar duration of coupon bond*)

$$\Delta P(y) \approx -\mathrm{DD}\Delta y$$

(approximation of the change of price $\Delta P(y)$ by means of the dollar duration and the corresponding change of yield to maturity Δy)

$$\begin{aligned} \mathrm{MD} &= \frac{1}{1+y}\frac{\sum_{t=1}^{n-1} tC(1+y)^{-t} + n(C+F)(1+y)^{-n}}{P(y)} \\ &= -\frac{1}{P(y)}\frac{\mathrm{d}P(y)}{\mathrm{d}y} = \frac{D}{1+y} = \frac{DD}{P(y)} \end{aligned}$$

(*modified duration of coupon bond*)

$$\Delta P(y)/P(y) \approx -\mathrm{MD}\Delta y, \text{resp.} \Delta P(y) \approx -P(y)\mathrm{MD}\Delta y$$

(approximation of the change of price $\Delta P(y)$ by means of the modified duration and the corresponding change of yield to maturity Δy)

$$\mathrm{CX} = \frac{\sum_{t=1}^{n-1} t(t+1)C(1+y)^{-t} + n(n+1)(C+F)(1+y)^{-n}}{P} = \frac{(1+y)^2}{P(y)}\frac{\mathrm{d}^2 P(y)}{\mathrm{d}y^2}$$

((*Macaulay*) *convexity of coupon bond* (and its modifications, see Sect. 6.5): improves the approximations based only on durations)

$$\Delta P(y) \approx -P(y)D\frac{\Delta y}{1+y} + \frac{1}{2}P(y)\mathrm{CX}\frac{(\Delta y)^2}{(1+y)^2}$$

(approximation of the change of price $\Delta P(y)$ by means of duration and convexity)

$$D = \frac{P_1 D_1 + \cdots + P_N D_N}{P_1 + \cdots + P_N}$$

(*duration of portfolio* constructed from bonds of N types: P_k is the price of bonds of the kth type in portfolio and D_k is their duration ($k = 1, \ldots, N$))

$$P^A(y_0) = P^P(y_0),$$

where $P^A(y_0)$ is the price of assets and $P^P(y_0)$ is the price of liabilities with internal rate of return y_0

(*matching assets and liabilities*: an investor balances assets and liabilities so that their present values calculated with the required internal rate of return y_0 are equal; in the bond context, the asset side can be just an investment bond portfolio with the price $P^A(y_0)$ and the yield to maturity y_0)

$$\begin{cases} P^A(y_0) = P^P(y_0) \\ D^A(y_0) = D^P(y_0) \\ \mathrm{CX}^A(y_0) > \mathrm{CX}^P(y_0) \end{cases},$$

where $P^A(y_0)$, $D^A(y_0)$ and $\mathrm{CX}^A(y_0)$ is the price, duration and convexity of assets and $P^B(y_0)$, $D^B(y_0)$ and $\mathrm{CX}^B(y_0)$ is the price, duration and convexity of liabilities with internal rate of return y_0

(*matching assets and liabilities* including *immunization*: an investor immunizes the balanced assets and liabilities to become immune against changes of the yield curve in a vicinity of the required internal rate of return y_0; the last inequality guarantees (under suitable assumptions) that then $P^A(y) > P^P(y)$ holds in this vicinity; various problems of construction of asset portfolios for given liability portfolios can be solved under these conditions (in addition, if the minimal acquisition costs of the investment bond portfolio is required, then the solution is called the *dedicated* bond portfolio)

- *Floater* (see *thereinbefore*): is a coupon bond with a variable coupon rate that is linked to a reference interest rate (e.g. to LIBOR or swap interest rates (when LIBOR is not quoted across the given period)) plus a margin (involving credit and liquidity spreads); this reference interest rate plus margin may be applied also for discounting cash flows corresponding to the floater

  Denotation:

F	face value of floater
i	annual reference interest rate plus margin: it serves (1) as the coupon rate for the current annual coupon period (i.e. the next forthcoming annual coupon is $C = F \cdot i$), and (2) as the interest rate for discounting cash flows
τ	time period (a part of year) since the last preceding coupon (with respect to the current date of pricing): $0 \le \tau < 1$
PV	price of floater (with respect to the current date of pricing)

$$\mathrm{PV} = \frac{(1+i)F}{1 + i(1-\tau)}$$

(*price of floater*: for pricing a floater, its maturity n is obviously irrelevant; when pricing a floater immediately after a coupon payment (i.e. $\tau = 0$), then $\mathrm{PV} = F$ (i.e. the floater is priced by its face value); in practice, the discounting rate for a floater can differ from its coupon rate (mainly due to different margins) so that the previous formula is only approximate)

9.3 Stocks

- *Stocks* (*shares*): are securities that represent an ownership position in a company (the property of stockholders is called *equity*); the stocks give their holders the right (1) to engage in the decisions concerning the company (*voting right*), (2) to participate in company's profit distributed to stockholders (*right to dividends*), (3) to participate in the residual value in the case of company's liquidation, (4) to make use of the *rights issue* (*privileged subscription*) which is an offer to existing stockholders to subscribe cash for new stocks in proportion to their existing holdings (see *thereinafter*); *stock companies* (*corporations*) collect equities by underwriting stocks which is a well proven way how to diversify enterprise risks over more subjects
- *Dividends*: are payments made to stockholders based on company's profit (stocks are so called *dividend securities* unlike the *fixed-income securities* represented mainly by bonds); the amount distributed to stockholders as dividends equals company's earnings less its *retained earnings* (the part of earnings intended for reserves and reinvestment); if a new stockholder buys the stock after the *ex-dividend date* (shortly before the *dividend date*), the next forthcoming dividend remains with the previous stockholder

- *Stock split*: increases the number of stocks in issue without any change in company's assets representing the equity, thereby reducing the stock price (e.g. in a 5 for 1 stock split, the company replaces each existing stock by five new ones); the stock split may be argued as the prospect of increased marketability due to increased divisibility of stock holdings, and the like

- Types of stocks:
  - *common stock* (US), *ordinary share* (UK): see *thereinbefore*
  - *preferred stock*: is a hybrid form of security that has characteristics of both common stocks and bonds (preferred stockholders receive their dividends prior to common stockholders, but they do not have voting right; moreover, the dividends are usually fixed similarly to bond coupons)

- Stock pricing:
  - *market price* (*market value*) P: is given by the current supply and demand on the capital market
  - *present value* (*fair value*) PV: is calculated as the present value of future cash flows corresponding to the stock (see Sect. 6.1); in the framework of the fundamental analysis (see *thereinafter*), it can be taken as a benchmark for the market price of the given stock

- Stock analysis:
  - *fundamental analysis*: consists in studying company's financial statements (the balance sheet and the income (i.e. profit and loss) statement) and related trends (in *efficient* capital markets, "any new information is immediately and fully reflected in prices")
  - *technical analysis*: is based on price movements and pattern trends
  - *psychological analysis*: studies behaviour of people dealing with stocks

- Ratios in fundamental analysis of stocks:

  (1) *earnings ratios*:

$$\text{ROA} = \text{earnings after taxes (net income)}/\text{total assets}$$

(*return on* (*total*) *assets*)

$$\text{ROE} = \text{earnings after taxes (net profit)}/\text{equity}$$

(*return on equity*)

  (2) *leverage ratios* (for long-term debts):

$$\text{debt}/(\text{equity} + \text{debt})$$

(*debt ratio*)

$$\text{debt/equity}$$

(*debt-equity ratio*)

(3) *investment ratios*:

$$\text{DPS} = \text{total dividends/number of stocks}$$

(*dividend per stock*)

$$\text{EPS} = \text{earnings after taxes (net income)/number of stocks}$$

(*earnings per stock*)

$$\text{DPS/EPS}$$

(*payout ratio*)

$$\text{EPS/DPS} = 1/\text{payout ratio}$$

(*dividend cover*)

$$(\text{EPS} - \text{DPS})/\text{EPS} = 1 - \text{payout ratio}$$

(*retention ratio*)

$$\text{DPS}/P$$

(*dividend yield*)

$$\text{earnings per stock}/P$$

(*earnings yield*)

$$\text{P/E} = P/\text{EPS} = 1/\text{earnings yield}$$

(P/E *ratio*, *price/earning ratio*: means (1) "which is the price of unit net income of the stock"; (2) "which is the payback period of the stock (measured in years)")

- Discount dividend model in fundamental analysis of stocks:

Denotation:

PV	present value of stock
P_t	expected market price of stock at the end of tth year (in particular, P_0 is the current market price of stock at time $t = 0$)
D_t	expected dividend paid out per stock at the end of tth year
E_t	expected earnings reported per stock at the end of tth year (in particular, E_0 are the last earnings per stock reported prior time $t = 0$)
d_t	payout ratio (i.e. $d_t = D_t / E_t$)
b_t	measure of retained earnings (i.e. $b_t = 1 - d_t$)
P/E	P/E ratio (i.e. P/E $= P_0 / E_0$ at time $t = 0$)

i annual valuation (reference) interest rate (see Sect. 6.1) for discounting cash flows corresponding to the stock (in practice, it can be e.g. the required by stockholders); mostly $i = i_f + i_r$, where i_f is the risk-free interest rate and i_r is the corresponding risk margin (see Sect. 5.1)

i_t annual valuation spot interest rate for t years (see Sect. 5.2)

$$\mathrm{PV} = D_1(1+i)^{-1} + D_2(1+i)^{-2} \cdots = \sum_{t=1}^{\infty} D_t(1+i)^{-t}$$

(general *discount dividend model*)

$$\mathrm{PV} = D_1(1+i_1)^{-1} + D_2(1+i_2)^{-2} \cdots = \sum_{t=1}^{\infty} D_t(1+i_t)^{-t},$$

(general *discount dividend model* with valuation spot interest rates)

$$\mathrm{PV} = D(1+i)^{-1} + D(1+i)^{-2} \cdots = \frac{D}{i}$$

(*zero-growth discount dividend model*; it is used sometimes to approximate calculation of the *P/E* ratio by means of the payout ratio: $P/E \approx d/i$)

$$\mathrm{PV} = D_1(1+i)^{-1} + D_1(1+g)(1+i)^{-2} + D_1(1+g)^2(1+i)^{-3} \ldots = \frac{D_1}{i-g},$$

where g is the annual growth rate ($i > g$) (*multiple-growth discount dividend model* (*Gordon's model*); in practice, one can estimate the annual growth rate as $g \approx b \cdot \mathrm{ROE}$)

$$\begin{aligned}\mathrm{PV} &= \sum_{t=1}^{T} D_1(1+g_1)^{t-1}(1+i)^{-t} + \sum_{t=T+1}^{\infty} D_1(1+g_1)^{T-1}(1+g_2)^{t-T}(1+i)^{-t} \\ &= \sum_{t=1}^{T} D_1(1+g_1)^{t-1}(1+i)^{-t} + \frac{D_1(1+g_1)^{T-1}(1+g_2)}{(1+i)^T(i-g_2)},\end{aligned}$$

where g_1 is the annual growth rate during the initial period of length T and g_2 is the annual growth rate during the following infinite period ($i > g_2$)

(*two-stage multiple-growth discount dividend model*; it can be generalized to more stages of growth)

- Earnings model in fundamental analysis of stocks:

$$\mathrm{PV} = (P/E)_{\mathrm{norm}} E_1,$$

where $(\mathrm{P/E})_{\mathrm{norm}}$ is an estimate of average P/E ratio of the given stock (*earnings model*: $(\mathrm{P/E})_{\mathrm{norm}}$ can be estimated by various methods (see *thereinafter*))

$$(\mathrm{P/E})_{\mathrm{norm}} = \frac{d_1}{i-g}$$

(estimate of $(\mathrm{P/E})_{\mathrm{norm}}$ (for an earnings model, see *thereinbefore*) based on the multiple-growth discount dividend model with growth rate g)

$$(\mathrm{P/E})_{\mathrm{norm}} = b_0 + b_1 g + b_2 d + b_3 \sigma,$$

where g is the annual growth rate, d is the payout ratio and σ is the risk measured as the volatility of the stock price returns, see Sect. 12.2)

(regression estimate of $(\mathrm{P/E})_{\mathrm{norm}}$ (see Sect. 27.11) (for an earnings model, see *thereinbefore*) based on the linear regression model
$(\mathrm{P/E})_{\mathrm{norm}} = \beta_0 + \beta_1 g + \beta_2 d + \beta_3 \sigma + \varepsilon$)

$$(\mathrm{P/E})_{\mathrm{norm}} = (\mathrm{P/E})_M,$$

where $(\mathrm{P/E})_M$ is an aggregate P/E ratio over all stocks of the given branch of economy

(comparative estimate of $(\mathrm{P/E})_{\mathrm{norm}}$ (for an earnings model, see *thereinbefore*))

- Indicators in technical analysis of stocks:

Denotation:

P_t expected market price of stock at the end of tth
$\hat{P}_t$ value P_t smoothed by a smoothing time series method (see Sect. 31.2)

(1) *moving averages and exponential smoothing* (see Sect. 31.2): e.g.

$$\hat{P}_t = \alpha P_t + (1-\alpha)\hat{P}_{t-1},$$

where α ($0 \le \alpha < 1$) is the smoothing constant (*simple exponential smoothing*) (recommended instruments of technical analysis of stock prices are based on exponential smoothing of time series (see Sect. 31.2); using suitable choices of α, one can achieve a higher degree of smoothing (the so-called *long averages*: e.g. for $\alpha = 0.1$), or a lower degree of smoothing (the so-called *short averages*: e.g. for $\alpha = 0.3$); if at a time point of the corresponding chart the short averages intersect the long averages from below (or from above), then it indicates to buy (or to sell) given stocks, respectively)

(2) *moments*: e.g.

$$M = \frac{\hat{P}_t}{\hat{P}_{t-m}},$$

where m is a preset fixed period (*moment of stock price*)

(3) *indicators of volatility*: e.g.

$$\mathrm{SD}_t = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(\hat{P}_{t-i+1} - P_{t-i+1}\right)^2},$$

where m is a preset fixed period (*standard deviation of stock price*)

(4) *oscillators*: e.g.

$$O_t = \hat{P}_t^{(1)} - \hat{P}_t^{(2)},$$

where $\hat{P}_t^{(1)}$ and $\hat{P}_t^{(2)}$ are values P_t smoothed by two different smoothing methods (usually by means of moving averages with different lengths, see Sect. 31.2) (*oscillator of stock price*)

(5) *indices*: e.g.

$$\mathrm{CFI} = \frac{\mathrm{PCF/NCF}}{1 + \mathrm{PCF/NCF}},$$

where PCF (or NCF) is the cumulative amount of all positive (or negative) cash flows, respectively, caused by changes of stock price during a given period (*cash flow index* CFI)

- *Rights issue* (*privileged subscription*): is an offer to existing stockholders to subscribe cash for new stocks in proportion to their existing holdings (in addition, this offer includes an advantageous *subscription price* for the stock); the rights are distributed to the existing stockholders similarly as the dividends, namely one usually allocates one right to each stock; in particular, if a stockholder buys the stock after the *ex-right date* (shortly before the *right date*), the distributed right remains with the previous stockholder

Denotation:

R	price of right
P_{before}	price of stock before the ex-right date
P_{after}	price of stock after the ex-right date
S	subscription price (see *thereinbefore*)
N	number of rights that entitle to purchase one stock for the subscription price S

$$R = \begin{cases} \dfrac{P_{\text{before}} - S}{N+1} & \text{before ex-right date} \\ \dfrac{P_{\text{after}} - S}{N} & \text{after ex-right date} \end{cases}$$

(*price of right*)

9.4 Currencies

- *Exchange rate*: is the number of units of domestic currency A which can be exchanged for one unit of foreign currency B (more generally, instead of domestic and foreign currencies one can have basic and non-basic ones, respectively); exchange rates can be
  - *fixed*: the currency value is fixed through regulation mechanisms to an acceptable standard (gold, US dollar, currency basket, and the like)
  - *floating*: the currency value is based through market mechanisms upon supply and demand for the currencies

- *Foreign exchange market* (FOREX or FX): is the place of currency trading

- *Exchange quotation* in bank practice:
  - *bid rate*: is used if the bank buys the foreign currency B and pays the exchange equivalent in the domestic currency A
  - *ask rate* (*offer rate*): is used if the bank sells the foreign currency B and receives the exchange equivalent in the domestic currency A

$$A/B_{\text{bid}} - A/B_{\text{ask}} \text{ (e.g. EUR/USD 0.7099} - \text{0.7107)}$$

where A is the domestic (i.e. basic) currency and B is the foreign (i.e. non-basic) currency

(*quoted exchange rate*; it holds mostly $A/B_{\text{bid}} < A/B_{\text{ask}}$; the range $A/B_{\text{bid}} - A/B_{\text{ask}}$ is called *bid/ask spread*)

$$B/C_{\text{bid}} = \frac{A/C_{\text{bid}}}{A/B_{\text{ask}}};\ B/C_{\text{ask}} = \frac{A/C_{\text{ask}}}{A/B_{\text{bid}}}$$

(calculation of *cross exchange rates*)

Further Reading

Adams, A.T. et al.: Investment Mathematics. Wiley, Chichester (2003)
Bosch, K.: Finanzmathematik für Banker. Oldenbourg Verlag, München (2001)
Brealey, R.A., Myers, S.C.: Principles of Corporate Finance. McGraw-Hill, New York (1988)
Brigham, E.F.: Fundamentals of Financial Management. The Dryden Press, Forth Worth, TX (1992)
Cissell, R., Cissell, H., Flaspohler, D.C.: Mathematics of Finance. Houghton Mifflin, Boston, MA (1982)
Dupacova, J., Hurt, J., Stepan, J.: Stochastic Modeling in Economics and Finance. Kluwer, Dordrecht (2002)
Grundmann, W., Luderer, B.: Formelsammlung: Finanzmathematik, Versicherungsmathematik, Wertpapieranalyse. Teubner, Stuttgart (2001)
Knox, D.M., Zima, P., Brown, R.L.: Mathematics of Finance. McGraw-Hill, Sydney (1984)

Luderer, B., Nollau, V., Vetters, K.: Mathematische Formeln für Wirtschaftswissenschaftler. Teubner, Stuttgart (2000)

McCutcheon, J.J., Scott, W.F.: An Introduction to the Mathematics of Finance. Heinemann, London (1986)

Rosenberg, J.M.: Dictionary of Banking and Financial Services. Wiley, New York (1980)

Santomero, A.M., Babbel, D.F.: Financial Markets, Instruments, and Institutions. McGraw-Hill (Irwin), Chicago, IL (1997)

Sears, R.S., Trennepohl, G.L.: Investment Management. The Dryden Press, Forth Worth, TX (1993)

Sharpe, W.F., Alexander, G.J.: Investments. Prentice Hall, Englewood Cliffs, NJ (1990)

Thomsett, M.C.: The Mathematics of Investing. Wiley, New York (1989)

Van Horne, J.C.: Financial Market Rates and Flows. Prentice-Hall, Englewood Cliffs, NJ (1978)

Chapter 10
Derivative Securities

Abstract Chapter 10 deals with term trading and derivative securities: 10.1. General Classification, 10.2. Forwards, 10.3. Futures, 10.4. Swaps, 10.5. Options.

10.1 General Classification

- *Derivative securities*: are securities whose values are dependent on (derived from) the values of other more basic underlying variables which may be prices of traded securities (bonds, stocks), currencies, commodities or other cash market instruments, stock indices, and the like (one deals here with *contingent claims* whose amounts are determined by the behaviour of market securities up to maturity dates they are paid)
  - from theoretical point of view: derivative securities are financial contracts whose value F_T at the maturity date T is determined exactly by the market price S_T of the underlying cash instrument at the time T
  - in contrast to *prompt trading*, derivative securities represent *term trading* with the agreed term between the contract date (on which the contract is entered into) and the maturity date; typically, all determinants of such a contract are agreed at the contract date, i.e.
    - type and volume of *underlying assets* (bonds, commodities, and the like); the term *financial derivatives* means that they are underlain by financial assets (i.e. not by commodities)
    - type of contract: obligation or right to buy or sell (the party of the contract that agrees to buy the underlying asset assumes the so-called *long position*, and the party that agrees to sell the underlying asset assumes the so-called *short position*)
    - *maturity date* (*delivery date*, *exercise date*, *expiration date*) in future
    - *term price* (*delivery price*, *exercise price*, *strike price*) of underlying assets
    - other relevant conditions (e.g. the type of underlying bonds, the delivery place of underlying commodities)

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_10,

- *Hedging*: means to protect the position of an investor (hedger) against the risk of market movement; an approximate hedging consists in taking opposite positions in two assets which are highly negatively correlated; however, a perfect hedging can be achieved just using derivatives, which enable to "lock" the price of the underlying asset
- *Arbitrage*: enables to form a riskless position by means of a suitable combination of derivative securities with the underlying asset so that such a configuration continues to earn riskless returns
- *Trading*: is a term business by speculators that intend to profit of the assumed price development; if one believes that the price of underlying assets will grow (or will drop) then one enters the term purchases (or term sales), respectively
- Types of derivative securities:
  - according to the type of the underlying assets:
    - *commodity derivatives*: contract future purchase or sale of material commodities
    - *currency derivatives*: contract future purchase or sale of some currencies
    - *interest rate derivatives*: contract future purchase or sale of interest instruments (deposits, credits, short-term or long-term bonds, etc.)
    - *stock derivatives*: contract future purchase or sale of stocks
    - *stock index derivatives*: bet on future development of stock indices
    - *financial derivatives* and *non-financial derivatives*: non-financial derivatives are commodity derivatives with exception of derivatives for precious metals; all others are financial derivatives
  - according to the symmetry in positions of both parties participating in the term trading:
    - *symmetric* positions: both parties enter the contract free of charge, but they are obliged to accomplish the dealing regardless of the real market situation; typically, it concerns the derivative contracts (e.g. *forwards* or *swaps*) at *over-the-counter* markets (OTC markets) outside exchanges
    - *asymmetric* positions: one party pays an *option premium* and obtains the possibility of option, once the contract matures (the counterparty must await in a passive position the decision of the active party); it concerns the derivative securities called *options*, *caps*, *floors* or *collars*

10.2 Forwards

- *Forward contract*: is an obligation to buy (for the party assuming the long position) or to sell (for the party assuming the short position) an underlying asset at a specified delivery price on a known maturity date (settlement day); on the contract date the delivery price is chosen so that it costs nothing to take either a long or a short position ("the value of the contract to both parties is zero", see *thereinafter*); the forward contract is settled at maturity: the holder of the short position delivers the asset to the holder of the long position in return for the amount equal to the delivery price; in practice, the forward is usually contracted between two financial institutions or between a financial institution and one of its corporate clients (it is not normally traded on an exchange)

- *Currency forward* (FX-*forward*): serves for the term purchase or sale of a currency at the *term exchange rate* agreed on the contract date (it differs mostly from the *spot exchange rate* valid on the contract date); typically, such a contract is usual between the bank and its client that endeavours to secure a reasonable exchange rate for intended future purchase or sale of a currency (or speculates at the foreign exchange market)

Denotation:

$\mathrm{TK}_{A/B}^{\mathrm{bid}}$ or $\mathrm{TK}_{A/B}^{\mathrm{ask}}$	bid or ask term exchange rate, if the bank buys or sells the foreign currency B and pays or receives the exchange equivalent in the domestic currency A, respectively (see Sect. 9.4)
$\mathrm{SK}_{A/B}^{\mathrm{bid}}$ or $\mathrm{SK}_{A/B}^{\mathrm{ask}}$	bid or ask spot exchange rate, if the bank buys or sells the foreign currency B and pays or receives the exchange equivalent in the domestic currency A, respectively (see Sect. 9.4)
i_A^d or i_A^c	annual interest rate of bank deposits (a borrowed money) or bank credits (a lent money) in the currency A over the term corresponding to the considered forward term, respectively (similarly for the currency B)
t	current date
T	maturity date of forward

$$\mathrm{TK}_{A/B}^{\mathrm{bid}} = \mathrm{SK}_{A/B}^{\mathrm{bid}} \frac{1 + i_A^d (T-t)/360}{1 + i_B^c (T-t)/360}$$

(*bid term exchange rate*)

$$\mathrm{TK}_{A/B}^{\mathrm{bid}} \sim \mathrm{SK}_{A/B}^{\mathrm{bid}} [1 + (i_A^d - i_B^c)(T-t)/360]$$

(*approximate* bid term exchange rate)

$$\mathrm{TK}_{A/B}^{\mathrm{ask}} = \mathrm{SK}_{A/B}^{\mathrm{ask}} \frac{1 + i_A^c (T-t)/360}{1 + i_B^d (T-t)/360}$$

(*ask term exchange rate*)

$$\mathrm{TK}_{A/B}^{\mathrm{ask}} \sim \mathrm{SK}_{A/B}^{\mathrm{ask}} [1 + (i_A^c - i_B^d)(T-t)/360]$$

(*approximate* ask term exchange rate)

- *FX-swap* (*foreign exchange swap*): modifies the currency forward (with applications for currency trading or for short-term asset/liability management); it is formed by spot and term currency transactions, which are inverse, but with different exchange rates (e.g. the spot purchase of dollars at euros and then the term sale of dollars at euros); here the key factor is the *swap rate* (in practice, it is used also when man quotes the FX-forwards):

$$\mathrm{SW}_{A/B}^{\mathrm{bid}\to\mathrm{ask}} = \mathrm{TK}_{A/B}^{\mathrm{ask}} - \mathrm{SK}_{A/B}^{\mathrm{bid}};$$

$$\mathrm{SW}_{A/B}^{\mathrm{ask}\to\mathrm{bid}} = \mathrm{TK}_{A/B}^{\mathrm{bid}} - \mathrm{SK}_{A/B}^{\mathrm{ask}}$$

(*swap rate* of FX-swap)

- *Interest rate forward* (*forward rate agreement* FRA): enables for a short future period to fix the interest rate for a received credit (i.e. a bank or a firm borrows money) or a provided credit (i.e. a bank or a firm lends money or invests in short terms); typically, FRA consists in the *interest balancing* for the difference really existing on the FRA maturity date (this date must be chosen as the initial date of the received or provided credit) between a fixed interest rate (so called FRA-*rate*) and a floating interest rate (so called *reference rate*, e.g. 2% above LIBOR) agreed for the *notional amount* of the credit (the term "notional" denotes that this amount serves only to calculation of the *balancing payment*):
  - *FRA buyer* (i.e. the long-position party) receives from the FRA seller the interest difference between reference rate and FRA-rate (if the difference is negative, then this balancing payment goes in fact from the FRA buyer to the FRA seller); *the FRA buyers hedge against the risk of increasing interest rates for a future borrowed money* ("they lock the future paid interests")
  - *FRA seller* (i.e. the short-position party) receives from the FRA buyer the interest difference between FRA-rate and reference rate; *the FRA sellers hedge against the risk of decreasing interest rates for a future lent or invested money* ("they lock the future received interests")

Denotation:

P	notional amount of the underlying credit
t	current date
T	maturity date of FRA (initial date of the underlying credit)
T^*	maturity date of the underlying credit ($t < T < T^*$)
i_{ref}	annual reference interest rate
i_{FRA}	annual FRA-rate
i_T	annual risk-free interest rate (see Sect. 5.1) at time t with maturity at time T
i_{T*}	annual risk-free interest rate (see Sect. 5.1) at time t with maturity at time T^*

$$K_{\mathrm{FRA}}^{\mathrm{buy}} = -K_{\mathrm{FRA}}^{\mathrm{sell}} = P\frac{(i_{\mathrm{ref}} - i_{\mathrm{FRA}})(T^* - T)/360}{1 + i_{\mathrm{ref}}(T^* - T)/360}$$

(*balancing payment* for the FRA buyer; balancing payment for the FRA seller has the opposite sign only)

$$i_{\mathrm{FRA}} = \frac{i_{T^*}(T^* - t) - i_T(T - t)}{[1 + i_T(T - t)/360](T^* - T)} (\textit{FRA-rate})$$

- *Cost-of-carry model*: is in general the relationship between term (i.e. future) prices and spot prices; the cost of carry measures the interest that is paid to finance an asset (plus the storage cost) less the income earned on the asset; in particular, if one opens a forward position, then it must correspond to refinancing costs that would be necessary to achieve the same result at the spot market
- *Delivery price*: is a specified price K of the asset (underlying the given forward) valid on the maturity date, but agreed on the contract date
- *Long forward value at time t* ($t \le T$): is a potential momentary profit f_t due to the long forward position assumed at time t; *short forward value at time t* has just the opposite value $-f_t$; long and short forward values on the contract date are zero (it costs nothing to take either a long or a short position); if S_T denotes the actual price of the underlying asset on the maturity date T, then $f_T = S_T - K$
- *Forward price at time t* ($t \le T$): is such a delivery price denoted as F_t which would make the forward to have zero value at time t; on the contract date, the forward price is equal to the delivery price K (however, as time passes, the forward price is liable to changes, while the delivery price remains the same)

Denotation:

K	delivery price
t	current date
T	maturity date of forward
S_t	price of underlying asset at time t (*spot quotation*)
S_T	price of underlying asset at time T (i.e. on the maturity date)
f_t	long forward value at time t
F_t	forward price at time t (*term quotation*)
i	annual riskless interest rate (see Sect. 5.1) at the period from t up to T
V_t	present value (at time t) of such bond coupons which are to be paid up to the forward maturity (one applies corresponding riskless interest rates for its calculation)
d	constant annual rate of dividend yield paid continuously
i_c	annual riskless interest rate for a foreign currency at the period from t up to T
s	constant annual rate of storage costs

$$f_t = S_t - Ke^{-i(T-t)}$$

(long forward value for *zero-coupon bond*)

$$F_t = K = S_t e^{i(T-t)}$$

(forward price for *zero-coupon bond* on the contract date)

$$f_t = S_t - V_t - Ke^{-i(T-t)}$$

(long forward value for *coupon bond*)

$$F_t = K = (S_t - V_t)e^{i(T-t)}$$

(forward price for *coupon bond* on the contract date)

$$f_t = S_t e^{-d(T-t)} - Ke^{-i(T-t)}$$

(long forward value for *stock*)

$$F_t = K = S_t e^{(i-d)(T-t)}$$

(forward price for *stock* on the contract date)

$$f_t = S_t e^{-i_c(T-t)} - Ke^{-i(T-t)}$$

(long forward value for *foreign currency*)

$$F_t = K = S_t e^{(i-i_c)(T-t)}$$

(forward price for *foreign currency* on the contract date)

$$f_t = S_t e^{s(T-t)} - Ke^{-i(T-t)}$$

(long forward value for *precious metal*)

$$F_t = K = S_t e^{(i+s)(T-t)}$$

(forward price for *precious metal* on the contract date)

10.3 Futures

- *Futures*: are forward contracts standardized in such a way that they can be normally traded on exchanges; such a mass trading opens further possibilities:
  - any profit or loss is daily recorded in the account of each contract holder (*marking to market*) as if each future contract has been daily settled and

simultaneously a new contract has been written (the term *futures price* is used instead of the forward price)
  - system of *margins* prescribed for participants eliminates financial (default) risks
  - one can open and close particular futures positions in an effective way by means of *clearing houses* that stand between the parties and guarantee contractual provisions
  - underlying assets can be replaced by intangible security indices (e.g. S&P 100)

- Standardization of futures includes e.g.:
  - *standardized type of underlying asset* (e.g. any government bond with rating at least AA and with maturity from 8 to 10 years)
  - *standardized quantity of underlying asset* (e.g. multiples of 125 000 EUR for FX-futures)
  - *standardized maturity date* (e.g. the third Friday in March)
  - minimal possible change of future price (*tick size*)

- Subject with an *open* (long or short) *position* can:
  - keep the position up to the maturity date and perform it through a *physical delivery* or a *cash settlement*
  - *close out* the position, i.e. enter into an opposite trade to the original one (e.g. if an investor went long one futures position, he or she can close out by going short one futures position of the same type); majority of futures contracts are closed out in this way (the total number of contracts outstanding at the end of a business day is reported as *open interest*)

$$N = \beta \frac{S}{F}$$

(number N of futures contracts for a stock index (see Sect. 10.1) that are necessary to hedge stocks (or stock portfolio) with the price S and the measure β (see Sect. 13.3) against the given stock index with the term quotation F per one contract)

$$\text{gross basis} = \text{spot quotation} - \text{actual term quotation}$$

(*gross basis*: is the difference between the spot quotation (i.e. the price of underlying asset at the current time) and the actual term quotation (i.e. the actual future price at the current time, see Sect. 10.2))

$$\text{carry basis} = \text{spot quotation} - \text{theoretical term quotation}$$

(*carry basis*: reflects net refinancing costs between the spot quotation and the theoretical term quotation (the theoretical term quotation follows from the cost-of-carry model, see Sect. 10.2))

$$\begin{aligned}\text{value basis} &= \text{theoretical term quotation – actual term quotation}\\ &= \text{gross basis – carry basis}\end{aligned}$$

(*value basis*: reflects market factors that affect the actual term quotation, but are different from the theoretical "cost-of-carry" factors; if the value basis is zero, then the situation with *fair value* occurs, i.e. the theoretical term quotation can be applied directly in practice)

10.4 Swaps

- *Swaps*: are agreements between two parties to exchange cash flows across a future period according to a prearranged formula

- *Interest rate swap IRS*: is a contract between two parties to exchange interest streams with different characteristics (e.g. fixed interests against floating interests) based on an underlying principal, which serves only to calculation of the interest payments (*notional amount*, *volume of swap*); typically, IRS's are contracted at OTC markets (outside exchanges) with following motivations:
  – speculation consisting in different expectations on future development of interest rates by both parties
  – hedging against the interest rate risk by exchanging the given interest schedules
  – better accessibility of credits or investments when man transfers to another interest schedule
  – possibility to acquire cheaper capital resources by reducing the costs of their acquisition

- *Cross currency swap CCS* (be careful to distinguish properly the CCS's and the *FX-swaps*, see Sect. 10.2)): is a swap contract, where, in addition, the interest payments exchanged by parties are nominated in different currencies due to exchange of notional amounts nominated in these currencies

- Special types of swaps:
  – *coupon swap* (fixed-to-floating-swap): means that one party pays the other fixed interest rate coupons, and the other, in return, pays floating interest rate coupons (e.g. based on the 6 months LIBOR)
  – *basis swap* (floating-to-floating-swap): means that both parties apply different floating interest schemes (e.g. the 6 months LIBOR against the 3 months LIBOR)
  – *step-up-swap*: increases gradually by means of a given scenario either the notional amount, or the corresponding interest rates
  – *depreciated swap*: reduces gradually the principal by its notional depreciation

- *up-front-payment-swap*: is initiated by a one-off payment that balances the present values of interests streams for both parties
- *forward-swap*: defers the initiation of the swap transaction to a future date
- *multi-leg-swap*: involves more swap parties mixing several types of basic swaps

10.5 Options

- *Options*: are contracts agreed at time t giving their holders the right (but not the obligation, which is the case of forwards and futures) to do a transaction by a future date T (*exercise date* or *expiration date*), where $t < T$; options contracts are mostly standardized and then traded on derivative exchanges but they can also have OTC forms
- *Call* (*option*): gives its buyer (*holder* in *long position*) the right to buy an underlying asset S_t by an expiration date for a preset price X (*exercise price*, *strike price*); the call option can be purchased for a price C_t (*call option price*, *call option premium*); the seller of the call option (*writer* in *short position*) must sell the underlying asset by the expiration date according to the holder's decision
- *Put* (*option*): gives its buyer (*holder* in *long position*) the right to sell an underlying asset S_t by an expiration date for a preset price X (*exercise price*, *strike price*); the put option can be purchased for a price P_t (*put option price*, *put option premium*); the seller of the put option (*writer* in *short position*) must buy the underlying asset by the expiration date according to the holder's decision
- *European option*: can be exercised only on the expiration date itself
- *American option*: can be exercised at any time up to the expiration date
- *Commodity option*, *currency option*, *interest rate option*, *stock option*, *stock index option* and others: is the classification of options according to the type of the underlying assets
- *Clearing houses* stand between the holders and writers on exchanges; they organize the option trading in an effective way by means of margins (similarly as for futures, see Sect. 10.3)

Denotation:

X exercise price (strike price)
t current date
S_t price of underlying asset at time t (*spot price*)

- Option is at time t:
  - *at-the-money*, if the exercise price is equal to the price of underlying asset ($X = S_t$)
  - *in-the-money*, if the exercise price is for the option holder more convenient than the price of underlying asset (i.e. $X < S_t$ for call, and $X > S_t$ for put)

– *out-of-the-money*, if the exercise price is for the option holder less convenient than the price of underlying asset (i.e. $X > S_t$ for call, and $X < S_t$ for put)

Denotation:

S price of underlying asset on the exercise date
X exercise price
C call option price (premium)
P put option price (premium)

$$Z^C = \max(0,\ S - X) - C$$

(*payoff of exercised call*: is the profit of the long call position, i.e. the profit of the call holder)

$$Z^P = \max(0,\ X - S) - P$$

(*payoff of exercised put*: is the profit of the long put position, i.e. the profit of the put holder)

- Call options fulfil:
  - profit of the long call position = loss of the short call position (*zero-sum game*)
  - the long call position produces the profit for $S > X + C$, the constrained loss for $X < S < X + C$ and the maximal loss for $S \leq X$
  - profit of the long call position is not constrained, but its loss is constrained (by the option price)
  - call holders and put writers often speculate on a later price increase of the underlying assets (such speculators are called *bulls*)
- *Put options* fulfil:
  - profit of the long put position = loss of the short put position (*zero-sum game*)
  - the long put position produces the profit for $S < X – P$, the constrained loss for $X – P < S < X$ and the maximal loss for $S \geq X$
  - profit of the long put position is constrained (by the value $X – P$) and its loss is also constrained (by the option price)
  - put holders and call writers often speculate on a later price decrease of the underlying assets (such speculators are called *bears*)
- *Warrants* (see Sect. 9.2): are typically call options issued by a stock company which give the holder the right to buy a specified number (so called *option ratio*) of common stocks for a fixed price; unlike the classical call options, (1) the warrants are usually long-term securities, (2) the writer of the warrants is simultaneously the issuer of the underlying stocks, (3) the warrants use to be parts of other securities (usually bonds), from which they can be separated
- (*Interest rates*) *caps*: provide corporate borrowers with protection against the rate of interest on a floating-rate loan going above some level called *cap rate*; if the

interest rate goes above the cap rate, the writer (seller) of the cap provides to the holder (buyer) of the cap the difference between the interest and the cap rate; in practice, the cap takes the form of consequent payments called *caplets*

- (*Interest rates*) *floors*: provide corporate lenders with protection against the rate of interest on a floating-rate loan going below some level called *floor rate* analogously as for caps; in practice, the floor takes the form of consequent payments called *floorlets*
- (*Interest rates*) *collars*: are combinations of caps and floors specifying both the upper and the lower limit for the interest rate that will be charged
- *Options on futures* (*futures options*): are options, whose underlying asset is a futures contract (see Sect. 10.3); if the holder of such a call (or put) option exercises then he or she acquires from its writer a long (or short) position in the underlying futures contract plus a cash amount equal to the excess of the futures price over the exercise price (or the excess of the exercise price over the futures price), respectively
- *Compound options*: are options on options
- *Swaptions*: are options on swaps (see Sect. 10.4)
- *Captions*: are options on caps (see *thereinbefore*)
- *Exotic options*: have more complicated payoffs than the standard European or American calls and puts; they are usually traded over the counter (i.e. OTC) and *path-dependent* (i.e. their payoffs depend on the history of the underlying asset); examples are (1) *Asian options* (the payoff is defined in terms of an average value of the underlying asset); (2) *barrier options* (the payoff depends on whether the underlying asset price reaches a certain level during a certain time period); (3) "*as-you-like-it*" *options* (the holder can decide after a certain time whether they will be calls or puts); (4) *binary options* (the payoff is a fixed amount if the underlying asset price rises or falls bellow the exercise price), and others

- *Option pricing*: in practice, the option price (*option premium*, see *thereinbefore*) is usually calculated by means of *Black-Scholes formula* and its modifications; the construction of such formulas is based on principles and instruments of stochastic financial analysis (see Sects. 15.8 and 15.9)
- *Intrinsic value* at time t: is the corresponding payoff when the option should exercise at time t (compare with the long forward value at time t, see Sect. 10.2); this value has only a theoretical meaning and depends only on the exercise price of the option and on the spot price of the underlying asset (see below the table)

	Change of intrinsic value of option	
	call	put
S_t $\uparrow$	$\downarrow$	$\downarrow$
X $\uparrow$	$\downarrow$	$\uparrow$

- *Time value* at time t: is the award which the holder of option is willing to pay to its writer in view of the possibility that by the exercise date the price of the underlying asset will change to his benefit; the total value of the option (i.e. the option price) can be thought as the sum of its intrinsic value and time value; the time value declines to zero value with forthcoming expiration date and depends on several factors (see below the table)

		Change of time value of option	
		call	put
$\lvert S_t - X\rvert$	$\uparrow$	$\downarrow$	$\downarrow$
$T-t$	$\uparrow$	$\uparrow$	$\uparrow$
σ	$\uparrow$	$\uparrow$	$\uparrow$
i	$\uparrow$	$\uparrow$	$\downarrow$

Denotation:

X	exercise price
t	current date
T	exercise date of option
S_t	price of underlying asset at time t (spot price)
σ	price volatility of underlying asset (see Sect. 12.2)
i	risk-free interest rate (see Sect. 5.1)
i_C	risk-free interest rate for currency bought by means of currency call option
i_P	risk-free interest rate for currency sold by means of currency put option
C_t	call option price (premium) at time t
P_t	put option price (premium) at time t
$\mathrm{IV}_t{}^C$	intrinsic value of call option at time t
$\mathrm{IV}_t{}^P$	intrinsic value of put option at time t
D_t	present value (at time t) of stock dividends paid up to the exercise date of option
d	constant annual rate of dividend yield paid continuously (see Sect. 10.2)
F_t	futures price at time t (see Sect. 10.3)
$\Phi(\cdot)$	distribution function of normal distribution $N(0, 1)$ (see Sect. 26.2)
$\varphi(x)$	probability density of normal distribution $N(0, 1)$ (see Sect. 26.5)

$$\mathrm{IV}_t^C = \max(S_t - X,\ 0) = (S_t - X)^+$$

(*intrinsic value of call option*)

$$\mathrm{IV}_t^P = \max(X - S_t,\ 0) = (X - S_t)^+$$

(*intrinsic value of put option*)

$$\text{option price} = \text{intrinsic value of option} + \text{time value of option}$$

- Lower and upper bounds for option prices (option premiums): hold in various particular cases, e.g.:

$$\max(S_t - X\mathrm{e}^{-i(T-t)},\ 0) \le C_t \le S_t$$

(bounds for price of *European call option*)

$$\max(X\mathrm{e}^{-i(T-t)} - S_t,\ 0) \le P_t \le X\mathrm{e}^{-i(T-t)}$$

(bounds for price of *European put option*)

$$\max(S_t - X\mathrm{e}^{-i(T-t)} - D_t,\ 0) \le C_t \le S_t - D_t$$

(bounds for price of *European call option on dividend-paying stock*)

$$\max(X\mathrm{e}^{-i(T-t)} - S_t + D_t,\ 0) \le P_t \le X\mathrm{e}^{-i(T-t)} + D_t$$

(bounds for price of *European put option on dividend-paying stock*)

$$\max(S_t - X\mathrm{e}^{-i(T-t)},\ 0) \le C_t \le S_t$$

(bounds for price of *American call option*)

$$\max(X - S_t,\ 0) \le P_t \le X$$

(bounds for price of *American put option*)

- *Put-call parity*: is the relation between option prices of mutually corresponding calls and puts (i.e. with the same underlying asset, exercise price and exercise date), e.g.:

$$P_t = C_t + X\mathrm{e}^{-i(T-t)} - S_t$$

(put-call parity for *European options*)

$$P_t = C_t + X\mathrm{e}^{-i(T-t)} - S_t + D_t$$

(put-call parity for *European options on dividend-paying stock*)

$$S_t - X \le C_t - P_t \le S_t - X\mathrm{e}^{-i(T-t)}$$

(put-call parity for *American options*)

$$S_t - X - D_t \leq C_t - P_t \leq S_t - Xe^{-i(T-t)}$$

(put-call parity for *American options on dividend-paying stock*)

- *Black-Scholes formula* (B-S formula): expresses analytically the option price as a function of five factors: S_t (spot price of underlying asset), X (exercise price of option), $T - t$ (time remaining to exercise date), σ (price volatility of underlying asset), i (risk-free interest rate) for the period of length $T - t$; in various modifications one must add further factors:

$$C_t = S_t\Phi(d_1) - Xe^{-i(T-t)}\Phi(d_2), \text{ where}$$

$$d_1 = \frac{\ln(S_t/X) + (i + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}; d_2 = \frac{\ln(S_t/X) + (i - \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} = d_1 - \sigma\sqrt{T-t}$$

(B-S formula for price of *European call option*)

$$P_t = Xe^{-i(T-t)}\Phi(-d_2) - S_t\Phi(-d_1) = Xe^{-i(T-t)}[1 - \Phi(d_2)] - S_t[1 - \Phi(d_1)]$$

(B-S formula for price of *European put option*)

$$C_t = S_t\Phi(d_1) - Xe^{-i(T-t)}\Phi(d_2)$$

(B-S formula for price of *American call option*; an analogous analytical formula for American put option does not exist, and one must apply numerical procedures)

$$C_t = S_te^{-d(T-t)}\Phi(d_1) - Xe^{-i(T-t)}\Phi(d_2)$$

(B-S formula for price of *European call option on dividend-paying stock*)

$$P_t = Xe^{-i(T-t)}\Phi(-d_2) - S_te^{-d(T-t)}\Phi(-d_1)$$

(B-S formula for price of *European put option on dividend-paying stock*)

$$C_t = S_te^{-i_C(T-t)}\Phi(d_1) - Xe^{-i_P(T-t)}\Phi(d_2), \text{ where}$$

$$d_1 = \frac{\ln(S_t/X) + (i_P - i_C + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}; \quad d_2 = d_1 - \sigma\sqrt{T-t}$$

(*Garman-Kohlhagen formula* for price of *currency call option*; analogously for currency put option)

$$C_t = \mathrm{e}^{-i(T-t)}[F_t\Phi(d_1) - \mathrm{XN}(d_2)]; \qquad P_t = C_t + \mathrm{e}^{-i(T-t)}(X - F_t)\,, \text{ where}$$

$$d_1 = \frac{\ln(F_t/X) + (\sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}\,; \qquad d_2 = d_1 - \sigma\sqrt{T-t}$$

(B-S formula for price of *European option on futures*, see Sect. 10.3)

- *Greeks*: are special measures for sensitivity (or risk) of the option price (option premium) to factors which affect this price (due to Greek symbols, one calls such measures simply "Greeks"); the Greeks are mostly (mathematical) derivatives of the option price (evaluated by the B-S formula) with respect to the corresponding factor; the resulting measure quantifies an approximate linear dependence of the option price on the given factor:
  - *delta*: describes sensitivity of the option price to changes of the price of the underlying asset
  - *gamma*: describes sensitivity of the measure delta to changes of the price of the underlying asset
  - *lambda*: describes sensitivity of the option price to changes of the price volatility of the underlying asset
  - *rho*: describes sensitivity of the option price to changes of the risk-free interest rate
  - *theta*: describes sensitivity of the option price to changes of the time up to exercise date
- Examples of Greeks for European options (see *thereinbefore*):

$$\mathrm{delta}_t^C = \frac{\partial C_t}{\partial S_t} = \Phi(d_1); \qquad \mathrm{delta}_t^P = \frac{\partial P_t}{\partial S_t} = \mathrm{delta}_t^C - 1 = -\Phi(-d_1)\ (\textit{delta})$$

$$\mathrm{gamma}_t^C = \frac{\partial^2 C_t}{\partial S_t^2} = \frac{\varphi(d_1)}{\sigma S_t\sqrt{T-t}}; \qquad \mathrm{gamma}_t^P = \frac{\partial^2 P_t}{\partial S_t^2} = \mathrm{gama}_t^C\ (\textit{gamma})$$

$$\mathrm{theta}_t^C = \frac{\partial C_t}{\partial t} = -\frac{\sigma S_t}{2\sqrt{T-t}}\varphi(d_1) - iX\mathrm{e}^{-i(T-t)}\Phi(d_2);$$

$$\mathrm{theta}_t^P = \frac{\partial P_t}{\partial t} = \mathrm{theta}_t^C + iX\mathrm{e}^{-i(T-t)}\ (\textit{theta})$$

$$\mathrm{lambda}_t^C = \frac{\partial C_t}{\partial \sigma} = S_t\sqrt{T-t}\cdot\varphi(d_1); \qquad \mathrm{lambda}_t^P = \frac{\partial P_t}{\partial \sigma} = \mathrm{vega}_t^C\ (\textit{lambda})$$

$$\mathrm{rho}_t^C = \frac{\partial C_t}{\partial i} = (T-t)X\mathrm{e}^{-i(T-t)}\Phi(d_2); \mathrm{rho}_t^P = \frac{\partial P_t}{\partial i} = -(T-t)X\mathrm{e}^{-i(T-t)}\Phi(-d_2)\ (\textit{rho})$$

- *Strategic combinations based on options*: in practice, one frequently combines options of various types or combines options with commodities or other securities (e.g. *hedging* of investment portfolio by means of long put options on underlying assets from this portfolio); in this way, one can realize various strategies according to expected development of market (e.g. a suitable combination will be profitable under low price volatility of underlying assets, and the like); popular strategies combine options of different type, but with the same underlying asset and the same exercise date:

$$\text{long synthetic stock} = \text{long call } (X) + \text{short put } (X)$$

(*synthetic stock*: combines reverse call and put positions (i.e. long and short) which have the same exercise price: e.g. the given long synthetic stock combining long call and short put replicates a real purchase of the underlying asset for an investor that expects increasing price of the underlying asset)

$$\text{long straddle} = \text{long call } (X) + \text{long put } (X)$$

(*straddle*, *V-combination*: combines the same call and put positions which have the same exercise price: e.g. the given long straddle combining long call and long put is suitable for an investor that expects increasing price volatility of the underlying asset)

$$\text{long bull spread} = \text{long call } (X_1) + \text{short call } (X_2), \text{ where } X1 < X_2$$

(*bull spread*: combines the reverse call positions or reverse put positions which have different exercise prices: e.g. the given long bull spread combining long call with exercise price X_1 and short call with exercise price X_2 for $X_1 < X_2$ is suitable for "bulls" that speculate on increasing prices)

$$\text{long bear spread} = \text{short call } (X_1) + \text{long call } (X_2), \text{ where } X1 < X_2$$

(*bear spread*: is analogous to the long bull spread: e.g. the given long bear spread combining short call with exercise price X_1 and long call with exercise price X_2 for $X_1 < X_2$ is suitable for "bears" that speculate on decreasing prices)

$$\text{long risk reversal} = \text{short put } (X_1) + \text{long call } (X_2), \text{ where } X1 < X_2$$

(*risk reversal*: combines the reverse call and put positions which have different exercise prices: e.g. the given long risk reversal combining short put with exercise price X_1 and long call with exercise price X_2 for $X_1 < X_2$ is suitable for "bulls" that speculate on increasing prices)

Further Reading

Baxter, M., Rennie, A.: Financial Calculus. An Introduction to Derivative Pricing. Cambridge University Press, Cambridge (1996)

Black, F., Scholes, M.: The pricing of options and corporate liabilities. Journal of Political Economy 81, 637–654 (1973)

Brealey, R.A., Myers, S.C.: Principles of Corporate Finance. McGraw-Hill, New York (1988)

Duffie, D.: Security Markets: Stochastic Models. Academic Press, New York (1988)

Dupacova, J., Hurt, J., Stepan, J.: Stochastic Modeling in Economics and Finance. Kluwer, Dordrecht (2002)

Elliot, R.J., Kopp, P.E.: Mathematics of Financial Markets. Springer, New York (2004)

Hull, J.: Options, Futures, and Other Derivative Securities. Prentice Hall, Englewood Cliffs, NJ (1993)

Karatzas, I., Shreve, S.E.: Methods of Mathematical Finance. Springer, New York (1999)

Knox, D.M., Zima, P., Brown, R.L.: Mathematics of Finance. McGraw-Hill, Sydney, NSW (1984)

Kwok, Y.-K.: Mathematical Models of Financial Derivatives. Springer, Singapore (1998)

Malliaris, A.G., Brock, W.A.: Stochastic Methods in Economics and Finance. North-Holland, Amsterdam (1982)

Musiela, M., Rutkowski, M.: Martingale Methods in Financial Modelling. Springer, New York (2004)

Neftci, S.N.: Mathematics of Financial Derivatives. Academic Press, San Diego, CA (2000)

Pelsser, A.: Efficient Methods for Valuing Interest Rate Derivatives. Springer, London (2000)

Roman, S.: Introduction to the Mathematics of Finance. Springer, New York (2004)

Santomero, A.M., Babbel, D.F.: Financial Markets, Instruments, and Institutions. McGraw-Hill (Irwin), Chicago, IL (1997)

Sharpe, W.F., Alexander, G.J.: Investments. Prentice Hall, Englewood Cliffs, NJ (1990)

Steele, J.M.: Stochastic Calculus and Financial Applications. Springer, New York (2001)

Wilmott, P., Howison, S., Dewynne, J.: The Mathematics of Financial Derivatives. Cambridge University Press, Cambridge (1995)

Chapter 11
Utility Theory

Abstract Chapter 11 deals briefly with fundamentals of utility theory.

- *Utility*: means a degree of satisfaction or welfare coming from an economic activity (in finance mainly from an investment activity)

$$\mathbf{x} \succ \mathbf{y}, \text{ where } \mathbf{x} = (x_1, \ldots, x_N)' \in X; \mathbf{y} = (y_1, \ldots, y_N)' \in X$$

(*preference* of investment strategy $\mathbf{x}$ to investment strategy $\mathbf{y}$: is an ordering of the elements of a set of investment strategies X (from the mathematical point of view, an ordering relation on $X \times X$) corresponding to preferences of a given investor (x_i is usually interpreted as the volume of the ith investment ($i = 1, \ldots, N$) in the investment strategy $\mathbf{x}$); *weak preference* $\mathbf{x}$ to $\mathbf{y}$ admits also the equivalence of $\mathbf{x}$ and $\mathbf{y}$, i.e. simultaneously the weak preference of $\mathbf{y}$ to $\mathbf{x}$)

$$\mathbf{x} \sim \mathbf{y}, \text{ where } \mathbf{x} = (x_1, \ldots, x_N)' \in X; \mathbf{y} = (y_1, \ldots, y_N)' \in X$$

(*equivalence* of investment strategies $\mathbf{x}$ and $\mathbf{y}$: means indifference of investment strategies from the point of view of a given investor)

- *Utility function*: is a real function on the set of investment strategies X, which indicates by means of relations of its values a given preference ordering:

$$(U(\mathbf{x}) > U(\mathbf{y}) \Leftrightarrow \mathbf{x} \succ \mathbf{y}) \ \wedge \ (U(\mathbf{x}) = U(\mathbf{y}) \Leftrightarrow \mathbf{x} \sim \mathbf{y})$$

(*utility function* on the set of investment strategies; any increasing function of a utility function is again a utility function)

$$I_c = \{\mathbf{x} \in X : \quad U(\mathbf{x}) = c\} \quad \text{where } c \in \mathrm{R} \text{ (\textit{indifference curve})}$$

- *Examples of utility functions on the set of investment strategies X*: are increasing, continuous, twice differentiable functions (a possible concavity reflects the

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_11,

law of diminishing marginal utility, according to which the additional utility will decrease with the increasing amount of investment):

$$U(\mathbf{x}) = \sum_{i=1}^{N} a_i U_i(x_i), \quad \text{where } a_i > 0 \text{ and } U_i(x_i) \text{ are utility functions } (i = 1, \ldots, N)$$

(*separable* utility function)

$$U(\mathbf{x}) = \frac{1}{\gamma} \sum_{i=1}^{N} a_i x_i^{\gamma}, \quad \text{where } 0 < \gamma \leq 1 \text{ and } a_i > 0 \ (i = 1, \ldots, N)$$

(separable *power* utility function: is linear for $\gamma = 1$ and concave for $0 < \gamma < 1$)

$$U(\mathbf{x}) = \sum_{i=1}^{N} a_i \ln x_i, \quad \text{where } a_i > 0 \ (i = 1, \ldots, N)$$

(separable *logarithmic* utility function: is the limit case of the power one for $\gamma \to 0_+$)

$$U(\mathbf{x}) = \prod_{i=1}^{N} x_i^{a_i}, \quad \text{where } a_i > 0 \ (i = 1, \ldots, N)$$

(*Cobb-Douglas production function*: is increasing and concave for $0 < a_i < 1$ ($i = 1, \ldots, N$))

- Examples of utility functions in the portfolio theory (see Sect. 13):

$$U(\mu, \sigma; \kappa) = \mu - \kappa \cdot \sigma^2$$

(*quadratic* utility function: parameter κ is the *measure of investor's risk aversion*)

- Examples of utility functions of the cumulated wealth W:

$$U(W) = \max_{\mathbf{x}} \{U(\mathbf{x}) : \mathbf{p}'\mathbf{x} = W\}, \quad \text{where } \mathbf{p} = (p_1, \ldots, p_N)' \text{ is a given vector of prices}$$

(utility function of wealth W defined by means of a utility function $U(\mathbf{x})$ on X)

$$U(W) = -\frac{1}{\eta} \mathrm{e}^{-\eta \cdot W}, \quad \text{where } \eta > 0 \text{ (\textit{exponential} utility function)}$$

$$U(W) = \frac{1}{\gamma} W^{\gamma}, \quad \text{where } \gamma > 0 \text{ (\textit{power} utility function (see \textit{thereinbefore}))}$$

$$U(W) = \ln W$$

(*logarithmic* utility function: is the limit case of the power one for $\gamma \to 0_+$ (see *thereinbefore*))

$$U(W) = \frac{1-\gamma}{\gamma}\left(\frac{\beta \cdot W}{1-\gamma} + \eta\right)^{\gamma}, \text{where } \beta > 0, \gamma \neq 0$$

(HARA utility function (hyperbolic absolute risk aversion))

- Measures of risk aversion:

$$A(W) = -\frac{U''(W)}{U'(W)}$$

(*Arrow-Pratt measure of absolute risk aversion*)

$$R(W) = W \cdot A(W)$$

(*Arrow-Pratt measure of relative risk aversion*)

Further Reading

Dupacova, J., Hurt, J., Stepan, J.: Stochastic Modeling in Economics and Finance. Kluwer, Dordrecht (2002)

Ingersoll, J.E.: Theory of Financial Decision Making. Rowman and Littlefield, Savage (1987)

Jarrow, R.A., Maksimovic, V., Ziemba, W.T. (eds.): Finance. Handbooks in Operations Research and Management Science, Volume 9. Elsevier, Amsterdam (1995)

Chapter 12
Rate of Return and Financial Risk

Abstract Chapter 12 is devoted to problems related to financial risk: 12.1. Rate of return, 12.2. Financial risk, 12.3. Value at risk VaR, 12.4. Credit at risk CaR.

12.1 Rate of Return

- *Rate of return ROR*: is difference between the capital (in monetary units) at the end and at the beginning of the given period (typically, expressed relatively to the capital at the beginning of the given period):

$$ROR = \frac{\text{Captial at end of period} - \text{Capital at beginning of period}}{\text{Captial at beginning of period}}$$

Denotation:

P_t price of asset (e.g. a security or a foreign currency) at time t measured in chosen monetary units (if the time is measured in months, then P_t is the price at the end of month t, while $P_{t\text{–}1}$ is the price at the end of previous month t–1)
D_t possible payment for the period from time t–1 up to time t (e.g. a coupon in the case of bonds or a dividend in the case of stocks); if the time is measured e.g. in months and D is the dividend paid out yearly, then D_t=D/12)
p_t logarithmic price of asset: $p_t = \ln P_t$

$$\Delta_t = P_t - P_{t-1}$$

(*absolute price change* at time t: is the absolute change of price from time t–1 up to time t)

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_12,

$$R_t = ROR_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{P_t}{P_{t-1}} - 1$$

(*rate of return* (*relative price change*) at time t: is the relative change of price from time t–1 up to time t (typically, given as a percent value); in practice, the relative price change is preferred to the absolute one since it respects the given price level)

$$P_t = P_{t-1} \cdot (1 + R_t)$$

(the next forthcoming price expressed by means of the previous price and the rate of return, see Sect. 2.1)

$$R_t = ROR_t = \frac{P_t + D_t - P_{t-1}}{P_{t-1}} = \frac{P_t - P_{t-1}}{P_{t-1}} + \frac{D_t}{P_{t-1}}$$

(*rate of return* at time t taking into account the coupon or the dividend)

$$r_t = \ln(1 + R_t) = \ln\left(\frac{P_t}{P_{t-1}}\right) = p_t - p_{t-1}$$

(*logarithmic rate of return* (also see Sect. 4): can be interpreted as the absolute change of logarithmic prices from time t–1 up to time t)

$$P_t = P_{t-1} \cdot \mathrm{e}^{r_t}$$

(the next forthcoming price expressed by means of the previous price and the logarithmic rate of return, see Sect. 4)

$$r_t = R_t + O(R_t^2)$$

(approximation of the logarithmic rate of return by the rate of return: just due to this relation, the both rates of return are usually looked upon as equivalent)

Denotation:

$P_{A/B,\,t}$ exchange rate of currencies A and B at time t (the price of the foreign (i.e. non-basic) currency B unit priced by the domestic (i.e. basic) currency A units, see Sect. 9.4)

w_{it} proportion of funds invested at time t into ith asset (see the portfolio aggregation of N assets in Sect. 13.1); $w_{it} \geq 0$ (however, the short position in an asset can have its weight negative); $w_{1t} + \ldots + w_{Nt} = 1$

P_{pt} price of portfolio at time t

$$r_{A/B,\,t} = \ln\left(\frac{P_{A/B,\,t}}{P_{A/B,\,t-1}}\right)$$

(logarithmic rate of return of currencies A and B)

$$r_{A/B,\,t} = -\,r_{B/A,\,t}$$

$$r_{A/B,\,t} = r_{A/C,\,t} - r_{B/C,\,t}$$

$$R_t(k) = \frac{P_t - P_{t-k}}{P_{t-k}}$$

(*time aggregated* rate of return over k previous periods)

$$R_t(k) = (1 + R_t)\cdot(1 + R_{t-1})\cdot\ldots\cdot(1 + R_{t-k+1}) - 1$$

$$r_t(k) = \ln(1 + R_t(k)) = \ln\left(\frac{P_t}{P_{t-k}}\right) = p_t - p_{t-k}$$

(*time aggregated* logarithmic rate of return)

$$r_t(k) = r_t + r_{t-1} + \ldots + r_{t-k+1}$$

$$P_{pt} = w_{1,\,t-1}\cdot P_{p,\,t-1}\cdot(1 + R_{1t}) + \ldots + w_{N,\,t-1}\cdot P_{p,\,t-1}\cdot(1 + R_{Nt})$$

(*price of portfolio* at time t)

$$R_{pt} = \frac{P_{pt} - P_{p,\,t-1}}{P_{p,\,t-1}}$$

(*portfolio aggregated* rate of return (i.e. rate of return of the whole portfolio))

$$R_{pt} = w_{1,\,t-1}\cdot R_{1t} + \ldots + w_{N,\,t-1}\cdot R_{Nt}$$

$$r_{pt} = \ln\left(1 + R_{pt}\right) = \ln\left(\frac{P_{pt}}{P_{p,\,t-1}}\right)$$

(*portfolio aggregated* logarithmic rate of return)

	Time aggregation	Portfolio aggregation
Rate of return	$R_t(k) = \prod_{j=0}^{k-1} (1 + R_{t-j}) - 1$	$R_{pt} = \sum_{i=1}^{N} w_{i,\,t-1} \cdot R_{it}$
Logarithmic rate of return	$r_t(k) = \sum_{j=0}^{k-1} r_{t-j}$	$r_{pt} = \ln\left(\sum_{i=1}^{N} w_{i,\,t-1} \cdot \mathrm{e}^{r_{it}}\right)$

$$r_{pt} = \ln\left(w_{1,\,t-1} \cdot \mathrm{e}^{r_{1t}} + \ldots + w_{N,\,t-1} \cdot \mathrm{e}^{r_{Nt}}\right) \approx w_{1,\,t-1} \cdot r_{1t} + \ldots + w_{N,\,t-1} \cdot r_{Nt}$$

12.2 Financial Risk

- *Financial risk*: concerns potential price changes of financial assets, where the corresponding price change (expressed mainly as the rate of return, see Sect. 12.1) is looked upon as a random variable (see Sect. 26.3)
- *Volatility*: is the financial risk of corresponding financial asset measured as the standard deviation (or sometimes as the variance, see Sect. 26.3) of the rate of return for the asset

Denotation:

r_t logarithmic rate of return: is abbreviated simply as *rate of return* since it is an approximation of R_t (see Sect. 12.1)

$$\sigma_t = \sigma(r_t) = \sqrt{var(r_t)} \quad (\textit{volatility} \text{ of rate of return for a financial asset})$$

$$\sigma_t \approx \sqrt{\mathrm{E}(r_t^2)}$$

(approximation of volatility (typically in the case of daily rate of return with mean value near to zero))

$$\sigma(P_t) = \sigma_t \cdot P_{t-1} \quad (\text{volatility of } \textit{price } P_t \text{ for a financial asset})$$

$$\sigma_t(k) = \sqrt{var(r_t + r_{t-1} + \ldots + r_{t-k+1})} = \sqrt{\sigma_t^2 + \sigma_{t-1}^2 + \ldots + \sigma_{t-k+1}^2}$$

(volatility of *time aggregated* rate of return over k previous periods under assumption of uncorrelated rates (see Sect. 26.6) in particular periods)

$$\sigma_{\text{annual}} \approx \sqrt{252} \cdot \sigma_{\text{daily}}$$

(approximation of *annual volatility* by *daily volatility* aggregating over 252 business days per year (which is the average number of business days per year frequently used in practice); the approximation may be used only under assumptions that daily values of rate of return are sufficiently uncorrelated and homoscedastic)

Denotation:

w_{it} proportion of funds invested at time t into the ith asset (see the portfolio aggregation over N assets in Sect. 13.1); $w_{it} \geq 0$ (however, the short position in an asset has its weight negative); $w_{1t} + \ldots + w_{Nt} = 1$; $w_t = (w_{1t}, \ldots, w_{Nt})'$ is the weight vector

σ_{it} volatility of rate of return r_{it} for the ith asset in the portfolio:

$$\sigma_{it} = \sqrt{\sigma_{ijt}} = \sqrt{var(r_{it})}$$

σ_{ijt} covariance (see Sect. 26.6) of rates of return for ith and jth assets in the portfolio:

Σ_t matrix of covariances σ_{ijt} (see Sect. 26.6)

P_{it} price of ith asset at time t (i.e. funds invested at time t into ith asset called the position in the ith asset of the portfolio)

r_{pt} rate of return of portfolio (i.e. portfolio aggregated rate of return) at time t

P_{pt} price of portfolio at time t

$$\sigma_{pt} = \sqrt{var(r_{pt})} = \sqrt{var\left(\sum_{i=1}^{N} w_{i,\,t-1} \cdot r_{it}\right)} = \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{i,\,t-1} \cdot w_{j,\,t-1} \cdot \sigma_{ijt}}$$

$$= \sqrt{\mathbf{w}'_{t-1} \cdot \mathbf{\Sigma}_t \cdot \mathbf{w}_{t-1}}$$

(volatility of *rate of return for portfolio* (i.e. the volatility of portfolio aggregated rate of return): takes into account the covariances between rates of return of particular assets (such correlations can be significant in practice)

$$\beta_{it} = \frac{1}{\sigma_{pt}} \cdot \frac{\partial \sigma_{pt}}{\partial w_{i,\,t-1}} = \frac{1}{2\sigma_{pt}^2} \cdot \frac{\partial \sigma_{pt}^2}{\partial w_{i,\,t-1}} = \frac{\mathrm{cov}(r_{it}, r_{pt})}{\sigma_{pt}^2}$$

(sensitivity coefficient of relative changes in the volatility of portfolio aggregated rate of return at time t to changes in the weight of the ith asset in the portfolio)

$$\sigma_{pt}\beta_{it}\Delta w_{i,\,t-1}$$

(*incremental volatility* at time t due to the increment $\Delta w_{i,\,t-1}$ of weight $w_{i,\,t-1}$)

$$\boldsymbol{\beta}_t = \frac{\boldsymbol{\Sigma}_t \cdot \mathbf{w}_{t-1}}{\mathbf{w}'_{t-1} \cdot \boldsymbol{\Sigma}_t \cdot \mathbf{w}_{t-1}} \quad \text{(sensitivity vector } \boldsymbol{\beta}_t = (\beta_{1t}, \ldots, \beta_{Nt})'\text{)}$$

$$\sigma_{pt} = \sum_{i=1}^{N} \sigma_{pt} \cdot \beta_{it} \cdot w_{i,\, t-1}$$

(decomposition of volatility of portfolio aggregated rate of return at time t: the component $\sigma_{pt}\ \beta_{it}\ w_{i,\,t\text{–}1}$ represents the part of volatility of portfolio aggregated rate of return corresponding to the ith asset in the portfolio)

$$\beta_{it} = \frac{1}{\sigma(P_{pt})} \cdot \frac{\partial \sigma(P_{pt})}{\partial P_{it}}$$

(sensitivity coefficient of relative changes in the volatility of portfolio price at time t to changes in the price (i.e. in the position) of ith asset in the portfolio)

$$\beta_t = \frac{\Sigma_t \cdot P_t}{P'_t \cdot \Sigma_t \cdot P_t}$$

(sensitivity vector for $\mathrm{P}_t = (P_{1t}, \ldots, P_{Nt})'$)

$$\sigma(P_{pt}) = \sum_{i=1}^{N} \sigma(P_{pt}) \cdot \beta_{it} \cdot P_{it}$$

(decomposition of volatility of portfolio price at time t: the component $\sigma(P_{pt})\ \beta_{it}\ P_{it}$ represents the part of volatility of portfolio price corresponding to the ith asset in the portfolio)

- *Prediction of volatility*: plays key role in the *market risk management*; it is important for the VaR methodology (see Sect. 12.3) or in the cases, where the price of financial assets depends on volatility by a predictable way (e.g. through B-S formula, see Sect. 10.5); there are various methods for volatility prediction based on time series analysis (see *thereinafter*); a quite different approach (popular in practice) is the method of *implied volatility* based on the fact that the price of options (see Sect. 10.5) depends analytically also on the volatility of particular rates of return so that the corresponding volatilities can be calculated (and then quoted) by means of the market prices of these options

Denotation:

$\sigma_{t+1|t}$ prediction of volatility for time t+1 at time t (i.e. prediction is based on the observed values of rate of return r_t, r_{t-1}, r_{t-2},. . .)

$$\sigma_{t+1|t} = \sqrt{\sigma^2_{t+1|t}} = \sqrt{\frac{1}{M}\sum_{i=0}^{M-1} r^2_{t-i}}, \quad \text{where } M \in \text{N is a fixed length}$$

(prediction of volatility by *method of moving averages* (see Sect. 31.2); the prediction makes use only of M previous values of rate of return r_t, $r_{t-1}, \ldots, r_{t-M+1}$; the choice of the suitable length M of moving averages ("length of window") is very important)

$$\sigma^2_{t+1|t} = (1-\lambda) \cdot \sum_{i=0}^{\infty} \lambda^i \cdot r^2_{t-i} = (1-\lambda) \cdot r^2_t + \lambda \cdot \sigma^2_{t\,|\,t-1}, \quad \text{where } 0 < \lambda < 1$$

(prediction of volatility by *method of* (*simple*) *exponential smoothing* (see Sect. 31.2); the prediction makes use of all previous values of rate of return r_t, r_{t-1}, r_{t-2}, ... with weights, which decrease exponentially with the age of observations; typically, the recursive form of the corresponding formula is used in practice; the choice of the suitable smoothing constant $1-\lambda$ is important, e.g. some commercial systems use $\lambda = 0.94$ for daily rates of return and $\lambda = 0.97$ for monthly rates of return)

$$\sigma^2_{t+1|t} = \alpha_0 + \alpha_1 r^2_t + \beta_1 \sigma^2_{t\,|\,t-1}$$

(prediction of volatility by *model GARCH*(1, 1) (i.e. by the principle of conditional heteroscedasticity in financial time series, see Sect. 31.5); in particular, the choice $\alpha_0 = 0$, $\alpha_1 = 1-\lambda$, $\beta_1 = \lambda$ corresponds to the volatility prediction by the method of exponential smoothing (see *thereinbefore*); the model parameters are estimated usually by the maximum likelihood method (see Sect. 27.9))

$$\sigma^2_{t+1|t} = \alpha_0 + \alpha_1 r^2_t$$

(prediction of volatility by *model ARCH*(1, 1); it is a special case of the model *GARCH*(1, 1) lacking some suitable properties of the model *GARCH* (see Sect. 31.5))

$$\sqrt{\mathrm{E}[(r_{t+T} + r_{t+T-1} + \ldots + r_{t+1})^2 | r_t, r_{t-1}, \ldots)]} \approx \sqrt{T} \cdot \sigma_{t+1|t}$$

(approximation of volatility prediction of *time aggregated rate of return* under assumption of uncorrelated values of rate of return in particular future periods)

$$\rho_{12\ \ t+1|t} = (1-\lambda) \cdot r_{1t} \cdot r_{2t} + \lambda \cdot \rho_{12\ \ t\,|\,t-1}, \quad \text{where } 0 < \lambda < 1$$

(prediction of *correlation coefficient* (see Sect. 26.6) of rates of return r_{1t} and r_{2t} of two financial assets by the method of (simple) exponential smoothing (see *thereinbefore*))

12.3 Value at Risk VaR

- *Methodology VaR* (*value at risk*): is based on an estimate of the worst loss that can occur with a given probability (confidence) in a given future period; it is one of the best used approach to set up *capital requirements* when regulating *capital adequacy* (e.g. in so called *internal models* for banks: *VaR* represents the smallest capital amount that guarantees the bank solvency with a given confidence); more generally, *VaR* is the key instrument for financial risk management (e.g. by means of commercial systems of the type RiskMetrics)

- Methodology VaR is specified by the following factors:
  - *holding period*: is the period in which a potential loss can occur; accordingly, the used terms are the daily value at risk (over one business day, e.g. in *RiskMetrics*) or the 10 days value at risk (over two calendar weeks with ten business day, e.g. according to recommendation of *Basle Committee on Banking Supervision*)
  - *confidence level*: is the probability that the actual loss does not exceed the value at risk (during the given holding period); in practice, one applies e.g. the confidence 95% (in RiskMetrics), or 99 % (according to Basle Committee on Banking Supervision)

Denotation:

X	random variable representing profit (if X is positive) or loss (if X is negative) accumulated during the given holding period (e.g. for a controlled investment portfolio)
c	required confidence (e.g. $c = 0.95$)
P	price (e.g. price of a controlled investment portfolio)
r	rate of return over the given holding period looked upon as a random variable with mean value μ
r_{1-c}	(1–c)-quantile of random variable r (see Sect. 26.3)
u_{1-c}	(1–c)-quantile of standard normal distribution $N(0, 1)$ (see Sect. 26.5: $u_{1-c} = -u_c$)
$C(X)$	*risk measure*: mapping which assigns real values $C(X)$ to random variables X (a typical example of a risk measure is *VaR*); a risk measure is called *coherent*, if it possesses following properties (for bounded random variables X and Y):
(*i*)	subadditivity: $C(X+Y) \leq C(X)+C(Y)$
(*ii*)	monotony: if $X \leq Y$, then $C(X) \leq C(Y)$
(*iii*)	positive homogeneity: $C(\lambda X)=\lambda C(X)$ for an arbitrary constant $\lambda>0$
(*iv*)	translation invariance: $C(X+a)=C(X)+a$ for an arbitrary constant $a>0$

$$\mathrm{P}(X < -\mathit{VaR}^{abs}) = \mathrm{P}(-X > \mathit{VaR}^{abs}) = 1 - c$$

(*absolute value at risk* VaR^{abs}: the value VaR^{abs} is the worst loss, which occurs within a100·c% portion of all possible outcomes; the opposite value $-\mathit{VaR}^{abs}$ is the (1–c)-quantile x_{1-c} of random variable X (see Sect. 26.3))

$$VaR = VaR^{rel} = VaR^{abs} + \mathrm{E}(X)$$

((*relative*) *value at risk* VaR^{rel}: is related to the mean value $\mathrm{E}(X)$ of random variable X (see Sect. 26.3) as the distance between the absolute value at risk (i.e. VaR^{abs}) and the mean loss (i.e. $-\mathrm{E}(X)$); in practice, it is called simply *value at risk* and denoted as *VaR*, since one uses it frequently)

$$VaR^{abs} = -P \cdot r_{1-c}$$

(absolute value at risk expressed by means of the rate of return)

$$VaR = VaR^{rel} = -P \cdot (r_{1-c} - \mu)$$

(value at risk expressed by means of the rate of return)

- Distribution of random variable X (profit or loss) or r (rate of return):
  - one assumes usually that it is *normal* (see *thereinafter*)
  - however, such a normality assumption may not be compatible with reality, where the loss tail of financial data (i.e. the left negative tail of X) can be much heavier than the normal one (in such a case the normal approximation produces a smaller value at risk than it should be actually); therefore various alternatives to the normal approximation have been suggested, e.g. the *t distribution* (see Sect. 26.5), a mixture of normal distributions, and the like

$$VaR^{abs} = -P \cdot (\mu + \sigma \cdot u_{1-c})$$

(absolute value at risk expressed by means of the rate of return with *normal distribution* $r \sim N(\mu,\sigma^2)$ (see Sect. 26.5))

$$VaR = -P \cdot \sigma \cdot u_{1-c}$$

(value at risk expressed by means of the rate of return with *normal distribution* $r \sim N(\cdot,\sigma^2)$)

$$VaR = -P \cdot \sigma \cdot u_{1-c} \cdot \sqrt{\Delta t}$$

(value at risk expressed by means of the rate of return with *normal distribution* $r \sim N(\cdot, \sigma^2)$ over holding period Δt (the rate of return r relates to the corresponding time unit, so that e.g. for annual rate of return r and daily holding period with 252 business days per year one must substitute $\Delta t = 1/252$))

$$VaR = 1.65 \cdot P \cdot \sigma \cdot \sqrt{\Delta t}$$

(value at risk expressed by means of the rate of return with *normal distribution* $r \sim N(\cdot,\sigma^2)$ over holding period Δt and with confidence level 95% (see RiskMetrics))

$$VaR = 2.33 \cdot P \cdot \sigma \cdot \sqrt{\Delta t}$$

(value at risk expressed by means of the rate of return with *normal distribution* $r \sim N(\cdot, \sigma^2)$ over holding period Δt and with confidence level 99% (see Basle Committee on Banking Supervision))

$$VaR_{t+\,\Delta t\,|\,t} = -P \cdot \sigma_{t+\,\Delta t\,|\,t} \cdot u_{1-c} \cdot \sqrt{\Delta t}$$

(value at risk constructed as prediction at time t for future holding period Δt using volatility prediction $\sigma_{t+\,\Delta t\,|\,t}$ (see Sect. 12.2))

$$DEaR = VaR_{t+1|t} = 1.65 \cdot P \cdot \sigma_{t+1|t}$$

(*daily earnings at risk*: are daily predictions of the value at risk (t denotes particular business days) with confidence level 95%)

$$\Delta VaR_p = VaR_p \cdot \beta_i \cdot \Delta P_i$$

(*incremental value at risk IVaR*): provides the increment ΔVaR_p of the portfolio's value at risk VaR_p corresponding to the increment ΔP_i of price P_i of the ith asset in the portfolio; β_i is the sensitivity coefficient of relative changes in the volatility of portfolio price to changes in the price of the ith asset in the portfolio (see Sect. 12.2))

$$VaR_p = \sum_{i=1}^{N} VaR_p \cdot \beta_i \cdot P_i$$

(decomposition of value at risk of portfolio)

$$VaR^{BC} = \frac{2.33}{1.65} \cdot \sqrt{10} \cdot VaR^{RM} = 4.47 \cdot VaR^{RM}$$

(example of conversion of the value at risk (under assumptions that the rates of return are *iid* normal random variables) for 1 day holding period and confidence level 95% (see RiskMetrics) to the value at risk for 2 weeks holding period (i.e. ten business days) and confidence level 99% (see Basle Committee); the table shows various combinations of holding periods and confidence levels, which give the same *VaR* per 1,000 monetary units of the standard deviation, e.g. 4 weeks holding period and confidence level 95% gives the same *VaR* as 2 weeks holding period and confidence level 99%)

Holding period	Confidence (%)	Δt	$P{\cdot}\sigma$	*VaR*
1 year	67.85	1	1,000	463.42
3 months	82.30	3/12	1,000	463.42
4 weeks	95.00	21/252	1,000	463.42
2 weeks	99.00	10/252	1,000	463.42
1 week	99.95	5/252	1,000	463.42
1 day	100.00	1/252	1,000	463.42

$$TVaR = \mathrm{E}\,(X \mid X < -\,VaR)$$

(*tail value at risk*, *expected tail loss*, *expected shortfall ES*)

(*TVaR* is the conditional mean value of potential loss higher than the value at risk for the same confidence level; in particular, $TVaR = \mathrm{E}(X)$ when the confidence level is 100%; unlike *VaR*, the *TVaR* is the coherent (and more sensitive) risk measure; another alternative to the value at risk is the *maximum possible loss*)

Application of VaR methodology: requires to take into account a lot of practical aspect, e.g.:

- various conventions (e.g. the capital requirements when regulating capital adequacy according to Basle II may be set up as the treble up to the quadruple of the 10 days *VaR*)
- the calculation of *VaR* in practice: namely
  - *variance-covariance method* applied to particular cash flows of the evaluated financial system
  - *historical simulation method* simulates rates of returns using the history of the evaluated financial system
  - *structured method Monte Carlo* involves simulations of an explicit parametric model for risk-factor changes
  - *method of observed losses over several periods* and *scaling*
  - *back testing* makes use of statistical tests to verify credibility of *VaR* approaches
  - *stress testing* checks the vulnerability of *VaR* approaches to various (hypothetical) catastrophic phenomena

12.4 Credit at Risk CaR

- *Credit risk*: is the risk that the *creditor* (*lender*) may not receive promised repayments on outstanding investments (such as loans, credits, bonds, etc.), because of the *default* of the *debtor* (*borrower*); defaults can consist in insolvency or reluctance of the debtor, in his refusal to deliver or to buy underlying assets in the case of options, and the like)

- Quantitative components of credit risk:
  - *Probability of default PD*: is the probability that the debtor may not fulfil contract liabilities (obligations agreed in the contract); it is quantified by means of *credit rating* using an *internal* appreciation of client's ability to pay or an *agency* rating (Standard & Poor's, Moody's, Fitch and others)
  - *Exposure at default EAD*: quantifies the momentary credit exposure that is subject to be lost in the case of default (it is equal directly to the unpaid amount of claims in the case of balance sheet items or to suitable credit equivalents in the case of off-balance sheet entries (e.g. financial derivatives))
  - *Loss given default LGD*: describes the portion of loss the creditor will really suffer in the case of default; it is a fraction of *EAD* determined approximately as "1 minus *recovery rate*" (*recoveries* are e.g. collaterals, residual values in the case of company's liquidation, and the like); e.g. $LGD = 0.6$ means that the creditor hopes to regain 40 % outstanding debt)
  - *Expected loss EL*: is the mean value of loss:

$$EL = PD \cdot EAD \cdot LGD$$

- *Credit at risk CaR*: is the methodology applying the *VaR* approach (see Sect. 12.3) to the credit risk; the credit exposure (adjusted by recoveries) is looked upon as a random variable *X*:

  Denotation:

X credit exposure adjusted by recoveries (i.e. such a random variable that its positive values represent creditor's loss in the case of debtor's default)
f(*x*) probability density (see Sect. 26.3) of the random variable *X*

- Some measures of methodology *CaR*:

$$ECE = \int_{-\infty}^{\infty} \max(x,\ 0) \cdot f(x)\,\mathrm{d}x$$

(*expected credit exposure ECE*: is the probability-weighted credit exposure (adjusted by recoveries) with negative values replaced by zero values)

$$ECE = \frac{\sigma}{\sqrt{2\pi}}$$

(expected credit exposure for $X \sim N(0, \sigma^2)$ (see Sect. 26.3))

$$EDL \approx ECE \cdot \frac{PD}{2}$$

(*expected default loss EDL* (compare with *EL*, see *thereinbefore*): is an evaluation of the average creditor's loss caused by the debtor's default (*PD/2* is an approximate estimate of the probability that the debtor defaults when the momentary credit exposure is positive))

$$EDL \approx \frac{\sigma \cdot PD}{\sqrt{8\pi}}$$

(expected default loss in the case of debtor's default for $X \sim N(0, \sigma^2)$)

$$CaR = 1,65 \cdot \sigma$$

(*credit at risk* with confidence level 95% for $X \sim N(0,\sigma^2)$): is an estimate of the worst credit exposure that can occur with a given probability (confidence); it equals to the corresponding quantile of the credit exposure)

$$DVaR = CaR \cdot PD$$

(*default-related value at risk DVaR*: is equivalent to *VaR* for the credit risk evaluating with a given confidence the worst creditor's loss caused by the debtor's default)

Further Reading

Bluhm, C., Overbeck, L., Wagner, C.: An Introduction to Credit Risk Modeling. Chapman & Hall/CRC (2003)
Bollerslev, T.: Generalized autoregressive conditional heteroscedasticity. Journal of Econometrics 31, 307–327 (1986)
Brealey, R.A., Myers, S.C.: Principles of Corporate Finance. McGraw-Hill, New York (1988)
CreditMetrics – Technical Document. J.P. Morgan, New York (1997) (www.riskmetrics.com)
Das, S. (ed.): Risk Management and Financial Derivatives. MacMillan, Houndmills, UK (1998)
Dowd, K.: Beyond Value at Risk. Wiley, Chichester, UK (1998)
Dupacova, J., Hurt, J., Stepan, J.: Stochastic Modeling in Economics and Finance. Kluwer, Dordrecht (2002)
Engle, R.: Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50, 987–1007 (1982)
Gourieroux, C.: The Econometrics of Individual Risk. Princeton University Press, Princeton and Oxford (2007)
Jorion, P.:Value at Risk. McGraw-Hill, New York (1997)
Lawrence, D.: Measuring and Managing Derivative Market Risk. Thomson, London (1996)
McNeil, A.J., Frey, R., Embrechts, P.: Quantitative Risk Management. Princeton University Press, Princeton and Oxford (2005)
Sharpe, W.F., Alexander, G.J.: Investments. Prentice Hall, Englewood Cliffs, NJ (1990)

Chapter 13
Portfolio Analysis and CAPM Model

Abstract Chapter 13 deals with portfolio analysis including the CAPM theory: 13.1. Construction of Portfolio, 13.2. Portfolio with a Risk-Free Asset, 13.3. CAPM Model.

13.1 Construction of Portfolio

- *Portfolio*: is a set of assets (investment entries) of various types; rational investors construct their portfolios so as *to minimize the risk and to maximize the return* of investment activities; therefore the portfolio construction consists in looking for an optimal trade-off between minimum risk and maximum return values of the resulting portfolio; the rates of return are systematically looked upon as random variables with mean values and standard deviations denoted as *expected returns* and (*financial*) *risk*, respectively; typically, investment portfolios are constructed for funds of big institutional and private clients (including pension funds)
- *Theory of portfolio*: is an important part of financial theory; its basic form (Lintner, Markowitz, Sharpe, Tobin and others) makes use of abstract assumptions of *efficient market*:

  – the investors make decisions exclusively on the information based on the expected returns and covariance structure of the returns; such an information is equally available to all the investors
  – no investor can affect the returns of particular assets in the portfolio
  – the investors have different preferences for the resulting trade-off between the minimum risk and maximum return values (in particular, they have different utility functions and indifference curves, see Chap. 11)
  – the investors choose portfolios with the highest expected return among those with the same risk, and portfolios with the smallest risk among those with the same expected return (see Sect. 12.2)
  – the assets for portfolio construction are infinitely divisible and marketable
  – the investment horizon is one period of time

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_13,

– there are no transaction costs and taxes
– there exists just one *risk-free interest rate* (see Sect. 5.1), and all the investors can lend or borrow any amount of necessary funds at this risk-free interest rate

- *Portfolio construction using N risk assets* (i.e. combining suitable amounts of N assets with nonzero risk): is the basic task of portfolio theory (with preset optimality requirements); each portfolio is determined uniquely by the system of weights w_i, which give proportions of particular assets in the constructed portfolio ($i = 1, \ldots, N$: $w_i \geq 0$ (the case of short positions in some assets with $w_i < 0$ will be considered later); $w_1 + \ldots + w_N = 1$)

Denotation:

r_i rate of return of ith asset ($i = 1, \ldots, N$): is looked upon as *random variable* (see Sect. 12.2)
$\bar{r}_i$ *expected return* of ith asset: $\bar{r}_i = \mathrm{E}(r_i)$
σ_i *risk* of ith asset: $\sigma_i^2 = \sigma_{ii} = \mathrm{var}(r_i)$; $\sigma_i = \sigma_{ii} = \sigma(r_i)$
σ_{ij} covariance (see Sect. 26.6) between rates of return of ith asset and jth asset: $\sigma_{ij} = \mathrm{cov}(r_i, r_j)$
ρ_{ij} correlation coefficient (see Sect. 26.6): $\rho_{ij} = \rho(r_i, r_j) = \sigma_{ij}/(\sigma_i \cdot \sigma_j)$
$\mathbf{r}$ vector of rates of return: $\mathbf{r} = (r_1, \ldots, r_N)'$
$\mathbf{w}$ vector of weights: $\mathbf{w} = (w_1, \ldots, w_N)'$, $\mathbf{w} \geq \mathbf{0}$; $\mathbf{e}' \cdot \mathbf{w} = 1$
$\boldsymbol{\Sigma}$ covariance matrix (see Sect. 26.6): $\boldsymbol{\Sigma} = (\sigma_{ij})$
r_P rate of return of portfolio
$\bar{r}_P$ expected return of portfolio: $\bar{r}_P = \mathrm{E}(r_P)$
σ_P risk of portfolio: $\sigma_P^2 = \mathrm{var}(r_P)$

$$\bar{r}_P = w_1 \cdot \bar{r}_1 + \ \ldots \ + w_N \cdot \bar{r}_N = \mathbf{w}' \, \bar{\mathbf{r}}$$

(*expected return of portfolio*)

$$\sigma_P = \sqrt{w_1^2 \cdot \sigma_{11} + w_1 \cdot w_2 \cdot \sigma_{12} + w_1 \cdot w_3 \cdot \sigma_{13} + \ldots + w_N^2 \cdot \sigma_{NN}}$$
$$= \sqrt{\sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \sigma_{ij}} = \sqrt{\mathbf{w}' \, \boldsymbol{\Sigma} \, \mathbf{w}}$$

(*risk of portfolio*)

$$\bar{r}_P = w \cdot \bar{r}_1 + \ (1 - w) \cdot \bar{r}_2, \text{ where } 0 \leq w \leq 1$$

(expected return of portfolio for $N = 2$)

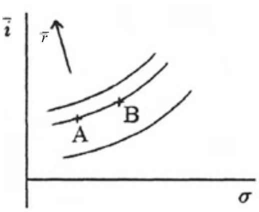
$$\sigma_P = \sqrt{w^2 \cdot \sigma_{11} + (1-w)^2 \cdot \sigma_{22} + 2 \cdot w \cdot (1-w) \cdot \sigma_{12}}$$
$$= \sqrt{w^2 \cdot \sigma_1^2 + (1-w)^2 \cdot \sigma_2^2 + 2 \cdot w \cdot (1-w) \cdot \sigma_1 \cdot \sigma_2 \cdot \rho_{12}}$$

(risk of portfolio for $N = 2$)

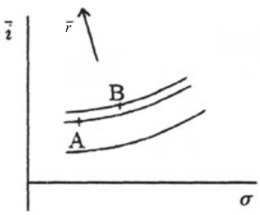
$$\sigma_P < \min(\sigma_1, \ldots, \sigma_N)$$

(*diversification of portfolio*: is a positive phenomenon from the investor's point of view, since the risk of portfolio is then smaller than the risk of each of N risk assets making up this portfolio; diversification can be achieved easier, if there are high negative correlations among particular assets in the portfolio (obviously, if the price of an asset declines, then the increasing price of another one may help, and vice versa)

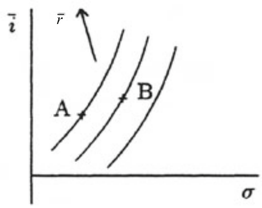
- $(\sigma, \bar{r})$-*plane*: is the plane, where the horizontal axis indicates risk as measured by the standard deviation and the vertical axis indicates reward as measured by the expected return; if the investor's utility function is a function of risk and expected return (e.g. $U(\bar{r}, \sigma; \kappa) = \bar{r} - \kappa \cdot \sigma^2$ with a parameter κ, see Chap. 11), then one can plot in this plane the corresponding *indifference curves* (see Chap. 11); a standard investor (see the assumptions of efficient market) has increasing convex indifference curves and *northwest direction of preferences*:



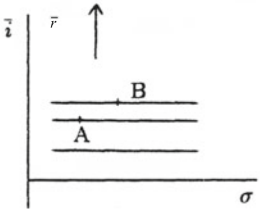
moderately risk-averse investor:
A ~ B



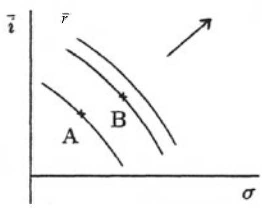
slightly risk-averse investor:
A < B



highly risk-averse investor:
A > B

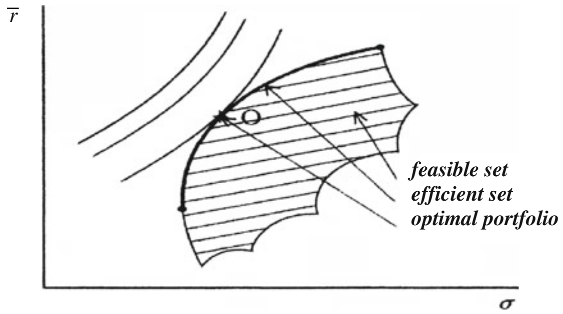


risk-neutral investor:
A < B



risk-seeking investor:
A < B

- *feasible set*: represents the set of all portfolios which can be formed from the given N assets (i.e. applying all the possible systems of weights); it has an "umbrella shape" in the $(\sigma, \bar{r})$-plane



- *efficient set* (*efficient frontier*): is a subset of the feasible set that offers maximum expected return for varying level of risk, and minimum risk for varying level of expected return; typically, it is the *northwest frontier* of the feasible set

- *Feasible portfolios*: are elements of the feasible set; they enable to construct further feasible portfolios:

  Denotation:

r_{Pj}	rate of return of jth feasible portfolio ($j = 1, \ldots, J$)
$\bar{r}_{Pj}$	expected return of jth feasible portfolio: $\bar{r}_{Pj} = \mathrm{E}(r_{Pj})$
σ_{Pj}	risk of jth feasible portfolio: $\sigma_{Pj} = \sigma(r_{Pj})$
σ_{Pij}	covariance between rates of return of ith and jth feasible portfolio
w_{Pj}	weight for jth feasible portfolio ($w_{Pj} \geq 0$; $w_{P1} + \ldots + w_{PJ} = 1$)

$$\bar{r}_P = w_{P1} \cdot \bar{r}_{P1} + \quad \ldots \quad + w_{PJ} \cdot \bar{r}_{PJ}$$

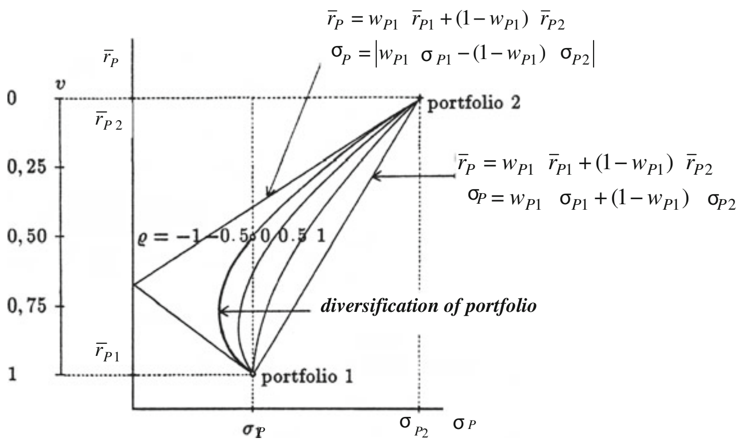
(expected return of portfolio made up of feasible portfolios)

$$\sigma_P \leq w_{P1} \cdot \sigma_{P1} + \quad \ldots \quad + w_{PJ} \cdot \sigma_{PJ}$$

(inequality for risk of portfolio made up of feasible portfolios; for $J = 2$ it implies the concavity of the curve joining two feasible points (portfolios) in the$(\sigma, \bar{r})$-plane with coordinates $(\sigma_{P1}, \bar{r}_{P1})$ and $(\sigma_{P2}, \bar{r}_{P2})$: in particular, such a curve is the abscissa for $\rho_{12} = 1$ and the broken-abscissa for $\rho_{12} = -1$)

$$\sigma_P^2 = \left(\frac{\bar{r}_P - \bar{r}_{P2}}{\bar{r}_{P1} - \bar{r}_{P2}}\right)^2 \cdot \sigma_{P1}^2 + \left(\frac{\bar{r}_P - \bar{r}_{P1}}{\bar{r}_{P1} - \bar{r}_{P2}}\right)^2 \cdot \sigma_{P2}^2 - 2 \cdot \frac{(\bar{r}_P - \bar{r}_{P1}) \cdot (\bar{r}_P - \bar{r}_{P2})}{(\bar{r}_{P1} - \bar{r}_{P2})^2} \cdot \sigma_{P12}$$

(relation between risk σ_P and expected return $\bar{r}_P$ of portfolio made up of two feasible portfolios)



- *Efficient portfolios*: are elements of the efficient set (the efficient frontier); one can construct them:

– by means of optimization methods (e.g. quadratic programming)
– in some special cases, by means of explicit formulas (see *thereinafter*) or by means of the capital market line CML (see Sect. 13.3)

Denotation:

r_0	prescribed constant rate of return for the constructed efficient portfolio
$\mathbf{r}$	vector of rates of return $\mathbf{r} = (r_1, \ldots, r_N)'$ of particular risk assets for the constructed efficient portfolio (exclusive of the case $\mathbf{r} = k \cdot \mathbf{e}$, where k is a constant)
$\bar{\mathbf{r}}$	vector of expected returns $\bar{\mathbf{r}} = (\bar{r}_1, \ldots, \bar{r}_N)'$ for vector $\mathbf{r}$
$\boldsymbol{\Sigma}$	positively definite covariance matrix of vector $\mathbf{r}$
$\mathbf{w}^*$	vector of weights $\mathbf{w}^* = (w_1{}^*, \ldots, w_N{}^*)'$ for the constructed efficient portfolio: $\mathbf{e}' \cdot \mathbf{w}^* = 1$ (negative weights, i.e. short positions in some assets, are possible for the constructed efficient portfolio)

$$\mathbf{w}^* = \delta(r_0) \cdot \mathbf{w}_1 + (1 - \delta(r_0)) \cdot \mathbf{w}_2, \text{where}$$

$$A = \mathbf{e}'\ \boldsymbol{\Sigma}^{-1}\mathbf{e}\ ; \quad B = \mathbf{e}'\ \boldsymbol{\Sigma}^{-1}\bar{\mathbf{r}} \neq 0\ ; \quad C = \bar{\mathbf{r}}'\ \boldsymbol{\Sigma}^{-1}\bar{\mathbf{r}}\ ; \quad \Delta = A \cdot C - B^2;$$

$$\mathbf{w}_1 = \frac{\boldsymbol{\Sigma}^{-1}\mathbf{e}}{A}\ ; \quad \mathbf{w}_2 = \frac{\boldsymbol{\Sigma}^{-1}\bar{\mathbf{r}}}{B}\ ; \quad \delta(r_0) = \frac{A \cdot (C - r_0 \cdot B)}{\Delta}$$

(vector of weights (negative weights are possible) for construction of *efficient portfolio with prescribed expected return*; in the excluded case $\mathbf{r} = k \cdot \mathbf{e}$, the resulting efficient portfolio may be formed by a single asset, namely by the one with minimal risk; moreover, the case $B = 0$ has also explicit solution:

$$\mathbf{w}^* = \frac{\Sigma^{-1}\mathbf{e}}{\mathbf{e}'\Sigma^{-1}\mathbf{e}} + \frac{r_0\Sigma^{-1}\bar{\mathbf{r}}}{\bar{\mathbf{r}}'\Sigma^{-1}\bar{\mathbf{r}}}\Bigg)$$

13.2 Portfolio with a Risk-Free Asset

- *Portfolio construction using N risk assets and a risk-free asset* (i.e. combining suitable amounts of *N* assets with nonzero risk and a suitable amount of one asset with zero risk): is a substantial extension of the theory from Sect. 13.1; the given risk-free asset in the portfolio can be either in the *long* position (with a positive weight), or in the *short* position (with a negative weight, which means that the investor borrows at the risk-free interest rate r_f to purchase some of the risk assets; the case of the short position in a risk asset is not usual in the basic formulation of portfolio theory)

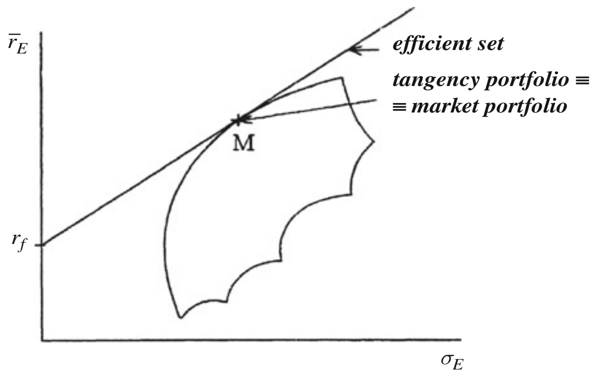
Denotation:

r_f risk-free rate of return of a risk-free asset (see Sect. 5.1): is a *nonrandom* variable (since it has the zero risk, i.e. the zero standard deviation, so that $\bar{r}_f = r_f$)

$\mathbf{w}_T$ vector of weights $\mathbf{w}_T = (w_{T1}, \ldots, w_{TN})'$ which determines the tangency portfolio (see *thereinafter*)

$\bar{r}_T$ expected return of tangency portfolio

σ_T risk of tangency portfolio



$$\begin{cases} \bar{r}_E = w \cdot r_f + (1-w) \cdot \bar{r}_T \\ \sigma_E = (1-w) \cdot \sigma_T \end{cases}$$

(*tangency portfolio*: an arbitrary *efficient* portfolio with expected return $\bar{r}_E$ and risk σ_E constructed of *N* risk assets and a risk-free asset can be alternatively made up of this risk-free asset included with a weight w (w may be negative, see *thereinbefore*) and of the so called tangency portfolio; the tangency portfolio is made up entirely of the considered risk assets and does not vary with varying efficient portfolios (only its weight $1 - w$ varies in such a case); this mechanism is sometimes called *separation theorem*, since the effect of the risk assets is separated just to the invariant tangency portfolio; the attribute "tangency" reminds the fact that the set of efficient portfolios (efficient frontier in the case of N risk assets combined with a risk-free asset) coincides with the tangent launched from the risk-free asset point in the $(\sigma, \bar{r})$-plane to the feasible set made up only of the risk assets (i.e. without the risk-free asset): the corresponding adherent point is just the tangency portfolio)

$$\alpha \cdot \boldsymbol{\Sigma} \cdot \mathbf{w}_T' = \bar{\mathbf{r}} - r_f \cdot \mathbf{e}$$

(relation for construction of tangency portfolio: if all the weights in the vector $\mathbf{w}_t$ which determines the tangency portfolio are positive, then there exists a constant $\alpha > 0$, for which these weights fulfil the given system of linear equations)

Denotation:

- r_0 — prescribed constant rate of return for the constructed efficient portfolio, if a risk-free asset is also used for the construction
- $\mathbf{r}$ — vector of rates of return $\mathbf{r} = (r_1, \ldots, r_N)'$ of particular risk assets for the constructed efficient portfolio (exclusive of the case $\mathbf{r} = k \cdot \mathbf{e}$, where k is a constant)
- $\bar{\mathbf{r}}$ — vector of expected returns $\bar{\mathbf{r}} = (\bar{r}_1, \ldots, \bar{r}_N)'$ for vector $\mathbf{r}$
- $\boldsymbol{\Sigma}$ — positively definite covariance matrix of vector $\mathbf{r}$
- $\mathbf{w}^*$ — vector of weights $\mathbf{w}^* = (w_0{}^*, w_1{}^*, \ldots, w_N{}^*)'$ for the constructed efficient portfolio: $\mathbf{e}' \cdot \mathbf{w}^* = 1$ (negative weights, i.e. short positions in some assets, are possible for the constructed efficient portfolio)

$$\mathbf{w}^* = \delta(r_0) \cdot \mathbf{w}_1 + (1 - \delta(r_0)) \cdot \mathbf{w}_2,$$

where $A = \mathbf{e}'\, \boldsymbol{\Sigma}^{-1} \mathbf{e}\,; \quad B = \mathbf{e}'\, \boldsymbol{\Sigma}^{-1} \bar{\mathbf{r}}\,; \quad r_f < \dfrac{B}{A}; \quad C = \bar{\mathbf{r}}'\, \boldsymbol{\Sigma}^{-1} \bar{\mathbf{r}}\,;$

$$\mathbf{w}_1 = (1,\ 0,\ \ldots,\ 0)'\,; \quad \mathbf{w}_2 = (0,\ \mathbf{w}_t')'\,; \quad \mathbf{w}_t = \frac{\boldsymbol{\Sigma}^{-1}(\mathbf{r} - r_f \cdot \mathbf{e})}{B - A \cdot r_f};$$

$$\delta(r_0) = 1 - \frac{(r_0 - r_f) \cdot (B - A \cdot r_f)}{A \cdot r_f^2 - 2B \cdot r_f + C}$$

(vector of weights (negative weights are possible) for the construction of *efficient portfolio with prescribed expected return*, if a *risk-free asset* is also used for the construction; the assumption $r_f < B/A$ concerns the fact that B/A is equal to the expected return of the (global) minimum-variance portfolio made up of risk assets)

- *Market portfolio M*: denotes a hypothetical portfolio with expected return $\bar{r}_M$ and risk σ_M, which would consist of all the investment assets, where the proportion to be invested in each asset corresponds to its relative market value (i.e. to the aggregate market value of the asset divided by the sum of the aggregate market values of all the investment assets); under standard assumptions of the efficient market (see Sect. 13.1), the market portfolio coincides with the tangency portfolio (if the risk-free asset is not a component of the market portfolio); in practice, the market portfolio is usually approximated by a suitable market indicator (e.g. a stock exchange indicator, see Sect. 29.3)

13.3 CAPM Model

- *Capital asset pricing model CAPM*: is the basic model of the contemporaneous investment theory, since it relates the analyzed investment strategy to the market portfolio (i.e. in fact, to a suitable market indicator, see Sect. 29.3) and to a risk-free asset (i.e. actually, to government securities, see Sect. 9.1)

Denotation:

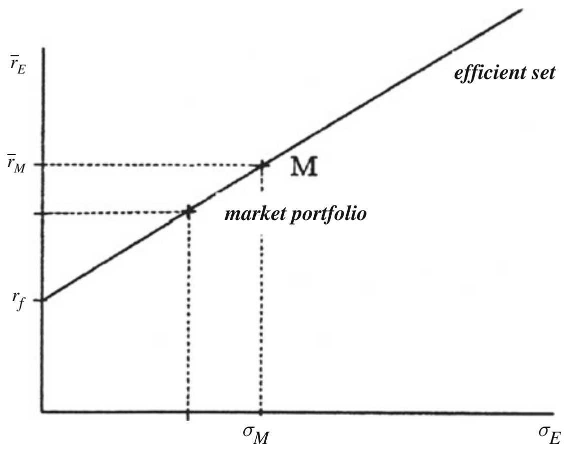
$\bar{r}_E, \sigma_E$	expected return and risk of efficient portfolio
$\bar{r}_T, \sigma_T$	expected return and risk of tangency portfolio
$\bar{r}_M, \sigma_M$	expected return and risk of market portfolio
r_f	risk-free rate of return
$\bar{r}_i, \sigma_i$	expected return and risk of ith risk asset ($i = 1, \ldots, N$)
σ_{iM}	covariance (see Sect. 26.6) between rate of return of ith risk asset and rate of return of market portfolio: $\sigma_{iM} = \text{cov}(r_i, r_M)$
$\bar{r}_R$	expected return of portfolio made up of risk assets
σ_{RM}	covariance (see Sect. 26.6) between rate of return of portfolio made up of risk assets and rate of return of market portfolio: $\sigma_{RM} = \text{cov}(r_R, r_M)$

$$\bar{r}_E = r_f + \frac{\bar{r}_M - r_f}{\sigma_M} \cdot \sigma_E$$

(*capital market line CML*: concerns an arbitrary efficient portfolio)

$$\frac{\bar{r}_E - r_f}{\sigma_E} = \frac{\bar{r}_M - r_f}{\sigma_M} = \frac{\bar{r}_T - r_f}{\sigma_T}$$

(*market price of risk*, *Sharpe's measure of portfolio*: is the same for an arbitrary efficient portfolio (i.e. also for the market or tangency portfolio))



$$\frac{\bar{r}_i - r_f}{\sigma_i} \le \frac{\bar{r}_E - r_f}{\sigma_E}, i = 1, \ldots, N$$

(market price of risk for *risk asset*: cannot exceed the market price of risk for an arbitrary efficient portfolio (i.e. also for the market or tangency portfolio))

$$\bar{r}_i - r_f = (\bar{r}_M - r_f) \cdot \beta_i, \quad \text{where } \beta_i = \frac{\sigma_{iM}}{\sigma_M^2}$$

(*security market line SML*: is suitable for financial analysis and pricing not only of particular risk assets (mostly stocks), but also for analysis of the whole enterprise branches; β_i are called *factors beta*, and one publishes them in financial periodical titles (so called Beta Books) for particular companies and branches; they are very important when evaluating exposure of capital: assets with $\beta > 1$ (or $\beta < 1$) are more risky (or less risky) than the market average, respectively (efficient portfolios including the market portfolio have $\beta = 1$, risk-free assets have $\beta = 0$)

$$\bar{r}_R - r_f = (\bar{r}_M - r_f) \cdot \beta_R, \quad \text{where } \beta_R = \frac{\sigma_{RM}}{\sigma_M^2}$$

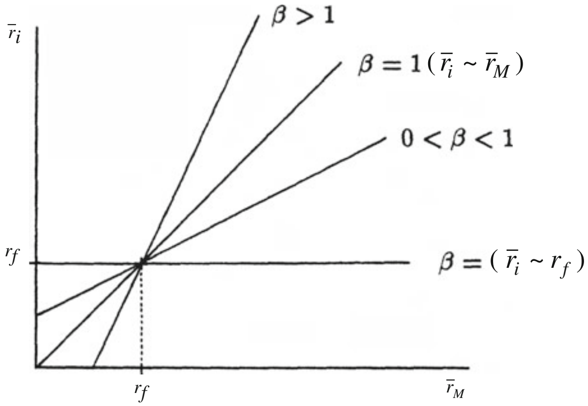
(generalization of SML for an arbitrary portfolio constructed of risk assets)

$$r_{it} - r_{ft} = \alpha_i + (r_{Mt} - r_{ft}) \cdot \beta_i + \varepsilon_{it}, t = 1, \ldots, T$$

(linear regression model (see Sect. 27.11) for estimation of factor beta (see *therein-before*) and *factor alpha* (*factor of disequilibrium*) of ith risk asset: one substitutes observed values ($t = 1, \ldots, T$) for the rate of return of the ith risk asset r_i, the risk-free asset r_f and the market portfolio r_M (ε_i is the residual of the model); the factor alpha of the ith risk asset represents the difference between the expected return according to the observed reality and the (equilibrium) expected return according to the CAPM theory: if the estimated alpha is significantly positive or negative, then the given risk asset produces returns that are over or below the appropriate values following from the theory, so that the asset seems to be at the market underestimated of overestimated, respectively)

$$\sigma_i^2 = \beta_i^2 \cdot \sigma_M^2 + \sigma_{\varepsilon_i}^2$$

(decomposition of (squared) risk of *i*th risk asset $\sigma_i{}^2$ to *market risk* $\beta_i{}^2 \cdot \sigma_M{}^2$ and *specific* (*unique*) *risk* $\sigma_{\varepsilon_i}^2$: the market risk is also denoted as *systematic risk* and the specific risk as *nonsystematic risk*, since only the specific (nonsystematic) risk is diversifiable in the sense that by holding a risk asset in a sufficiently large portfolio, the prevailing part of the risk of the whole portfolio is that of market (systematic) risk; while a higher systematic risk is compensated by a higher expected return (according to the relation SML due to the higher factor beta, see *thereinbefore*), it is not the case for nonsystematic risk; in practice, the systematic and nonsystematic risks are key concepts taken into account in the framework of the capital adequacy)



Further Reading

Brealey, R.A., Myers, S.C.: Principles of Corporate Finance. McGraw-Hill, New York (1988)
Dupacova, J., Hurt, J., Stepan, J.: Stochastic Modeling in Economics and Finance. Kluwer, Dordrecht (2002)
Elton, E.J., Gruber, M.J.: Modern Portfolio Theory and Investment Analysis. Wiley, New York (1991)
Ingersoll, J.E.: Theory of Financial Decision Making. Rowman & Littlefield, Savage (1987)
Markowitz, H.M.: Portfolio selection. Journal of Finance 6, 77–91 (1952)
Sharpe, W.F., Alexander, G.J.: Investments. Prentice Hall, Englewood Cliffs, NJ (1990)

Chapter 14
Arbitrage Theory

Abstract Chapter 14 is an introduction to basic formulas of arbitrage theory.

- *Arbitrage opportunity*: is a possibility of a risk-free profit through simultaneous long and short positions in suitable investment assets that ensure free of risk a higher rate of return than the risk-free one (on the other hand, the arbitrage opportunity enabled by different prices of assets at different places at the same time is almost impossible in the contemporary financial world); the characterization as "money pump" is sometimes used
- *Arbitrage-free principle*: is a hypothetical situation that excludes any arbitrage opportunity ("there is no such thing like a free lunch"); this principle is a theoretical background of contemporary approach to mathematical pricing of some assets (e.g. financial derivatives, see Chap. 10)

Denotation:

$\mathbf{S}(t)$	vector $\mathbf{S}(t) = (S_1(t),\ldots, S_N(t))'$ of prices of unit amounts for N investment assets at time t
$\mathbf{x}(t)$	vector $\mathbf{x}(t) = (x_1(t),\ldots, x_N(t))'$ of numbers of unit amounts for N investment assets at time t
k	index indicating one of K possible states of financial market ($k = 1, \ldots, K$)
$\mathbf{D}(t+1)$	payoff matrix $\mathbf{D}(t+1) = (d_{ik}(t+1))$ of type $N \times K$: $d_{ik}(t+1)$ is the payoff at time $t+1$ for an investor that held the unit amount of the ith investment asset ($i = 1, \ldots, N$) from time t to time $t+1$ under the condition that the financial market was during this period at the state k ($k = 1, \ldots, K$); the first one of the given assets ($i = 1$) is a risk-free asset with the rate of return r_f (see Sects. 5.1 and 13.2: therefore the first row of the matrix $\mathbf{D}(t+1)$ contains identical elements $1 + r_f$, since the payoff for the investor that held the risk-free asset does not depend on the state of financial market during the holding period)

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_14,

$$\mathbf{S}(t)' \cdot \mathbf{x}(t) \text{ (price of investment portfolio at time } t)$$

$$\mathbf{D}(t+1)' \cdot \mathbf{x}(t)$$

(vector of payoffs at time t+1 for particular states of financial market k (k = 1, ..., K), if the investor held the amounts $\mathbf{x}(t)$ of investment assets from time t to time t +1)

$$\mathbf{S}(t)' \cdot \mathbf{x}(t) \leq 0 \text{ and } \mathbf{D}(t+1)' \cdot \mathbf{x}(t) > \mathbf{0}$$

(*arbitrage opportunity of the first type*: with no cost at time t, profits are achieved in all states at time t +1 (see **List of Symbols** for the meaning of inequalities between vectors))

$$\mathbf{S}(t)' \cdot \mathbf{x}(t) < 0 \text{ and } \mathbf{D}(t+1)' \cdot \mathbf{x}(t) \geq \mathbf{0}$$

(*arbitrage opportunity of the second type*: with a negative cost at time t, the investment is not loss-making at time t +1 (the unspecified term *arbitrage opportunity* means the arbitrage opportunity of the first or second type in this text))

$$\mathbf{S}(t)' \cdot \mathbf{x}(t) = 0 \quad \text{and} \quad P(\mathbf{D}(t+1)' \cdot \mathbf{x}(t) \geq \mathbf{0}) = 1 \quad \text{and} \quad \mathrm{E}\,(\mathbf{D}(t+1)' \cdot \mathbf{x}(t)) > 0$$

(*arbitrage opportunity through probability concepts*: the last condition can be replaced by an equivalent one: $P(\mathbf{D}(t+1)' \cdot \mathbf{x}(t) > \mathbf{0}) > 0$)

$$\begin{pmatrix} S_1(t) \\ S_2(t) \\ \vdots \\ S_N(t) \end{pmatrix} = \begin{pmatrix} 1+r_f & 1+r_f & \ldots & 1+r_f \\ d_{21}(t+1) & d_{22}(t+1) & \ldots & d_{2K}(t+1) \\ \vdots & \vdots & \vdots & \vdots \\ d_{N1}(t+1) & d_{N2}(t+1) & \cdots & d_{NK}(t+1) \end{pmatrix} \cdot \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix}$$

(*iff condition of nonexistence of arbitrage opportunities* (see *thereinbefore*): there exist positive numbers $\psi_1, \ldots, \psi_N$ so that the given relation is fulfilled)

$$\begin{pmatrix} 1 \\ S(t) \end{pmatrix} = \begin{pmatrix} 1+r_f & 1+r_f \\ S_1(t+1) & S_2(t+1) \end{pmatrix} \cdot \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$$

(special case of the "iff condition" of nonexistence of arbitrage opportunities for $N = 2$ (i.e. only two investment assets are used: the first of them is the risk-free one, and its amount at time t is normed to unit) and $K = 2$ (i.e. only two states of the financial market are possible, where the price per unit amount of the second (risk) asset at time t+1 is denoted $S_1(t+1)$ at the first state and $S_2(t+1)$ at the second state): there exist positive numbers ψ_1 and ψ_2 so that the given relation is fulfilled)

$$r_1 < 1 + r_f < \ r_2,$$

where $\quad r_1 = \dfrac{S_1(t+1)}{S(t)}; \quad r_2 = \dfrac{S_2(t+1)}{S(t)}; \quad S_1(t+1) < S_2(t+1)$

(special case of the "iff condition" of nonexistence of arbitrage opportunities for $N = 2$ and $K = 2$ (see *thereinbefore*))

$$S(t) = \frac{1}{1+r_f} \cdot \mathrm{E}_{\tilde{P}}\left(S(t+1)\right) = \frac{1}{1+r_f} \cdot \left(\tilde{P}_1 \cdot S_1(t+1) + \tilde{P}_2 \cdot S_2(t+1)\right),$$

where $\tilde{P}_i = (1 + r_f) \cdot \psi_i$ (for $i = 1, 2$) is the so-called *risk-neutral probability* ($\tilde{P}_i > 0,\ \tilde{P}_1 + \tilde{P}_2 = 1$)

(*present price of risk asset* (under the condition of nonexistence of arbitrage opportunities) expressed as the *discount mean value of future payoffs of risk asset* using the *risk-free rate of return for discounting* and the *risk-neutral probability for mean value calculation*; it is the model situation when evaluating some assets (e.g. financial derivatives, see Sect. 15.8) or when evaluating options and guaranteed values in insurance contracts (see Sects. 18.1 and 19.6))

$$\tilde{P}_1 \cdot r_1 + \tilde{P}_2 \cdot r_2 = 1 + r_f, \text{ where } \tilde{P}_i = (1+r_f) \cdot \psi_i;\ r_i = \frac{S_i(t+1)}{S(t)}(i = 1, 2)$$

(*the expected return of risk asset calculated by means of risk-neutral probability* (see *thereinbefore*) *coincides with the risk-free rate of return* under the condition of nonexistence of arbitrage opportunities)

$$\mathrm{E}_{\tilde{P}}\left(X(t+s) \mid I_t\right) = X(t), \quad \text{where} \quad X(t+s) = \frac{1}{(1+r_f)^s} \cdot S(t+s),\ s \geq 0$$

(normed price $X(t + s)$ of risk asset is the *martingale* (see Sect. 26.9) with respect to the *risk-neutral probability*: the conditional mean value (see Sect. 26.8) of this normed price $X(t + s)$ at time $t + s$ with respect to the total of information I_t on the given asset observed up to time t is namely equal to the price $X(t) = S(t)$ at time t; the martingales are important instruments of financial stochastic analysis, see Sect. 15.8)

- *Arbitrage pricing model APM*: is the basic model of the so-called *arbitrage pricing theory* (*APT theory*); it is a multifactor model (the model CAPM in Sect. 13.3 as a one-factor model, where the corresponding factor is a suitable market indicator, is its special case); moreover, an assumption that there are no arbitrage opportunities at the balanced market is usually assumed; the practical applications of the APM make use of the regression analysis (see Sect. 27.11) and the factor analysis.

Further Reading

Brealey, R.A., Myers, S.C.: Principles of Corporate Finance. McGraw-Hill, New York (1988)

Dupacova, J., Hurt, J., Stepan, J.: Stochastic Modeling in Economics and Finance. Kluwer, Dordrecht (2002)

Ross, S.A.: Arbitrage theory of capital asset pricing. Journal of Economic Theory 13, 341–360 (1976)

Sears, R.S., Trennepohl, G.L.: Investment Management. Dryden, Forth Worth (1993)

Sharpe, W.F., Alexander, G.J.: Investments. Prentice Hall, Englewood Cliffs, NJ (1990)

Chapter 15
Financial Stochastic Analysis

Abstract Chapter 15 deals with formulas of stochastic calculus: 15.1. Wiener Process in Finance, 15.2. Poisson Process in Finance, 15.3. Ito Stochastic Integral, 15.4. Stochastic Differential Equations SDE, 15.5. Ito's Lemma, 15.6. Girsanov Theorem on Equivalent Martingale Probability, 15.7. Theorem on Martingale Representation, 15.8. Derivatives Pricing by Means of Equivalent Martingale Probabilities, 15.9. Derivatives Pricing by Means of Partial Differential Equations PDE, 15.10. Term Structure Modeling.

15.1 Wiener Process in Finance

$$\{W_t, t \geq 0\}, \text{ where (also see Sect. 30.4)}$$

$$\begin{cases} (i) & W_0 = 0 \\ (ii) & \text{trajectories of the process are continuous in time} \\ (iii) & W_{t_2} - W_{t_1}, \ \ldots, \ W_{t_n} - W_{t_{n-1}} \text{ are independent for arbitrary } 0 \leq t_1 < \ldots < t_n \\ (iv) & W_t - W_s \sim N(0, |s-t|) \text{ with respect to probability } P \text{ for arbitrary } s, \ t \geq 0 \end{cases}$$

(*Wiener process* (*Brownian motion*) with respect to probability P: is a homogeneous Markov process (i.e. in continuous time) with continuous states on a probability space with probability P (see Sect. 30.3))

$$\{W_t, t \geq 0\}, \text{ where}$$

$$\begin{cases} (i) & W_0 = 0 \\ (ii) & \text{trajectories of the process are continuous in time} \\ (iii) & W_t \text{ is square integrable martingale with respect to } \Im_t \text{ and probability } P \\ (iv) & E\left((W_t - W_s)^2\right) = |s-t| \text{ for arbitrary } s, \ t \geq 0 \end{cases}$$

(equivalent definition of Wiener process as a continuous martingale on a pro-bability space with probability P (see Sect. 26.9): $\Im_t$ are σ-algebras of the *filtration*, to

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_15,

which the martingale is *adapted* ($\Im_t$ can be interpreted as a summary information on the past and present of the process at time t: if one knows the information coded in $\Im_t$ then one also exactly knows the value W_t, which becomes conditionally on $\Im_t$ an observable nonrandom value; if the filtration is altered then the considered Wiener process is altered as well); in particular according to martingale properties, W_t has unpredictable increments; the normality of the increments in this equivalent definition of Wiener process is guaranteed by Lévy theorem)

$$\mathrm{cov}(W_s, W_t) = \min(s, t)$$

- Some properties of Wiener process (the properties of trajectories hold only with probability one):
  - trajectories (see Sect. 30.1) of the process are nowhere differentiable (i.e. these functions of time have nowhere their first derivatives, although they are continuous)
  - trajectories of the process hit any and every real value no matter how large, or how negative
  - the process has a fractal form (i.e. if one chooses a trajectory it looks just the same in any scale)
  - the process after suitable transformations (e.g. in order to achieve a necessary trend, volatility and non-negativity) can be used to model *continuous* motions of interest rates, exchange rates and prices of financial assets (see Sect. 15.8); in the case of *jumps* in the motion one must combine Wiener process with Poisson process (see Sect. 15.2)

$$\{X_t = \mu \cdot t + \sigma \cdot W_t,\ t \geq 0\}$$

(Wiener process with *drift* (*trend parameter*) μ and *volatility* (*diffusion parameter*) σ, i.e. $\mu \cdot t$ is a *trend component* and $\sigma \cdot W_t$ is a *diffusion component* of X_t: in general the drift μ and the volatility σ can be stochastic processes (see Sect. 15.3); $\mathrm{E}(X_t) = \mu \cdot t$ and $\mathrm{var}(X_t) = \sigma^2 \cdot t$)

$$\{S_t = \mathrm{e}^{X_t} = \mathrm{e}^{\mu \cdot t + \sigma \cdot W_t},\ t \geq 0\}$$

(*exponential* Wiener process or *geometric* Brownian motion with *logarithmic drift* μ and *logarithmic volatility* σ: these are namely “linear drift” and “linear volatility” of the process $\ln S_t$; in general, the drift μ and the volatility σ can be stochastic processes (see Sect. 15.3); $\mathrm{E}(S_t) = \exp(\mu_t + \sigma_t^2/2)$ and $\mathrm{var}(S_t) = \exp[(2\mu_t + \sigma_t^2)\cdot t]\cdot[\exp(\sigma_t^2 \cdot t) - 1]$; the exponential transformation is applied frequently in order to achieve the non-negativity of the process)

15.2 Poisson Process in Finance

$$\{N_t, t \geq 0\},\ \text{where (also see Sect. 30.4)}$$

$$\begin{cases} (i) \quad N_0 = 0 \\ (ii) \quad N_{t_2} - N_{t_1},\ ...,\ N_{t_n} - N_{t_{n-1}} \text{ are independent for arbitrary } 0 \leq t_1 < ... < t_n \\ (iv) \quad N_t - N_s \sim P\left(\lambda \cdot (t-s)\right) \text{ for arbitrary } 0 \leq s < t \end{cases}$$

Poisson process with *intensity* $\lambda > 0$: is a homogeneous Markov process, i.e. in continuous time, with discrete state space $S = \mathrm{N}_0 = \{0, 1, \ldots\}$ (see Sect. 30.3), where

$$p_{ij}(h) = \begin{cases} \lambda \cdot h + o(h) & \text{for } j = i+1 \\ 1 - \lambda \cdot h + o(h) & \text{for } j = i \\ o(h) & \text{for } j > i+1 \\ 0 & \text{for } j < i \end{cases};$$

N_t at time $t \geq 0$ gives the number of occurrences of a random event in time interval $\langle 0,\ t\rangle$ and is suitable to model random variables developing in jumps, i.e. to model so called *rare events* which are e.g. stock exchange crashes in finance; it holds

$$p_i(t) = P(N_t = i) = e^{-\lambda \cdot t}\frac{(\lambda \cdot t)^i}{i!},\ t \geq 0;\ i \in \mathrm{N}_0$$

(see Poisson distribution in Sect. 26.4 with mean value $\lambda \cdot t$); the lengths of intervals between particular occurrences of the given event are *iid* random variables with distribution $Exp(\lambda)$ (see exponential distribution in Sect. 26.5 with mean value $1/\lambda$ and variance $1/\lambda$); it holds $\mathrm{E}(N_t) = \lambda \cdot t$ and $\mathrm{var}(N_t) = \lambda \cdot t$; in general, one can have an intensity λ_t depending on time)

$$\{M_t = N_t - \lambda \cdot t,\ t \geq 0\}$$

(*compensated* Poisson process with intensity $\lambda > 0$: is a quadratically integrable martingale continuous from the right (see Sect. 26.9), i.e. in particular, it has unpredictable increments; it holds $\mathrm{E}(M_t) = 0$ and $\mathrm{var}(M_t) = \lambda \cdot t$)

- Comparison of Wiener and Poisson processes from the point of view of financial applications:
  - increments of Wiener process $W_{t+h} - W_t$ (for $h \to 0_+$ denoted as $\mathrm{d}W_t$, see Sect. 15.3) describe "*regular events of insignificant size*" (for small h), while increments of Poisson process $N_{t+h} - N_t$ (for $h \to 0_+$ denoted as $\mathrm{d}N_t$) describe "*irregular* (*rare*) *events of significant size* (*jumps*)"

– Wiener process is suitable to model *continuous* motions of interest rates, exchange rates and prices of financial assets, while Poisson process (mostly in combination with Wiener process) is suitable to model their motions with *jumps*

15.3 Ito Stochastic Integral

$$\int_0^T \sigma_t \, \mathrm{d}W_t, \text{ where } \sigma_t \text{ is a nonanticipating stochastic process; } \mathrm{E}\left(\int_0^T \sigma_t^2 \, \mathrm{d}t\right) < \infty,$$
$$0 \leq T < \infty$$

(*Ito stochastic integral*: is an analogy of Riemann-Stieltjes integral with a "random function" W_t (i.e. Ito integral is a random variable); values σ_t can be also random (but in this case independent on the future) and must not be "explosive" (in square) in the given interval; *nonanticipativity* (or also *previsibility*) at time t is independence on the future in the sense of measurability of σ_t with respect to the past and present information included in $\Im_t$, see Sect. 15.1; in practice, $\sigma_t = \sigma(S_t, t)$ is often the volatility of a given financial variable S_t for a suitable deterministic function $\sigma(\cdot\,,\cdot)$; in particular, if $\sigma_t = \sigma(W_t, t)$ for a continuous function $\sigma(\cdot\,,\cdot)$ then Ito stochastic integral exists)

Assumption: Each Ito stochastic integral given in the following text exists

$$\sum_{k=1}^{n} \sigma_{t_{k-1}} \cdot (W_{t_k} - W_{t_{k-1}}) \xrightarrow{L_2} \int_0^T \sigma_t \, \mathrm{d}W_t,$$
$$\text{i.e. } \lim_{n\to\infty} \mathrm{E}\left\{\left[\sum_{k=1}^{n} \sigma_{t_{k-1}} \cdot (W_{t_k} - W_{t_{k-1}}) - \int_0^T \sigma_t \, \mathrm{d}W_t\right]^2\right\} = 0$$

(definition of Ito stochastic integral: this integral is the limit in the mean square (see Sect. 26.14) of the areas of approximate Riemann-Stieltjes rectangles of the form $\sigma_{t_{k-1}} \cdot (W_{t_k} - W_{t_{k-1}})$ for finer and finer partitions $0 = t_0 < t_1 < \ldots < t_n = T$ of the integration interval (the limit value remains the same for arbitrarily chosen partitions); one applies sometimes more generally the limit in probability (see Sect. 26.14) instead of the limit in the mean square)

$$\int_0^T (a_{1t} \cdot \sigma_{1t} + a_{2t} \cdot \sigma_{2t}) \, \mathrm{d}W_t = a_{1t} \cdot \int_0^T \sigma_{1t} \, \mathrm{d}W_t + a_{2t} \cdot \int_0^T \sigma_{2t} \, \mathrm{d}W_t,$$

where a_{1t} and a_{2t} are deterministic functions (*linearity*)

$$\int_s^u \sigma_t \,\mathrm{d}W_t = \int_0^T I_{\langle s,\,u\rangle}(t)\cdot\sigma_t \,\mathrm{d}W_t;\quad \int_s^u \mathrm{d}W_t = W(u) - W(s),\ \text{where } 0 \le s < u \le T$$

$$\mathrm{E}\left(\int_0^T \sigma_t \,\mathrm{d}W_t\right) = 0;\ \text{in particular}\ \mathrm{E}\left(\int_0^T W_t \,\mathrm{d}W_t\right) = 0$$

$$\mathrm{E}\left(\int_0^T \sigma_t \,\mathrm{d}W_t\right)^2 = \mathrm{E}\left(\int_0^T \sigma_t^2 \,\mathrm{d}W_t\right) = \mathrm{E}\left(\int_0^T \sigma_t^2 \,\mathrm{d}t\right)$$

$$= \int_0^T \mathrm{E}(\sigma_t^2)\,\mathrm{d}t,\ \text{if} \int_0^T \mathrm{E}(\sigma_t^2)\,\mathrm{d}t < \infty$$

$$\mathrm{E}\left(\int_0^T \sigma_{1t} \,\mathrm{d}W_t \cdot \int_0^T \sigma_{2t} \,\mathrm{d}W_t\right) = \mathrm{E}\left(\int_0^T \sigma_{1t}\cdot\sigma_{2t} \,\mathrm{d}W_t\right) = \mathrm{E}\left(\int_0^T \sigma_{1t}\cdot\sigma_{2t} \,\mathrm{d}t\right)$$

$$= \int_0^T \mathrm{E}(\sigma_{1t}\cdot\sigma_{2t})\,\mathrm{d}t,\ \text{if} \int_0^T \mathrm{E}(\sigma_{1t}\cdot\sigma_{2t})\,\mathrm{d}t < \infty$$

$$X_t = \int_0^t \sigma_s \,\mathrm{d}W_s,\ \text{where } 0 \le t \le T$$

(stochastic process X_t defined by means of Ito stochastic integral is the quadratically integrable continuous martingale (see Sect. 26.9) with respect to $\Im_t$)

$$\int_0^t W_s \,\mathrm{d}W_s = \frac{1}{2}(W_t^2 - t)\ ,\ \text{where } 0 \le t \le T$$

$$\{X_t = \int_0^t \mu_s \mathrm{d}s + \int_0^t \sigma_s \,\mathrm{d}W_s,\ t \ge 0\}$$

(Wiener process with *drift* μ_t and *volatility* σ_t (also see Sect. 15.1): drift μ_t and volatility σ_t are nonanticipating stochastic processes (see *thereinbefore*) and have often the form $\mu_t = \mu\ (X_t, t)$ and $\sigma_t = \sigma\ (X_t, t)$ for suitable deterministic functions $\mu\ (\cdot\ ,\ \cdot)$ and $\sigma\ (\cdot\ ,\ \cdot)$)

$$\{S_t = \mathrm{e}^{X_t} = \exp\left(\int_0^t \mu_s \mathrm{d}s + \int_0^t \sigma_s \,\mathrm{d}W_s\right)\ ,\ t \ge 0\}$$

(exponential Wiener process or geometric Brownian motion with *logarithmic drift* μ_t and *logarithmic volatility* σ_t (also see Sect. 15.1))

$\int_0^T \sigma_t \, \mathrm{d}M_t$, where $M_t = N_t - \lambda \cdot t$ is a compensated Poisson process (see Sect. 15.2)

(Ito stochastic integral with respect to process M_t: definition and assumptions are the same as for $\int_0^T \sigma_t \, \mathrm{d}W_t$; the martingale $X_t = \int_0^t \sigma_s \, \mathrm{d}M_s$ is continuous only from the right in this case since it moves in jumps)

15.4 Stochastic Differential Equations SDE

Denotation and assumptions:

$\{\Im_t\,, t \geq 0\}$	filtration (see Sect. 15.1)
$\{X_t\,, t \geq 0\}$	stochastic process (more generally, $0 \leq t \leq T$ with possibility $T = \infty$)
μ_t and σ_t	nonanticipating stochastic processes (see Sect. 15.3, in particular, one can put $\mu_t = \mu\,(X_t, t)$ and $\sigma_t = \sigma\,(X_t, t)$ for suitable deterministic functions $\mu\,(\cdot\,,\cdot)$ and $\sigma\,(\cdot\,,\cdot)$) such that $\int_0^t \lvert\mu_s\rvert \, \mathrm{d}s < \infty$ and $\int_0^t \sigma_s^2 \, \mathrm{d}s < \infty$ with probability one for arbitrary $0 \leq t \leq T$
$\mathrm{d}X_t = \mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t$	*integral* relation $X_t = X_0 + \int_0^t \mu_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s$ (see Sect. 15.3) rewritten as *differential* one with *stochastic differentials* $\mathrm{d}X_t$ and $\mathrm{d}W_t$

$$X_{t+h} - X_t = \int_t^{t+h} \mu_s \mathrm{d}s + \int_t^{t+h} \sigma_s \mathrm{d}W_s \approx \mu_t \cdot h + \sigma_t \cdot (W_{t+h} - W_t)$$

(relation $\mathrm{d}X_t = \mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t$ rewritten in this approximate form is suitable for models of interest rates (but not for models of market prices of stock))

$$\mathrm{d}X_t = \mu(X_t, t)\,\mathrm{d}t + \sigma(X_t, t)\,\mathrm{d}W_t$$

((*Ito*) *stochastic differential equation* (SDE): its solution is a stochastic process X_t which under initial condition X_0 fulfils relation)

$$X_t = X_0 + \int_0^t \mu(X_s, s)\mathrm{d}s + \int_0^t \sigma(X_s, s)\mathrm{d}W_s\Big)$$

$$\mathrm{d}X_t = \mu \, \mathrm{d}t + \sigma \, \mathrm{d}W_t$$

(SDE with *linear constant coefficients* and initial condition $X_0 = 0$: $\rightarrow$ its solution is $X_t = \mu \cdot t + \sigma \cdot W_t$)

$$\mathrm{d}X_t = \mu \cdot X_t \mathrm{d}t + \sigma \cdot X_t \mathrm{d}W_t \quad \text{or equivalently} \quad \frac{\mathrm{d}X_t}{X_t} = \mu \ \mathrm{d}t + \sigma \ \mathrm{d}W_t$$

(*geometric* SDE with constant coefficients and initial condition $X_0 = 1$: it is applied frequently to model development of relative changes of financial variables (e.g. development of a rate of return if the price is X_t):

$$\rightarrow \text{ its solution is } X_t = \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right) \cdot t + \sigma \cdot W_t\right\}\Bigg)$$

$$\mathrm{d}X_t = \sigma \cdot X_t \mathrm{d}W_t \quad \text{or equivalently} \quad \frac{\mathrm{d}X_t}{X_t} = \sigma \ \mathrm{d}W_t$$

(special case of geometric SDE with $\mu = 0$ and initial condition $X_0 = 1$: relative changes of a financial variable X_t are calibrated increments of Wiener process:

$$\rightarrow \text{ its solution is } X_t = \exp\left\{-\frac{1}{2}\sigma^2 \cdot t + \sigma \cdot W_t\right\}\Bigg)$$

$$\mathrm{d}X_t = \left(\mu + \frac{1}{2}\sigma^2\right) \cdot X_t \mathrm{d}t + \sigma \cdot X_t \mathrm{d}W_t \quad \text{or equivalently} \quad \frac{\mathrm{d}X_t}{X_t} = \left(\mu + \frac{1}{2}\sigma^2\right) \mathrm{d}t + \sigma \ \mathrm{d}W_t$$

(special case of geometric SDE with initial condition $X_0 = 1$: the natural logarithm of its solution is Wiener process with drift μ and volatility σ (see Sect. 15.1):

$$\rightarrow \text{ its solution is } X_t = \exp\{\mu \cdot t + \sigma \cdot W_t\})$$

15.5 Ito's Lemma

Denotation and assumptions:

$\mathrm{d}X_t = \mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t$	stochastic process X_t in differential form (see Sect. 15.4)
$f(X_t, t)$	continuous (nonrandom) function of the stochastic process X_t and time t with continuous partial derivatives $f_x = \partial f/\partial X_t, f_{xx} = \partial^2 f/\partial X_t{}^2, f_t = \partial f/\partial t$

$$\mathrm{d}f(X_t, t) = \left(f_x \cdot \mu_t + f_t + \frac{1}{2} f_{xx} \cdot \sigma_t^2\right) \mathrm{d}t + f_x \cdot \sigma_t \mathrm{d}W_t, \text{ where } 0 \le t \le T$$

(*Ito's lemma*)

$$f(X_t, t) = f(X_0, 0) + \int_0^t \left(f_x \cdot \mu_s + f_s + \frac{1}{2} f_{xx} \cdot \sigma_s^2\right) \mathrm{d}s + \int_0^t f_x \cdot \sigma_s \mathrm{d}W_s, \text{ where } 0 \le t \le T$$

(integral form of Ito's lemma)

$$\mathrm{d}f(\mathbf{X}_t, t) = \left(\mathbf{f}_x' \cdot \boldsymbol{\mu}_t + f_t + \frac{1}{2}\mathrm{tr}(\boldsymbol{\Sigma}_t \boldsymbol{\Sigma}_t' \mathbf{f}_{xx})\right)\ \mathrm{d}t + \mathbf{f}_x' \cdot \boldsymbol{\Sigma}_t \mathrm{d}\mathbf{W}_t, \text{ where } 0 \le t \le T$$

(*multivariate* Ito's lemma: $\mathbf{W}_t$ is an m-variate stochastic process formed by independent Wiener processes; $\mathbf{X}_t$ is a d-variate stochastic process with differential form $\mathrm{d}\mathbf{X}_t = \boldsymbol{\mu}_t \mathrm{d}t + \boldsymbol{\Sigma}_t \mathrm{d}\mathbf{W}_t$ (i.e. in particular, $\boldsymbol{\mu}_t$ is a d-variate and $\boldsymbol{\Sigma}_t$ is a $(d \times m)$-variate stochastic process of appropriate properties); $f(\mathbf{X}_t, t)$ is a continuous (nonrandom) function of the stochastic process $\mathbf{X}_t$ and time t with d-variate vector $\mathbf{f}_x = (\partial f/\partial \mathbf{X}_{ti})$, $(d \times d)$-variate matrix $\mathbf{f}_{xx} = (\partial^2 f/\partial \mathbf{X}_{ti}\, \partial \mathbf{X}_{tj})$ and scalar $f_t = \partial f/\partial t$ of continuous partial derivatives)

$$\mathrm{d}\,(\exp(\mu \cdot t + \sigma \cdot W_t)) = \left(\mu + \frac{1}{2}\sigma^2\right) \cdot \exp(\mu \cdot t + \sigma \cdot W_t)\ \mathrm{d}t + \sigma \cdot \exp(\mu \cdot t + \sigma \cdot W_t)\ \mathrm{d}W_t$$

(example of an application of Ito's lemma, also see Sect. 15.4)

$$\mathrm{d}(W_t^2) = \mathrm{d}t + 2W_t\, \mathrm{d}W_t$$

(identity derived using Ito's lemma or integral relation $\int_0^t W_s\, \mathrm{d}W_s = (W_t^2 - t)/2$ (see Sect. 15.3): it differs significantly from the nonrandom differential $\mathrm{d}(x^2) = 2x\, \mathrm{d}x$)

$$\mathrm{d}(X_t \cdot Y_t) = X_t \mathrm{d}Y_t + Y_t \mathrm{d}X_t + \sigma_t \cdot \rho_t\, \mathrm{d}W_t \text{ for } \begin{cases} \mathrm{d}X_t &= \mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t \\ \mathrm{d}Y_t &= \nu_t \mathrm{d}t + \rho_t \mathrm{d}W_t \end{cases}$$

$$\mathrm{d}(X_t \cdot Y_t) = X_t \mathrm{d}Y_t + Y_t \mathrm{d}X_t \text{ for } \begin{cases} \mathrm{d}X_t &= \mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t \\ \mathrm{d}Y_t &= \nu_t \mathrm{d}t + \rho_t \mathrm{d}\tilde{W}_t \end{cases},$$

where processes W_t and $\tilde{W}_t$are independent

15.6 Girsanov Theorem on Equivalent Martingale Probability

Denotation and assumptions:

W_t	Wiener process with respect to probability P (see Sect. 15.1)
P'	probability *equivalent* to probability P: $P(A) = 0$ if and only if $P'(A) = 0$ (i.e. probabilities P and P' dominate each other, see Sect. 26.8)
γ_t	nonanticipating (see Sect. 15.3) stochastic process for arbitrary $0 \le t \le T < \infty$ which fulfils $\mathrm{E}_P\{\exp[(1/2) \int_0^T \gamma_t^2\ \mathrm{d}t]\} < \infty$ (the index P denotes the calculation of the mean value for P)
r_f	risk-free interest rate (see Sects. 5.1 and 13.2)

$$\mathrm{d}\tilde{W}_t = \gamma_t \mathrm{d}t + \mathrm{d}W_t \quad \text{or equivalently} \quad \tilde{W}_t = \int_0^t \gamma_s \mathrm{d}s + W_t$$

(*Girsanov theorem*: for P and γ_t (see *thereinbefore*), there exists probability Q equivalent to P such that $\tilde{W}_t$ is Wiener process with respect to Q; Q is called *equivalent martingale probability* since it enables to transform the stochastic process $\int_0^t \gamma_t \mathrm{d}t + W_t$, which in general is not a martingale (with respect to P), to a martingale (with respect to Q); Radon-Nikodym derivative of Q with respect to P (see Sect. 26.8) is $\frac{\mathrm{d}Q}{\mathrm{d}P} = \exp\left(-\frac{1}{2}\int_0^T \gamma_t^2 \mathrm{d}t - \int_0^T \gamma_t \mathrm{d}W_t\right)\Bigg)$

$$\mathrm{d}X_t = \mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t \text{ with respect to } P \rightarrow \mathrm{d}X_t = \nu_t \mathrm{d}t + \sigma_t \mathrm{d}\tilde{W}_t \text{ with respect to } Q$$
$$\text{for } \gamma_t = \frac{\mu_t - \nu_t}{\sigma_t}$$

(transformation of Wiener process with respect to probability P with drift μ_t and volatility σ_t (written in the form of SDE, see Sects. 15.1 and 15.4) to Wiener process with respect to equivalent martingale probability Q (see *thereinbefore*) with *another* drift ν_t, but still maintaining the original volatility σ_t (assuming that $\mathrm{E}_P\{\exp[(1/2)\int_0^T \gamma_t^2\, \mathrm{d}t]\} < \infty$); in particular, for $\nu_t = 0$ the transformation result is Wiener process with respect to Q without drift and with the original volatility)

$$\mathrm{e}^{-r_f \cdot t} S_t = \mathrm{e}^{-r_f \cdot t} S_0 \mathrm{e}^{X_t} = S_0 \mathrm{e}^{(\mu - r_f)\cdot t + \sigma \cdot W_t} \text{ with respect to } P$$
$$\rightarrow \frac{\mathrm{d}S_t}{S_t} = r_f \mathrm{d}t + \sigma\, \mathrm{d}\tilde{W}_t \quad \text{with respect to } Q \text{ for} \quad \gamma = \frac{\mu + \sigma^2/2 - r_f}{\sigma}$$

(transformation of exponential Wiener process S_t (in general μ and σ may be nonanticipating stochastic processes, see Sect. 15.3) which is discounted by risk-free rate of return r_f (see Sects. 5.1 and 13.2) to geometric SDE (see Sect. 15.4) describing relative changes of the process S_t by means of increments formed by the risk-free drift r_f and Wiener process with the original logarithmic volatility σ; the equivalent martingale probability Q is called the *risk-neutral probability* in this context (since the drift is risk-free, also see Chap. 14): it is applied when pricing financial derivatives (see Sect. 15.8); γ is called the *market price of risk*: it is an excess of the drift over r_f per volatility unit or risk unit (since it holds $\mathrm{d}S_t/S_t = (\mu + \sigma^2/2)\, \mathrm{d}t + \sigma\, \mathrm{d}W_t$ with respect to the original P); when no arbitrage opportunity exists (see Chap. 14) then all tradable assets must have the same market price of risk; the corresponding SDE has the form $\mathrm{d}U_t/U_t = \sigma \mathrm{d}\tilde{W}_t$, where $U_t = \exp(-r_f \cdot t) \cdot S_t$)

15.7 Theorem on Martingale Representation

Denotation and assumptions:

M_t — Wiener process with drift and volatility σ_t (see Sect. 15.3), which is a martingale (see Sect. 26.9) with respect to probability Q (the volatility σ_t is positive with probability one for each t)

K_t — arbitrary martingale with respect to probability Q

$$K_t = K_0 + \int_0^t \varphi_s \mathrm{d}M_s \quad \text{or equivalently} \quad \mathrm{d}K_t = \varphi_t \mathrm{d}M_t$$

(*theorem on martingale representation*: for M_t and K_t (see *thereinbefore*) there exists a nonanticipating (see Sect. 15.3) stochastic process ϕ_t ($\int_0^T \varphi_t^2 \cdot \sigma_t^2 \,\mathrm{d}t < \infty$) fulfilling the given relation: the so-called martingale representation of K_t by means of M_t; it is a special case of Doob-Meyer submartingale decomposition (see Sect. 26.9))

$$\mathrm{d}X_t = \mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t, \text{ where } \mathrm{E}\left(\int_0^T \sigma_t^2 \mathrm{d}t\right)^2 < \infty$$

(Wiener process X_t (see Sects. 15.3 and 15.4) with drift μ_t and volatility σ_t fulfilling the given assumption on σ_t is a *martingale*, if and only if its *drift* is zero ($\mu_t = 0$); the martingale property of the process X_t without the given assumption on σ_t is fulfilled only locally so that one has a so-called *local martingale*)

$$\frac{\mathrm{d}X_t}{X_t} = \sigma_t \mathrm{d}W_t, \text{ where } \mathrm{E}\left(\exp\left(\frac{1}{2}\int_0^T \sigma_t^2 \mathrm{d}t\right)\right) < \infty$$

(exponential Wiener process X_t (see Sects. 15.3 and 15.4) without drift and with logarithmic volatility σ_t fulfilling the given assumption on σ_t is a *martingale*; in particular, the process U_t fulfilling $\mathrm{d}U_t/U_t = \sigma \mathrm{d}\tilde{W}_t$ (see Sect. 15.6) is a martingale)

15.8 Derivatives Pricing by Means of Equivalent Martingale Probabilities

Denotation and assumptions:

$S_t = S_0 \mathrm{e}^{\mu \cdot t + \sigma \cdot W_t}$ — price of an asset unit underlying a financial derivative at time t: it has the form of exponential Wiener process with logarithmic drift μ and logarithmic volatility σ on a probability space with σ-algebras $\Im_t$ of the filtration (to which the martingale W is adapted) and with probability P (see Sect. 15.1)

$B_t = \mathrm{e}^{r_f \cdot t}$ interest factor of continuous compounding (see Chap. 4) with risk-free rate of return r (see Sects. 5.1 and 13.2): it can be interpreted as the price of the unit ($B_0 = 1$) of a risk-free zero-coupon bond at time t (see Sect. 9.2)

Z_T claim guaranteed by the financial derivative in maturity date T (e.g. for a call option with strike price X one puts $Z_T = (S_T - X)^+$, see Sect. 10.5)

V_t price (or value) of the financial derivative (if no arbitrage opportunity exists, see Chap. 14) at time t ($0 \le t \le T$)

ϕ_t number of the underlying asset units (see *thereinbefore*) in the portfolio, which replicates the price motion of the financial derivative (if no arbitrage opportunity exists) at time t ($0 \le t \le T$)

ψ_t number of units of the zero-coupon bond (see *thereinbefore*) in the replicating portfolio at time t ($0 \le t \le T$)

$$V_t = \varphi_t \cdot S_t + \psi_t \cdot B_t$$

(price of financial derivative at time t: one can find it as the price of the replicating portfolio at time t using ϕ_t units of the underlying asset with unit price S_t and ψ_t units of the risk-free zero-coupon bond with unit price B_t)

$$\mathrm{d}V_t = \varphi_t \cdot \mathrm{d}S_t + \psi_t \cdot \mathrm{d}B_t$$

(*self-financing portfolio*: any change of its price can be realized exploiting price changes of its components alone without injecting a cash to the portfolio)

- Construction of replicating portfolio to price a financial derivative (see *thereinbefore*):
  (1) $U_t = \mathrm{e}^{-r_f \cdot t} S_t = B_t^{-1} S_t$ with respect to $P \to \frac{\mathrm{d}U_t}{U_t} = \sigma\ \mathrm{d}\tilde{W}_t$ with respect to Q (output: *equivalent martingale* (*risk-neutral*) *probability* Q such that the process S_t discounted with respect to the risk-free rate of return is a martingale (see Sects. 15.6 and 15.7))
  (2) $V_t = \mathrm{e}^{-r_f \cdot (T-t)} \mathrm{E}_Q\,(Z_T \,|\Im_t) = B_t \cdot \mathrm{E}_Q \left(B_T^{-1} \cdot Z_T \,|\Im_t\right)$ with respect to Q (output: the *price of the financial derivative at time t* (see *thereinbefore*) is equal (if no arbitrage opportunity exists, see Chap. 14) to the conditional mean value of the maturity claim Z_T (see *thereinbefore*) discounted using the risk-free rate of return to time t with respect to the probability Q (see (1)) and the information included in the σ-algebra $\Im_t$ (see Sect. 26.8); if one aims to price the financial derivative only (i.e. an explicit construction of replicating portfolio is not required) then the procedure can be finished by calculating V_t)

(3) $\mathrm{d}E_t = \varphi_t \mathrm{d}U_t$, where $E_t = B_t^{-1} \cdot V_t = \mathrm{E}_Q\left(B_T^{-1} \cdot Z_T \,|\Im_t\right)$ with respect to Q (output: *number of units of the underlying asset* ϕ_t according to the theorem on martingale representation (ϕ_t is nonanticipating stochastic process, see Sect. 15.7) since the processes E_t and U_t are martingales with respect to Q (see Sect. 15.7))

(4) $\psi_t = E_t - \varphi_t \cdot U_t = B_t^{-1} \cdot (V_t - \varphi_t \cdot S_t)$ (output: *number of units of the bond* ψ_t; the corresponding replicating portfolio has indeed the price $V_t = \varphi_t \cdot S_t + \psi_t \cdot B_t$ and is self-financing (see *thereinbefore*))

$$V_t = C_t = S_t \cdot \Phi(d_1) - X \cdot \mathrm{e}^{-r_f(T-t)} \cdot \Phi(d_2),)$$

where $d_1 = \dfrac{\ln(S_t/X) + (r_f + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}$; $d_2 = d_1 - \sigma\sqrt{T-t}$

(*Black-Scholes formula* for the price of a *European call option* with maturity date T and strike price X at time t (also see Sect. 10.5): an example of application of the formula (2) for

$Z_T = (S_T - X)^+$ and $S_t = S_0 \cdot \exp(\mu \cdot t + \sigma \cdot W_t) = S_0 \cdot \exp\left((r_f - \sigma^2/2) \cdot t + \sigma \cdot \tilde{W}_t\right)$,

where $\tilde{W}_t$ has the distribution $N(0, t)$ with respect to the appropriate equivalent martingale (risk-neutral) probability Q, i.e. S_t/S_0 has with respect to the same Q the distribution $LN\left((r_f - \sigma^2/2)\cdot t,\ \sigma^2 \cdot t\right)$ (see Sect. 26.5))

15.9 Derivatives Pricing by Means of Partial Differential Equations PDE

- Alternative approach to derivatives pricing consists in solving specific *partial differential equations* (PDE) derived using the principle of risk-free portfolio with the risk-free rate of return r_f

Denotation and assumptions:

$\mathrm{d}S_t = \mu(S_t, t) \cdot \mathrm{d}t + \sigma(S_t, t) \cdot \mathrm{d}W_t$	price motion of an asset underlying a financial derivative
$V_t = F(S_t, t)$	price of the financial derivative (if no arbitrage opportunity exists, see Chap. 14)) at time t ($0 \leq t \leq T$); it is an (unknown) function F of the underlying asset, and one looks for F solving the corresponding PDE
$F(S_T, T) = G(S_T, T)$	boundary condition of the PDE with a given function G (e.g. it is $G(S_T, T) = (S_T - X)^+$ for a call option with strike price X and maturity date T, see Sect. 10.5)

$$-r_f \cdot F(S_t,t) + r_f \cdot \frac{\partial F(S_t,t)}{\partial S_t} \cdot S_t + \frac{\partial F(S_t,t)}{\partial t} + \frac{1}{2} \cdot \frac{\partial^2 F(S_t,t)}{\partial S_t^2} \cdot \sigma^2 \ (S_t,t) = 0,$$

$$0 \leq t \leq T$$

(PDE for the price of a financial derivative $F(S_t,\ t)$ with a suitable boundary condition $F(S_T,\ T) = G(S_T,\ T)$)

$$-r_f \cdot F(S_t,t) + r_f \cdot \frac{\partial F(S_t,t)}{\partial S_t} \cdot S_t + \frac{\partial F(S_t,t)}{\partial t} + \frac{1}{2} \cdot \frac{\partial^2 F(S_t,t)}{\partial S_t^2} \cdot \sigma^2 \cdot S_t^2 = 0,$$

$$0 \leq t \leq T$$

(PDE for the price of a *European call option* with maturity date T and strike price X (the price motion of the underlying asset is modeled as the exponential Wiener process $\mathrm{d}S_t/S_t = \mu \cdot \mathrm{d}t + \sigma \cdot \mathrm{d}W_t$ and the boundary condition is $G(S_T,\ T) = (S_T - X)^+$): the solution is Black-Scholes formula (see Sects. 10.5 and 15.8))

15.10 Term Structure Modeling

- *Term structure of interest rates*: is dependence of interest rates on the time to maturity (also see yield curve); its modeling is important for analysis and pricing of various interest instruments (in particular, bonds (see Sect. 9.2) and interest derivatives (see Chap. 10)); unlike the classical *discrete* approach to these problems (e.g. see yield curves in Sect. 9.2 or rates of return in Sect. 12.1), the *continuous* approach (see *thereinafter*) makes use of methods of stochastic financial analysis

Denotation:

$P(t,T)$	price of a risk-free zero-coupon bond (see Sect. 9.2) with time to maturity T and unit face value ($P(T,T) = 1$) at time t
$R(t,T) = -\frac{\ln P(t,T)}{T-t}$	*rate of return* (*yield curve*, *yield-to-maturity*) to time T at time t ($t < T$) in the sense of continuous compounding
$F(t,T,U) = -\frac{\ln P(t,U) - \ln P(t,T)}{U-T}$	*forward rate of return* from time T to time U at *spot* time t ($t < T < U$) in the sense of continuous compounding
$r_t = R(t,t) = -\frac{\partial \ln P(t,t)}{\partial T}$	*instantaneous rate of return* (*short rate*) at time t in the sense of continuous compounding
$f(t,T) = -\frac{\partial \ln P(t,T)}{\partial T}$	*instantaneous forward rate of return* from time T at spot time t ($t \leq T$); $f(t,\ t) = r_t = R(t,\ t)$

- It holds:

$$P(t,T) = \mathrm{e}^{-R(t,T)\cdot(T-\,t)} \text{ (equivalent form with rate of return } R\,(t,T))$$

$$P(T,U) = \mathrm{e}^{-F(t,T,U)\cdot(U-T)} \text{(equivalent form with forward rate of return } F(t,T,U))$$

$$P(t,T) = \exp\left(-\int_t^T f(t,u)\,\mathrm{d}u\right)$$

(equivalent form with instantaneous forward rate of return $f(t, u)$)

$$R(t,T) = \frac{1}{T-t}\int_t^T f(t,u)\,\mathrm{d}u \text{ (equivalent form for rate of return } R\,(t,T))$$

- *Yield curve* (also see Sect. 9.2): is a plot of $R(t,\cdot)$ against T for fixed t
- *Discount curve*: is a plot of $P(t,\cdot)$ against T for fixed t
- (*Credit*) *spread curve*: is a plot of $s(t,\cdot) = R(t,\cdot) - R_f\,(t,\cdot)$ against T for fixed t where $s(t,T)$ is the *spread* between the rate of return $R(t,T)$ of a risk zero-coupon bond and the rate of return $R_f(t,T)$ of a risk-free zero-coupon bond

Denotation and assumptions:

$\mathrm{d}r_t = a(r_t,t)\,\mathrm{d}t + b(r_t,t)\,\mathrm{d}W_t$ *single-factor interest rate model*: describes the motion of instantaneous rate of return r_t (see *thereinbefore*) by means of SDE (see Sect. 15.4); it includes only one interest factor r_t; special forms of this SDE (see *thereinafter*) are applied to modeling term structure of interest rates $R(r_t, t, T)$ and prices of interest instruments $P(r_t, t, T)$ assuming *non-occurrence of arbitrage opportunities* (see Chap. 14); for simplicity one often omits arguments in the notation, e.g. $P(r_t, t, T) = P$

$$\mathrm{d}P = P\cdot\mu\ \mathrm{d}t + P\cdot\sigma\ \mathrm{d}z, \text{ where}$$

$$\mu = \mu(r_t,t,T) = \tfrac{1}{P}\cdot\left(\tfrac{\partial P}{\partial t} + a\cdot\tfrac{\partial P}{\partial r} + \tfrac{1}{2}b^2\cdot\tfrac{\partial^2 P}{\partial r^2}\right); \quad \sigma = \sigma(r_t,t,T) = \tfrac{1}{P}\cdot b\cdot\tfrac{\partial P}{\partial r}$$

(SDE for $P(r_t, t, T) = P$ in the single-factor interest rate model (see *thereinbefore*))

$$q = q(r_t, t) = \frac{\mu(r_t, t, T) - r_t}{\sigma(r_t, t, T)}$$

(*market price of risk* (also see Sects. 13.3 and 15.6): if no arbitrage opportunities exist then q does not depend on the time to maturity T (i.e. bonds with different times to maturity have the same market price of risk); the equality $\mu - r = q \cdot \sigma$ can be interpreted in such a way that the expected yield μ exceeding r compensates the risk $q \cdot \sigma$)

$$\frac{\partial P}{\partial t} + (a - b \cdot q) \cdot \frac{\partial P}{\partial r} + \frac{1}{2} b^2 \cdot \frac{\partial^2 P}{\partial r^2} = P \cdot r \text{ for } t < T; \; P(r_T, T, T) = 1$$

((*Vasicek*) PDE for $P(r_t,\ t,\ T) = P$ in the single-factor interest ratemodel)

$$P(r_t, t, T) = \mathrm{E}\left(\exp\left(-\int_t^T r_u \,\mathrm{d}u - \frac{1}{2}\int_t^T q^2(r_u, u)\,\mathrm{d}u - \int_t^T q(r_u, u)\,\mathrm{d}W_u \right) \middle| \, \Im_t \right) \text{ for } t \le T$$

(solution of Vasicek PDE (see *thereinbefore*) by means of Ito stochastic integral, see Sect. 15.3)

- In practice one applies:
  - first of all some special cases of the single-factor interest rate model: *mean-reverting Vasicek model* (see *thereinafter*), *Cox-Ingersoll-Ross model* (see *thereinafter*) and others
  - *binomial tree models* (e.g. models Rendleman-Bartter, Jarrow-Rudd and others)
  - *multi-factor interest rate models* including more interest factors simultaneously (e.g. models Brennan-Schwartz, Fong-Vasicek, Longstaff-Schwartz and others)

$$\mathrm{d}r_t = \alpha \cdot (\gamma - r_t)\,\mathrm{d}t + b\,\mathrm{d}W_t, \text{ where } \alpha > 0;\ \gamma \in \mathrm{R};\ b \in \mathrm{R};\ q(r_t,\ t) = q = \text{const}$$

(*mean-reverting Vasicek model* (*Ornstein-Uhlenbeck process*): r_t fluctuates around a constant level γ but the trend coefficient $\alpha \cdot (\gamma - r_t)$ reverts the values r_t that deviated too much from γ back to this level (remind of the fact that the volatility b is constant); one assumes for simplicity that the market price of risk q is constant)

$$\mathrm{d}P = P \cdot \mu\,\mathrm{d}t + P \cdot \sigma\,\mathrm{d}W_t, \text{ where } \mu = \mu(r_t, t, T) = r_t - \frac{b \cdot q}{\alpha} \cdot \left(1 - \mathrm{e}^{-\alpha \cdot (T-t)}\right);$$
$$\sigma = \sigma(r_t, t, T) = \frac{b}{\alpha} \cdot \left(1 - \mathrm{e}^{-\alpha \cdot (T-t)}\right)$$

(SDE for $P(r_t,\ t, T) = P$ in the mean-reverting Vasicek model (see *thereinbefore*))

$$P(r_t, t, T) = \exp\left[\tfrac{1}{\alpha} \cdot \left(1 - e^{-\alpha \cdot (T-t)}\right) \cdot \left(\gamma + \tfrac{b \cdot q}{\alpha} - \tfrac{b^2}{2\alpha^2} - r_t\right) - \right.$$
$$\left. - \left(\gamma + \tfrac{b \cdot q}{\alpha} - \tfrac{b^2}{2\alpha^2}\right) \cdot (T-t) - \tfrac{b^2}{4\alpha^3} \cdot \left(1 - \mathrm{e}^{-\alpha \cdot (T-t)}\right)^2\right] \text{ for } t \le T$$

(price of zero-coupon bond in the mean-reverting Vasicek model (see *thereinbefore*))

$$R(r_t, t, T) = \gamma + \frac{b \cdot q}{\alpha} - \frac{b^2}{2\alpha^2} - \frac{1}{\alpha \cdot T} \cdot \left(1 - \mathrm{e}^{-\alpha \cdot T)}\right) \cdot \left(\gamma + \frac{b \cdot q}{\alpha} - \frac{b^2}{2\alpha^2} - r_t\right)$$
$$+ \frac{b^2}{4\alpha^3 T} \cdot \left(1 - \mathrm{e}^{-\alpha \cdot T}\right)^2$$

(time structure of interest rates in the mean-reverting Vasicek model (see *thereinbefore*); $R(r_t, t, 0) = r_t$; $R(r_t, t, \infty) = \gamma + b \cdot q/\alpha - b^2/(2\alpha^2)$; $R(r_t, t, T)$ increases, or reverses from increase to decrease, or decreases if $r_t \le \gamma + b \cdot q/\alpha - 3b^2/(4\alpha^2)$, or $\gamma + b \cdot q/\alpha - 3b^2/(4\alpha^2) < r_t < \gamma + b \cdot q/\alpha$, or $r_t \ge \gamma + b \cdot q/\alpha$, respectively)

$\mathrm{d}r_t = \alpha \cdot (\gamma - r_t)\,\mathrm{d}t + b \cdot \sqrt{r_t}\,\mathrm{d}W_t$, where $\alpha > 0$; $\gamma \in \mathrm{R}$; $b \in \mathrm{R}$; $q(r_t, t) = q =$ const

(*Cox-Ingersoll-Ross model*: in contrast to Vasicek model the volatility in this model is not constant but proportional to the value r_t (consequently the corresponding solution $P(r_t, t, T)$ does not attain negative values); one assumes again for simplicity that the market price of risk q is constant)

$$P(r_t, t, T) = A(t, T) \cdot \exp[-B(t, T) \cdot r_t)], \text{ where } \psi = \alpha + b \cdot q\ ;\ \delta = \sqrt{\psi^2 + 2b^2};$$
$$A(t, T) = \left[\frac{2\delta \cdot \mathrm{e}^{(\delta + \psi) \cdot (T-t)/2}}{2\delta + (\delta + \psi) \cdot \left(\mathrm{e}^{\delta \cdot (T-t)} - 1\right)}\right]^{2\alpha \cdot \gamma / b^2};$$
$$B(t, T) = \frac{2\left(\mathrm{e}^{\delta \cdot (T-t)} - 1\right)}{2\delta + (\delta + \psi) \cdot \left(\mathrm{e}^{\delta \cdot (T-t)} - 1\right)}$$

(price of zero-coupon bond in Cox-Ingersoll-Ross model (see *thereinbefore*))

- *No arbitrage models*: is another class of models for time structure of interest rates which make use of information on the initial structure of interest rates, e.g. information included in the function $f(0, t)$ (i.e. in the initial instantaneous forward rate of return, see *thereinbefor*e); examples are models Ho-Lee, Hull-White, Heath-Jarrow-Morton, Black-Derman-Toy, Black-Karasinski and others; in practice, one applies also frequently the so-called *LIBOR market model* LMM estimated by means of the current yield curve and the current prices of interest derivatives

Further Reading

Baxter, M., Rennie, A.: Financial Calculus. An Introduction to Derivative Pricing. Cambridge University Press, Cambridge (1996)

Black, F., Scholes, M.: The pricing of options and corporate liabilities. Journal of Political Economy 81, 637–654 (1973)

Duffie, D.: Security Markets: Stochastic Models. Academic, New York (1988)

Dupacova, J., Hurt, J., Stepan, J.: Stochastic Modeling in Economics and Finance. Kluwer, Dordrecht (2002)

Elliot, R.J., Kopp, P.E.: Mathematics of Financial Markets. Springer, New York (2004)

Hull, J.: Options, Futures, and Other Derivative Securities. Prentice Hall, Englewood Cliffs, NJ (1993)

Karatzas, I., Shreve, S.E.: Methods of Mathematical Finance. Springer, New York (1999)

Knox, D.M., Zima, P., Brown, R.L.: Mathematics of Finance. McGraw-Hill, Sydney (1984)

Kwok, Y.-K.: Mathematical Models of Financial Derivatives. Springer, Singapore (1998)

Malliaris, A.G., Brock, W.A.: Stochastic Methods in Economics and Finance. North-Holland, Amsterdam (1982)

Musiela, M., Rutkowski, M.: Martingale Methods in Financial Modelling. Springer, New York (2004)

Neftci, S.N.: Mathematics of Financial Derivatives. Academic Press, London (2000)

Pelsser, A.: Efficient Methods for Valuing Interest Rate Derivatives. Springer, London (2000)

Roman, S.: Introduction to the Mathematics of Finance. Springer, New York (2004)

Steele, J.M.: Stochastic Calculus and Financial Applications. Springer, New York (2001)

Wilmott, P., Howison, S., Dewynne, J.: The Mathematics of Financial Derivatives. Cambridge University Press, Cambridge (1995)

Part II
Insurance Formulas

Chapter 16
Insurance Classification

Abstract Chapter 16 provides various classifications of insurance industry.

- *Insurance*: is an instrument of financial elimination of negative consequences of contingency
- *Insurance risk*: is a potential possibility of occurrence of an *insured event*, where according to the *insurance contract* (*insurance policy*) the insurance company pays out an *insurance benefit*
- Characterization of insurance: the *insured* cedes his or her risks (potential loss consequences of such risks from the insured's individual point of view may be catastrophic or winding up) to the *insurer* (*insurance company*), that is capable owing to incoming *insurance premiums* (if the *insurance portfolio*, i.e. the set of insurance contracts of similar character, is sufficiently large) not only to master the accepted risks (as a whole), but even to make business profit of them

- *Classification of insurance risks*: coincides often with particular insurance products:
  - *objective risk*: is given by objective factors (e.g. by age, gender, health conditions, profession, characteristics of insured object or environment)
  - *subjective risk*: is given by subjective factors (e.g. by tendency of insured to retain one's life, health and property, to avoid conflicts against law)
  - *moral hazard*: appears in such a situation, where the insured does not prefer unconditionally the loss prevention to the loss occurrence
  - *pure risk*: is not incited artificially (as it is the case e.g. in lottery)
  - *personal risk*: is the risk of early exit, physical hazard (→ *accident* (or *bodily injury*) *insurance*), morbidity risk (→ *sickness insurance*) or longevity risk (→ *life annuities*)
  - *natural peril*: is the risk of a direct loss in consequence of a natural catastrophe (e.g. fire, flood, earthquake, and the like)
  - *traffic risk*: is the risk of loss in connection to vehicles (→ *hull* (or *casco*) *insurance*) or in connection to goods in transit (→ *transport* (or *cargo*) *insurance*)

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_16,

- *theft and vandalism risk*
- *engineering risk*: is the risk of accident or break-down of a machinery in consequence of unqualified servicing, material defect, faulty technology, and the like
- *business interruption risk* (including *loss of profits risk*): is the risk of an indirect loss in consequence of a natural disaster, accident, break-down, and the like
- *liability risk* (*third party risk*): is the risk of loss caused through conduct of the insured to life, health and property of another subject ("third party" in addition to the insured and to the insurer); the corresponding insurance product is e.g. *motor third party liability insurance*)
- *socio-political risk*: includes wars and war-like events, embargos, strikes and similar restraints of trade
- *business and financial risks*: result from varying economic and business conditions; an important special case is e.g. *credit risk* (see Sect. 12.4)
- *modern risks*: are e.g. nuclear risk, ecological hazard (environmental risk), risk of AIDS, and the like
- *underwriting risk*: consists in a potentiality that the insurer does not manage to achieve the balance between the incoming premium and the paid claims; if the insurance portfolios enlarge, then the underwriting risk decreases

- *Classification of insurance*:
  - *private insurance*:
    - *life insurance*: can be further classified to (1) *capital life insurance* with savings premium which is such a part of the premium that accumulates capital amounts, (2) *risk life insurance* (in particular, *term insurance*) with risk premium covering only mortality risk (i.e. without savings premium) and (3) *life annuities* (*insurance*)
    - *property insurance*
    - *accident insurance*
    - *insurance for private medical treatment*
    - *liability insurance* (*third party insurance*) (globally, the property insurance, accident insurance, insurance for private medical treatment and liability insurance are referred to as *non-life insurance* or as *general insurance business* or in the US as *property and casualty insurance*)
  - *social insurance* (*social security*): guarantees benefits in the case of incapacity for employment that can be either temporary (→ e.g. *sickness insurance*) or permanent mainly in consequence of age or disablement (→ (*retirement* or *invalidity*) *pension insurance* guaranteed by the state or by pension funds)
  - *health* (or *medical*) *insurance*: is guaranteed by the state or has a contractual form as the insurance for private medical treatment (see *thereinbefore*)

- *Classification of insurance from the point of law*:
  - *voluntary* (or *optional*) *insurance*: is effected as a free decision of clients (in the form of insurance contracts (insurance policies))
  - *compulsory* (or *obligatory*) *insurance*:
    - *compulsory contractual insurance*: a statutory regulation requires to effect this insurance (in the form of insurance policy) as the necessary condition for pursuing a given activity (e.g. motor third party liability insurance)
    - *mandatory insurance*: is the insurance required by law (no insurance policies are effected)

- *Classification of insurance from the point of insurable interest*:
  - *indemnity insurance* (*insurance against loss and damage*): the insurance indemnity just covers the loss caused by the insurance event, i.e. the insurance benefit is constrained by the extent of insurable interest
  - *sum insurance*: the insurance benefit serves to cover "abstract" needs without specifying an extent of insurable interest (it concerns mainly the life insurance, but also e.g. the *accident insurance for the case of death*)

- *Classification of life insurance*:
  - *risk life insurance*: the insurance event is the insured's death; in particular, the *term insurance* means that this event must occur during a stipulated period (e.g. during 10 years since the inception of insurance); if it is not the case, one calls it the *whole life insurance*
  - *pure endowment*: the insurance event is the insured's survival till a stipulated age
  - *endowment*: the insurance event is both the insured's death and the insured's survival till a stipulated age according as which event occurs earlier
  - *life annuity*: can be looked upon as a special case of pure endowment, where the insurance benefits repeat regularly as long as the insured is alive
  - *fixed term insurance* (*children's insurance*): e.g. educational insurance, dowry insurance, and the like
  - *tontines*: mean that participants join together in order to capitalize jointly their contributions and to distribute continually the accumulated capital among survivors
  - *modern life insurance products*: e.g.
    - *critical illness* (or *dread disease*) *insurance*: extends the mortality risk in such a way that the sum insured (or its percentage part) is paid out, as soon as a given illness is diagnosed (usually cancer, heart attack or stroke of given seriousness)
    - *long-term care insurance*: the insurance benefit is paid out regularly, when the insured's common skills (mobility, hygiene, food, and the like) necessary for life are reduced; it has the form of daily allowances or care expenditures payments
    - *permanent health insurance* (called *disability insurance* in the US): provides benefits on sickness and disability, the amounts of which are related

to the loss of earnings suffered by the insured due to being unable to carry his or her normal occupation
- *unit-linked insurance* (called *variable life insurance* in the US): is a combination of life insurance (usually the term insurance) and investment funds; the insurance benefit may depend on the momentary price of investment units owned by the policyholder in investment portfolios of the insurer (it implies that the client bears the whole investment risk of the policy)
- *index-linked insurance*: the insurance indemnity may depend on the average development of a market factor (see Sect. 29.3, e.g. a stock exchange indicator, the LIBOR interest rate, and the like)
- *universal life insurance*: is characterized first of all by significant *flexibility* of collected insurance premiums and by various alternatives of insurance benefits; moreover, it separates strictly the risk premium and the savings premium so that the mechanism of the product is transparent even for laymen; it is frequently combined with unit-linked insurance (see *thereinbefore*) and then called the *variable universal life*; particular cases are products, which enable benefit changes always after a change of insurable interest without evaluating health conditions
- *bancassurance*: makes use of the fact that nowadays insurance companies and banks cooperate closely in the framework of financial groups (holding companies): the policyholders are mainly clients of the bank, the allocated insurance products (using often bank sale desks) may support bank products (e.g. credits), and the like

- *Classification of life insurance from the point of establishment of technical provisions* (see Sect. 18.3):
  - *capital life insurance* (*capitalizing type of life insurance*, *life insurance with policy value*): uses the savings premium which is such a part of the premium, which generates a substantial capital amount called the *premium reserve* (the premium reserve as the most important technical provision in life insurance represents the *policy value*)
  - *risk life insurance*: uses only the risk premium covering mortality risk (i.e. there is no savings premium) so that the premium reserve of risk life insurance products is mostly negligible (see e.g. the term insurance)

- *Classification of life insurance from the point of profit benefits*:
  - *participating life insurance* (or *with-profit benefits product*): means that policyholders participate significantly in profits earned by the insurer (under the with-profit contract the insurer increases the return to the policyholders by means of additional periodic distributions (dividends), which are determined from time to time to cover individual policyholders' shares of the insurer's profits; it is popular in the UK
  - *non-participating life insurance* (or *non-profit benefits product*): means that policyholders participate only in the *technical gain* that consists of (1)

investment gain due to difference between the actually earned interest rate and the *technical interest rate* used for insurance calculations, (2) *mortality gain* due to analogous difference in mortalities, and (3) *expense gain* due to analogous difference in expenses

- *Classification of non-life insurance* (general insurance business):
  - *personal lines of business*:
    - *household* (*contents*) *insurance*
    - *building insurance*
    - *private motor insurance* (includes both the *hull insurance* and the *motor third party liability insurance*)
    - *accident insurance*: pays out benefit for (1) *period of necessary medical treatment*, (2) *permanent consequences of injury* and (3) *fatality death*
    - *insurance for private medical treatment*
    - *travel insurance*
    - *liability insurance*:
      - *voluntary*: e.g. *personal liability insurance*
      - *contractual compulsory*: e.g. *motor third party liability insurance*, *professional malpractice insurance*, *hunting liability insurance*, etc.
  - *commercial insurance* (*business and industrial insurance*):
    - *motor fleet* or *commercial vehicles insurance*
    - *commercial fire* (or *multi-peril*) *insurance*: e.g. *FLEXA* (fire, lightning, explosion, aircraft)
    - *aviation, marine* and *transport insurance*
    - *liability insurance*: e.g. *general liability insurance*, *product liability insurance*, *professional indemnity*
    - *agriculture insurance*: e.g. *crop-hail insurance*, *crop insurance*, *forest insurance*

- *Classification of pension insurance*:
  - *social insurance* (*social security*): is guaranteed by the state (the so-called *state pensions*) and financed mostly through the *pay-as-you-go* (*PAYGO*) system
  - *pension funds*

- *Classification of pension insurance according to contribution and benefit calculations*:
  - *defined contribution plan*: one calculates benefits with respect to prescribed contributions
  - *defined benefit plan*: one calculates contributions with respect to required benefits

- *Classification of benefits of pension plans*:
  - *retirement pension*: is paid out after achievement of the *retirement age*
  - *pension due to participation*: is paid out after achievement of a stipulated period of participation
  - *disability pension*: is paid out after admission of the *partial* or *total disability*
  - *survivor's pension*: e.g. widow's pension, widower's pension and orphan's pension
  - *lump sum benefit*: is paid out as a single amount instead of pension payments
  - *surrender*: is paid out if the pension is cancelled (lapsed)

Further Reading

Booth, P., Chadburn, R., Cooper, D., Haberman, S., James, D.: Modern Actuarial Theory and Practice. Chapman and Hall/CRC, London (1999)

Bowers, N.L. et al.: Actuarial Mathematics. The Society of Actuaries, Itasca, IL (1986)

Daykin, C.D., Pentikäinen, T., Pesonen, M.: Practical Risk Theory for Actuaries. Chapman and Hall, London (1994)

Hart, D.G., Buchanan, R.A., Howe, B.A.: The Actuarial Practice of General Insurance. The Institute of Actuaries of Australia, Sydney (1996)

Straub, E.: Non-Life Insurance Mathematics. Springer, Berlin (1988)

Teugels, J., Sundt, B. (eds.): Encyclopedia of Actuarial Science. Wiley, New York (2004)

Chapter 17
Actuarial Demography

Abstract Chapter 17 provides formulas of actuarial demography that are important for life and pension insurance: 17.1. Selected Population Indicators, 17.2. Life Tables, 17.3. Mortality and Survival Modeling, 17.4. Multiple Decrement Models, 17.5. Multiple Life Functions, 17.6. Commutation Functions.

17.1 Selected Population Indicators

- Unfortunately the terminology for the population (or demographic) indicators is not uniform in particular territories; a simple verbal description is sometimes preferred to symbolic formulas in this brief survey:

mid-population	is the population of the given territory (males, females, both genders (unisex)) in middle of the given period (e.g. the mid-population in year 2009 is the population as of 1 July of 2009); sometimes it may be also e.g. the arithmetic mean of the initial and final population (see *thereinafter*)
final population	is the population of the given territory (males, females, both genders (unisex)) at the end of the given period (e.g. the final population in year 2009 is the population as of 31 December of 2009); similarly for the *initial population*
(*individual*) *age*	is the completed age attained by the given person at the given time (i.e. the age of the last preceding birthday)
age-specific population	is the classification of population by age according to particular age groups which may be 1-year groups (i.e. the classification by age units), 5-year groups (the so-called *abridged age classification*)

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_17,

or others; sometimes the classification by *generations* (*cohorts*) is more convenient; the term *age pyramid* is also used in this context

natural increase of population is the difference between the numbers of live births and deaths in the given period and territory; it may be negative

total increase of population is the difference between the final and initial population in the given period and territory (see *thereinbefore*); it is the aggregate of the natural increase (see *thereinbefore*) and the *net migration* (*migration balance*)

crude birth rate is the number of live births in the given period per 1,000 inhabitants from the mid-population for this period (see *thereinbefore*); if also dead births are included, then the indicator becomes the *total birth rate*

age-specific fertility rate is the number of live births in the given period per 1,000 women in the given age group from the mid- population for this period (see *thereinbefore*)

total fertility rate is the average number of live births delivered by a woman during her *reproductive age span* (childbearing ages are often taken from 15 to 49 years old, but sometimes from 15 to 45) under the assumptions that the age-specific fertility rates (see *thereinbefore*) would be fixed at level of the given period, for which the total fertility rate is calculated, and the female mortality during the reproduction age span would be zero; the total fertility rate under 2.1 means a long-term population decrease

gross reproduction rate is the average number of live-born daughters delivered by a woman during her reproductive age span (see *thereinbefore*) under the assumption that the age-specific fertility rates (see *thereinbefore*) would be fixed at level of the given period, for which the gross reproduction rate is calculated; if only the daughters are taken into account who survive till such an age of their mothers, when these daughters have been born, then the indicator becomes the *net reproduction rate*

crude mortality rate is the number of deaths in the given period per 1,000 inhabitants from the mid-population for this period (see *thereinbefore*)

age-specific mortality rate	is the number of deaths in the given period per 1,000 inhabitants in the given age group from the mid-population for this period (see *thereinbefore*); in addition, this rate may be further classified by *causes of death*
infant mortality rate	is the number of infant deaths (i.e. the deaths during the first year of life) in the given period per 1,000 live births in this period; if only neonate deaths during the first 28 days of life are taken into account, then the indicator becomes the *neonatal mortality rate*
age-specific life expectancy	is the expected number of years that a person (usually distinguishing males and females) in the given age group will live under the assumption that the age-specific mortality rates (see *thereinbefore*) would remain at level of the year for which the life expectancy is calculated; the *life expectancy* is the expected future lifetime at birth (i.e. at zero age)
crude marriage rate	is the number of marriages occurring in the given period per 1,000 inhabitants from the mid-population for this period (see *thereinbefore*)
age-specific marriage rate	is the number of marriages occurring in the given period per 1,000 men (or women, respectively) in the given age group from the mid-population for this period (see *thereinbefore*)
crude divorce rate	is the number of divorces occurring in the given period per 1,000 inhabitants from the mid-population for this period (see *thereinbefore*)
age-specific divorce rate	is the number of divorces occurring in the given period per 1,000 men (or women, respectively) in the given age group from the mid-population for this period (see *thereinbefore*)
divorce ratio	is the percentage of ultimately divorced marriages in the given period

Denotation:

K_t	final population in year t (see *thereinbefore*)
S_t, $S_t{}^m$, $S_t{}^f$	mid-population (both genders, males, females) in year t (see *thereinbefore*)
S_{xt}, $S_{xt}{}^m$, $S_{xt}{}^f$	mid-population (both genders, males, females) of age x in year t (see *thereinbefore*)

$$r_t = \frac{K_t - K_{t-1}}{K_{t-1}} \quad (\textit{growth rate of population} \text{ in year } t)$$

$$\bar{r} = \left(\frac{K_t}{K_{t-n}}\right)^{1/n} - 1 \quad (\textit{mean annual growth rate of population} \text{ during } n \text{ years})$$

$$\rho_t = \ln\left(\frac{K_t}{K_{t-1}}\right) \quad (\textit{growth intensity of population} \text{ in year } t)$$

$$\bar{\rho} = \frac{1}{n}\ln\left(\frac{K_t}{K_{t-n}}\right) \quad (\textit{mean annual growth intensity of population} \text{ during } n \text{ years})$$

$$\frac{S_t^m}{S_t^f}$$

(*masculinity index* or *sex ratio* in year t; its inverse value is *femininity index* in year t)

$$\frac{S_{xt}^m}{S_{xt}^f} \quad (\textit{masculinity index} \text{ or } \textit{sex ratio} \text{ of age } x \text{ in year } t)$$

$$\frac{\sum_{x=0}^{14} S_{xt}}{S_t} \quad (\textit{proportion of persons under age } 15 \text{ or } \textit{child proportion} \text{ in year } t)$$

$$\frac{\sum_{x \geq 65} S_{xt}}{S_t} \quad (\textit{proportion of aged persons} \text{ in year } t)$$

$$\frac{\sum_{x \geq 65} S_{xt}}{\sum_{x=0}^{14} S_{xt}} \quad (\textit{aged-child ratio} \text{ in year } t)$$

$$\frac{\sum_{x=0}^{14} S_{xt} + \sum_{x \geq 65} S_{xt}}{\sum_{x=15}^{64} S_{xt}}$$

(*age dependency ratio* in year t: represents the ratio of the combined child and aged population to the population of intermediate "working" age; the age limits for child and aged population may be different in particular countries and may vary in time)

$$\bar{x} = \frac{1}{S_t}\sum_x (x+0.5)\cdot S_{xt} \quad (\textit{mean age of population} \text{ in year } t)$$

$$\tilde{x}: \sum_{x=0}^{\tilde{x}-1} S_{xt} < \frac{S_t}{2} \leq \sum_{x=0}^{\tilde{x}} S_{xt} \quad (\textit{median age of population} \text{ in year } t\text{, see Sect. 27.5})$$

$$\hat{x}: S_{\hat{x}t} = \max_x S_{xt} \quad (\textit{modal age of population} \text{ in year } t\text{, see Sect. 27.5})$$

17.2 Life Tables

- *Life tables*: serve as decrement instruments (in addition to financial instruments) for calculations in life and pension insurance; they are mostly constructed separately for males and females; one distinguishes the following types of life tables:
  - *complete*: have 1-year age intervals (i.e. items for age 0, 1, ...)
  - *abridged*: have multiyear age intervals (mostly 0, 1–4, 5–9, 10–14, ...)
  - *current* (or *period*): are snapshots of the current mortality of population; they are based on mortality experience of population during a short time (mostly 1-year) period
  - *generation* (or *cohort*): are records of the actual lifetime of a given generation; they are based on mortalities experienced by particular birth cohorts

- Description of columns of a *complete current life tables* (such life tables are common in insurance practice):
  - *Age x* (or *life aged x*):
    - $x = 0, 1, \ldots, \omega$ (the last age interval means "at age ω or more"); insurance companies usually use $x = 15, 16, \ldots$ or $x = 18, 19, \ldots$
  - *Probability of death at age x*: q_x
    - is the probability of dying within 1 year, given that the age x has been attained (i.e. the probability that a life aged x will die within 1 year)
  - *Probability of survival at age x*: $p_x = 1 - q_x$
    - is the probability of surviving 1 year, given that the age x has been attained (i.e. the probability that a life aged x will survive to age $x + 1$); $p_x = 1 - q_x$
    - more generally (but not given explicitly in the life tables): ${}_nq_x$ (or ${}_np_x$) is the probability that a life aged x will die within n years (or will survive to age $x + n$), respectively
  - *Number of survivors to age x*: l_x
    - is the expected number of lives surviving to age x from l_0 newborns
    - the sequence $l_0 \geq l_1 \geq l_2 \geq \ldots$ is called *mortality decrement order* (or simply *life table*): it represents a hypothetical lifetime of chosen l_0 newborns under the assumption that the age-specific mortality rates (see Sect. 17.1) would remain at level of the period for which the life table is constructed; l_0 is called *radix* of the life table with a usual value $l_0 = 100{,}000$ (but sometimes also e.g. $l_{15} = 100{,}000$ or $l_{18} = 100{,}000$)
    - the information contained in $q_0, q_1, q_2, \ldots$ is fully equivalent to the information contained in $l_0, l_1, l_2, \ldots$

- *Number of deaths at age x*: d_x
  - is the expected number of deaths at age x by survivors of the initial group of l_0 newborns ($d_\omega = l_\omega$)
- *Number of years lived at age x*: L_x
  - is the expected total number of years lived between ages x and $x + 1$ by survivors of the initial group of l_0 newborns; one applies frequently the approximations of the type $L_x = l_{x+1} + 0.5 \cdot d_x = (l_x + l_{x+1})/2$ (with exceptions $L_0 = l_1 + 0.1 \cdot d_0$, $L_\omega = l_\omega /2$)
- *Number of years lived beyond age x*: T_x
  - is the expected total number of years lived beyond age x by survivors of the initial group of l_0 newborns; the usual approximation is simply $T_x = L_x + L_{x+1} + \ldots$
- *Life expectancy at age x*: $^\circ e_x$
  - is the expected future lifetime at age x; the usual approximation is $^\circ e_x = T_x / l_x$
  - in particular, $^\circ e_0$ is the *life expectancy* (*at birth*) (i.e. at the age zero)

- Selected relations for life table values:

$$_np_x = p_x \cdot p_{x+1} \cdot \ \ldots \ \cdot p_{x+n-1} = (1 - q_x) \cdot (1 - q_{x+1}) \cdot \ \ldots \ \cdot (1 - q_{x+n-1})$$

$$_mp_x \cdot {}_np_{x+m} = {}_{m+n}p_x$$

$$d_x = l_x - l_{x+1}$$

$$q_x = \frac{d_x}{l_x} = \frac{l_x - l_{x+1}}{l_x}; \qquad p_x = \frac{l_{x+1}}{l_x}; \qquad l_{x+1} = p_x \cdot l_x = (1 - q_x) \cdot l_x$$

$$_{n|}q_x = \frac{d_{x+n}}{l_x} \text{ (probability that a life aged } x \text{ will die just at age } x + n)$$

$$^\circ e_x = \frac{1}{2} + {}_1p_x + {}_2p_x + \ldots = \left(\frac{l_x + l_{x+1}}{2} + \frac{l_{x+1} + l_{x+2}}{2} + \ldots + \frac{l_{\omega-1} + l_\omega}{2} + \frac{l_\omega}{2} \right) \Big/ l_x$$

$$x + {}^\circ e_x = \frac{1}{2} + \frac{\sum_{t=x} t \cdot d_t}{\sum_{t=x} d_t}$$

$q_x = 1 - \mathrm{e}^{-m_x}$ for mortality rate at age x (see Sect. 17.1) estimated as $m_x = \dfrac{M_x^{III}}{S_x}$

(*estimate of life table value* q_x according to a usual methodology of life table construction in practice: S_x is the mid-population (males or females) at age x (see Sect. 17.1); ${M_x}^{III}$ is the number of "deaths of the third order" at age x (e.g. for year 2010 at age 40 is M_{40}^{III} the number of deaths in year 2010 at completed age 40 from generations 1969 and 1970); the estimated q_x enables to find the other life table values p_x, l_x, d_x, L_x, T_x, $^\circ e_x$)

- *Graduation of life tables*: is a smoothing of fluctuations in the sequence q_x or l_x due to the fact that the table values are statistical estimates constructed using observed data (e.g. without applying graduations, the necessary inequalities $l_0 \geq l_1 \geq l_2 \geq \ldots$ may not hold); in practice, there exist various method of graduation (though modern statistical offices publish life tables in a properly smoothed form), e.g.:

$$\hat{l}_x = \hat{k} \cdot \hat{s}^x \cdot \hat{g}^{\hat{c}^x}$$

(*analytical graduation* of life tables by means of (analytical) *laws of mortality* (see e.g. *Gompertz-Makeham's law* in Sect. 17.3); $\hat{k}$, $\hat{s}$, $\hat{g}$, $\hat{c}$ are estimated parameters k, s, g, c obtained by means of the regression analysis (see Sect. 27.11) for the values l_x prior to graduation; sometimes graduations for several age spans must be combined using spline methods)

$$\hat{q}_x = \frac{1}{25}(q_{x-4} + 2q_{x-3} + 3q_{x-2} + 4q_{x-1} + 5q_x + 4q_{x+1} + 3q_{x+2} + 2q_{x+3} + q_{x+4})$$

(*mechanical graduation* of life tables by means of *9-point Wittstein's method*; it is a special case of the moving averages (see Sect. 31.2))

$$\hat{q}_x = \frac{1}{27}(-q_{x-4} + 2q_{x-2} + 8q_{x-1} + 9q_x + 8q_{x+1} + 2q_{x+2} - q_{x-4})$$

(*mechanical graduation* of life tables by means of *9-point Schärtlin's method*; it is again a special case of the moving averages (see Sect. 31.2); other methods of this type are used in practice, e.g. *Whittaker-Henderson's*, *King-Hardy's*, *graphical smoothing*, and the like)

- Some aspects of life tables applied in life and pension insurance:
  (1) *male and female mortality in life tables*: due to the significantly higher male mortality in comparison to the female mortality accross the whole insured age span, the insurers apply various strategies, e.g. they calculate insurance premiums:
    - *separately for males and females* using male and female life tables
    - *without distinguishing gender* using unisex life tables
    - *by shifting male insurance premiums* (in order to obtain female premiums): e.g. 5-year shifts are popular so that a female effecting a life

insurance at age 40 pays the same premiums as a male effecting the same type of insurance at age 35

(2) *security loading implicitly included in life tables*: the insurers adjust life tables by loadings (margins) in their favour, since the loadings increase the corresponding insurance premiums; e.g. one uses:
- an artificial aging by 1 year in the case of mortality risk (then e.g. the value q_{40} equals to the actual value q_{41} prior to the age shift)
- an artificial rejuvenation by 2 years in the case of longevity risk (then e.g. the value q_{40} equals to the actual value q_{38} prior to the age shift)
- other modifications, e.g. of the type $\hat{q}_x \pm c/{}^{\circ}e_x$, where ${}^{\circ}e_x$ is the life expectancy at age x (see *thereinbefore*) and c is a suitable constant (e.g. 0.015, if a higher security loading is necessary)

(3) *select life tables*: unlike the classical (*aggregate*) life tables, the probabilities q_x are graded according to the age at entry so that the tables contain values of the type $q_{[x-t]+t}$, which is the probability of death at age x, given that the insurance has been effected t years ago; typically, a person who has just bought insurance will be (due to subjective reasons or medical tests) of better health than a person who has bought insurance several years ago; hence the selection leads to inequalities $q_{[x]} < q_{[x-1]+1} < q_{[x-2]+2} < \ldots$

(4) *antiselection in life tables*: reflects the effort of policyholders to effect an insurance in their favour (i.e. inconvenient for the insurer); unhealthy persons want to have a policy covering the mortality risk, so that the insurer protects oneself through entry medical tests; on the contrary, healthy persons want to have a policy covering the longevity risk (mainly life annuities), so that the insurer protects oneself through *reduction coefficients* r_x for life annuity portfolios, where the calculated probability of death at age x is not q_x, but $r_x \cdot q_x$ for $0 < r_x < 1$

(5) *multiple life tables*: are important instruments for *multiple life insurance* (e.g. widow's, widower's or orphan's pensions, insurance of boards of directors, and others, see Sect. 18.8); these tables apply to couples of persons (x, y) of the type "male at age x and female at age y" (or generally, to n-tuples of persons), where under the assumption of independent mortality behaviour of male and female population one can put

$$p_{xy} = p_x \cdot p_y; \qquad q_{xy} = 1 - p_{xy}; \qquad l_{xy} = l_x \cdot l_y$$

p_{xy}	probability that a couple (x, y) will survive to ages $(x+1, y+1)$
q_{xy}	probability that a couple (x, y) will fail prior to ages $(x+1, y+1)$
l_{xy}	number of couples surviving to ages (x, y) (the *joint-life status*)

p_x, l_x values from a male life table (see *thereinbefore*)
p_y, l_y values from a female life table (see *thereinbefore*)

(6) *combination of current and generation life tables*: unlike the classical current life tables, the probabilities q_x are graded according to the calendar years so that the tables contain the values $q_x(t)$ (the probability of death at age x in calendar year t) or $q_x{}^\tau$ (the probability of death at age x for generation $\tau = t - x$):

$$q_x(t) = T(x,t) \cdot q_x^B = e^{-F(x)\cdot(t-t^*)} \cdot q_x^B; \quad q_x^\tau = e^{-F(x)\cdot(x+\tau-t^*)} \cdot q_x^B = q_{x+h(\tau)}^B$$

$T(x, t)$ trend function
$F(x)$ trend factor
$q_x{}^B$ probability of death at age x from a *basic life table* that corresponds to a preset calendar year t^* (the basic values $q_x{}^B$ must be tabulated in advance for a fixed t^*)
$h(\tau)$ shift of current age of a life aged x from generation τ to the adjusted age $x + h(\tau)$ proper for this generation (the shifts $h(\tau)$ must be tabulated in advance)

(7) *multiple decrement life tables*: are graded according to different causes of decrement, due to which a person leaves an initial status (e.g. in the disability pension insurance, the initial status is “active” and the causes of decrement may be “disablement” and “death”); the tables contain values of the type $q_x{}^{(j)}$, which is the probability of leaving the initial status at age x by cause j (the so-called *probability of decrement*, see Sect. 17.4)

17.3 Mortality and Survival Modeling

- In life insurance one must combine financial calculations with mathematical modeling of mortality, since insurance events in the context of life insurance consist in dying within a given age span or surviving to a given age:

$$F_x(t) = P(T_x \le t) = P(T_x < t)$$

(random variable T_x represents the *future lifetime at age x*; the sharp inequality when calculating the distribution function $F_x(t)$ is justified by the continuity of the random variable T_x)

$$x + T_x \text{ (whole lifetime (past and future) of a life aged } x)$$

$$S_x(t) = P(T_x > t) \ (\textit{survival function at age } x)$$

$$q_x = F_x(1) = P(T_x \le 1)$$

(*probability of death at age x* (see Sect. 17.2): is the probability that a life aged x will die within 1 year)

$$p_x = S_x(1) = P(T_x > 1)$$

(*probability of survival at age* x (see Sect. 17.2): is the probability that a life aged x will survive to age $x + 1$)

$${}_{s|}q_x = F_x(s+1) - F_x(s) = P(s < T_x \le s+1)$$

(probability that a life aged x will die at age $x + s$; it simplifies to q_x for $s = 0$)

$${}_tq_x = F_x(t) = P(T_x \le t)$$

(*t-year probability of death at age* x: is the probability that a life aged x will die within t years; it simplifies to q_x for $t = 1$)

$${}_tp_x = S_x(t) = P(T_x > t)$$

(*t-year probability of survival at age* x: is the probability that a life aged x will survive to age $x + t$; it simplifies to p_x for $t = 1$)

$${}_{s|t}q_x = F_x(s+t) - F_x(s) = P(s < T_x \le s+t)$$

(probability that a life aged x will survive to age $x + s$, but will die within further t years; it simplifies to ${}_{s|}q_x$ for $t = 1$)

$${}_{s+t}p_x = {}_sp_x \cdot {}_tp_{x+s}$$

$${}_{s+t}q_x = 1 - (1 - {}_sq_x) \cdot (1 - {}_tq_{x+s})$$

$${}_{s|}q_x = {}_sp_x \cdot q_{x+s}; \qquad {}_{s|\,t}q_x = {}_sp_x \cdot {}_tq_{x+s}$$

$${}_np_x = p_x \cdot p_{x+1} \cdot \ldots \cdot p_{x+n-1}$$

Assumption: the random variable T_x (see *thereinbefore*) has probability density (see Sect. 26.3):

$$f_x(t) = \frac{\mathrm{d}}{\mathrm{d}t} F_x(t) = -\frac{\mathrm{d}}{\mathrm{d}t}\, {}_tp_x$$

$\mu_{x+t} = \dfrac{f_x(t)}{{}_tp_x} = \dfrac{f_x(t)}{S_x(t)} = -\dfrac{\mathrm{d}}{\mathrm{d}t} \ln({}_tp_x)$ (*force of mortality* at age $x + t$ of a life aged x)

$$\mu_x \cdot \Delta x \approx P(T_x \le \Delta x) = {}_{\Delta x}q_x$$

($\mu_x \cdot \Delta x$ can be interpreted as an approximation of the probability that a life aged x will die within age interval $(x, x + \Delta x)$ of a (small) length Δx)

$$ {}_tq_x = \int_0^t {}_sp_x \cdot \mu_{x+s} \mathrm{d}s \; ; \qquad {}_tp_x = \exp\left(- \int_0^t \mu_{x+s} \mathrm{d}s \right) $$

$$ {}^\circ e_x = \mathrm{E}(T_x) = \int_0^\infty t \cdot f_x(t)\, \mathrm{d}t = \int_0^\infty (1 - F_x(t))\ \mathrm{d}t = \int_0^\infty t \cdot {}_tp_x \cdot \mu_{x+t}\, \mathrm{d}t = \int_0^\infty {}_tp_x\, \mathrm{d}t $$

(*life expectancy at age* x: is the expected future lifetime at age x)

$$ \mathrm{var}(T_x) = \int_0^\infty t^2 \cdot {}_tp_x \cdot \mu_{x+t}\, \mathrm{d}t - \left(\int_0^\infty t \cdot {}_tp_x \cdot \mu_{x+t}\, \mathrm{d}t \right)^2 = \int_0^\infty 2t \cdot {}_tp_x\, \mathrm{d}t - \left(\int_0^\infty {}_tp_x\, \mathrm{d}t \right)^2 $$

$$ l_x = l_0 \cdot \exp\left(- \int_0^x \mu_t\, \mathrm{d}t \right); \quad d_x = \int_x^{x+1} l_t \cdot \mu_t\, \mathrm{d}t; \quad L_x = \int_x^{x+1} l_t\, \mathrm{d}t; $$

$$ T_x = \int_x^\omega l_t\, \mathrm{d}t; \quad {}^\circ e_x = \frac{1}{l_x} \cdot \int_x^\omega l_t\, \mathrm{d}t $$

(values in life tables (see Sect. 17.2) calculated with respect to continuous arguments)

- *Laws of mortality*: represent mortality decrement orders expressed by means of analytical functions; they can be classified according to the corresponding force of mortality:
  - *constant force of mortality*:

$$ \mu_x = \lambda; \qquad {}_tp_x = \mathrm{e}^{-\lambda \cdot t} $$

  - *De Moivre's law of mortality*: has T_x with the uniform distribution (see Sect. 26.5):

$$ \mu_x = \frac{1}{\omega - x}; \qquad {}_tp_x = \frac{\omega - x - t}{\omega - x} $$

  - *Gompertz's law of mortality*: has μ_x that increases exponentially:

$$ \mu_x = B \cdot c^x; \qquad {}_tp_x = g^{c^x(c^t - 1)} \quad (B > 0; \quad c > 1; \quad g = \exp\{-B/\ln(c)\}) $$

– *Gompertz-Makeham's law of mortality*: modifies Gompertz's law:

$$\mu_x = A + B \cdot c^x; \quad {}_tp_x = s^t \cdot g^{c^x(c^t - 1)}$$

$$(A > 0; \quad B > 0; \quad c > 1; \quad g = \exp\{-B/\ln(c)\}; \quad s = \exp(-A))$$

– *Weibull's law of mortality*: has μ_x that increases polynomially:

$$\mu_x = k \cdot x^n; \quad {}_tp_x = w^{(x+t)^{n+1} - x^{n+1}} (k > 0; \quad n > 0; \quad w = \exp\{-k/(n+1)\})$$

$$T_x = K_x + S_x, \text{ where } K_x = [T_x] \text{ and } 0 \le S_x < 1$$

(K_x is the *curtate future lifetime at age x* (i.e. a random variable with values 0, 1, 2, ... representing the *integer* number of completed future years lived by a person aged x); S_x is the fraction of the year, during which the person aged x is alive in the year of death)

$$P(K_x = k) = {}_kp_x - {}_{k+1}p_x = {}_kp_x \cdot q_{x+k}, \quad k = 0, 1, 2, \ldots$$

$$e_x = \mathrm{E}(K_x) = \sum_{k=0}^{\infty} k \cdot P(K_x = k) = \sum_{k=1}^{\infty} k \cdot {}_kp_x \cdot q_{x+k} = \sum_{k=1}^{\infty} {}_kp_x$$

(*life expectancy at age x* as the expected *curtate* future lifetime at age x (in contrast to $^\circ e_x = \mathrm{E}(T_x)$))

$$^\circ e_x \approx e_x + \frac{1}{2}$$

(approximation of $\mathrm{E}(T_x)$, given that S_x (i.e. the fraction of the year of death, see *thereinbefore*) has the uniform distribution (see Sect. 26.5) in interval (0, 1), see also Sect. 17.2)

$$\mathrm{var}(T_x) \approx \mathrm{var}(K_x) + \frac{1}{12} = \sum_{k=1}^{\infty} k^2 \cdot {}_kp_x \cdot q_{x+k} - \left(\sum_{k=1}^{\infty} {}_kp_x\right)^2 + \frac{1}{12}$$

(approximation of $\mathrm{var}(T_x)$, given that S_x (i.e. the fraction of the year of death, see *thereinbefore*) has the uniform distribution (see Sect. 26.5) in interval (0, 1))

- Some approximations of the *distribution of* S_x (i.e. the fraction of the year of death, see *thereinbefore*):

  – *assumption of linearity*: ${}_t q_x = t \cdot q_x$ for $0 \le t \le 1$:
  then (1) ${}_t p_x = 1 - t \cdot q_x$ for $0 \le t \le 1$; (2) K_x and S_x are independent; (3) S_x has the uniform distribution in (0, 1)

- *assumption of constant force of mortality*: $\mu_{x+t} = \mu$ for $0 \leq t \leq 1$:
  then (1) ${}_t p_x = (p_x)^t$ for $0 \leq t \leq 1$; (2) K_x and S_x are not independent; (3) S_x has the censored exponential distribution in (0, 1)
- *rounding of S_x to next forthcoming multiple of* $1/m$: $S_x^{(m)} = [m \cdot S_x + 1]/m$:
  then (1) if K_x and S_x are independent, then also K_x and $S_x^{(m)}$ are independent; (2) if S_x has the uniform distribution in (0, 1), then also $S_x^{(m)}$ has the uniform (but discrete) distribution in (0, 1)

17.4 Multiple Decrement Models

- In actuarial models the person under consideration is in a specified status (called *initial status*) at age x and leaves that status at time T_x due to one of m mutually exclusive *causes of decrement* (numbered from 1 to m); such a situation generalizes the classical model with the only one cause of decrement (namely death), since one must study the pair of random variables: (1) the *remaining lifetime in the initial status* T_x and (2) the *cause of decrement* J ($J = 1, \ldots, m$); e.g. in a simple disability pension model, the initial status is "active" and the causes of decrement may be "disablement" and "death" (see multiple decrement life tables in Sect. 17.2):

$$ {}_t q_x^{(j)} = P(T_x \leq t,\ J = j),\ j = 1, \ldots, m $$

(*t-year probability of decrement at age x by cause j*: is the probability that a life aged x will leave the initial status within t years by cause j)

$$ {}_t q_x = {}_t q_x^{(1)} + \ldots + {}_t q_x^{(m)} = P(T_x \leq t)\ ; \qquad {}_t p_x = 1 - {}_t q_x $$

(*t-year probability of decrement at age x* without distinguishing causes of decrement: is the probability that a life aged x will leave the initial status within t years)

Assumption: there exists probability density $f_x^{(j)}(t)$ (see Sect. 26.3):

$$ f_x^{(j)}(t) = \frac{\mathrm{d}}{\mathrm{d}t} {}_t q_x^{(j)} $$

$$ \mu_{x+t}^{(j)} = \frac{f_x^{(j)}(t)}{{}_t p_x} $$

(*force of decrement at age $x + t$ by cause j* for a life that is at age $x + t$ still in the initial status)

$$ \mu_{x+t} = \mu_{x+t}^{(j)} + \ldots + \mu_{x+t}^{(j)} $$

(*force of decrement at age $x + t$* without distinguishing causes of decrement)

$$P(K_x = k,\ J = j) = {}_k p_x \cdot q_{x+k}^{(j)}, \quad k = 0, 1, 2, \ldots; j = 1, 2, \ldots, m$$

(K_x is the *curtate remaining lifetime in the initial status at age* x: $K_x = [T_x]$)

17.5 Multiple Life Functions

- *Multiple life insurance* (see Sect. 18.8): are insurance products, in which insurance benefits depend on life status (mostly "alive" or "dead") of several lives (married couples, whole families, boards of directors, and others); in such situations one applies the actuarial *model of couples* (x, y) (usually a male aged x at entry and a female aged y at entry), *model of triplets* (x, y, z) (in addition, a child aged z at entry), or more generally, *model of n-tuples* with entry ages $(x_1, \ldots, x_n)$ (see multiple life tables in Sect. 17.2):

q_{xy}	probability that a couple (x, y) will fail prior to ages $(x + 1, y + 1)$
$p_{xy} = p_x \cdot p_y$	probability that a couple (x, y) will survive to ages $(x + 1, y + 1)$ (the opportunity to multiply probabilities from individual life tables (e.g. for males and females) follows from the assumption of *independent mortality behaviour of male and female population*, see Sect. 17.2)
$l_{xy} = l_x \cdot l_y$	*number of couples surviving to ages* (x, y) (the so-called *joint-life status* again under the assumption of independence, see *thereinbefore*)
d_{xy}	*number of couples failing at ages* (x, y) (i.e. prior to achieving ages $(x + 1, y + 1)$)

$$d_{xy} = l_{xy} - l_{x+1,\,y+1} = l_x \cdot l_y - l_{x+1} \cdot l_{y+1}$$

$$q_{xy} = \frac{d_{xy}}{l_{xy}} = \frac{l_{xy} - l_{x+1,\,y+1}}{l_{xy}}; \qquad p_{xy} = \frac{l_{x+1,\,y+1}}{l_{xy}}$$

17.6 Commutation Functions

- *Commutation functions*: are constructed by (financial) discounting (see Sect. 3.2) of life table functions (see Sect. 17.2); these actuarial instruments achieved great popularity, since (1) they simplify numerical calculations of many actuarial values and (2) various expected values (applied e.g. to calculation of insurance premiums or reserves) may be derived within *deterministic models*, which are closely related just to the commutation functions (though nowadays, *stochastic models* based on probability theory are preferred in actuarial practice):

- *Commutation functions of zero order*:

$$D_x = l_x v^x \ (\textit{discounted number of survivors at age } x)$$

$$C_x = d_x v^{x+1} \ (\textit{discounted number of deaths at age } x)$$

- *Commutation functions of first order*:

$$N_x = D_x^{[2]} = \sum_{j=0}^{\omega - x} D_{x+j} = D_x + D_{x+1} + \ \dots \ + D_\omega$$
$$M_x = C_x^{[2]} = \sum_{j=0}^{\omega - x} C_{x+j} = C_x + C_{x+1} + \ \dots \ + C_\omega$$

- *Commutation functions of second order*: are applied mostly to variable insurance products as e.g. to increasing life annuities in Chap. 18 (the commutation functions of higher than the second order have mostly no applications in practice):

$$S_x = D_x^{[3]} = \sum_{j=0}^{\omega - x} N_{x+j} = N_x + N_{x+1} + \ \dots \ + N_\omega$$
$$R_x = C_x^{[3]} = \sum_{j=0}^{\omega - x} M_{x+j} = M_x + M_{x+1} + \ \dots \ + M_\omega$$

- Relations among commutation functions (where $v = 1/(1 + i)$; $d = 1 - v$), e.g.:

$$\sum_{j=0}^{n-1} D_{x+j} = D_x + D_{x+1} + \ \dots \ + D_{x+n-1} = N_x - N_{x+n}$$
$$\sum_{j=0}^{n-1} C_{x+j} = C_x + C_{x+1} + \ \dots \ + C_{x+n-1} = M_x - M_{x+n}$$
$$\sum_{j=0}^{\omega - x} (j+1) \cdot D_{x+j} = S_x; \quad \sum_{j=0}^{\omega - x} (j+1) \cdot C_{x+j} = R_x$$
$$\sum_{j=0}^{n-1} (j+1) \cdot D_{x+j} = S_x - S_{x+n} - n \cdot N_{x+n}$$
$$\sum_{j=0}^{n-1} (j+1) \cdot C_{x+j} = R_x - R_{x+n} - n \cdot M_{x+n}$$
$$C_x = v \cdot D_x - D_{x+1}$$
$$M_x = v \cdot N_x - N_{x+1} = D_x - d \cdot N_x; \quad R_x = v \cdot S_x - S_{x+1} = N_x - d \cdot S_x$$

$$N_x = N_{x+1} + D_x; \quad N_\omega = D_\omega\,; \quad M_x = M_{x+1} + C_x; \quad M_\omega = C_\omega$$

$$S_x = S_{x+1} + N_x; \quad S_\omega = N_\omega\,; \quad R_x = R_{x+1} + M_x; \quad R_\omega = M_\omega$$

- *Commutation functions in models of couples* (see Sects. 17.5 and 18.8):

$$D_{xy} = l_{xy} \cdot v^{(x+y)/2} = l_x \cdot l_y \cdot v^{(x+y)/2}; \qquad N_{xy} = D_{xy} + D_{x+1,\,y+1} + \ldots$$

$$C_{xy} = d_{xy} \cdot v^{(x+y)/2+1} = (l_x \cdot l_y - l_{x+1} \cdot l_{y+1})v^{(x+y)/2+1};$$

$$M_{xy} = C_{xy} + C_{x+1,\,y+1} + \ldots$$

Further Reading

Anderson, A.W.: Pension Mathematics for Actuaries. Wellesley, MA (1992)

Batten, R.W.: Mortality Table Construction. Prentice-Hall, Englewood Cliffs, NJ (1978)

Booth, P., Chadburn, R.., Cooper, D., Haberman, S., James, D.: Modern Actuarial Theory and Practice. Chapman and Hall/CRC, London (1999)

Bowers, N.L. et al.: Actuarial Mathematics. The Society of Actuaries, Itasca, IL (1986)

Brown, R.L.: Introduction to Mathematics of Demography. ACTEX Publications, Winsted and Avon, CT (1991)

Gerber, H.U.: Lebensversicherungsmathematik. Springer, Berlin (1986) (*English translation*: Life Insurance Mathematics. Springer, Berlin (1990))

Heubeck, K.: Richttafeln für die Pensionsversicherung. Verlag Heubeck, Köln, Germany (1983)

Keyfitz, N.: Applied Mathematical Demography. Wiley, New York (1977)

Koller, M.: Stochastische Modelle in der Lebensversicherung. Springer, Berlin (2000)

Lee, E.M.: An Introduction to Pension Schemes. The Institute of Actuaries and the Faculty of the Actuaries, London (1986)

London, D.: Graduation: The Revision of Estimates. ACTEX Publications, Winsted, CT (1985)

Milbrodt, H., Helbig, M.: Mathematische Methoden der Personenversicherung. DeGruyter, Berlin (1999)

Neill, A.: Life Contingencies. Heinemann, London (1977)

Parmenter, M.M.: Theory of Interest and Life Contingencies, with Pension Applications. ACTEX Publications, Winsted and New Britain, CT (1988)

Teugels, J., Sundt, B. (eds.): Encyclopedia of Actuarial Science. Wiley, New York (2004)

Wolff, K.-H.: Versicherungsmathematik. Springer, Wien (1970)

Wolfsdorf, K.: Versicherungsmathematik (Teil 1: Personenversicherung). Teubner, Stuttgart, Germany (1997)

Chapter 18
Classical Life Insurance

Abstract Chapter 18 contains formulas of classical life insurance: 18.1. Basic Concepts of Life Insurance, 18.2. Symbols and Calculation Principles of Life Insurance, 18.3. Technical Provisions in Life Insurance, 18.4. Pure Endowments, 18.5. Whole Life and Term Insurance, 18.6. Further Products of Capital Life Insurance, 18.7. Life Annuities, 18.8. Multiple Life Insurance, 18.9. Premium Reserve and Its Implications, 18.10. Medical Underwriting.

18.1 Basic Concepts of Life Insurance

- *Insured event*: is a contingency specified in the insurance contract (insurance policy), under which the insurer pays out a benefit insured; under a life insurance contract the benefit insured consists either of a single payment (the *sum insured*), or repeated payments (e.g. in life annuities)

- *Term of insurance*: is the period of insurance cover, according to which one classifies:

  - *temporary insurance*: has a limited period of cover stipulated in the insurance contract
  - *perpetual insurance*: has a period of cover, which in not limited by the insurance contract (e.g. a whole life annuity)
  - *deferred insurance*: has a *period of deferment*, by which the given insurance cover is postponed (e.g. a deferred life annuity); in the case of short-term deferment (in order to reduce moral hazards of clients, see Chap. 16), the concept of *waiting period* is used instead

- *Participants in life insurance* (classification):

  - *insurer*: is a legal entity that is entitled according to law to carry out insurance
  - *policyholder*: is a natural person or a legal entity that concluded an insurance contract with the insurer (an important obligation of policyholders is to pay insurance premiums)

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_18,

  - *insured* (*person*): is a natural person, the life and health of which is the subject matter of the insurance (in non-life insurance the insured may be a legal entity, as well)
  - *beneficiary*: is a natural person or a legal entity that has right to the benefit insured as the consequence of the insured event

- *Insurance premium* (classification):

  - *net premium*: is calculated by means of the so-called *equivalence principle*, according to which the expected loss of insurer should be zero
  - *gross premium* (*office premium*): is the expense-loaded net premium (usually including the security loading, see Sect. 7.2), which covers *expenses* of the insurer
  - classification according to premium payments:

    - *single premium*
    - periodic premiums of a constant amount (*level premiums*) usually at the beginning of stipulated periods (months, quarters, years)
    - *periodic premiums of varied amounts*

  - *written* versus *collected premium*: is classification according to payment status
  - *adjusted premium*: is periodically increased in accordance with the inflation
  - special modes of premium payments: e.g.

    - *paid-up policy*: means that the policy is converted to a mode with *reduced benefits* (mainly due to the premature cessation of premium payments)
    - *waiver*: is an exemption from premium payments (mainly due to disability of the policyholder)

- *Surrender*: means that withdrawing life insurance policyholders are entitled to *non-forfeiture benefits* (which are benefits that are not lost because of the premature cessation of premium payments); the surrender amount is usually a given part of the corresponding premium reserve (see Sect. 18.9)
- *Insurance with return of premiums in the case of death*: is appropriate in situations, where the insured's death would mean the cancellation of the given insurance without remuneration of beneficiaries (e.g. in a pure endowment with periodic premiums, see Sect. 18.4); this amendment raises the price of the original insurance product
- *Insurance tariff*: lists gross premiums for particular insurance products offered by insurance companies; the tariffs take into account:

  - insured's gender (see Sect. 7.2)
  - insured's age at entry: usually the difference between the calendar year of insurance contract and the calendar year of insured's birth
  - term of insurance (see *thereinbefore*)

- *Technical interest rate*: is the valuation interest rate (see Sect. 6.1) used in life insurance when pricing cash flows in calculations of premiums, premium reserves, and the like; its level may be regulated by the state (*state insurance* *supervision*); the low (or high) value of the technical interest rate increases (or decreases) the insurance price, respectively; typically, a conservative choice of this interest rate is typical in practice

- *Profit sharing*: is an allocation of a given part of insurer's technical gain to policyholders (see Sect. 18.9); the technical gain consists of (1) *investment gain* due to a positive difference between the actually earned interest rate and the technical interest rate used for insurance calculations, (2) *mortality gain* due to an appropriate difference in assumed and actual mortalities, and (3) *expense gain* due to an appropriate difference in assumed and actual expenses; the values assumed a priori in calculations are usually called *statutory values*
- *Life insurance options*: mean that in various phases of life insurance contracts clients have a possibility of choice among several alternatives without additional charges (e.g. an option to convert a sum insured into a life annuity)
- *Multiple life insurance* (see Sect. 18.8): are insurance products, in which insurance benefits depend on life status (mostly "alive" or "dead") of several lives (married couples, whole families, boards of directors, and the like)
- *Group insurance*: relates to a group of insured persons that exists because of another reason than the given insurance (mainly a life insurance for employees of an employer)
- *Comprehensive policy* (*insurance package*): is a multiple risk cover by a single insurance policy; various *riders* of basic policies are popular in practice (e.g. a classical life insurance policy with a rider for accident insurance or for dread diseases)

18.2 Symbols and Calculation Principles of Life Insurance

$\lozenge$	position of the corresponding actuarial symbol (e.g. A, P, a)
$\ddot{\lozenge}$, $\lozenge$	payment-due and immediate payment, respectively
$\lozenge_x$, $\lozenge_y$	one person aged x (mostly a male) and one person aged y (mostly a female)
$\lozenge_{x\,y\,\ldots}$	several persons aged x, y, ...
$\lozenge_{x:\overline{n}\rceil}$, $\lozenge_x$	temporary and perpetual payments, respectively
${}_{k\,\vert}\lozenge_x$, $\lozenge_x$	deferred payments and payments without deferment, respectively
$\lozenge^{(m)}$	payments payable $m$thly (mostly within 1 year)
$\bar{\lozenge}$	continuous payments (i.e. $m \to \infty$)
$I\,\lozenge$	unit increasing payments of the type 1, 2, ...
$D\,\lozenge$	unit decreasing payments of the type n, $n-1$, ...

- Cash flows (see Chap. 6) in classical life insurance:
  - *insurance benefits* (*settlements*) from insurers:
    - lump sum benefit (e.g. on survival, on death, on a stipulated date)
    - multiple (periodic) benefits within a stipulated period (e.g. life annuities)
    - constant or increasing or decreasing or general or continuous benefits, and the like

– *insurance premiums* from policyholders:

  – single premium (on the commencement date of policy issue, i.e. at time $t = 0$)
  – periodic premiums within a stipulated period (e.g. whole life or temporarily)
  – constant or increasing or decreasing or general or continuous premiums, and the like

– *net approach*: takes into account (as basis of calculation) only insurance benefits
– *gross approach*: takes into account not only insurance benefits, but also expenses by the insurer and other facts relevant in practice

• *Equivalence principle*: means that the insurance premiums are calculated in such a way that the expected present value (calculated at the time of policy issue, see Sect. 6.1) of the premiums is equal to the expected present value of the benefits; this principle can be also formulated by means of the zero expected total loss by the insurer, where the *total loss* is the difference between the present value of benefits and the present value of premiums; the equivalence principle is appropriate since (1) the discounting of cash flows may solve the problem of long-term horizons in life insurance and (2) the expected values may solve the problem of randomness of corresponding variables

Denotation:

Z	random variable representing the *present value of insurance benefits* (see Sect. 6.1) calculated at the time of policy issue ($t = 0$) by means of the technical interest rate (see Sect. 18.1)
$PV = \mathrm{E}(Z)$	*expected present value of insurance benefits* (see *thereinbefore*)
PV_P	*expected present value of insurance premiums* calculated at the time of policy issue ($t = 0$) by means of the technical interest rate; the expected value is necessary since periodic premium payments may be also of random character (e.g. they usually cease after the insured's death)
$L = PV - PV_P$	net expected loss by insurer
PV_E	*expected present value of expenses by insurer* calculated at the time of policy issue ($t = 0$) by means of the technical interest rate
$L^{gross} = (PV + PV_E) - PV_P$	gross expected loss by insurer

$$L = 0$$

(*equivalence principle for net values*: $PV = PV_P$; more generally, one can write $\mathrm{E}\{L(u)\} = 0$, where u is the difference between the present value of benefits and the present value of premiums calculated at the time of policy issue and $L(\cdot)$ is a suitable (e.g. asymmetric) loss function)

$$L^{gross} = 0$$

(*equivalence principle for gross values*: $PV + PV_E = PV_P$; in life insurance practice, the actual incomes mostly exceed the actual outcomes so that the accrued (unscheduled) profit is (partly) allocated to clients as their profit sharing (see Sect. 18.1)

$$\sigma(Z) = \sqrt{\mathrm{var}(Z)} = \left\{ \mathrm{E}(Z - PV_P)^2 \right\}^{1/2} = \left\{ \mathrm{E}(Z^2) - (PV_P)^2 \right\}^{1/2}$$

(*underwriting risk*: consists in a potentiality that the insurer does not achieve the balance between the incoming premium and the outcoming benefits (see Chap. 16))

$${}_hE_x = {}_hp_x \cdot v^h = \frac{l_{x+h} \cdot v^h}{l_x} = \frac{D_{x+h}}{D_x}$$

(*actuarial discounting*: the coefficient for multiplying a given actuarial value corresponding to age $x + h$ in order to shift it to age x, i.e. by h years backwards)

$$\frac{1}{{}_hE_x} = \frac{1}{{}_hp_x} \cdot (1+i)^h = \frac{l_x \cdot (1+i)^h}{l_{x+h}} = \frac{D_x}{D_{x+h}}$$

(*actuarial compounding*: the coefficient for multiplying a given actuarial value corresponding to age x in order to shift it to age $x + h$, i.e. by h years forwards)

- Expenses by insurer:

  α *acquisition expenses*: are charged against the policy as the percentage α of the sum insured, respectively as the percentage α of the annuity level; one usually distinguishes:

  - α^N *new business commission*: is the first-year provision for insurance agents, medical examination, and the like
  - α^C *collecting commission*: are provisions in further years for maintaining the policy in force

  β *collection expenses*: are charged at the beginning of every year in which a premium is to be collected as the percentage β of the expense-loaded (periodic) premium

  γ *administration expenses*: are charged against the policy at the beginning of every year during the entire contract period as the percentage γ of the sum insured, respectively as the percentage γ of the annuity level; various expenses are included in this item, such as wages, data processing costs, investment costs, taxes, licence fees, and the like; if the period in which premiums are collected is shorter than the entire contract period, then one sometimes distinguishes:

γ_1 administration expenses during period when premiums are collected
γ_2 administration expenses during period when premiums are no longer collected

δ *life annuity expenses*: are charged at the beginning of every year in which an annuity payment is to be paid as the percentage δ of the annuity level

ε *integrated expenses*: are charged by some insurers as the percentage ε of the expense-loaded premium; they include all types of expenses (see *thereinbefore*)

18.3 Technical Provisions in Life Insurance

- *Technical provisions* (*insurance reserves*): are established by insurers (mostly as the book costs according to law) to fulfil obligations arising from insurance activities (such obligations are probable or certain, though their amount or time may be still uncertain); the technical provisions are important *statutory liabilities* of each insurance company that are conform to special accounting principles and tax regulations; the assets covering the technical provisions are subject to strict investment restrictions since their *financial placement* should fulfil principles of *prudence*, *diversification*, *profitability* and *liquidity*; according to particular insurance legislations, various technical provisions may be established in life insurance (see also Sect. 21.4 for non-life insurance):

  - *premium reserve*: must be established due to the fact that at time t ($t > 0$) following the policy issue ($t = 0$) there is no longer an equivalence between future financial obligations of the insurer and the policyholder: the expected present value of future benefits will always exceed the expected present value of future level premiums at time t; this fact implies a positive difference which is a significant life insurer's liability called the premium reserve; moreover, one distinguishes (1) *gross premium reserve* (also called *expense-loaded premium reserve*), if the expenses by insurer (see Sect. 18.2) are included, and (2) *net premium reserve*, if it is not the case; the difference between the sum insured and the premium reserve (net or gross) is called the *amount at risk*
  - *reserve for unearned premium*: corresponds to such a part of the written premium that relates to future accounting periods; e.g. a quarter of an annual premium paid at the beginning of October covers the rest of the current year of account (the so-called *earned premium*), while the remaining three quarters (the so-called *unearned premium*) relate to the first 9 months of the next year of account so that one must establish the reserve for unearned premium in the current year of account for this purpose
  - *claim reserve*: covers obligations due to insured events (claims) which in the current accounting period have been:

  - *reported but not settled* (the so-called *RBNS reserve*)
  - *incurred but not reported* (the so-called *IBNR reserve*)

  (the claim reserves use not to be significant in life insurance (unlike the non-life insurance, see Sect. 21.4))

- *reserve for bonuses and rebates*: covers costs of bonuses and rebates guaranteed by insurance policies
- *life insurance reserve where investment risk is borne by policyholder*: applies to unit-linked insurance products (see Sect. 19.3)
- other life insurance reserves approved by authorities (by the state insurance supervision)

Denotation:

${}_tZ$	random variable representing the present value of insurance benefits (see Sect. 6.1) calculated at time t ($t > 0$) following the policy issue ($t = 0$) by means of the technical interest rate (see Sect. 18.1)
${}_tPV = \mathrm{E}({}_tZ)$	expected present value of insurance benefits calculated at time t (see *thereinbefore*)
${}_tPV_P$	expected present value of insurance premiums calculated at time t
${}_tV_x = {}_tPV - {}_tPV_P$	*net premium reserve* at time t for a life aged x at entry (*prospective method*: the reserve is calculated as the net expected loss by insurer at time t (${}_0V_x = 0$))
${}_tFV$	expected future (i.e. final) value of insurance benefits (see Sect. 6.1) calculated at time t
${}_tFV_P$	expected future (i.e. final) value of premiums calculated at time t
${}_tV_x = {}_tFV_P - {}_tFV$	*net premium reserve* at time t for a life aged x at entry (*retrospective method*: the reserve is calculated as the net past profit by insurer at time t (${}_0V_x = 0$); the method gives the same result as the prospective one (see *thereinbefore*))
${}_tPV_E$	expected present value of expenses by insurer calculated at time t
${}_tV_x{}^{gross} = ({}_tPV + {}_tPV_E) - {}_tPV_P$	*gross premium reserve* (also called *expense-loaded premium reserve*) at time t for a life aged x at entry (prospective method: the retrospective one is analogous)

- Formulas of the net and gross premium reserves for particular insurance products (see *thereinafter*) are presented in the *prospective form* (in practice, the prospective formulas are preferred since (1) they enable to carry out comfortably various future changes in the insurance policy; (2) they may be simpler at time t than the retrospective ones when premiums are no longer collected at time t (in particular, this scenario holds for products with single premiums)):

$$ {}_tV_{x:\overline{n}|} = \frac{\sum_{j=t+1}^{n} (a_j \cdot D_{x+j} + b_j \cdot C_{x+j-1})}{D_{x+t}} - \frac{P_{x:\overline{n}|} \cdot \sum_{j=t+1}^{n} D_{x+j-1}}{D_{x+t}} $$

(*net premium reserve* at the end of the year t of an n-year insurance for a life aged x at entry (analogously for perpetual insurances): $P_{x:\overline{n}|}$ is the annual premium paid always at the beginning of further year of insurance; a_t is the stipulated benefit paid on survival of the end of the year t of insurance; b_t is the stipulated benefit paid at the end of the year t of insurance on death within this year; when premiums are no longer collected at time t (in particular, in products with single premiums), then the second term (the subtrahend) in the given formula is dropped out; some insurance products (e.g. the term insurance, see Sect. 18.5) establish so small premium reserves that such reserves may be ignored in practice: hence the insurance products may be classified to *capitalizing* and *non-capitalizing* ones)

$$ {}_tV_{x:\overline{n}|} = \frac{P_{x:\overline{n}|} \cdot \sum_{j=1}^{t} D_{x+j-1}}{D_{x+t}} - \frac{\sum_{j=1}^{t} (a_j \cdot D_{x+j} + b_j \cdot C_{x+j-1})}{D_{x+t}} $$

(*retrospective form of net premium reserve*: is equal to the prospective one (see *thereinbefore*); the retrospective form is not usual in practice)

$$ {}_tV_{x:\overline{n}|} \cdot D_{x+t} = ({}_{t-1}V_{x:\overline{n}|} + P_{x:\overline{n}|}) \cdot D_{x+t-1} - (a_t \cdot D_{x+t} + b_t \cdot C_{x+t-1}) $$

(*recursive form of net premium reserve*: is used to derive some relations, e.g. to decompose the period premium to the saving premium and the risk premium (see *thereinafter*))

$$ P^s_{x:\overline{n}|}(t) = {}_tV_{x:\overline{n}|} \cdot v - {}_{t-1}V_{x:\overline{n}|}; \qquad P^r_{x:\overline{n}|}(t) = \frac{a_t \cdot D_{x+t} + (b_t - {}_tV_{x:\overline{n}|}) \cdot C_{x+t-1}}{D_{x+t-1}} $$

(*saving premium* at time t: serves to increase the net premium reserve at the beginning of the year t in addition to the interest compounding; *risk premium* at time t: covers on average the risk that the insurer will pay off at the end of tth year the benefit amounting $a_t \cdot p_{x+t-1} \cdot v + b_t \cdot q_{x+t-1} \cdot v = a_t \cdot (D_{x+t}/D_{x+t-1}) + b_t \cdot (C_{x+t-1}/D_{x+t-1})$ making use of the fund ${}_tV_{x:\overline{n}|} \cdot q_{x+t-1} \cdot v = {}_tV_{x:\overline{n}|} \cdot (C_{x+t-1}/D_{x+t-1})$ released from the net premium reserve due to death event within the year t (see e.g. Sect. 18.6 for the endowment); if the death event occurs in a capitalizing product (see *thereinbefore*), then the sum insured required from the insurer as the benefit is composed by two sources: (1) by the net premium reserve established by saving premiums of the given policyholder and (2) by the *amount at risk* (which is the difference between the sum insured and the premium reserve at the given time) established by risk premiums of all policyholders (the mechanism works in such a way that the amounts at risk for the policies with death events are covered by risk premiums across the whole insurance portfolio))

$$ {}_tV^{gross}_{x:\overline{n}|}(P^{gross}) \;=\; {}_tV_{x:\overline{n}|}(P) - \alpha^N \cdot \frac{\ddot{a}_{x+t:\,\overline{n-t}|}}{\ddot{a}_{x:\overline{n}|}} \;=\; {}_tV_{x:\overline{n}|}(P) - \alpha^N \cdot (1 - \,{}_tV_{x:\overline{n}|}(P)) $$

(*gross premium reserve for periodic premiums* (analogously for perpetual insurances): in contrast to the net premium reserve, the new business commission α^N (i.e. the first-year provision, see Sect. 18.2) symbols in the formula; the term which is subtracted from the net premium reserve in the formula is called the *zillmerising term* (or the negative acquisition expenses reserve); its subtraction from the net premium reserve is called *zillmerisation* and has the following interpretation: (1) the new business commission α^N (amounting significant values nowadays) is expended by the insurer immediately at the time of policy issue $\rightarrow$ (2) however, this amount due is paid back gradually in particular periodic premium payments (hence, the insurer becomes the "creditor" of the policyholder) $\rightarrow$ (3) therefore at the given time t the insurer always reduces the net premium reserve by the non-amortized part of the new business commission at time t, which is just the zillmerising term at time t; the zillmerisation may produce negative values of the gross premium reserve (usually in initial years of long-term policies), which are mostly replaced by zero values in practice)

$$ {}_tV^{gross}_{x:\overline{n}|}(SP^{gross}) \;=\; {}_tV_{x:\overline{n}|}(SP^{gross}) + (\alpha^C + \gamma) \cdot \ddot{a}_{x+t:\,\overline{n-t}|} $$

(*gross premium reserve for single premium* (analogously for perpetual insurances): in contrast to the net premium reserve, the expenses α^C and γ (see Sect. 18.2) symbol in the formula; the term which is added to the net premium reserve in the formula should just cover these future costs)

$$ {}_{t+\frac{13-m}{12}}V_{x:\overline{n}|} = \frac{m-1}{12}\,{}_tV_{x:\overline{n}|} + \frac{13-m}{12}\,{}_{t+1}V_{x:\overline{n}|} $$

(net premium reserve *at fractional durations*: *monthly* values of the net premium reserve with *monthly* premiums (analogously for perpetual insurances or for gross premium reserves at fractional durations))

$$ {}_{t+\frac{13-m}{12}}V_{x:\overline{n}|} = \frac{m-1}{12}\,{}_tV_{x:\overline{n}|} + \frac{13-m}{12}\,{}_{t+1}V_{x:\overline{n}|} + \frac{m-1}{12}\,P_x $$

(net premium reserve *at fractional durations*: *monthly* values of the net premium reserve with *annual* premiums (analogously for perpetual insurances or for gross premium reserves at fractional durations))

18.4 Pure Endowments

Denotation (see also Sects. 17.3 and 18.2, it holds for further insurance products in Chap. 18, as well):

Z	random variable representing the present value of insurance benefits calculated at the time of policy issue by means of the technical interest rate (see Sect. 18.2)
T_x	future lifetime at age x (see Sect. 17.3)
K_x	curtate future lifetime at age x (see Sect. 17.3)
$\sigma(Z)$	underwriting risk (see Sect. 18.2)
SP_x	net single premium
$P_{x:\overline{n}\rceil}$, P_x	net annual premium
SP_x^{gross}	gross single premium
$P_{x:\overline{n}\rceil}^{gross}$, P_x^{gross}	gross annual premium
${}_tV_{x:\overline{n}\rceil}$, ${}_tV_x$	net premium reserve (see Sect. 18.3)
${}_tV_{x:\overline{n}\rceil}^{gross}$, ${}_tV_x^{gross}$	gross premium reserve (see Sect. 18.3)

- *Pure endowment*: provides the payment of the sum insured only if the insured aged x at entry is alive at the end of the stipulated term of n years; if the insured dies during the term of insurance, then the policy ceases without remuneration (therefore in the case of periodic premiums, the return of the premiums on death is usual in practice, though this amendment raises the price of the original pure endowment (see *thereinafter*)):

$$SP_{x:\overline{n}\rceil} = {}_nE_x = A_{x:\overline{n}\rceil}^{\;\;1} = \frac{D_{x+n}}{D_x} \text{ (}n\text{-year pure endowment)}$$

$$= \mathrm{E}(Z) = {}_np_x \cdot v^n, \text{ where } Z = \begin{cases} 0\,, & K_x = 0,\ 1,\ \ldots,\ n-1 \\ v^n, & K_x = n,\ n+1,\ \ldots \end{cases} \text{ according to the table:}$$

Value of K_x	Value of Z	Probability
0	0	${}_{0\mid}q_x = q_x$
1	0	${}_{1\mid}q_x$
2	0	${}_{2\mid}q_x$
$\vdots$	$\vdots$	$\vdots$
$n-1$	0	${}_{n-1\mid}q_x$
n, $n+1$, $\vdots$	v^n	${}_np_x$

$$\sigma(Z) = \left\{ {}_np_x \cdot v^{2n} - ({}_np_x \cdot v^n)^2 \right\}^{1/2} = \{{}_np_x \cdot {}_nq_x\}^{1/2} \cdot v^n = \left\{ {}_n^2E_x - ({}_nE_x)^2 \right\}^{1/2},$$

where ${}_n^2E_x$ is as ${}_nE_x$ calculated with the discount factor v^2 instead of v

$$P_{x:\overline{n}|} = \frac{{}_nE_x}{\ddot{a}_{x:\overline{n}|}} = \frac{D_{x+n}}{N_x - N_{x+n}}$$

$$P^{(m)}_{x:\overline{n}|} = \frac{{}_nE_x}{m \cdot \ddot{a}^{(m)}_{x:\overline{n}|}} \approx \frac{{}_nE_x}{m \cdot \left[\ddot{a}_{x:\overline{n}|} - \frac{m-1}{2m} \cdot \left(1 - \frac{D_{x+n}}{D_x} \right) \right]}$$

$$SP^{gross}_x = {}_nE_x + \alpha^N + (\alpha^C + \gamma) \cdot \ddot{a}_{x:\overline{n}|}$$

$$P^{gross}_{x:\overline{n}|} = \frac{{}_nE_x + \alpha^N + (\alpha^C + \gamma) \cdot \ddot{a}_{x:\overline{n}|}}{(1-\beta) \cdot \ddot{a}_{x:\overline{n}|}} = \frac{1}{1-\beta} \cdot \left({}_nP_x + \frac{\alpha^N}{\ddot{a}_{x:\overline{n}|}} + \alpha^C + \gamma \right)$$

$$P^{gross+remuneration}_{x:\overline{n}|} = \frac{{}_nE_x + \alpha^N + (\alpha^C + \gamma) \cdot \ddot{a}_{x:\overline{n}|}}{(1-\beta) \cdot \ddot{a}_{x:\overline{n}|} - (IA)^1_{x:\overline{n}|}}$$

($(IA)^1_{x:\overline{n}|}$ is defined in Sect. 18.5)
(pure endowment with return of level premiums on insured's death during insurance term)

$$\begin{aligned}
{}_tV_{x:\overline{n}|}(SP_x) &= {}_{n-t}E_{x+t} = \frac{D_{x+n}}{D_{x+t}} \\
{}_tV_{x:\overline{n}|}(P_{x:\overline{n}|}) &= {}_{n-t}E_{x+t} - P_{x:\overline{n}|} \cdot \ddot{a}_{x+t:\,\overline{n-t}|} = \frac{D_{x+n}}{D_{x+t}} \cdot \frac{N_x - N_{x+t}}{N_x - N_{x+n}} \\
{}_tV^{gross}_{x:\overline{n}|}(SP^{gross}_x) &= {}_tV_{x:\overline{n}|}(SP_x) + (\alpha^C + \gamma) \cdot \ddot{a}_{x+t:\,\overline{n-t}|} \\
{}_tV^{gross}_{x:\overline{n}|}(P^{gross}_{x:\overline{n}|}) &= {}_tV_{x:\overline{n}|}(P_{x:\overline{n}|}) - \alpha^N \cdot \frac{\ddot{a}_{x+t:\,\overline{n-t}|}}{\ddot{a}_{x:\overline{n}|}}
\end{aligned}$$

$$\begin{aligned}
{}_tV^{gross}_{x:\overline{n}|}(B^{gross+remuneration}_{x:\overline{n}|}) &= {}_tV_{x:\overline{n}|}(P_{x:\overline{n}|}) - \alpha^N \cdot \frac{\ddot{a}_{x+t:\,\overline{n-t}|}}{\ddot{a}_{x:\overline{n}|}} \\
&\quad + P^{gross}_{x:\overline{n}|} \left[\frac{R_{x+t} - R_{x+n} + t \cdot M_{x+t} - n \cdot M_{x+n}}{D_{x+t}} \right. \\
&\quad \left. - \frac{R_x - R_{x+n} - n \cdot M_{x+n}}{D_{x+t}} \cdot \frac{N_{x+t} - N_{x+n}}{N_x - N_{x+n}} \right]
\end{aligned}$$

18.5 Whole Life and Term Insurance

- *Whole life insurance*: provides the payment of the sum insured at the end of the year of death for the insured aged x at entry (see Sects. 18.2 and 18.4 for denotation):

$$SP_x = A_x = \frac{M_x}{D_x} \quad (\textit{whole life})$$
$$= \mathrm{E}(Z) = \sum_{k=0}^{\infty} {}_{k|}q_x \cdot v^{k+1} = \sum_{k=0}^{\infty} {}_k p_x \cdot q_{x+k} \cdot v^{k+1}$$

where $Z = v^{K_x+1}$ according to the table:

Value of K_x	Value of Z	Probability
0	v	${}_{0\vert}q_x = q_x$
1	v^2	${}_{1\vert}q_x$
2	v^3	${}_{2\vert}q_x$
$\vdots$	$\vdots$	$\vdots$

$$\sigma(Z) = \left\{ {}^2A_x - (A_x)^2 \right\}^{1/2},$$

where ${}^2A_x$ is as A_x calculated with the discount factor v^2 instead of v

$$\sigma\left(v^{K_x+1} - P_x \cdot \ddot{a}_{\overline{K_x+1}|}\right) = \left(1 + \frac{P_x}{d}\right) \cdot \left\{ {}^2A_x - (A_x)^2 \right\}^{1/2} > \left\{ {}^2A_x - (A_x)^2 \right\}^{1/2}$$

(underwriting risk for whole life insurance with annual premiums)

$$SP_x = \bar{A}_x = \mathrm{E}(Z) = \mathrm{E}(v^{T_x}) = \int_0^{\infty} {}_t p_x \cdot \mu_{x+t} \cdot v^t \, \mathrm{d}t \approx (1+i)^{1/2} \cdot A_x$$

(*continuous approach*: provides the payment of the sum insured at the moment of death (without waiting for the end of the year of death, see *thereinbefore*))

$$\sigma(Z) = \left\{ {}^2\bar{A}_x - (\bar{A}_x)^2 \right\}^{1/2}$$

(*continuous approach*: ${}^2\bar{A}_x$ is as $\bar{A}_x$ calculated with the discount factor v^2 instead of v)

$$P_x = \frac{A_x}{\ddot{a}_x} = \frac{1}{\ddot{a}_x} - d = \frac{M_x}{N_x}; \quad P_x^{(m)} = \frac{A_x}{m \cdot \ddot{a}_x^{(m)}} \approx \frac{A_x}{m \cdot \left(\ddot{a}_x - \frac{m-1}{2m}\right)}$$

$$P_x = \frac{\bar{A}_x}{\ddot{a}_x} \text{ (net annual premium applying continuous approach)}$$

$$P_{x:\overline{k}|} = \frac{A_x}{\ddot{a}_{x:\overline{k}|}} = \frac{M_x}{N_x - N_{x+k}} \quad \text{(net annual premium payable } k \text{ years } (k \le n))$$

$$SP_x^{gross} = A_x + \alpha^N + (\alpha^C + \gamma) \cdot \ddot{a}_x$$

$$P_x^{gross} = \frac{A_x + \alpha^N}{(1-\beta)\cdot \ddot{a}_x} + \frac{\alpha^C + \gamma}{1-\beta}; \quad P_{x:\overline{k}|}^{gross} = \frac{A_x + \alpha^N + (\alpha^C + \gamma_1)\cdot \ddot{a}_{x:\overline{k}|} + \gamma_2 \cdot {}_{k|}\ddot{a}_x}{(1-\beta)\cdot \ddot{a}_x}$$

$${}_tV_x(SP_x) = A_{x+t} = \frac{M_{x+t}}{D_{x+t}}$$

$$\begin{aligned} {}_tV_x(P_x) &= A_{x+t} - P_x \cdot \ddot{a}_{x+t} = 1 - (P_x + d)\cdot \ddot{a}_{x+t} = 1 - \frac{\ddot{a}_{x+t}}{\ddot{a}_x} = \left(1 - \frac{P_x}{P_{x+t}}\right)\cdot A_{x+t} \\ &= \left(1 - \frac{P_x}{P_{x+t}}\right)\cdot A_{x+t} = (P_{x+t} - P_x)\cdot \ddot{a}_{x+t} = \frac{P_{x+t} - P_x}{P_{x+t} + d} = 1 - \frac{D_x}{D_{x+t}}\cdot \frac{N_{x+t}}{N_x} \end{aligned}$$

${}_tV_x(P_{x:\overline{k}|}) = A_{x+t} - P_{x:\overline{k}|} \cdot \ddot{a}_{x+t:\,\overline{k-t}|}$ (for annual premiums payable k years $(k \le n)$)

$$\sigma\left(v^{K_{x+t}+1} - P_x \cdot \ddot{a}_{\overline{K_{x+t}+1}|}\right) = \left(1 + \frac{P_x}{d}\right)\cdot \left\{{}^2A_{x+t} - (A_{x+t})^2\right\}^{1/2}$$

(net reserve risk for whole life insurance with annual premiums)

$$\begin{aligned} {}_tV_x^{gross}(SP_x^{gross}) &= {}_tV_x(SP_x) + (\alpha^C + \gamma)\cdot \ddot{a}_{x+t} \\ V_x^{gross}(P_x^{gross}) &= {}_tV_x(P_x) - \alpha^N \cdot \frac{\ddot{a}_{x+t}}{\ddot{a}_x} \end{aligned}$$

$$A_x = 1 - d\cdot \ddot{a}_x = v\cdot \ddot{a}_x - a_x$$

$$A_{x+1} = \frac{D_x}{D_{x+1}}\cdot\left(A_x - \frac{C_x}{D_x}\right)$$

$${}_{k|}A_x = {}_kE_x \cdot A_{x+k} = \frac{M_{x+k}}{D_x} \quad \text{(deferment by } k \text{ years)}$$

$$(IA)_x = \sum_{k=0}^{\infty}(k+1)\cdot {}_kp_x \cdot q_{x+k}\cdot v^{k+1} = \frac{R_x}{D_x}$$

(whole life insurance with increasing sum insured of the type 1, 2, . . .)

$$\begin{aligned} (I_{\overline{h}|}A)_x &= \sum_{k=0}^{h-1}(k+1)\cdot {}_kp_x \cdot q_{x+k}\cdot v^{k+1} + h\cdot \sum_{k=h}^{\infty} {}_kp_x \cdot q_{x+k}\cdot v^{k+1} \\ &= \frac{R_x - R_{x+h} - h\cdot M_{x+h}}{D_x} \end{aligned}$$

(whole life insurance with varied sum insured of the type 1, 2, . . ., h, h, h, . . .)

$$(D_{h\rceil}A)_x = \sum_{k=0}^{h-1} (h-k) \cdot {}_kp_x \cdot q_{x+k} \cdot v^{k+1} + \sum_{k=h}^{\infty} {}_kp_x \cdot q_{x+k} \cdot v^{k+1}$$
$$= \frac{(h+1) \cdot M_x - M_{x+h} - R_x + R_{x+h}}{D_x}$$

(whole life insurance with varied sum insured of the type h, $h-1$, ..., 1, 1, 1, . . .)

- *Term insurance*: provides the payment of the sum insured at the end of the year of death for the insured aged x at entry, only if death occurs within the first n years (if the insured survives n years, then the policy ceases without remuneration); in practice, (mortgage) banks make use of it to cover the mortality risk during amortization of credits, i.e. the debtor is the term insured and the bank becomes the beneficiary (see Sects. 18.2 and 18.4 for denotation):

$$SP_x = A^1_{x:n\rceil} = \frac{M_x - M_{x+n}}{D_x} \ (n\text{-}year\ term\ insurance)$$
$$= \mathrm{E}(Z) = \sum_{k=0}^{n-1} {}_{k|}q_x \cdot v^{k+1} = \sum_{k=0}^{n-1} {}_kp_x \cdot q_{x+k} \cdot v^{k+1},$$

$$\text{where } Z = \begin{cases} v^{K_x+1}, & K_x = 0,\ 1,\ \ldots,\ n-1 \\ 0 & K_x = n,\ n+1,\ \ldots \end{cases} \text{ according to the table:}$$

Value of K_x	Value of Z	Probability
0	v	${}_{0\|}q_x = q_x$
1	v^2	${}_{1\|}q_x$
2	v^3	${}_{2\|}q_x$
$\vdots$	$\vdots$	$\vdots$
$n-1$	v^n	${}_{n-1\|}q_x$
n, $n+1$, $\vdots$	0	${}_np_x$

$$\sigma(Z) = \left\{ {}^2A^1_{x:n\rceil} - (A^1_{x:n\rceil})^2 \right\}^{1/2},$$

where ${}^2A^1_{x:n\rceil}$ is as $A^1_{x:n\rceil}$ calculated with the discount factor v^2 instead of v

$$SP_x = \bar{A}^1_{x:\overline{n}|} = \int_0^n {}_tp_x \cdot \mu_{x+t} \cdot v^t \, \mathrm{d}t \approx (1+i)^{1/2} \cdot A^1_{x:\overline{n}|}$$

(*continuous approach*: provides for payment of the sum insured at the moment of death)

$$P_{x:\overline{n}|} = \frac{A^1_{x:\overline{n}|}}{\ddot{a}_x} = \frac{M_x - M_{x+n}}{N_x - N_{x+n}}$$

$$P^{(m)}_{x:\overline{n}|} = \frac{A^1_{x:\overline{n}|}}{m \cdot \ddot{a}^{(m)}_{x:\overline{n}|}} \approx \frac{A^1_{x:\overline{n}|}}{m \cdot \left[\ddot{a}_{x:\overline{n}|} - \frac{m-1}{2m} \cdot \left(1 - \frac{D_{x+n}}{D_x}\right)\right]}$$

$$SP^{gross}_x = A^1_{x:\overline{n}|} + \alpha^N + (\alpha^C + \gamma) \cdot \ddot{a}_{x:\overline{n}|}$$

$$P^{gross}_{x:\overline{n}|} = \frac{A^1_{xn} + \alpha^N}{(1-\beta) \cdot \ddot{a}_{x:\overline{n}|}} + \frac{\alpha^C + \gamma}{1-\beta}$$

$${}_tV_{x:\overline{n}|}(SP_x) = A^1_{x+t:\,\overline{n-t}|} = \frac{M_{x+t} - M_{x+n}}{D_{x+t}}$$

$${}_tV_{x:\overline{n}|}(P_{x:\overline{n}|}) = A^1_{x+t:\,\overline{n-t}|} - P_{x:\overline{n}|} \cdot \ddot{a}_{x+t:\,\overline{n-t}|}$$

$$= \frac{M_{x+t} - M_{x+n}}{D_{x+t}} - \frac{M_x - M_{x+n}}{D_{x+t}} \cdot \frac{N_{x+t} - N_{x+n}}{N_x - N_{x+n}}$$

$${}_tV^{gross}_{x:\overline{n}|}(SP^{gross}_x) = {}_tV_{x:\overline{n}|}(SP_x) + (\alpha^C + \gamma) \cdot \ddot{a}_{x+t:\,\overline{n-t}|}$$

$${}_tV^{gross}_{x:\overline{n}|}(P^{gross}_{x:\overline{n}|}) = {}_tV_{x:\overline{n}|}(P_{x:\overline{n}|}) - \alpha^N \cdot \frac{\ddot{a}_{x+t:\,\overline{n-t}|}}{\ddot{a}_{x:\overline{n}|}}$$

$$A^1_{x:\overline{n}|} = 1 - d \cdot \ddot{a}_{x:\overline{n}|} - {}_nE_x = v \cdot \ddot{a}_{x:\overline{n}|} - a_{x:\overline{n}|}$$

$$A^1_{x+1:\,\overline{n}|} = \frac{D_x}{D_{x+1}} \cdot \left(A^1_{x:\overline{n}|} - \frac{C_x - C_{x+n}}{D_x}\right)$$

$${}_{k|}A^1_{x:\overline{n}|} = {}_kE_x \cdot A^1_{x+k:\overline{n}|} = \frac{M_{x+k+n} - M_{x+k}}{D_x} \quad \text{(deferment by } k \text{ years)}$$

$$A_x = A^1_{x:\overline{n}|} + {}_{n|}A_x$$

(decomposition of whole life insurance to term and deferred ones)

$$(IA)^1_{x:\overline{n}|} = \ddot{a}_{x:\overline{n}|} - d \cdot (I\ddot{a})_{x:\overline{n}|} = \frac{R_x - R_{x+n} - n \cdot M_{x+n}}{D_x}$$

(term insurance with increasing sum insured of the type 1, 2, …, n)

$$(DA)^{1}_{x:\overline{n}|} = \frac{n \cdot M_x - R_{x+1} + R_{x+n+1}}{D_x}$$

(term insurance with decreasing sum insured of the type n, $n-1$, . . ., 1)

18.6 Further Products of Capital Life Insurance

- *Endowment*: provides the payment of the sum insured at the end of the year of death for the insured aged x at entry, only if death occurs within the first n years, otherwise at the end of the stipulated term of n years (i.e. the sum insured is payable on death or survival according to which alternative occurs earlier); products of this type belong to the most popular ones in Europe (see Sects. 18.2 and 18.4 for denotation):

$$SP_x = A_{x:\overline{n}|} = \frac{M_x - M_{x+n} + D_{x+n}}{D_x} \; (n\text{-}year\ endowment)$$

$$= \mathrm{E}(Z) = \sum_{k=0}^{n-1} {}_{k|}q_x \cdot v^{k+1} + {}_{n}p_x \cdot v^n = \sum_{k=0}^{n-2} {}_{k|}q_x \cdot v^{k+1} + {}_{n-1}p_x \cdot v^n,$$

$$\text{where } Z = \begin{cases} v^{K_x+1}, & K_x = 0,\ 1,\ \ldots,\ n-1 \\ v^n, & K_x = n,\ n+1,\ \ldots \end{cases} \quad \text{according to the table :}$$

Value of K_x	Value of Z	Probability
0	v	${}_{0\|}q_x = q_x$
1	v^2	${}_{1\|}q_x$
2	v^3	${}_{2\|}q_x$
$\vdots$	$\vdots$	$\vdots$
$n-2$	v^{n-1}	${}_{n-2\|}q_x$
$n-1$, n, $n+1$, $\ldots$	v^n	${}_{n-1\|}q_x + {}_n p_x = {}_{n-1} p_x$

$$\sigma(Z) = \left\{ {}^2A_{x\overline{n}|} - (A_{x\overline{n}|})^2 \right\}^{1/2},$$

where ${}^2A_{x\overline{n}|}$ is as $A_{x\overline{n}|}$ calculated with the discount factor v^2 instead of v

$$SP_x^{S/D} = \frac{S \cdot (M_x - M_{x+n}) + D \cdot D_{x+n}}{D_x}$$

(endowment with the sum insured S on death and the sum insured D on survival)

$$SP_x = \bar{A}_{x:\overline{n}|} = \bar{A}^{\,1}_{x:\overline{n}|} + {}_nE_x = \int_0^n {}_tp_x \cdot \mu_{x+t} \cdot v^t \,\mathrm{d}t \; + \; {}_np_x \cdot v^n \approx \frac{(1+i)^{1/2} \cdot (M_x - M_{x+n}) + D_{x+n}}{D_x}$$

(*continuous approach*: on death provides the payment of the sum insured at the very moment)

$$P_{x:\overline{n}|} = \frac{A_{x:\overline{n}|}}{\ddot{a}_{x:\overline{n}|}} = \frac{M_x - M_{x+n} + D_{x+n}}{N_x - N_{x+n}}$$

$$P^{(m)}_{x:\overline{n}|} = \frac{A_{x:\overline{n}|}}{m \cdot \ddot{a}^{(m)}_{x:\overline{n}|}} \approx \frac{A_{x:\overline{n}|}}{m \cdot \left[\ddot{a}_{x:\overline{n}|} - \frac{m-1}{2m} \cdot \left(1 - \frac{D_{x+n}}{D_x}\right)\right]}$$

$$P_{x:\overline{k}|} = \frac{A_{x:\overline{n}|}}{\ddot{a}_{x:\overline{k}|}} = \frac{M_x - M_{x+n} + D_{x+n}}{N_x - N_{x+k}} \quad \text{(net annual premium payable } k \text{ years)}$$

$$P^{gross}_{x:\overline{n}|} = \frac{A_{x:\overline{n}|} + \alpha^N + (\alpha^C + \gamma) \cdot \ddot{a}_{x:\overline{n}|}}{(1-\beta) \cdot \ddot{a}_{x:\overline{n}|}} = \frac{1}{1-\beta} \cdot \left(P_{x:\overline{n}|} + \frac{\alpha^N}{\ddot{a}_{x:\overline{n}|}} + \alpha^C + \gamma\right)$$

$$P^{gross\,(m)}_{x:\overline{n}|} = \frac{A_{x:\overline{n}|} + \alpha^N + (\alpha^C + \gamma) \cdot \ddot{a}_{x:\overline{n}|}}{(1-\beta) \cdot m \cdot \ddot{a}^{(m)}_{x:\overline{n}|}}$$

$$\approx \frac{A_{x:\overline{n}|} + \alpha^N + (\alpha^C + \gamma) \cdot \ddot{a}_{x:\overline{n}|}}{(1-\beta) \cdot m \cdot \left[\ddot{a}_{x:\overline{n}|} - \frac{m-1}{2m} \cdot \left(1 - \frac{D_{x+n}}{D_x}\right)\right]}$$

$$P^{gross}_{x:\overline{k}|} = \frac{A_{x:\overline{n}|} + \alpha^N + (\alpha^C + \gamma_1) \cdot \ddot{a}_{x:\overline{k}|} + \gamma_2 \cdot {}_{k|}\ddot{a}_{x:\overline{n}|}}{(1-\beta) \cdot \ddot{a}_{x:\overline{k}|}}$$

(gross annual premium payable k years)

$${}_tV_{x:\overline{n}|}(SP_x) = A_{x+t:\,\overline{n-t}|} = \frac{M_{x+t} - M_{x+n} + D_{x+n}}{D_{x+t}}$$

$$\begin{aligned}
{}_tV_{x:\overline{n}|}(P_{x:\overline{n}|}) &= A_{x+t:\,\overline{n-t}|} - P_{x:\overline{n}|} \cdot \ddot{a}_{x+t:\,\overline{n-t}|} \\
&= \frac{1}{{}_tE_x}(P_{x:\overline{n}|} \cdot \ddot{a}_{x:\overline{t}|} - A_{x:\overline{t}|}) = 1 - \frac{\ddot{a}_{x+t:\,\overline{n-t}|}}{\ddot{a}_{x:\overline{n}|}} \\
&= \left(\frac{A_{x+t:\,\overline{n-t}|}}{\ddot{a}_{x+t:\,\overline{n-t}|}} - P_{x:\overline{n}|}\right) \cdot \ddot{a}_{x+t:\,\overline{n-t}|} = (P_{x+t:\,n-t} - P_{x:\overline{n}|}) \cdot \ddot{a}_{x+t:\,\overline{n-t}|} \\
&= \left(1 - P_{x:\overline{n}|} \cdot \frac{\ddot{a}_{x+t:\,\overline{n-t}|}}{A_{x+t:\,\overline{n-t}|}}\right) \cdot A_{x+t:\,\overline{n-t}|} \\
&= \left(1 - \frac{P_{x:\overline{n}|}}{P_{x+t:\,\overline{n-t}|}}\right) \cdot A_{x+t:\,\overline{n-t}|} \\
&= 1 - \frac{D_x}{D_{x+t}} \cdot \frac{N_{x+t} - N_{x+n}}{N_x - N_{x+n}}
\end{aligned}$$

$$P^s_{x:\overline{n}|}(t) = v \cdot {}_tV_{x:\overline{n}|}(P_{x:\overline{n}|}) - {}_{t-1}V_{x:\overline{n}|}(P_{x:\overline{n}|});\ P^r_{x:\overline{n}|}(t) = v \cdot q_{x+t-1} \cdot (1 - {}_tV_{x:\overline{n}|}(P_{x:\overline{n}|}))$$

(saving and risk premium at time t: $P_{x:\overline{n}|} = P^s_{x:\overline{n}|}(t) + P^r_{x:\overline{n}|}(t)$)

$$ {}_tV_{x:\overline{n}|}(P_{x:\overline{k}|}) = \begin{cases} A_{x+t:\,\overline{n-t}|} - P_{x:\overline{n}|} \cdot \ddot{a}_{x+t:\,\overline{n-t}|}, & t < k \\ A_{x+t:\,n-t}, & t \geq k \end{cases} $$

(for annual premiums payable k years)

$$ {}_t\bar{V}_{x:\overline{n}|}(\bar{P}_{x:\overline{n}|}) = \bar{A}_{x+t:\,\overline{n-t}|} - \bar{P}_{x:\overline{n}|} \cdot \bar{a}_{x+t:\,\overline{n-t}|} $$

$$ = 1 - \frac{\bar{a}_{x+t:\,\overline{n-t}|}}{\bar{a}_{x:\overline{n}|}} = 1 - \frac{\int_0^{n-t} {}_sp_{x+t} \cdot v^s \,\mathrm{d}s}{\int_0^{n} {}_sp_{x} \cdot v^s \,\mathrm{d}s} $$

$$ {}_tV^{gross}_{x:\overline{n}|}(SP^{gross}_x) = {}_tV_{x:\overline{n}|}(SP_x) + (\alpha^C + \gamma) \cdot \ddot{a}_{x+t:\,\overline{n-t}|} $$

$$ {}_tV^{gross}_{x:\overline{n}|}(P^{gross}_{x:\overline{n}|}) = {}_tV_{x:\overline{n}|}(P_{x:\overline{n}|}) - \alpha^N \cdot \frac{\ddot{a}_{x+t:\,\overline{n-t}|}}{\ddot{a}_{x:\overline{n}|}} $$

$$ A_{x:\overline{n}|} = 1 - d \cdot \ddot{a}_{x:\overline{n}|} = v \cdot \ddot{a}_{x:\overline{n}|} - a_{x:\,\overline{n-1}|} $$

$$ A_{x:\overline{n}|} = A^1_{x:\overline{n}|} + {}_nE_x = A^1_{x:\overline{n}|} + A_{x:\overset{1}{\overline{n}|}} $$

(decomposition of endowment insurance to term and pure endowment ones)

$$ (IA)_{x:\overline{n}|} = (IA)^1_{x:\overline{n}|} + n \cdot {}_nE_x = \frac{R_x - R_{x+n} - n \cdot M_{x+n} + n \cdot D_{x+n}}{D_x} $$

(endowment with increasing sum insured of the type 1, 2, . . ., n)

$$ (DA)_{x:\overline{n}|} = (DA)^1_{x:\overline{n}|} + {}_nE_x = \frac{n \cdot M_x - R_{x+1} + R_{x+n+1} + D_{x+n}}{D_x} $$

(endowment with decreasing sum insured of the type n, $n-1$, . . ., 1)

- *Fixed term insurance*: provides the payment of the sum insured at the end of the stipulated term of n years regardless as to whether the insured aged x at entry is alive at age $x + n$ or died in the meantime; the payments of periodical premiums cease on death of the insured within the stipulated term; one applies it as children's insurance, educational insurance, dowry insurance, and the like (see Sects. 18.2 and 18.4 for denotation):

$$P_{x:\overline{n}|} = \frac{v^n}{\ddot{a}_{x:\overline{n}|}} = \frac{v^n \cdot D_x}{N_x - N_{x+n}} \quad (\textit{n-year fixed term insurance})$$

$$P^{gross}_{x:\overline{n}|} = \frac{v^n + \alpha^N}{(1-\beta) \cdot \ddot{a}_{x:\overline{n}|}} + \frac{\alpha^C + \gamma}{1-\beta}$$

$$_tV_{x:\overline{n}|}(P_{x:\overline{n}|}) = v^{n-t} - P_{x:\overline{n}|} \cdot \ddot{a}_{x+t:\,\overline{n-t}|}$$

$$= v^{n-t} - v^n \cdot \frac{D_x}{D_{x+t}} \cdot \frac{N_{x+t} - N_{x+n}}{N_x - N_{x+n}}$$

18.7 Life Annuities

- *Whole life annuity*: provides the annual payments as long as the insured aged x at entry is alive; if the periodic payments are made at the beginning (or at the end) of each payment period, i.e. in advance (or in arrear), then such an annuity is called *annuity-due* (or *immediate annuity*), respectively (see Sects. 18.2 and 18.4 for denotation):

$$SP_x = \ddot{a}_x = \frac{N_x}{D_x} \quad (\textit{whole life annuity-due})$$

$$= \mathrm{E}(Z) = \sum_{k=0}^{\infty} {}_{k|}q_x \cdot \ddot{a}_{\overline{k+1}|} = \sum_{k=0}^{\infty} {}_kp_x \cdot v^k,$$

where $Z = \ddot{a}_{\overline{K_x+1}|}$ according to the table:

Value of K_x	Value of Z	Probability
0	$\ddot{a}_{\overline{1}|}$	${}_{0|}q_x = q_x$
1	$\ddot{a}_{\overline{2}|}$	${}_{1|}q_x$
2	$\ddot{a}_{\overline{3}|}$	${}_{2|}q_x$
$\vdots$	$\vdots$	$\vdots$

$$SP_x = a_x = \frac{N_{x+1}}{D_x} = \sum_{k=1}^{\infty} {}_{k|}q_x \cdot a_{\overline{k}|} = \sum_{k=1}^{\infty} {}_kp_x \cdot v^k = \ddot{a}_x - 1$$

(*whole life immediate annuity*)

$$\sigma(Z) = \frac{1}{d}\left\{ {}^2A_x - (A_x)^2 \right\}^{1/2}$$

(underwriting risk of whole life annuity-due or immediate annuity: ${}^2A_x$ is as A_x calculated with the discount factor v^2 instead of v)

$$\ddot{a}_x^{(m)} = \sum_{k=0}^{\infty} \frac{1}{m} \cdot {}_{\frac{k}{m}}p_x \cdot v^{\frac{k}{m}} \approx \ddot{a}_x - \frac{m-1}{2m}$$

$$a_x^{(m)} = \sum_{k=1}^{\infty} \frac{1}{m} \cdot {}_{\frac{k}{m}}p_x \cdot v^{\frac{k}{m}} \approx a_x + \frac{m-1}{2m}$$

(unit whole life annuity-due and immediate annuity payable $m$thly, i.e. the insurer pays $m$thly within each year the payments in amount of $1/m$ as long as the insured is alive)

$$\ddot{a}_x^{(m)} \approx \ddot{a}_x - \frac{m-1}{2m} - \frac{m^2-1}{6m^2} \cdot i\,; \qquad a_x^{(m)} \approx a_x + \frac{m-1}{2m} - \frac{m^2-1}{6m^2} \cdot i$$

(more precise approximations)

$$\bar{a}_x = \int_0^{\infty} {}_tp_x \cdot v^t \, \mathrm{d}t = \int_0^{\infty} {}_tp_x \cdot \mathrm{e}^{-\delta \cdot t} \, \mathrm{d}t \approx \ddot{a}_x - \frac{1}{2} = a_x + \frac{1}{2}$$

(*continuous* whole life annuity with unit rate of payment and force of interest δ (see also Sect. 7.4))

$$\sigma(Z) = \frac{2}{\delta} \cdot (\bar{a}_x - {}^2\bar{a}_x) - (\bar{a}_x)^2$$

(*continuous approach*: ${}^2\bar{a}_x$ is as $\bar{a}_x$ calculated with the force of interest 2δ instead of δ)

$$SP_x^{gross} = \frac{(1+\gamma+\delta) \cdot \ddot{a}_x}{1-\beta} \quad \text{(whole life annuity-due)}$$

$${}_tV_x(SP_x) = \ddot{a}_{x+t} \quad \text{(whole life annuity-due)}$$

$$\ddot{a}_x = \frac{1-A_x}{d}$$

$$\ddot{a}_{x+1} = \frac{D_x}{D_{x+1}} \cdot (\ddot{a}_x - 1)$$

$$(I\ddot{a})_x = \frac{S_x}{D_x} \quad \text{(\textit{increasing} whole life annuity-due of the type } 1, 2, 3, \ldots)$$

- *Temporary life annuity*: provides the annual payments during the first n years as long as the insured aged x at entry is alive (see Sects. 18.2 and 18.4 for denotation):

$$SP_x = \ddot{a}_{x:\overline{n}|} = \frac{N_x - N_{x+n}}{D_x} \text{ (\textit{n-year life annuity-due})}$$
$$= \mathrm{E}(Z) = \sum_{k=0}^{n-1} {}_{k|}q_x \cdot \ddot{a}_{\overline{k+1}|} + {}_n p_x \cdot \ddot{a}_{\overline{n}|}$$
$$= \sum_{k=0}^{n-2} {}_{k|}q_x \cdot \ddot{a}_{\overline{k+1}|} + {}_{n-1}p_x \cdot \ddot{a}_{\overline{n}|} = \sum_{k=0}^{n-1} {}_k p_x \cdot v^k,$$

$$\text{where } Z = \begin{cases} \ddot{a}_{\overline{K_x+1}|}, & K_x = 0,\ 1,\ \ldots,\ n-1 \\ \ddot{a}_{\overline{n}|}, & K_x = n,\ n+1,\ \ldots \end{cases} \text{ according to the table:}$$

Value of K_x	Value of Z	Probability
0	$\ddot{a}_{\overline{1}|}$	${}_{0|}q_x = q_x$
1	$\ddot{a}_{\overline{2}|}$	${}_{1|}q_x$
2	$\ddot{a}_{\overline{3}|}$	${}_{2|}q_x$
$\vdots$	$\vdots$	$\vdots$
$n-2$	$\ddot{a}_{\overline{n-1}|}$	${}_{n-2|}q_x$
$n-1$, n, $n+1$, $\vdots$	$\ddot{a}_{\overline{n}|}$	${}_{n-1|}q_x + {}_n p_x = {}_{n-1} p_x$

$$\sigma(Z) = \frac{1}{d}\left\{{}^2A_{x\overline{n}|} - (A_{x\overline{n}|})^2\right\}^{1/2}$$

(underwriting risk of n-year temporary life annuity-due: ${}^2A_{x:\overline{n}|}$ is as $A_{x:\overline{n}|}$ calculated with the discount factor v^2 instead of v)

$$SP_x = a_{x:\overline{n}|} = \frac{N_{x+1} - N_{x+n+1}}{D_x} = \sum_{k=1}^{n} {}_k p_x \cdot v^k = \ddot{a}_{x:\overline{n+1}|} - 1$$

(*n-year life immediate annuity*)

$$\sigma(Z) = \frac{1}{d}\left\{{}^2A_{x:\overline{n+1}|} - (A_{x:\overline{n+1}|})^2\right\}^{1/2}$$

(underwriting risk of n-year life immediate annuity)

$$\ddot{a}^{(m)}_{x:\overline{n}|} \approx \ddot{a}_{x:\overline{n}|} - \frac{m-1}{2m} \cdot \left(1 - \frac{D_{x+n}}{D_x}\right) \; ; \qquad a^{(m)}_{x:\overline{n}|} \approx a_{x:\overline{n}|} + \frac{m-1}{2m} \cdot \left(1 - \frac{D_{x+n}}{D_x}\right)$$

(unit n-year life annuity-due and immediate annuity payable $m$thly, i.e. the insurer pays $m$thly within each year the payments in amount of $1/m$)

$$\bar{a}_{x:\overline{n}|} = \int_0^n {}_tp_x \cdot v^t \, \mathrm{d}t = \int_0^n {}_tp_x \cdot \mathrm{e}^{-\delta \cdot t} \, \mathrm{d}t \approx \ddot{a}_{x:\overline{n}|} - \frac{1}{2}\left(1 - \frac{D_{x+n}}{D_x}\right)$$
$$= a_{x:\overline{n}|} + \frac{1}{2}\left(1 - \frac{D_{x+n}}{D_x}\right)$$

(*continuous* n-year life annuity with unit rate of payment and force of interest δ (see also Sect. 7.4))

$$\ddot{a}_{x:\overline{n}|} = \frac{1 - A_{x:\overline{n}|}}{d}$$
$$\ddot{a}_{x:\overline{n+1}|} = \ddot{a}_{x:\overline{n}|} + {}_nE_x$$
$$(I\ddot{a})_{x:\overline{n}|} = \frac{S_x - S_{x+n} - n \cdot N_{x+n}}{D_x}$$

(*increasing* temporary life annuity-due of the type 1, 2, . . ., n)

$$(D\ddot{a})_{x:\overline{n}|} = \frac{n \cdot N_x - S_{x+1} + S_{x+n+1}}{D_x}$$

(*decreasing* temporary life annuity-due of the type n, $n-1$, . . ., 1)

$$(I\ddot{a})_{x:\overline{n}|} + (D\ddot{a})_{x:\overline{n}|} = (n+1) \cdot \ddot{a}_{x:\overline{n}|}$$

- *Whole life annuity guaranteed for n years*: is the whole life annuity, which on insured's death during the first n years descends to beneficiaries so that the annuity persists at least n years in any case:

$$\ddot{a}_{\overline{x\overline{n}|}} = \ddot{a}_{\overline{n}|} + {}_{n|}\ddot{a}_x = \frac{1 - v^n}{1 - v} + \frac{N_{x+n}}{D_x} \text{ (whole life annuity-due guaranteed \textit{for n years})}$$

- *Whole life annuity extended by n years*: is the whole life annuity, which on insured's death descends to beneficiaries for the next n years forthcoming:

$$\ddot{a}_{x:\lceil} = \ddot{a}_x + A_x \cdot \ddot{a}_{\overline{n}|} = \frac{(D_x + C_x \cdot \ddot{a}_{\overline{n}|}) + (D_{x+1} + C_{x+1} \cdot \ddot{a}_n) + \ldots}{D_x}$$

$$= \frac{N_x + M_x \cdot \frac{1-v^n}{1-v}}{D_x}$$

(whole life annuity-due *extended* by n years)

- *Deferred life annuity*: defers the first payment by a fixed *deferment* of k years; during the deferment period the periodic premiums are usually collected; if the insured dies during the deferment, then the life annuity ceases without remuneration (therefore in the case of periodic premiums, the return of the premiums on death during deferment is usual in practice, though this amendment raises the price of the original deferred life annuity (see *thereinafter*)):

$$SP_x = {}_{k|}\ddot{a}_x = \frac{N_{x+k}}{D_x} \quad (\textit{whole life annuity-due deferred by k years})$$

$$\sigma(Z) = \frac{1}{d}\left\{{}^{2}_{k|}A_x - \left({}_{k|}A_x\right)^2\right\}^{1/2}$$

(underwriting risk of whole life annuity-due deferred by k years: ${}^{2}_{k|}A_x$ is as ${}_{k|}A_x$ calculated with the discount factor v^2 instead of v)

$$SP_x = {}_{k|}a_x = \frac{N_{x+k+1}}{D_x} = {}_{k+1|}\ddot{a}_x \; (\textit{whole life immediate annuity deferred by k years})$$

$$\sigma(Z) = \frac{1}{d}\left\{{}^{2}_{k+1|}A_x - \left({}_{k+1|}A_x\right)^2\right\}^{1/2}$$

(underwriting risk of whole life immediate annuity deferred by k years)

$${}_{k|}\ddot{a}_x^{(m)} \approx {}_{k|}\ddot{a}_x - \frac{m-1}{2m} \cdot \frac{D_{x+k}}{D_x}\,; \qquad {}_{k|}a_x^{(m)} \approx {}_{k|}a_x + \frac{m-1}{2m} \cdot \frac{D_{x+k}}{D_x}$$

(unit whole life annuity-due and immediate annuity payable $m$thly deferred by k years, i.e. the insurer pays $m$thly within each year following the deferment the payments in amount of $1/m$)

$${}_{k|}\bar{a}_x = \int_k^\infty {}_tp_x \cdot v^t \,\mathrm{d}t$$

(*continuous* deferred whole life annuity with unit rate of payment and force of interest δ (see also Sect. 7.4))

$$P_{x:\overline{k}|} = \frac{{}_{k|}\ddot{a}_x}{\ddot{a}_{x:\overline{k}|}} = \frac{N_{x+k}}{N_x - N_{x+k}}\,; \qquad P_{x:\overline{k}|} = \frac{{}_{k|}a_x}{\ddot{a}_{x:\overline{k}|}} = \frac{N_{x+k+1}}{N_x - N_{x+k}}$$

(whole life annuity-due and immediate annuity deferred by k years)

$$SP_x^{gross} = \frac{(1+\delta)\cdot {}_{k|}\ddot{a}_x + \gamma \cdot \ddot{a}_x}{1-\beta} \quad \text{(whole life annuity-due deferred by } k \text{ years)}$$

$$P_{x:\overline{k}|}^{gross} = \frac{(1+\delta+\gamma_2)\ \cdot {}_{k|}\ddot{a}_x + \alpha^N}{(1-\beta)\cdot \ddot{a}_{x:\overline{k}|}} + \frac{\alpha^C + \gamma_1}{1-\beta}$$

(whole life annuity-due deferred by k years)

$$P_{x:\overline{n}|}^{gross+remuneration} = \frac{(1+\delta)\ \cdot {}_{k|}\ddot{a}_x + \alpha + \gamma \cdot \ddot{a}_{x:\overline{k}|}}{(1-\beta)\ \cdot \ddot{a}_{x:\overline{k}|} - (IA)_{x:\overline{k}|}^1}$$

(whole life annuity-due deferred by k years with return of level premiums on insured's death during deferment)

$${}_tV_x(P_{x:\overline{k}|}) = \begin{cases} {}_{k-t|}\ddot{a}_{x+t} - P_{x:\overline{k}|} \cdot \ddot{a}_{x+t:\,\overline{k-t}|} = \frac{N_{x+k}}{D_{x+t}} \cdot \frac{N_x - N_{x+t}}{N_x - N_{x+k}}, & t < k \\ \ddot{a}_{x+t} = \dfrac{N_{x+t}}{D_{x+t}}, & t \geq k \end{cases}$$

(whole life annuity-due deferred by k years)

$${}_tV_x^{gross}(SP_x^{gross}) = \begin{cases} (1+\gamma+\delta)\ \cdot {}_{k-t|}\ddot{a}_{x+t} + (\alpha^C + \gamma)\cdot \ddot{a}_{x+t:\,\overline{k-t}|}, & t < k \\ (1+\gamma+\delta)\cdot \ddot{a}_{x+t}, & t \geq k \end{cases}$$

(whole life annuity-due deferred by k years)

$${}_tV_x^{gross}(P_{x:\overline{k}|}^{gross}) = \begin{cases} (1+\delta)\ \cdot {}_tV_x(P_{x:\overline{k}|}) - \alpha^N \cdot \dfrac{\ddot{a}_{x+t:\,\overline{k-t}|}}{\ddot{a}_{x:\overline{k}|}}, & t < k \\ (1+\delta)\ \cdot {}_tV_x(P_{x:\overline{k}|}), & t \geq k \end{cases}$$

(whole life annuity-due deferred by k years)

$${}_{k|}\ddot{a}_x = \frac{D_{x+t}}{D_x} \cdot \ddot{a}_{x+k}$$

$${}_{k|}\ddot{a}_{x+1} = \frac{D_x}{D_{x+1}} \cdot \left({}_{k|}\ddot{a}_x - {}_kE_x\right)$$

$$\ddot{a}_x = \ddot{a}_{x:\overline{n}|} + {}_{n|}\ddot{a}_x$$

(decomposition of whole life annuity to temporary and deferred ones)

18.8 Multiple Life Insurance

- *Multiple life insurance*: are insurance products, in which insurance benefits depend on life status (mostly "alive" or "dead") of several lives (married couples, whole families, boards of directors, and the like); in such situations one applies the actuarial *model of n-tuples* with entry ages $(x_1, \ldots, x_n)$ (in particular, see Sects. 17.5 and 17.6 for *model of couples* (x, y), which are usually a male aged x at entry and a female aged y at entry):
- *Risk insurance for couple*: provides the payment of the sum insured on the first death within the given couple (such a principle is called the *joint-life status*):

$$A_{xy} = \frac{M_{xy}}{D_{xy}} = 1 - d \cdot \ddot{a}_{xy}$$

- *Endowment for couple*: provides the payment of the sum insured on the first death in the given couple, if a death occurs within the first n years, otherwise at the end of the stipulated term of n years (i.e. on survival by the couple):

$$A_{xy:\overline{n}|} = \frac{M_{xy} - M_{x+n,\, y+n} + D_{x+n,\, y+n}}{D_{xy}} = 1 - d \cdot \ddot{a}_{xy:\overline{n}|}$$

- *Life annuity for couple to the first death*: provides the annual payments as long as the both persons of the given couple are alive:

$$\ddot{a}_{xy} = \frac{N_{xy}}{D_{xy}}$$

- *Temporary life annuity for couple to the first death*: provides the annual payments as long as the both persons of the given couple are alive, but not longer than n years:

$$\ddot{a}_{xy:\overline{n}|} = \frac{N_{xy} - N_{x+n,\, y+n}}{D_{xy}}$$

- *Life annuity for couple to the second death*: provides the annual payments as long as a persons of the given couple is alive (such a principle is called the *last-survivor status*):

$$\ddot{a}_{\overline{xy}} = \ddot{a}_x + \ddot{a}_y - \ddot{a}_{xy}$$

- *Life annuity for couple from the first death to the second death* (*annuity for survivor*): provides the annual payments as long as just one person of the given couple is alive (annuities payable during the existence of a status, but only after the failure of a second status, are called *reversionary annuities*):

$$\ddot{a}_{\overline{xy}}^{[1]} = \ddot{a}_x + \ddot{a}_y - 2\ddot{a}_{xy}$$

- *Asymmetric life annuity for survivor*: provides the annual payments as long as just the second person of the given couple is alive (e.g. widow's annuity):

$$\ddot{a}_{y|x} = \ddot{a}_y - \ddot{a}_{xy} \quad \text{(analogously for widower's annuity } \ddot{a}_{x|y}\text{)}$$

- *Family insurances*: are comprehensive insurances to protect family against combined risks (mainly death and accident of parents and children)

18.9 Premium Reserve and Its Implications

- *Surrender* (see Sect. 18.1): means that withdrawing life insurance policyholders are entitled to *non-forfeiture benefits* (which are benefits that are not lost because of the premature cessation of premium payments); the surrender amount is usually a given part of the corresponding premium reserve (see Sect. 18.3) so that only the capitalizing type of life insurance (i.e. the life insurance with policy value) can allow it (on the other hand, the term *lapse* is used when the policy ceases without any remuneration for clients); in practice, the surrender in amount of the (gross) premium reserve reduced by the so-called *surrender charge* is usual; the surrender charge is frequently given as a percentage of the reserve and mostly decreases with increasing duration of the policy, e.g. the surrender at the year t of a policy with the term $n \geq 20$ for a life aged x at entry may be constructed in such a way that the surrender charge decreases linearly from 10 to 2%:

$$ {}_tS_{x:\overline{n}|} = \begin{cases} 0.90\ {}_tV^{gross}_{x:n} & \text{for } t \leq 3 \\ (0.885 + 0.005\ t) \cdot {}_tV^{gross}_{x:\overline{n}|} & \text{for } 3 < t < 19 \\ 0.98\ {}_tV^{gross}_{x:\overline{n}|} & \text{for } t \geq 19 \end{cases}$$

- *Paid-up policy*: means that the policy is converted to a mode with reduced benefits (mainly due to premature cessation of premium payments); similarly as for the surrender, only the capitalizing type of life insurance admits this special mode: even if the policy holder does not pay the periodic premiums, the policy persists, but with reduced parameters; in practice, one reduces mainly the sum insured or annuity payments (preserving the original term of the insurance contract); e.g. for the endowment it may be:

$$ {}_tR_{x:\overline{n}|} = \frac{{}_tS_{x:\overline{n}|}}{A_{x+t:\,\overline{n-t}|} + \gamma \cdot \ddot{a}_{x+t:\,\overline{n-t}|}}$$

  (*reduced sum insured* corresponding to a unit of the original sum insured when an endowment policy is paid-up at time t: the surrender is used as the net single premium with administration expenses (see Sect. 18.2) charged for the remaining period of the policy)

- *Valorized periodic premium* (*adjusted periodic premium*): is offered each year (without any medical test) to policyholders as a counter-inflation *valorization*:

$$S' = S + \frac{g}{100} \cdot \frac{P^{gross}}{P^{gross}_{x+t:\, \overline{n-t|}}}$$

(adjusted sum insured S' ($S' > S$) in the case that in the year t +1 of the policy, when the insured aged x at entry should pay the gross annual premium P^{gross} for the sum insured S as in the previous year t, the insured will pay the gross annual premium valorized by g percent; it is necessary to repeat this formula with proper values, when one valorizes the premiums gradually in particular years)

- Increased sum insured for increased premium:

$$(P^{gross})' = S \cdot P^{gross}_{x:\overline{n|}} + (S' - S) \cdot P^{gross}_{x+t:\, \overline{n-t|}}$$

(adjusted annual premium $(P^{gross})'$ in the case that in the year t +1 of the policy the sum insured S originally stipulated by the insured aged x at entry is increased to a new value S' ($S' > S$) without any lump sum increase of the premium reserve at the moment of change)

- Increased sum insured for increased premium and increased premium reserve:

$$(P^{gross})' = S' \cdot P^{gross}_{x:\overline{n|}}; \qquad \Delta = (S' - S) \cdot {}_tV^{gross}_{x:\overline{n|}}$$

(adjusted annual premium $(P^{gross})'$ and lump sum adjustment of premium reserve Δ in the case that in the year t+1 of the policy the sum insured S originally stipulated by the insured aged x at entry is increased to a new value S' ($S' > S$))

- *Profit sharing* (see Sect. 18.1): is an allocation of a given part of insurer's *technical gain* to policyholders; the technical gain consists of (1) *investment gain* due to a positive difference between the actually earned interest rate and the technical interest rate used for insurance calculations, (2) *mortality gain* due to an appropriate difference in assumed and actual mortalities, and (3) *expense gain* due to an appropriate difference in assumed and actual expenses; the values assumed a priori in calculations are usually called *statutory values*; the practical forms of the profit sharing can be various, e.g.
  - *extra bonus* included in insurance benefits
  - *direct payments* of the profit sharing yearly to the client
  - *reduction of premiums* in amount of the profit sharing
  - *reduction of insured period* (only in the capitalizing type of life insurance) by means of the profit sharing added to the premium reserve
  - *another insurance* paid by the profit sharing of the original insurance

$${}_tG_{x:\overline{n}|} = {}_tG^{investment}_{x:\overline{n}|} + {}_tG^{mortality}_{x:\overline{n}|} + {}_tG^{expense}_{x:n}, \text{ where}$$

$${}_tG^{investment}_{x:\overline{n}|} = \frac{1}{1-q'_{x+t-1}} \cdot (i'-i) \cdot \left[{}_{t-1}V^{gross}_{x:\overline{n}|} + (1-\beta') \cdot P^{gross}_{x:\overline{n}|} - \delta(t) \cdot \alpha' - \gamma'\right]$$

$$G^{mortality}_{x:\overline{n}|} = \frac{1}{1-q'_{x+t-1}} \cdot (q_{x+t-1} - q'_{x+t-1}) \cdot \left(1 - {}_tV^{gross}_{x:\overline{n}|}\right)$$

$${}_tG^{expense}_{x:\overline{n}|} = \frac{1}{1-q'_{x+t-1}} \cdot (1+i) \cdot [\delta(t) \cdot (\alpha - \alpha') + (\gamma - \gamma') + (\beta - \beta') \cdot P^{gross}_{x:\overline{n}|}],$$

where $\delta(t) = 1$ for $t = 1$ and $\delta(t) = 0$ otherwise

(*contribution formulas*: provide the profit ${}_tG_{x:\overline{n}|}$ per unit sum insured in the year t of the policy for the statutory values i, q_x, α, β, γ (for simplicity, one puts $\alpha^N = \alpha$ and $\alpha^C = 0$) and the actual values i', $q_x{}'$, $\alpha\,'$, $\beta\,'$, $\gamma\,'$)

$${}_tG_{x:\overline{n}|} \approx k \cdot (i'-i) \cdot \frac{{}_{t-1}V^{gross}_{x:\overline{n}|} + {}_tV^{gross}_{x:\overline{n}|}}{2}$$

(*approximate contribution formulas* used in practice: takes into account only the technical profit due to investment, i.e. the difference between the actually earned interest rate *i*' and the technical interest rate *i* used for insurance calculations; *k* is a ratio guaranteed by the insurance contract (e.g. $k = 0.90$))

18.10 Medical Underwriting

- *Medical underwriting*: is a qualified evaluation of the health state of the client that wants to enter into an insurance contract covering (at least partly) mortality or morbidity risks; the corresponding procedure is based on results of the entry medical examination; the insurer's approach to a person with health problems may be different:
  - the person is insurable, but for extra premium (such an approach is frequent in practice, see *thereinafter*)
  - the person is insurable, but in the case of death the insurer may reduce the insurance benefits
  - the person is not insurable and is refused by the insurer

- *Medical underwriting manuals*: are special statistical tables (usually constructed by prestigious reinsures) that enable to evaluate quantitatively the increased health risk (*impaired life*) by means of medical reports, tests and diagnoses taking into account gender, age, suggested term of insurance, and the like; the output offered by the manual for an increased mortality risk can be as follows:

– *age rating*: consists in an artificial aging of the client by a given number of years
– *increased force of mortality*: one adds to the force of mortality (see Sect. 7.3) constants recommended by the manual:

$$
\begin{aligned}
{}_n p_x^c \cdot v^n &= \exp\left(-\int_0^n (\mu_{x+s} + c)\,\mathrm{d}s\right) \cdot \mathrm{e}^{-\delta \cdot n} \\
&= \exp\left(-\int_0^n \mu_{x+s}\,\mathrm{d}s\right) \cdot \left(\mathrm{e}^{-(c+\delta)}\right)^n = {}_n p_x \cdot (v \cdot \mathrm{e}^{-c})^n
\end{aligned}
$$

(this relation shows that the force of mortality increased in a given age interval by a constant c can be interpreted as the force of interest increased by the same constant)

– *multiplicative and additive excess mortality* is the most frequently used approach: one looks up in the manual for given gender, age at entry and various health characteristics of the potential client:
– *multiplicative excess mortality* m_m (%)
– *additive excess mortality* m_a (per mile): concerns frequently postoperative states
– then instead of classical probabilities of death q_x (see Sect. 7.2), all technical calculations for this client apply adjusted probabilities of death ${q_x}^H$, which take into account the client's health state:

$$q_x^H = \left(1 + \frac{m_m}{100}\right) \cdot q_x + \frac{m_a}{1\,000}$$

• *Health coefficient*: indicates the ratio of the premium adjusted due to the client's health state to the standard premium of the insurance tariff

Further Reading

Black, K., Skipper, H.D.: Life Insurance. Prentice-Hall, Englewood Cliffs, NJ (1994)
Booth, P., Chadburn, R.., Cooper, D., Haberman, S., James, D.:Modern Actuarial Theory and Practice. Chapman and Hall /CRC, London (1999)
Bowers, N.L. et al.: Actuarial Mathematics. The Society of Actuaries, Itasca, IL (1986)
Gerber, H.U.: Lebensversicherungsmathematik. Springer, Berlin (1986) (*English translation*: Life Insurance Mathematics. Springer, Berlin (1990))
Koller, M.: Stochastische Modelle in der Lebensversicherung. Springer, Berlin (2000)
Milbrodt, H., Helbig, M.: Mathematische Methoden der Personenversicherung. DeGruyter, Berlin (1999)

Neill, A.: Life Contingencies. Heinemann, London (1977)
Parmenter, M.M.: Theory of Interest and Life Contingencies, with Pension Applications. ACTEX Publications, Winsted and New Britain, CT (1988)
Reichel, G.: Mathematische Grundlagen der Lebensversicherung. Volumes 3, 5 and 9 of the series Angewandte Versicherungsmathematik der DGVM. Verlag Versicherungswirtschaft, Karlsruhe (1975, 1976, 1978)
Teugels, J., Sundt, B. (eds.): Encyclopedia of Actuarial Science. Wiley, New York (2004)
Wolff, K.-H.: Versicherungsmathematik. Springer, Wien, Austria (1970)
Wolfsdorf, K.: Versicherungsmathematik (Teil 1: Personenversicherung). Teubner, Stuttgart, Germany (1997)

Chapter 19
Modern Approaches to Life Insurance

Abstract Chapter 19 presents some modern forms and instruments of life insurance: 19.1. Critical Illness Insurance, 19.2. Flexible Products of Life Insurance, 19.3. Unit Linked, 19.4. Profit Testing, 19.5. Embedded Value, 19.6. Fair Value.

19.1 Critical Illness Insurance

- *Critical illness insurance* (*dread disease insurance*, *DD insurance*): extends the mortality risk in such a way that the sum insured (or its percentage part) is paid out, as soon as certain terminal or highly debilitating illnesses are diagnosed (usually cancer that is life threatening, heart attack and stroke, but also other diseases may be included, e.g. paralysis, Alzheimer's or Parkinson's disease, multiple sclerosis, blindness, kidney failure) or medical procedures must be performed (such as a coronary bypass or organ transplant)

- Protective measures of insurers:

  – *waiting period*: is a stipulated period (usually several months), from when the policy was written, before the coverage is in force (it is an anti-selective measure)
  – *survival period*: requires the insured to survive a minimum number of days (usually 28 or 30), from when the illness was first diagnosed (it is an administrative measure to distinguish the DD insurance benefit from the classical risk life insurance benefit)

- Types of DD insurance:

  – *accelerated benefit*: the benefit stipulated for a basic (ordinary) life insurance (e.g. for a term insurance, see Sect. 18.5) is accelerated in such a way that the sum insured (*face amount*) is paid out either fully (so called *one hundred percent DD acceleration*), or partly (so called *k percent DD acceleration*), already when the illness is first diagnosed; a *DD rider* of the basic policy (see Sect. 18.1) is the typical form in this case

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_19,

- *independent benefit*: the DD benefit is guaranteed without any link to other circumstances; a *DD stand-alone policy* is the typical practical form in this case

Denotation:

q_x	is the probability that a life aged x will die within 1 year (i.e. prior the age $x + 1$ has been attained, see Sect. 17.2)
$q_x{}^{acc}$	is the probability that a life aged x without DD diagnosis will die or will be DD diagnosed within 1 year (the index *acc* reminds "acceleration")
i_x	is the probability that a life aged x without DD diagnosis will be DD diagnosed within 1 year (the probability of the *first incidence*)
k_x	is ratio of the number of DD deaths at age x to overall number of deaths at age x
$q_x{}^{notDD/notDD}$	is the probability that a life aged x without DD diagnosis will die within 1 year, where the cause of death will not be the DD diagnosis
$q_x{}^{DD/notDD}$	is the probability that a life aged x with DD diagnosis will die within 1 year, where the cause of death will not be the DD diagnosis

$$q_x^{acc} \approx i_x + (1 - k_x) \cdot q_x = i_x + q_x - k_x \cdot q_x$$

(approximate formula for q_x^{acc} based on an approximation $q_x{}^{notDD/notDD} \approx q_x{}^{DD/notDD}$)

$$l_0^{acc} = 100\ 000$$

$$l_{x+1}^{acc} = l_x^{acc} \cdot (1 - q_x^{acc}); \quad D_x^{acc} = l_x^{acc} \cdot v^x; \quad n_x^{acc} = D_x^{acc} + D_{x+1}^{acc} + \ \dots$$

$$d_{x+1}^{acc} = l_x^{acc} \cdot q_x^{acc}; \quad C_x^{acc} = d_x^{acc} \cdot v^{x+1}; \quad M_x^{acc} = C_x^{acc} + C_{x+1}^{acc} + \ \dots$$

(commutation functions constructed for DD insurance with accelerated benefit (see *thereinbefore*))

$$P_{x:\overline{n}|}^{acc} = \frac{A_{x:\overline{n}|}^{1\ acc}}{\ddot{a}_{x:\overline{n}|}^{acc}} = \frac{M_x^{acc} - M_{x+n}^{acc}}{N_x^{acc} - N_{x+n}^{acc}}$$

(*term insurance with one hundred percent DD acceleration*: net annual premium)

$$P^{acc}_{x:\overline{n}|} = \frac{A^{acc}_{x:\overline{n}|}}{\ddot{a}^{acc}_{x:\overline{n}|}} = \frac{M^{acc}_x - M^{acc}_{x+n} + D^{acc}_{x+n}}{N^{acc}_x - N^{acc}_{x+n}}$$

(*endowment with one hundred percent DD acceleration*: net annual premium)

$$P^{acc\,(k\%)}_{x:\overline{n}|} = \frac{k}{100} \cdot P^{acc}_{x:\overline{n}|} + \left(1 - \frac{k}{100}\right) \cdot P_{x:\overline{n}|},$$

where $P^{acc}_{x:\overline{n}|}$ is the net annual premium for term insurance with one hundred percent DD acceleration (see *thereinbefore*) and $P_{x:\overline{n}|}$ is the net annual premium for term insurance (see Sect. 18.5)

(*term insurance with k percent DD acceleration*: net annual premium; an analogous formula holds for the endowment)

$$d^i_x = l^{acc}_x \cdot i_x; \qquad C^i_x = d^i_x \cdot v^{x+1}; \qquad M^i_x = C^i_x + C^i_{x+1} + \ \ldots$$

(commutation functions constructed for DD insurance with independent benefit (see *thereinbefore*))

$$P^{ind}_{x:\overline{n}|} = \frac{A^{1\ ind}_{x:\overline{n}|}}{\ddot{a}^{acc}_{x:\overline{n}|}} = \frac{M^i_x - M^i_{x+n}}{N^{acc}_x - N^{acc}_{x+n}}$$

(*DD term insurance with independent benefit*: net annual premium)

19.2 Flexible Products of Life Insurance

- *Flexible products of life insurance* (*flexible life insurance policies*): give policyholders numerous options in terms of premiums, sum insured (face amounts), and investment objectives so that they can change these components of the policies in response to changing needs and circumstances; flexible life policies include: (1) Universal Life, (2) Adjustable Life, (3) Variable Universal Life, and other types (see *thereinafter*)

- *Universal Life*: is a flexible premium, adjustable death benefit insurance which accumulates *cash value* in the individual policyholder's account, i.e. it combines a whole life insurance protection with a savings feature (however in contrast to the endowment, the both components of Universal Life are strictly separated):
  - the risk component of the premium $P^{risk}(x)$ for the case of death is presented as *mortality charges* and is calculated for the age x being and for the stipulated sum insured S (usually applying up-to-date mortalities) as:

    $$P^{risk}(x) = S \cdot q_x$$

  - the cash value is credited each month with premium payments and with interest at a rate specified by the company; on the other hand, the cash value is debited each month by cost of insurance (COI) charges, i.e. by mortality charges, administrative costs and fees for insurer; the policyholder is regularly informed on these items

- on insured's death the insurer pays out the sum insured (stipulated for the case of death) plus the amount remaining in the cash account of the given policy; the surrender value of the policy (see Sect. 18.1) is the amount remaining in the cash account less applicable surrender charges; on survival of a stipulated term (if any) the insurer pays out the amount remaining in the cash account
- the policyholder may choose one of two options regarding the death benefit:
  - the first option provides a level death benefit equal to the policy's sum insured (face value): this choice supposes that more of the premium is placed in the cash account making the cash value rise more quickly
  - the second option provides an increasing death benefit equal to the policy's sum insured plus the cash value: cash value does not increase so quickly because more of the premium is applied to the higher mortality charges of the increasing death benefit over the policy's life

- *Adjustable Life*: is a policy that gives the policyholder the options to adjust the sum insured (face value), premium, and length of the coverage (e.g. from term to whole life) without having to change the policy

- *Variable Universal Life*: is a combination of Universal Life (see *thereinbefore*) with Unit Linked (see *thereinafter*)

19.3 Unit Linked

- *Unit Linked* (*Variable Life*): is a combination of life insurance (usually the term insurance) and investment funds; the insurance benefit may depend on the momentary price of investment units owned by the policyholder in investment portfolios of the insurer; unlike the classical life insurance (see Chap. 18) or the Universal Life (see Sect. 19.2):
  - the policyholders bear the whole investment risk by determining investment strategy for individual policies
  - most insurers offer a wide range of investment funds to suit client's investment objectives, risk profile and time horizon, e.g.:
    - cash funds (money market funds that invest in cash, bank deposits and money market instruments)
    - fixed interest and bond funds (see Sect. 9.2)
    - equity funds (see Sect. 9.3)
    - balanced funds (they combine equity investments with fixed interest instruments)
    - real estate funds
    - regional funds, environmental funds, humanitarian funds, and the like

- the policyholder's benefits are defined by reference to the number and the price of shares (*units*) in the particular investment funds; these units are notionally owned by the policyholder that disposes of a high flexibility in payment of premiums similarly as in the Universal Life (i.e. the policyholders have individual accounts)
- on insured's death the insurer pays out the sum insured (stipulated for the case of death) plus the value of the client's units; on survival of a stipulated term (if any) or on surrender the insurer pays out the value of the client's units
- the death benefit due to the term insurance (see Sect. 18.5) involved in the Unit Linked is usually decreasing in time; various constructions of this insurance are frequent in practice, e.g.:
  - for level premiums with regularly decreasing sum insured (see Sect. 18.5)
  - for risk premium with sum insured decreasing in such a way that after adding the client's units it provides a stipulated minimal level
- modern approaches (e.g. *profit testing*, see Sect. 19.4) are used in Unit Linked calculations
- the insurer must solve a significant need of new business capital (the so-called *new business strain*), since particular insurance policies may show significant negative cash flows in the beginning

Basic concepts:

- *Unit*: is a basic share of the given investment fund; typically, the first few years' premiums are invested in *initial units*, from which higher charges may be deducted to redeem the initial expenses; later years' premiums are invested in *accumulation units*, from which annual charges (regular fees and risk premiums) are deducted (obviously, these charges are much lower than the initial ones)
- *Offer price* vs. *bid price* of unit: is the price, at which the unit can be purchased or sold, respectively; the difference between offer and bid prices (the so-called *bid/offer spread*) is credited to the insurer
- *Unit fund*: is the account of units that are owned by the given client
- *Sterling fund*: is the account, in which premiums unallocated in the unit fund plus explicit charges which may be deducted from the unit fund are accumulated ("sterling" refers simply to "cash" here, to distinguish it from the unit fund)
- *Allocation percentage*: is the allocated part of premiums, i.e. the part of premiums, for which the units are purchased
- *Allocation ratio*: describes the distribution of the allocated premiums among particular investment funds
- Profit of the Unit Linked insurer comes only from various charges and fees credited to the insurer, which are e.g.
  - fund management charges
  - bid/offer spread charges (see *thereinbefore*)
  - surrender charges

– fund switching charges for transfer of units to other investment funds of the insurer

Denotation:

F_t	unit fund (see *thereinbefore*) at the beginning of the year t
a_t	allocation percentage (see *thereinbefore*) at the year t
P_t	annual premium paid at the beginning of the year t
λ	bid/offer spread charges (see *thereinbefore*)
i_u	rate of return for the client's unit fund
m_t	rate of fund management charges deducted from the client's unit fund at the end of the year t
C_t	charges deducted from the client's unit fund at the end of the year t
CF_t	insurer's profit (cash flow) from a Unit Linked policy at the end of the year t
E_t	expenses expended by the insurer at the beginning of the year t
i_s	rate of return for the client's sterling fund
S_t	death benefit for the term insurance
q_{x+t-1}	probability of client's death at age $x + t - 1$

$$F_{t+1} = [F_t + a_t \cdot P_t \cdot (1 - \lambda)] \cdot (1 + i_u) \cdot (1 - m_t)$$

(*development of client's unit fund*: the amount $a_t \cdot P_t \cdot (1 - \lambda)$ is invested in client's favour)

$$c_t = [F_t + a_t \cdot P_t \cdot (1 - \lambda)] \cdot (1 + i_u) \cdot m_t$$

(*charges deducted from the client's unit fund*: the amount $a_t \cdot P_t \cdot (1 - \lambda)$ is invested in client's favour)

$$CF_t = [(1 - a_t)P_t + a_t \cdot P_t \cdot \lambda - E_t] \cdot (1 + i_s) + C_t - S_t \cdot q_{x+\,t-1}$$

(*insurer's profit*: the amount $(1 - a_t) \cdot P_t$ is credited to the sterling fund; the amount $a_t \cdot P_t \cdot \lambda$ is credited to the insurer; the expression in brackets is the amount of the sterling fund at the beginning of the year t)

$$S_t = \max\{S - F_{t+1}, 0\}$$

(*death benefit decreasing* in such a way that after adding the client's units it provides a stipulated minimal level S (see *thereinbefore*)

- *Index Linked*: means that a capital life insurance product is linked to development of a market indicator (e.g. a stock exchange indicator) of a given capital market (see Sect. 29.3)

19.4 Profit Testing

- *Profit testing*: is a modern approach to actuarial calculations carried out using cash flow techniques where one views particular life insurance contracts as cash flow systems (see Chap. 6); typically, profit tests generate expected *profit signatures* of particular tranches of business on the basis of premium rates assumptions with the goal to keep the profitability in acceptable bounds; convenience of such an approach consists in:
  - effective measurement of *profitability* (e.g. momentary losses will be compensated by future profits)
  - realistic analysis due to realistic basis of calculation
  - flexibility of *pricing* (i.e. the calculation of premiums) and *valuation* (i.e. the calculation of reserves)
  - effective exploitation of modern computing facilities (e.g. *simulations of scenarios*, *stress testing*, and the like)
- *Risk discount rate* (*rdr*): is the interest rate used for discounting future cash flows (see Sect. 6.1) in modern approaches to life insurance; it reflects the price of investment capital and the risk due to its investment in insurance business; the components of *rdr* are:
  - risk free rate *rfr* (see Sects. 5.1 and 13.2)
  - costs of investment capital
  - risk margins for given market and business
  - other margins (e.g. due to risk corresponding to the country of business)

$$rdr = rfr + (r_M - rfr) \cdot \beta$$

(*risk discount rate* according to the model CAPM (see Sect. 13.3): is constructed by means of risk free rate *rfr*, expected market return (expected return of market portfolio) r_M and factor β; the term $(r_M - rfr)\cdot\beta$ is called *risk margin*)

Denotation:

P_t	(gross) premium at the beginning of the year t
E'_t	actual expenses by insurer at the beginning of the year t
i'	actual investment rate of return by insurer (i.e. actually earned interest rate)
${}_tV_x^{stat}$	statutory premium reserve (i.e. the premium reserve calculated by means of mortalities and technical interest rate as were assumed for the premium calculations) at the end of the year t for a life aged x at entry (see Sect. 18.3)
S_t^{death}	death benefit at the end of the year t on death during this year
S_t^{surv}	survival benefit at the end of the year t on survival up to this moment
S_t^{surr}	surrender benefit at the end of the year t on surrender during this year

p_x' actual probability of survival at age x (see Sect. 17.2)
q_x' actual probability of death at age x (see Sect. 17.2)
w_t' actual probability of surrender (see Sect. 18.9) during the year t

$$PRO_t = (P_t - E'_t) \cdot (1 + i') + {}_{t-1}V_x^{stat} \cdot (1 + i') - {}_tV_x^{stat} \cdot p'_{x+t-1} - \\ -S_t^{death} \cdot q'_{x+t-1} - S_t^{surv} \cdot p'_{x+t-1} - S_t^{surr} \cdot w'_t$$

(annual *profit* per policy in force: is the profit expected to be earned during the year t per policy of the given type, which is in force at the beginning of the year t; if the values assumed for calculations (without apostrophes) are equal to the actual ones (with apostrophes) then $PRO_t = 0$)

$$PROS_t = PRO_t \cdot {}_{t-1}p'_x$$

(annual *profit signature* per policy: is the future profit expected at entry age x to be earned during the year t per policy of the given type)

$$PVFP = \sum_{t=1}^{n} \frac{PROS_t}{(1 + rdr)^t}$$

(*present value of future profits* expected per policy of the given type)

- *Criteria of profitability* applied in profit testing:
  - *profit as percentage of PV of premiums* (*profit margin*):

$$\frac{PVFP}{\sum_{t=1}^{n} P_t \cdot {}_{t-1}p'_x \cdot (1 + rdr)^{-(t-1)}} \cdot 100\%$$

  - *profit as percentage of PV of commissions* (commissions *Com* for insurance agents may be used as a basis if one measures profitability of the given insurance business):

$$\frac{PVFP}{\sum_{t=1}^{n} Com_t \cdot {}_{t-1}p'_x \cdot (1 + rdr)^{-(t-1)}} \cdot 100\%$$

  - *profit as internal rate of return IRR* (see Sect. 6.2):

$$\sum_{t=1}^{n} \frac{PROS_t}{(1 + IIR)^t} = 0$$

  - *profit according to payback of invested capital*: the criterion prefers the insurance business with the shortest payback periods (see Sect. 6.3) defined as the first k such that $PVFP_k \geq 0$:

$$PVFP_k = \sum_{t=1}^{k} \frac{PROS_t}{(1+rdr)^t}, \qquad k = 1, \ldots, n$$

19.5 Embedded Value

- *Embedded value EV*: is a modern methodology suggested primarily for the valuation of insurance liabilities (i.e. for the calculation of premium reserves); it consists in global profit testing analysis (see Sect. 19.4) for the whole insurance portfolio
- Valuation of insurance liabilities:
  - by means of *statutory reserves* (see Sect. 19.4): is based on conservative assumptions applied to the premium calculation; such assumptions do not frequent-ly correspond to reality; the statutory reserves are the standards used in the framework of the insurance business regulation
  - by means of *EV reserves* (see *thereinafter*): reduces the statutory reserves by the present value of future profits (see Sect. 19.4) respecting solvency requirements; such an approach is important for shareholders and management (e.g. it respects the fact that momentary losses may be compensated by future profits)
- Characteristics of EV:
  - one projects (i.e. discounts backwards to the valuation date) future (gross) profit cash flows (see Sect. 19.4) concerning the *whole insurance portfolio*
  - these projections are based on the *best estimates* of the corresponding future profits (the best estimated level of assumptions on mortality, lapses, commissions, expenses, investment return, etc. is understood to be their expected value)
  - the risk and uncertainty that real future cash flows will differ from their best estimates is covered in the discount operations using the *risk discount rate* (*rdr*, see Sect. 19.4)
  - costs of *embedded options* and guarantees (see Sect. 18.1) cannot be included objectively using EV approach

Denotation is analogous to the one for profit testing (see Sect. 19.4):

P'_t	total estimated premium income at the beginning of the year t (i.e. premiums of all policies in force estimated for the beginning of the year t in the investigated insurance portfolio)
E'_t	total estimated expenses by insurer at the beginning of the year t
i'	estimated annual rate of return by insurer in the investigated insurance portfolio
${}_tV'^{stat}$	total estimated statutory premium reserves at the end of the year t
$S_t'^{death}$	total estimated death benefit at the end of the year t
$S_t'^{surv}$	total estimated survival benefit at the end of the year t
$S_t'^{surr}$	total estimated surrender benefit at the end of the year t

$$PRO_t = (P'_t - E'_t) \cdot (1 + i') + {}_{t-1}V'^{stat} \cdot (1 + i') - {}_tV'^{stat} - S'^{death}_t - S'^{surv}_t - S'^{surr}_t$$

(total estimated *profit* in the investigated insurance portfolio during the year t)

$$PVFP = \sum_{t=1}^{n} \frac{PRO_t}{(1 + rdr)^t}$$

(*present value of future profits* in the investigated insurance portfolio; if one applies net profits (after tax) reduced by *future changes of required solvency capital* (the so-called *allocated capital AC*, see *thereinafter*) expressed as $(AC_{t+1} - AC_t) - AC_t \cdot i' \cdot (1 - tax)$, then one obtains the present value of future *distributable earnings* (see *thereinafter*), which could be paid out as dividends to shareholders of the insurance company)

$${}_tV^{EV} = {}_t V^{stat} - PVFP$$

(*EV reserves*: an indirect method of calculation of EV reserves that consists in reducing the statutory reserves by the present value of future *distributable earnings* (see *thereinbefore*); the concept of EV reserves is important for shareholders and management (see *thereinbefore*))

Denotation:

- FA *free assets* of insurer: is the capital that does not support any in-force business of insurer at the valuation date (i.e. the assets that back neither liabilities nor required solvency margin of the insurer so that the shareholders can use them without jeopardizing the company economy)
- NA *net assets* of insurer
- AC *allocated capital* (*assigned capital*) of insurer: is the required solvency capital (see also Sect. 24.3); it holds $NA = FA + AC$

$$EV = FA + PVFP$$

(*embedded value* (*implicit value*) of insurer: is used as an estimation of the value of the company (it is another application of the EV approach beside the liability valuation by means of EV reserves, see *thereinbefore*); *PVFP* is the present value of future *distributable earnings* (see *thereinbefore*) from *in-force business*; sometimes the present value of future changes of required solvency capital (see *thereinbefore*) is expressed alternatively as the present value of differences between rdr and net rate of return $i' \cdot (1 - tax)$ applied to allocated capital amounts, i.e. $\sum_{t=1}^{n} AC_{t-1} \cdot \big(rdr - i' \cdot (1 - tax)\big) / (1 + rdr)^t$, where $AC = AC_0$)

$$VBIF = PVFP - AC$$

(*value of business in-force* (*portfolio value*): is the best used estimate of the market price of insurance portfolio)

$$AV = EV + PVNB$$

(*appraisal value*: is the estimate of the market price of insurance company; it completes the EV reserves by the present value of future distributable earnings (see *thereinbefore*) from the estimated *new business* (*PVFP* covers only the in-force business); *AV* is applied e.g. for sales, acquisitions and mergers of insurance companies)

19.6 Fair Value

- *Fair value FV*: is an alternative to *EV* (see Sect. 19.5), which promotes itself in the framework of modern accounting principles (International Accounting Standards IAS and International Financial Reporting Standards IFRS); in general, the fair value is defined as *the amount for which an asset could be exchanged and a liability settled between knowledgeable, willing parties in an arm's transaction*; similarly as for the EV reserves (see Sect. 19.5), the valuation of liabilities by means of the FV reserves is important for shareholders and management of the company

- Characteristics of FV:
  - one projects (i.e. discounts backwards to the valuation date) *future cash flows* (see *thereinafter*) concerning the *whole insurance portfolio*
  - unlike EV approach, necessary assumptions are not used on the best estimation level, but they are adjusted by the so-called *market value margin MVM*, which should cover the risk and uncertainty of future development; in particular, the investment rate of return rfr^{MVM} used in this context is the risk-free interest rate adjusted by MVM (see *thereinafter*)
  - unlike EV approach, the discount operations are performed by means of the *risk-free interest rate* (*rfr*, see Sects. 5.1 and 13.2)
  - unlike EV approach, costs of embedded options and guarantees can be included objectively (e.g. by means of deflators, see thereinafter)

Denotation is analogous to the one for embedded value (see Sect. 19.5):

P_t^{MVM}	total estimated premium income (in the investigated insurance portfolio) adjusted by MVM at the beginning of the year t
E_t^{MVM}	total estimated expenses by insurer adjusted by MVM at the beginning of the year t
rfr^{MVM}	risk-free interest rate adjusted by MVM used as the investment rate of return
$S_t^{\,death,\,MVM}$	total estimated death benefit adjusted by MVM at the end of the year t

$S_t^{\,surv,\,MVM}$ total estimated survival benefit adjusted by MVM at the end of the year t

$S_t^{\,surr,\,MVM}$ total estimated surrender benefit adjusted by MVM at the end of the year t

$$CF_t = (P_t^{MVM} - E_t^{MVM}) \cdot (1 + rfr^{MVM}) - S_t^{death,\,MVM} - S_t^{surv,\,MVM} - S_t^{surr,\,MVM}$$

(total estimated *cash flow* in the investigated insurance portfolio at the end of the year t)

$$FV = \sum_{t=1}^{n} \frac{CF_t}{(1 + rfr)^t}$$

(*fair value FV*: is constructed as the present value of *future cash flows* (see *thereinbefore*) in the investigated insurance portfolio performing discount operations by means of the risk-free interest rate; unlike the future profits PRO_t in EV (see Sect. 19.5), the future cash flows CF_t in FV: (1) do not include increments of the statutory reserves; (2) their best estimated level is adjusted by MVM; (3) are discounted by means of the risk-free interest rate)

- *Deflators*: enable to evaluate *embedded options*, which are client's possibilities to choice among several alternatives of insurance contract development without additional charges (see Sect. 18.1); in the framework of modern accounting principles IFRS (see *thereinbefore*), the embedded options should be evaluated as the risks by insurers

Denotation:

$B_t = \mathrm{e}^{\,r_f \cdot\, t}$ interest factor of continuous compounding (see Chap. 4) using risk-free interest rate r_f : it can be interpreted as the price of unit ($B_0 = 1$) of a risk-free, zero-coupon bond at time t (see Sect. 9.2)

Z_T claim guaranteed by the evaluated embedded option (see *thereinbefore*) on maturity date T (compare with the claim guaranteed by a financial derivative, see Sect. 15.8)

V_t price (or value) of the embedded option at time t ($0 \le t \le T$)

Q equivalent martingale (risk-neutral) probability Q with respect to probability P and filtration $\Im_t$ describing dynamics of the risk events to be evaluated (compare with the financial derivatives pricing, see Sect. 15.8)

$$V_t = \mathrm{e}^{-\,r_f \cdot\, (T-\,t)} \mathrm{E}_Q\,(Z_T \mid \Im_t) = B_t \cdot \mathrm{E}_Q \left(B_T^{-1} \cdot Z_T \mid \Im_t\right)$$

(price of *embedded option* at time t (see *thereinbefore*): is equal (if no arbitrage opportunity exists, see Chap. 14) to the conditional mean value (with respect to probability Q and filtration $\Im_t$) of the claim Z_T discounted using the risk-free rate of return to time t)

$$D_t = \mathrm{e}^{-r_f \cdot t} \mathrm{E}_P \left(\frac{\mathrm{d}Q}{\mathrm{d}P} \mid \Im_t \right)$$

(*deflator* at time t: $\frac{\mathrm{d}Q}{\mathrm{d}P}$ is Radon-Nikodym derivative of Q with respect to P, see Sect. 26.8)

$$V_t = D_t^{-1} \cdot \mathrm{E}_P \left(D_T \cdot Z_T \mid \Im_t \right)$$

(price of embedded option at time *t by means of deflator*: application of the deflator enables to use the conditional mean value with respect to the original (risk) probability P)

Further Reading

Abbink, M., Saker, M.: Getting to grips with fair value. Staple Inn Actuarial Society, London (2002)

Baars, G., Bland, P.: Elements of dread disease pricing and design. Swiss Reinsurance Company, Zurich (1993)

Booth, P., Chadburn, R., Cooper, D., Haberman, S., James, D.: Modern Actuarial Theory and Practice. Chapman and Hall/CRC, London (1999)

Bowers, N.L. et al.: Actuarial Mathematics. The Society of Actuaries, Itasca, IL (1986)

Buechner, J.D., Eason, T.F., Manzler, D.L.: Why Universal Life. The National Underwriting Company, Cincinnati, OH (1983)

Dash, A., Grimshaw, D.: Dread disease cover. An actuarial perspective. Staple Inn Actuarial Society, London (1990)

Hairs, C.J. et al.: Fair valuation of liabilities. British Actuarial Journal 8, 203–340 (2002)

Jarvis, S., Southall, F., Varnell, E.: Modern valuation techniques. Staple Inn Actuarial Society, London (2001)

Koller, M.: Stochastische Modelle in der Lebensversicherung. Springer, Berlin (2000)

Milbrodt, H., Helbig, M.: Mathematische Methoden der Personenversicherung. DeGruyter, Berlin (1999)

Teugels, J., Sundt, B. (eds.): Encyclopedia of Actuarial Science. Wiley, New York (2004)

Wolfsdorf, K.: Versicherungsmathematik (Teil 1: Personenversicherung). Teubner, Stuttgart (1997)

Chapter 20
Pension Insurance

Abstract Chapter 20 is devoted to the topic of pensions: 20.1. Basic Concepts of Pension Insurance, 20.2. Defined Contribution Plan, 20.3. Defined Benefit Plan.

20.1 Basic Concepts of Pension Insurance

- *Pension*: is an arrangement to provide people with income when they are no longer earning a regular income from employment; typically, pensions concern major population groups defined in a specific way: e.g. the population of a whole region or country (*social pensions*, *state pensions*, *social security*), the employees or members of a trade association (*occupational pensions*, *employer-based pensions*) or other forms of *pension plans* (*retirement schemes*, *pension funds*) combined frequently with *personal savings*; from the point of view of insurance, the pensions may cover risks of longevity (social insufficiency), disability or sickness

- *Contributions*: are incomes of pension plans (mainly the payments by active participants, employers or state) in favour of participants; typically, the contribution is expressed as a percentage of participant's salary

- *Benefits*: are outcomes of pension plans (mainly the payments of pensions); the pensions may be classified according to circumstances on which they are contingent:
  - *retirement pension* (*old age pension*): on achievement of a (normal) *retirement age* (sometimes an *early retirement* or *late retirement* is also possible)
  - *past-service pension* (*merit pension*): on achievement of a prescribed number of active participation years, in which contributions have been paid
  - *disability pension*: in the case of (total) disability (on reinstatement of the disabled participant, the disability pension is suspended)
  - *survivor's pension*: in the case of participant's death (e.g. widow's pension)

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_20,

- *lump-sum settlement*: may be chosen instead of particular pensions
- *surrender*: is a reimbursement when the active participation is insufficient (sometimes such a participant obtains so called *vested right* to transfer the accrued amount e.g. to a pension fund of another employer)

- Types of pension plans:

  (a) *according to calculations of contributions and benefits*:

  - *defined contribution plan*: contributions are paid into an *individual account* for each participant with investment returns credited also to this individual's account; on retirement, the participant's account is used to provide benefits (pensions) that are dependent upon the amount of money contributed and the investment performance; in, particular it means that the contributions are known but the benefits are unknown until calculated by prescribed methods (see Sect. 20.2)
  - *defined benefit plan*: guarantees a certain benefit (pension) after retirement; this benefit is determined by a *benefit formula* that may incorporate (1) the participant's *salary*, (2) *service years* (i.e. the years of employment or the years of active participation with contributions paid), (3) *age at retirement* (and other factors), e.g.

    - *flat rate* pension plan: defines flat pensions for all pensioners
    - *money times service* pension plan: defines pensions based only on the service years (e.g. a plan offering 100 USD a month per 1 year of service will provide 2,000 USD a month to the retiree with twenty service years)
    - *final salary* pension plan: defines monthly pensions equal to the service years, multiplied by the participant's salary at retirement, multiplied by a factor known as *accrual rate*
    - *career average salary* pension plan: defines pensions similarly as the final salary plan except for averaging salary over all the service years instead of the final salary
    - *final average salary* pension plan: defines pensions similarly as the final salary plan except for averaging salary over the final service years prior to retirement instead of the final salary

    The calculation of contributions which will guarantee the defined benefit (in funded pension plans, see *thereinafter*) is mostly based on the *equivalence principle* (see Sect. 18.2), i.e. one must calculate the present value (for age at entry) of the corresponding benefit (probabilities of survival in active state to retirement age should be used for retirement pensions, probabilities of transition to disability state should be used for disability pensions, etc., see Sect. 20.3) and cover it by the present value of contributions, which are to be fixed (usually as a percentage of salary)

(b) *according to financing*: various strategies of liability capitalization are possible with two distinctive types:

- *unfunded pension plan* (*pay-as-you-go financing*, *PAYGO financing*): means that benefits are paid directly from current workers' contributions and taxes; it is a usual way how social security systems are financed all over the world
- *funded pension plan*: means that contributions are invested in funds to meet the benefits; typically, the contributions to be paid are regularly reviewed when evaluating the plan's assets and liabilities by actuaries to ensure that the funds will meet future payment obligations (see Sect. 20.3)

(c) *according to organizing aspects*, e.g.:

- *company pension scheme* (*staff provident scheme*): is organized directly by the employer; the pensions usually figure as liabilities in the employer's balance-sheet
- *professional pension scheme*: is organized by a professional operator accredited by the employer (a commercial insurance company, a provident society, etc.)
- *state pension institutions* (*governmental insurer*, *social security administration*)

20.2 Defined Contribution Plan

- Formulas for defined contribution plans (see Sect. 20.1) are represented here by possible examples how the payments of retirement and survivor's pensions can be calculated:

Denotation:

C^{RP}	participant's account for retirement pension at the moment when the pension becomes due
C^{SP}	participant's account for survivor's pension at the moment when the pension becomes due
x	participant's age at the moment when the retirement pension becomes due, or survivor's age at the moment when the survivor's pension becomes due
P	annual pension payment
$\ddot{a}_x$	present value of unit whole life annuity-due (see Sect. 18.7)
$\ddot{a}_{x:\overline{n}\rceil}$	present value of unit n-year annuity-due (see Sect. 18.7)
$\ddot{a}_{\overline{n}\rceil}$	present value of annuity-certain used in finance (see Sect. 7.1)
$\mathring{e}_x$	life expectancy at age x (see Sect. 17.2)

- Examples of formulas for *retirement pension*:

$$P = \frac{C^{RP}}{\ddot{a}_x}$$

(whole life retirement pension *by means of actuarial approach with guaranteed rate of return i* ($\sim$ technical interest rate); the investment gain over the guaranteed rate of return may be distributed to pensioners in addition to the benefits P)

$$P = \frac{C^{RP}}{\ddot{a}_{\overline{{}^\circ e_x}|}}$$

(whole life retirement pension *by means of life expectancy with guaranteed rate of return i*; it holds $\ddot{a}_{\overline{{}^\circ e_x}|} > \ddot{a}_x$)

$$P = \frac{C^{RP}}{{}^\circ e_x}$$

(whole life retirement pension *by means of life expectancy without guaranteed rate of return*; this type of annuity decomposition is sometimes applied in the tax context)

$$P_{x+n} = \frac{C^{RP}_{x+n}}{\ddot{a}_{x+n}}; \quad P_{x+n} = \frac{C^{RP}_{x+n}}{\ddot{a}_{\overline{{}^\circ e_{x+n}}|}}; \quad P_{x+n} = \frac{C^{RP}_{x+n}}{{}^\circ e_{x+n}}$$

where C_{x+n} and ${}^\circ e_{x+n}$ are up-to-date values for the year of calculation (whole life retirement pension with *varied payments according to up-to-date situation*)

- Examples of formulas for *survivor's pension*:

$$P = \frac{C^{SP}}{\ddot{a}_{x:\overline{n}|}}$$

(temporary survivor's pension *by means of actuarial approach with guaranteed rate of return*)

$$P = \frac{C^{SP}}{\ddot{a}_{\overline{n}|}}; \quad P = \frac{C^{SP}}{n}$$

(temporary survivor's pension in the form of *annuity-certain* (see Sect. 7.1) with guaranteed rate of return i, respectively without guaranteed rate of return)

$$P = \frac{C^{RP}}{\ddot{a}_{\overline{x:\overline{n}|}}} = \frac{C^{RP}}{\ddot{a}_{\overline{n}|} + {}_{n|}\ddot{a}_x} = \frac{C^{RP}}{\frac{1-v^n}{1-v} + \frac{N_{x+n}}{D_x}}$$

(retirement pension in combination with survivor's pension *guaranteed for n years*: combines the whole life retirement pension with the temporary survivor's pension in such a way that on death of the recipient of retirement pension during the first n

years, the payments of original amount go on as survivor's pension till the end of the guaranteed period n (see also Sect. 18.7))

$$P = \frac{C^{RP}}{\ddot{a}_{x\lceil n}} = \frac{C^{RP}}{\ddot{a}_x + A_x \cdot \ddot{a}_{\overline{n}|}} = \frac{C^{RP}}{\frac{N_x}{D_x} + \frac{M_x}{D_x} \cdot \frac{1-v^n}{1-v}}$$

(retirement pension in combination with survivor's pension *extended by n years*: combines the whole life retirement pension with the temporary survivor's pension in such a way that on death of the recipient of retirement pension, the payments of original amount go on as survivor's pension for the next n years forthcoming (see also Sect. 18.7))

20.3 Defined Benefit Plan

- Actuarial calculations for defined benefit plans:
  - are based on the *present value of* (*expected*) *future benefits PVFB* and contributions of the participant aged x applying the *equivalence principle* (see Sect. 18.2) and respecting possibly the salary development; one may use commutation functions to simplify numerical calculations similarly as in the life insurance (see Sect. 17.6)
  - take into account the service years distinguishing usually the *past service PS* and the *future service FS* in the pension plan (see *thereinafter*)
  - enable to construct (in order to fund the pension plan, i.e. to cover benefits gradually prior to their maturity by necessary financial funds):
    - *normal contributions NC* (*standard contribution rate SCR*): are designed to amortize *PVFB* (see *thereinbefore*) over the participants service years (working lifetime), the pattern of amortization being specified by particular actuarial funding methods; it is an analogy of premiums in life insurance
    - *actuarial liabilities AL*: are equal (prospectively, at age y for $y > x$) to *PVFB* at this age less the present value of (expected) future *NC* yet to be made; it is an analogy of premium reserves in the life insurance (see Sect. 18.3)

Denotation:

l_x	number of active participants of the pension plan at age x
d_x	number of deaths at age x
w_x	number of withdrawals at age x
i_x	number of disability incidences at age x
r_x	number of retirement pensions paid since age x

$$l_{x+1} = l_x - (d_x + w_x + i_x + r_x)$$

(*decrement order of active participants*: decrements are distributed uniformly over each year)

$$d_x = q_x \cdot l_x, \quad w_x = q_x^w \cdot l_x, \quad i_x = q_x^i \cdot l_x, \quad r_x = q_x^r \cdot l_x$$

(probability expressions: q_x^j is the decrement rate (decrement probability) of the type j at age x)

- *Present value of retirement pension independent on salary* (i.e. the *PVFB* corresponding to the retirement pension, see *thereinbefore*) assuming that
  - the participant is aged x with n service years
  - the annual retirement pension accrual per each service year is b (see Sect. 20.1)
  - the *normal retirement age* is u with possibility of early retirement, but without possibility of late retirement
  - the present value of unit pension paid since retirement at age x is $\ddot{a}_x^r$
  - the discount operations are performed by means of the discount factor v (see Sect. 3.2)

$$PV = PS + FS = n \cdot b \cdot \frac{M_x^{ra}}{D_x} + b \cdot \frac{R_x^{ra}}{D_x},$$

  where

$$D_x = l_x \cdot v^x, \quad C_x^{ra} = r_x \cdot \ddot{a}_x^r \cdot v^x, \quad M_x^{ra} = \sum_{j=0}^{u-x} C_{x+j}^{ra}, \quad R_x^{ra} = \sum_{j=1}^{u-x} M_{x+j}^{ra}$$

- *Present value of disability pension independent on salary* assuming that
  - the participant is aged x with n service years
  - the annual disability pension accrual per each service year is b
  - the normal retirement age is u
  - the present value of unit pension paid since disability incidence at age x is $\ddot{a}_x^i$
  - the discount operations are performed by means of the discount factor v (see Sect. 3.2)

$$PV = PS + FS = n \cdot b \cdot \frac{M_x^{ia}}{D_x} + b \cdot \frac{R_x^{ia}}{D_x},$$

  where

$$D_x = l_x \cdot v^x, \quad C_x^{ia} = i_x \cdot \ddot{a}_x^i \cdot v^x, \quad M_x^{ia} = \sum_{j=0}^{u-x-1} C_{x+j}^{ia}, \quad R_x^{ia} = \sum_{j=1}^{u-x-1} M_{x+j}^{ia}$$

- *Present value of retirement pension dependent on career average salary* assuming that
  - the participant is aged x with n service years
  - the *salary scale function* is $\{s_x\}$ (i.e. the ratio of salary increase between ages x and $x + t$ is s_{x+t}/s_x), the participant's cumulative salary during the past service is S, and the participant's annual salary at age x is s

- the annual retirement pension accrual per each service year is z_u/K for a fixed K, where z_u is the participant's average salary over the whole past service (over the whole career)
- the normal retirement age is u with possibility of early retirement, but without possibility of late retirement
- the present value of unit pension paid since retirement at age x is $\ddot{a}_x^r$
- the discount operations are performed by means of the discount factor v (see Sect. 3.2)

$$PV = PS + FS = \frac{S}{K} \cdot \frac{M_x^{ra}}{D_x} + \frac{s}{K} \cdot \frac{{}^sR_x^{ra}}{{}^sD_x},$$

$$\text{where } {}^sD_x = s_x \cdot D_x, \quad {}^sM_x^{ra} = s_{x-1} \cdot M_x^{ra}, \quad {}^sR_x^{ra} = \sum_{j=1}^{u-x} {}^sM_{x+j}^{ra}$$

- *Present value of retirement pension dependent on final average salary* (over m last salaries prior to retirement, the case $m = 1$ (i.e. with dependence on *final salary* only) being frequent in practice) assuming that
  - the participant is aged x with n service years
  - the salary scale function is $\{s_x\}$ (i.e. the ratio of salary increase between ages x and $x + t$ is s_{x+t}/s_x), the participant's average salary over m years prior to age x is $z_x = (s_{x-m} + \ldots + s_{x-1})/m$, and the participant's annual salary at age x is s
  - the annual retirement pension accrual per each service year is z_u/K for a fixed K, where z_u is the participant's average salary over m years prior to retirement
  - the normal retirement age is u with possibility of early retirement, but without possibility of late retirement
  - the present value of unit pension paid since retirement at age x is $\ddot{a}_x^r$
  - the discount operations are performed by means of the discount factor v (see Sect. 3.2)

$$PV = PS + FS = \frac{n \cdot s}{K} \cdot \frac{{}^zM_x^{ra}}{{}^sD_x} + \frac{s}{K} \cdot \frac{{}^zR_x^{ra}}{{}^sD_x},$$

where

$${}^zC_x^{ra} = z_x \cdot C_x^{ra}, \quad {}^zM_x^{ra} = \sum_{j=0}^{u-x} {}^zC_{x+j}^{ra}, \quad {}^zR_x^{ra} = \sum_{j=1}^{u-x} {}^zM_{x+j}^{ra}$$

- *Present value of contributions independent on salary* assuming that
  - the participant is aged x with n service years
  - the annual participant's contribution is C
  - the normal retirement age is u
  - the discount operations are performed by means of the discount factor v (see Sect. 3.2)

$$PV = C \cdot \frac{N_x}{D_x},$$

where $N_x = \sum_{j=0}^{u-x-1} D_{x+j}$

- *Present value of contributions dependent on salary* assuming that
  - the participant is aged x with n service years
  - the salary scale function is $\{s_x\}$ (i.e. the ratio of salary increase between ages x and $x + t$ is s_{x+t}/s_x), the participant's annual salary at age x is s, and the participant's annual contribution is c percent of annual salary
  - the normal retirement age is u
  - the discount operations are performed by means of the discount factor v (see Sect. 3.2)

$$PV = c \cdot s \cdot \frac{^sN_x}{^sD_x},$$

where $^sN_x = \sum_{j=0}^{u-x-1} {^sD_{x+j}}$

Denotation:

nc_t	normal contributions NC (see *thereinbefore*) at time t (relatively to the salary)
$\sum$	sum over all active participants of the pension plan at time t
$x(t)$	age of a participant at time t
x_0	entry age of a participant
$S_{x(t)}$	annual salary of a participant at age $x(t)$
$PV_{x:k}$	liability of the pension plan towards a participant aged x due to the participant's future service between ages x and $x + k$ respecting the future salary increase
F_t	actual value of the pension fund at time t

- *Methods of funding* by regularly collected contributions:
  - *Current Unit*: keeps the fund in such an amount, so that in any time the fund covers liabilities towards the participants (e.g. if the complete withdrawals might be immediately reimbursed); this objective implies the corresponding contribution percentage (it is the example of the so-called *fund-driven methods*, where the fund amount is primary, implying the contribution amount):

$$nc_t = \frac{\sum PV_{x(t)-1:\,\overline{1}|}}{\sum S_{x(t)}}$$

  - *Projected Unit*: is based on the same principle as the Current Unit (see *thereinbefore*), but one uses projections to the retirement age instead (it is again the example of fund-driven methods):

$$nc_t = \frac{\sum PV_{x(t)-\,1,:1\rceil}}{\sum S_{x(t)}} \qquad \text{(just as for the Current Unit, see } \textit{thereinbefore}\text{)}$$

– *Entry Age*: sets down the contribution in such an amount, so that right away at the entry age the present value of expected future contributions corresponds to the present value of expected future benefits *PVFB* (it is the example of the so-called *contribution-driven methods*, where the contribution level is primary, implying the amount of the cumulated fund):

$$nc_t = \frac{\sum \frac{PV_{x_0\,:\,u-x_0\rceil}}{{}^s\ddot{a}_{x_0\,:\,u-x_0\rceil}}}{\sum S_{x(t)}}, \quad \text{where } {}^s\ddot{a}_{x:u-x\rceil} = \frac{{}^sN_x}{{}^sD_x}$$

– *Attained Age*: sets down the contribution in such an amount, so that the value of future contributions expected since the attained age corresponds to the value of future benefits expected since the attained age (it is again the example of contribution-driven methods, see *thereinbefore*):

$$nc_t = \frac{\sum \frac{PV_{x(t)\,:\,u-\,x(t)\rceil}}{{}^s\ddot{a}_{x(t)\,:\,u-\,x(t)}}}{\sum S_{x(t)\rceil}}$$

– *Aggregate Method*: sets down the contribution in such an amount, so that the difference between the value of pension plan liabilities (for both the past and future service) and the value of the actual cumulated fund is covered by the amount of expected future contributions (again the example of contribution-driven methods, see *thereinbefore*)

$$nc_t = \frac{\sum PV_{x(t)\,:\,u-x(t)\rceil} - F_t}{\sum (S_{x(t)} \cdot {}^s\ddot{a}_{x(t)\,:\,u-x(t)\rceil})}$$

Further Reading

Anderson, A.W.: Pension Mathematics for Actuaries. Wellesley, Massachusetts (1992)
Booth, P., Chadburn, R., Cooper, D., Haberman, S., James, D.: Modern Actuarial Theory and Practice. Chapman and Hall / CRC, London (1999)
Bowers, N.L. et al.: Actuarial Mathematics. The Society of Actuaries, Itasca, IL (1986)
Koller, M.: Stochastische Modelle in der Lebensversicherung. Springer, Berlin (2000)
Lee, E.M.: An Introduction to Pension Schemes. The Institute of Actuaries and the Faculty of the Actuaries, London (1986)
Milbrodt, H., Helbig, M.: Mathematische Methoden der Personenversicherung. DeGruyter, Berlin (1999)
Teugels, J., Sundt, B. (eds.): Encyclopedia of Actuarial Science. Wiley, New York (2004)
Winklevoss, H.E.: Pension Mathematics with Numerical Illustrations. Irwin, Homewood, IL (1977)
Wolfsdorf, K.: Versicherungsmathematik (Teil 1: Personenversicherung). Teubner, Stuttgart, Germany (1997)

Chapter 21
Classical Non-Life Insurance

Abstract Chapter 21 contains formulas of classical non-Life insurance: 21.1. Basic Concepts of Non-Life Insurance, 21.2. Premium Calculations in Non-Life Insurance, 21.3. Forms of Non-Life Insurance and Deductibles, 21.4. Technical Provisions in Non-Life Insurance, 21.5. Bonus-Malus Systems.

21.1 Basic Concepts of Non-Life Insurance

- Some concepts of non-life insurance are used also in the context of life insurance (see Sect. 18.1)

- *Intensity of insurance protection*: is the ratio I of the insurance benefit to the corresponding claim ($0 \leq I \leq 1$):

$$I = \frac{\text{insurance benefit}}{\text{claim}}$$

- *Value insured V* (mainly in property and liability insurance, see Chap. 16) is
  - at time of the contract: the real value of the insured object
  - at time of the claim: the current value or the replacement value of the insured object (see *thereinafter*)

- *Current value insurance*: means that the insurance contract applies such a value of the insured object, which corresponds to the value of the new object reduced by the amortization
- *Replacement value insurance*: means that the insurance contract applies such a value of the insured object, which is necessary to replace the object at the time of the claim
- *Maximum possible loss MPL*: is the estimate of the maximum loss, which would occur as the result of a damage caused by the most destructive peril to be insured
- *Sum insured S* has technical motivation: (1) it is used for calculation of premiums; (2) it restricts insurance benefits from above; (3) it depends usually on the client's decision

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_21,

- *Underinsurance*: is a situation with $S < V$ (see Sect. 21.3); in such a case, the insurer may reduce insurance benefits using ratio S/V

- *Insurance premium*: may be classified as in the life insurance (see Sect. 18.1) but other classifications are also possible:
  - *natural premium*: covers the corresponding risk in a given insurance portfolio during 1 year (i.e. in such a portfolio there is a balance between collected premiums and paid insurance benefits within each year so that no premium reserves (see Sect. 18.3) are necessary
  - *multi-year insurance premium*: can be
    - *pre-paid premium*: with possibility of its reimbursed cancellation
    - *single premium*: without possibility of its cancellation
  - *written premium*: is given by insurance contracts
  - *collected premium*: is really paid by policyholders; it can be
    - *earned premium*: relates to the current accounting period (current accounting year)
    - *unearned premium*: relates to future accounting periods (see also the technical provision for unearned premium in Sect. 21.4)
  - *net premium*: is calculated so as to cover in average the corresponding insurance benefits (see also Sect. 18.1)
  - *risk premium*: is net premium with safety margin to cover adverse claim deviations
  - *gross premium* (*office premium*): is the expense-loaded net premium usually including also the security and profit loading:

$$\begin{aligned} &\text{gross premium} = \text{risk premium} + \text{expenses} + \text{calculated profit} = \\ &= \text{net premium} + \text{safety margin} + \text{expenses} + \text{calculated profit} \end{aligned}$$

  - *credibility premium*: is a combination of the premium based on insurer's past data and the premium based on global past data (over the global insurance market); this concept relates to *credibility*, which is attributed to insurer's data in comparison with global data (see Sects. 21.2 and 22.4)

- *Tariff groups*: are homogeneous groups of insurance contracts, for which the insured risk is approximately the same; therefore the insurer may fix the same premium rate within each tariff group; if the number of tariff groups is high then the risks are very strictly classified (e.g. in the accident insurance)

- *Cancellation* (*lapse*) of insurance contracts: can be
  - *voluntary*: due to decision of the client
  - *natural*: due to non-existence of insurable risk or interest
  - *due to the claim* (if the contract is automatically cancelled in such a case)

- *Deductible*: means (see Sects. 21.3 and 22.6) that the client participates in a given way in the claim settlement (it implies that the client's premium is lower when the client's deductible is stipulated)

- *Bonus-malus system* (see Sects. 21.5 and 22.7): means that the premium adjusts according to the claims experience (the reduction discounted from the premium is called *bonus*, and the penalty charged to the premium is called *malus* in such a case)

Denotation and statistical data (usually for particular tariff groups and calendar years):

N	number of policies
n	number of claims
total sum insured	i.e. sum over all policies
total claim	i.e. sum of claims over all policies (plus claim reserves (see Sect. 21.4) established in the given year for future years minus claim reserves established in previous years and used in the given year)
total premium	i.e. sum of premiums over all policies

- *Statistical indicators* (usually for particular tariff groups and calendar years):

$$\text{Average Insurance Benefit} = \frac{\text{total claim}}{N}$$

$$\text{Average Sum Insured} = \frac{\text{total sum insured}}{N}$$

$$\text{Average Claim} = \frac{\text{total claim}}{n}$$

$$\text{Loss Frequency} = q_1 = \frac{n}{N}$$

$$\text{Premium Rate} = \frac{\text{total premium}}{\text{total sum insured}}$$

$$\text{Claim Rate} = \frac{\text{total claim}}{\text{total sum insured}}$$

$$\text{Claim Ratio} = \frac{\text{total claim}}{\text{total premium}}$$

$$\text{Average Claim Degree} = q_2 = \frac{\text{Average Claim}}{\text{Average Sum Insured}}$$

$$\text{Average Insured Benefit} = q_1 \times \text{Average Claim} = \text{Claim Rate} \times \text{Average Sum Insured}$$

$$\text{Claim Rate} = \text{Claim Ratio} \times \text{Premium Rate} = q_1 q_2$$

- *Loss table*: is a frequency table (see Sect. 27.3) of the claim distribution for a given tariff group; it contains particular claim degrees (see *thereinbefore*) with corresponding relative frequencies; analogously as the life tables in life insurance (see Sect. 17.2), the loss tables in non-life insurance are constructed by means of real data, but for a hypothetical sample of claims (e.g. for a sample of 100,000 claims)

- Description of columns of a *loss table*:
  - z: *interval claim degrees* (e.g. if using ten interval claim degrees of the same width, then $z = 30\%$ denotes the interval claim degree $(20\%, 30\%\rangle$)
  - T_z: *number of claims* with the interval claim degree z (the sum of T_z over all z is equal to the total number of claims $n = 100{,}000$)
  - t_z: *relative frequency of claims* with the interval claim degree z: $t_z = T_z/n$
  - b_z: *cumulative relative frequency of claims* with the interval claim degrees at most z (i.e. the sum of relative frequencies of claims with the interval claim degrees not exceeding z)
  - Y_z: *weighted claim degree* z, where the relative frequency of claims with the interval claim degree z is used as the weight: $Y_z = t_z \times$ mid-point of interval z
  - G_z: *cumulative weighted claim degree* at most z (i.e. the sum of weighted claim degrees not exceeding z)

$$q_2 = G_{100\%}$$

(*average claim degree* for practical calculations of premiums (see Sect. 21.2) by means of a suitable loss table for a given tariff group)

$$q_2 = \int_0^1 t_z z \, \mathrm{d}z \quad (\textit{average claim degree}\text{: the continuous version})$$

21.2 Premium Calculations in Non-Life Insurance

Denotation:

S	sum insured
q_1	loss frequency (for a given tariff group, see Sect. 21.1)
q_2	average claim degree (for a given tariff group, see Sect. 21.1)
v	annual discount factor (see Sect. 3.2; it may be ignored when calculating annual natural premiums in non-life insurance)
e	average length of claim period (for a given tariff group, see *thereinafter*)

S_e	sum insured per unit of claim period
RP	risk premium (see Sect. 21.1)
N	number of policies (for a given tariff group)
s	estimated standard deviation (see Sect. 27.6) of claims (for particular policies of a given tariff group)
$\hat{\gamma}_1$	estimated skewness (see Sect. 27.7) of claims (for particular policies of a given tariff group)
P^{gross}	gross premium
prof	calculated profit (see Sect. 21.1) as a percentage of gross premium

$$P = vq_1q_2S$$

(*annual net premium*: it is the annual *natural* premium (see Sect. 21.1) per sum insured S; $v \cdot q_1 \cdot q_2$ is the corresponding *net premium rate* (i.e. the annual net premium per unit sum insured, see Sect. 21.1))

$$P = vq_1eS_e$$

(*annual net premium*: it is a particular case of the previous formula in such a situation that the sum insured S_e is paid per each time unit (e.g. per each day) of insured's stay in a specified *claim status* (e.g. in the status of sick persons in the context of sickness insurance, or in the status of persons unable to work in the context of accident insurance))

$$\Pi_n = P + Pv + \cdots + Pv^{n-1} = P\frac{1-v^n}{1-v} = P\ddot{a}_{\overline{n}|}$$

(*pre-paid net premium* (see Sect. 21.1) for n years, when the annual net premium is P; the reimbursement due to a possible cancellation after r years is
$\Pi_{n-\,r} = P\ddot{a}_{\overline{n-r}|}$)

$$\Pi_n = P\frac{1-v^n(1-a)^n(1-q_1)^n}{1-v(1-a)(1-q_1)}$$

(*single net premium* (see Sect. 21.1) for n years, when the annual net premium is P (a is the probability of natural cancellation, see Sect. 21.1))

$$\mathrm{RP} = (1+\lambda_1)P$$

(*risk premium* (i.e. the net premium with safety margin loaded, see Sect. 21.1) *on principle of mean value* ($\lambda_1 > 0$ is a constant))

$$\mathrm{RP} = P + \lambda_2 s$$

(*risk premium on principle of standard deviation* ($\lambda_2 > 0$ is a constant))

$$\mathrm{RP} = P + \lambda_3 s^2$$

(*risk premium on principle of variance* ($\lambda_3 > 0$ is a constant); the constants λ in the previous formulas for risk premiums can be obtained as confidence interval estimates (see Sect. 27.9) with prescribed confidence level assuming a suitable probability distribution of claims (see Sect. 21.1))

$$\mathrm{RP} = P + \lambda_3 s^2 + \lambda_4 \gamma_1$$

(*risk premium on principle of variance and skewness* ($\lambda_3 > 0$, $\lambda_4 > 0$ are constants))

$$\mathrm{RP} = \frac{1}{\lambda_5} \ln E\left(\mathrm{e}^{\lambda_5 X)}\right)$$

(*risk premium on exponential principle* ($\lambda_5 > 0$ is a constant, X has the claim distribution))

$$\mathrm{RP} = P + \frac{4}{\sqrt{N}} s$$

(*risk premium* on principle of standard deviation *used in practice*: the form given here shows a very high safety)

- *Expenses* in the context of non-life insurance can be classified to the expenses for
  - *acquisition* (mainly provisions of insurance agents)
  - *organization* (e.g. extensions of regional offices)
  - *administration*
  - *collection* of premiums
  - *cancellation* (lapses of policies)
  - *settlement* of claims
  - *independent* on premiums
  - *dependent* on premiums (e.g. they increase with increasing premiums)

$$P^{\text{gross}} = \frac{\mathrm{RP} + \mathrm{acq}_{\mathrm{ind}} + \mathrm{org}_{\mathrm{ind}} + \mathrm{adm}_{\mathrm{ind}} + \mathrm{coll}_{\mathrm{ind}} + \mathrm{canc}_{\mathrm{ind}} + \mathrm{settl}_{\mathrm{ind}}}{1 - (\mathrm{acq}_{\mathrm{dep}} + \mathrm{org}_{\mathrm{dep}} + \mathrm{adm}_{\mathrm{dep}} + \mathrm{coll}_{\mathrm{dep}} + \mathrm{canc}_{\mathrm{dep}} + \mathrm{settl}_{\mathrm{dep}}) - \mathrm{prof}}$$

(*gross premium*)

$$P^{\text{gross}} = \frac{\mathrm{RP}}{1 - \varepsilon - \mathrm{prof}}, \quad (\textit{gross premium})$$

where ε are combined expenses (in percents of gross premium)

$$P_{\text{credib}} = Z \cdot P_{\text{own}} + (1 - Z) \cdot P_{\text{global}}$$

(*credibility premium* (see Sect. 21.1): P_{own} is the premium based on insurer's past data; P_{global} is the premium based on global past data (over the global insurance market); $Z\,(0 \le Z \le 1)$ is a *credibility coefficient*, i.e. the weight corresponding to the own data; in the context of theory of risk (see Sect. 22.4), one solves the problem of an optimal value of Z (the so-called *theory of credibility*); typically, Z increases with increasing insurance portfolio, but decreases with increasing variability of claims)

$$P_t = \alpha X_t + (1 - \alpha) P_{t-1},$$

where α $(0 \le \alpha \le 1)$ is a *smoothing constant*

(*exponential smoothing* (see Sect. 31.2): is a method of the so-called *experience-rating*, which generalizes the credibility approach to the calculation of premiums in non-life insurance (see *thereinbefore*); the past premiums are all the time corrected using the a posteriori information on the current claim development of the given insurance product)

21.3 Forms of Non-Life Insurance and Deductibles

Denotation:

S	sum insured (in a given policy, see Sect. 21.1)
V	value insured (in a given policy, see Sect. 21.1)
X	loss amount (in a given policy)
B	insurance benefit (claim payment in a given policy)
q_1	loss frequency (in a given tariff group, see Sect. 21.1)
q_2	average claim degree (in a given tariff group, see Sect. 21.1)
v	annual discount factor (see Sect. 3.2; it may be ignored when calculating annual natural premiums in non-life insurance)
B_D, B_I	client's deductible and insurer's payment in a policy with insurance benefit B $(B = B_D + B_I)$

- *Form of non-life insurance*: is the function that shows how the insurance benefit (the claim payment) depends on the loss amount incurred in the given policy (usually through the sum insured); moreover, the form of insurance determines the intensity of insurance protection (see Sect. 21.1); the forms of non-life insurance can be classified to (1) *sum insurances*; (2) *loss insurances*; (3) *deductibles*:

  (1) *Sum insurances*: typically for this form, the insurance benefit is equal directly to the sum insured abstractedly from the loss incurred (e.g. the insurance benefit for accidental death in the accident insurance); unlike the life insurance, this form is not typical for the non-life insurance; the intensity of insurance protection I cannot be found for this form:

$$B_{\text{sum}} = S \quad \text{(insurance benefit in } \textit{sum insurance} \text{ policies)}$$

$$P_{\text{sum}} = vq_1 S \quad \text{(net premium in } \textit{sum insurance} \text{ policies)}$$

(2) *Loss insurances* (*indemnity insurances*): the insurance benefit depends on the loss amount X incurred in the given policy

$$B \leq X \quad \text{(insurance benefit in } \textit{loss insurance} \text{ policies)}$$

It is the typical form of property and liability insurances; in practice one applies various forms of the loss insurances:

- *insurance without sum insured*: the insurance benefit is equal directly to the loss incurred (the sum insured is not prescribed at all in such a case); this form is not usual in practice, though it provides the full insurance protection ($I = 1$):

$$B_{\text{without}} = X \quad \text{(insurance benefit in insurance policies } \textit{without sum insured}\text{)}$$

$$P_{\text{without}} = vq_1 q_2 V \quad \text{(net premium in insurance policies } \textit{without sum insured}\text{)}$$

- *full value insurance*: the client's choice of a suitable sum insured S ($S \leq V$), which is related to the value insured V through a ratio s ($s \leq 1$), implies the corresponding intensity of insurance protection (since it holds $I = s$); it is the preferred form of the property insurances:

$$B_{\text{full}} = sX,$$

where $s = \dfrac{S}{V}$ (insurance benefit in *full value insurance* policies)

$${}^{S}P^{V}_{\text{full}} = vq_1 q_2 S = vq_1 q_2 sV = sP_{\text{without}}$$

(net premium in *full value insurance* policies: if $S = V$ (i.e. $s = 1$), then the client has chosen the insurance without sum insured (see *thereinbefore*); if $S < V$ (i.e. $s < 1$), then the client has chosen the *underinsurance* (with an implicit deductible of the client, see Sect. 21.1))

- *first loss insurance*: this form has $B = X$ (similarly as for the insurance without sum insured, see *thereinbefore*), but the insurance benefit is upper bounded by the sum insured S (i.e. an implicit deductible of the client is again possible); it is popular e.g. in the liability insurance:

$$B_{\text{first}} = \begin{cases} X & \text{for } X \leq S \\ S & \text{for } X > S \end{cases} \text{(insurance benefit in } \textit{first loss insurance} \text{ policies)}$$

$$^{S}P^{V}_{\text{first}} = vq_1[G_s V + (1 - b_s)S] = vq_1[G_s + (1 - b_s)s]V, \quad \text{where } s = \frac{S}{V}$$

(net premium in *first loss insurance* policies: one may make use of the values from a suitable loss table for the given tariff group (see Sect. 21.1))

(3) *Deductibles*: mean (see also Sect. 22.6) that the clients participate in the claim settlements, which implies lower premiums than in the case without deductibles; the deductibles must be always combined with basic forms of insurance:

– *quota deductible*: the insured's deductible is a stipulated proportion $(1 - q)$ of the insurance claim (*quota* q gives the proportion of the insurer's participation):

$$B_D = (1 - q)B$$

(insured's *quota deductible*)

$$_{q}P_{\text{without}} = qP_{\text{without}}$$

(net premium combining insurance *without sum insured* (see *thereinbefore*) with *quota deductible*)

– *excess deductible*: the insured's deductible is the whole claim amount, which does not exceed a stipulated amount F_e (the so-called *priority*), or it is the amount F_e, if the claim amount is higher than F_e:

$$B_D = \begin{cases} B & \text{for } B \le F_e \\ F_e & \text{for } B > F_e \end{cases}$$

(insured's *excess deductible*)

$$^{S}_{F_e}P^{V}_{\text{first}} = vq_1[G_s + (1 - b_s)s - G_{f_e} - (1 - b_{f_e})f_e]V, \quad \text{where } s = \frac{S}{V}, f_e = \frac{F_e}{V}$$

(net premium combining *first loss insurance* (see *thereinbefore*) with *excess deductible*)

– *integral deductible*: the insured's deductible is the whole claim amount, which does not exceed a stipulated amount F_i, or it is zero, if the claim amount is higher than F_i:

$$B_D = \begin{cases} B & \text{for } B \le F_i \\ 0 & \text{for } B > F_i \end{cases} \quad \text{(insured's } \textit{excess deductible}\text{)}$$

$$ {}_{F_i}^{S}P_{\text{full}}^{V} = vq_1(q_2 - G_{f_i})S, \quad \text{where } f_i = \frac{F_i}{V} $$

(net premium combining *full value insurance* (see *thereinbefore*) with *integral deductible*)

21.4 Technical Provisions in Non-Life Insurance

- *Technical provisions* (*insurance reserves*): are established (similarly as in the life insurance, see Sect. 18.3) by insurers (mostly as the book costs according to law) to fulfil obligations arising from insurance activities (such obligations are probable or certain, though their amount or time may be still uncertain); the technical provisions are important statutory liabilities of each insurance company and are conform to special accounting principles and tax regulations; the assets covering the technical provisions are subject to strict investment restrictions in order that their financial placement fulfils principles of prudence, diversification, profitability and liquidity; according to particular insurance legislations, various technical provisions may be established in the non-life insurance (see also Sect. 18.3 for the life insurance):

  - *reserve for unearned premium*: similarly as in the life insurance (see Sect. 18.3), this reserve corresponds to such a part of the written premium, which relates to future accounting periods
  - *claim reserve*: unlike the life insurance (see Sect. 18.3), it is usually the most important technical provisions in the framework of non-life insurance due to the fact, that the final evaluation of insurance claims (e.g. the confirmation of the degree of disability) may sometimes take several years; the claim reserves cover obligations due to the claims, which in the current accounting period have been:

    - *reported but not settled* (the so-called *RBNS reserve*)
    - *incurred but not reported* (the so-called *IBNR reserve*):

    $$ \text{claim reserve} = \text{RBNS reserve} + \text{IBNR reserve} $$

    qualified estimates of the claim reserves can be obtained by means of actuarial methods based on *run-off triangles* (see *thereinafter*) respecting external factors (inflation, changes of legislative, and the like)

  - *reserve for bonuses and rebates*: covers costs of bonuses and rebates guaranteed by insurance policies (e.g. the bonus system in the motor third party liability insurance)
  - *equalization reserve*: should cover increased costs of insurance claims due to CR (claim ratio) fluctuations (see Sect. 21.1), which cannot be managed by

the insurer; this reserve is typical for special insurance products, e.g. for the credit insurance; in some countries it can have a prescribed form, e.g.

$$k\sqrt{\frac{1}{14}\sum_{i=1}^{15}(\mathrm{CR}_{t-i}-\overline{\mathrm{CR}}_t)^2 P_t}, \quad \text{where } \overline{\mathrm{CR}}_t = \frac{1}{15}\sum_{i=1}^{15}\mathrm{CR}_{t-i}$$

(amount of *equalization reserve* at time t (Germany): k-multiple (e.g. $k = 4.5$) of the sample standard deviation of the claim rate (see Sect. 21.1) over the last 15 years applied to the premium P_t at time t)

– other life insurance reserves approved by authorities (by state insurance supervisions similarly as for the life insurance (see Sect. 18.3))

- *Run-off triangle*: is a specific arrangement of past claim data of the given insurance company, which is used for qualified estimation of its claim reserves (mainly in the case of IBNR reserves, but sometimes also for the claim reserves in total (see *thereinbefore*)); the past claims paid off yearly by the insurer up to the current time t of estimation are arranged to rows according to the accident years and to columns according to the development years (see *thereinafter*, though sometimes other schemes are possible, e.g. with quarterly data reflecting seasonality, with rows according to the reporting years, with rows according to the underwriting years, and the like):
  – *accident year*: is the year, in which the given claim incurred (if the triangle concerns ten past years, then e.g. the first row contains total payments on the claims incurred in the first year of this past decade)
  – *development year*: means how many years (before the payment) the given claim has incurred (e.g. the first column denoted as "0" contains total claim payments, which have been made directly in the year, in which the corresponding claims incurred)
  – *calendar year*: the particular south-east diagonals of the run-off triangle correspond to the corresponding calendar years (e.g. the main diagonal contains total claim payments, that have been made in the most recent calendar year of the considered past period)

- *Cumulative run-off triangle* (see below table): is constructed by the sequential accumulation of the elements of particular rows in the run-off triangle:

$$X_{j,s} = \sum_{k=0}^{s} Y_{j,k}\,, \quad j = 1,\dots,t; \quad s = 0,\dots,t-1$$

(*cumulative run-off triangle*: $Y_{j,\,s}$ is the total payment for the claims incurred in the accident year j, which has been paid out in the development year s)

	Development year						
Accident year	0	1	...	s	...	$t-2$	$t-1$
1	$X_{1,0}$	$X_{1,1}$	...	$X_{1,s}$	...	$X_{1,t-2}$	$X_{1,t-1}$
2	$X_{2,0}$	$X_{2,1}$	...	$X_{2,s}$	...	$X_{2,t-2}$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		
j	$X_{j,0}$	$X_{j,1}$	...				
$\vdots$	$\vdots$	$\vdots$					
$t-1$	$X_{t-1,0}$	$X_{t-1,1}$					
t	$X_{t,0}$						

- *Run-off triangle methods*: are statistical methods (Bornhuetter-Ferguson, Cape Cod, Chain-Ladder, de Vylder, London Chain-Ladder, separation and others, see *thereinafter*) that enable to complete (cumulative) run-off triangles to rectangles by estimated future values (predictions) $\hat{X}_{j,\,s}$ and to obtain in this way the prediction $\hat{X}_{j,\infty}$ of the total payments on the claims incurred in the year j ($j = 1, \ldots, t$):

$$\hat{X}_{j,\infty} - X_{j,\,t-j} \approx \hat{X}_{j,\,t-1} - X_{j,\,t-j}, \quad j = 1, \ldots, t$$

(the part of the claim reserve, which at the end of the year t covers future payments on the claims incurred in the year j (it holds $\hat{X}_{j,\,t-1} \approx \hat{X}_{j,\infty}$ for sufficiently large t)

$$Y_{j,\,s} = X_{j,\,s} - X_{j,\,s-1}, \quad j = 1, \ldots, t; \quad s = 0, \ldots, t-1$$

(non-cumulative values: $Y_{j,\,s}$ is the total payment on the claims incurred in the accident year j, which has been paid out in the development year s (see *thereinbefore*))

$$\hat{c}_s = \frac{\sum_{j=1}^{t-s-1} X_{j,\,s+1}}{\sum_{j=1}^{t-s-1} X_{j,\,s}}, \quad s = 0, \ldots, t-2$$

(*development coefficients* for cumulative run-off triangle)

$$\hat{c}_s = \frac{\sum_{j=1}^{t-s-1} w_{j,\,s}\, c_{j,\,s}}{\sum_{j=1}^{t-s-1} w_{j,\,s}}, \quad \text{where} \quad c_{j,\,s} = \frac{X_{j,\,s+1}}{X_{j,\,s}}, \quad j = 1, \ldots, t-1; \quad s = 0, \ldots, t-2$$

(*modified* development coefficients for cumulative run-off triangle: $w_{j,\,s}$ are weights chosen in such a way that the weights of more recent data are higher than the weights of earlier data (one of possible choices is $w_{j,\,s} = j + s$))

$$\hat{X}_{j,\,s} = \hat{c}_{t-j} \cdot \ldots \cdot \hat{c}_{s-1} \cdot X_{j,\,t-j}, \quad j = 2, \ldots, t; \quad s = t - j + 1, \ldots, t - 1$$

(*Chain-Ladder* method and its various modifications: is based on the assumption that the ratio of cumulative claim amounts between neighbouring development years remains approximately same over particular years of accident; therefore the completion of the triangles to the rectangles can be done just by means of the *development coefficients* $\hat{c}_s$ (see *thereinbefore*))

$$\hat{X}_{j,\,s} = \hat{c}_{j,\,t-j} \cdot \ldots \cdot \hat{c}_{j,\,s-1} \cdot X_{j,\,t-j}, \quad j = 2, \ldots, t; \quad s = t - j + 1, \ldots, t - 1$$

(Chain-Ladder method with development coefficients *depending on accident year j*: the estimated development coefficients $\hat{c}_{j,\,s}$ can be obtained from $c_{j,\,s}$ (see *thereinbefore*) by means of extrapolation (to the rectangle) in columns of the triangle $c_{j,\,s}$ ($j = 1, \ldots, t - 1$; $s = 0, \ldots, t - 2$) using a trend regression; moreover, in the last column one can extrapolate by repeating the value $c_{1,\,t-2}$, in the last but one column by repeating the average $(c_{1,\,t-3} + c_{2,\,t-3})/2$, and the like)

- *Separation methods*: are special run-off triangle methods, which separate the stable (i.e. time invariant) distribution $\{r_0, r_1, \ldots, r_{t-1}\}$ of delayed claims from unstable factors $\{\lambda_1, \lambda_2, \ldots, \lambda_t\}$ that present dependence on time including the inflation (see below table); non-cumulative run-off triangles are applied in this context (see *thereinbefore*):

$$Y_{j,\,s} = r_s \lambda_{j+s}, \quad j = 1, \ldots, t; \quad s = 0, \ldots, t - 1; \quad j + s \le t$$

(principle of *separation methods*: $Y_{j,\,s}$ is the total payment on the claims incurred in the accident year j, which has been paid out in the development year s (see *thereinbefore*); moreover, these payments in particular rows are at first divided by values that can be taken as suitable risk measures for accident particular years (e.g. by the numbers of claims incurred in particular accident years)

	Development year				
Accident year	0	1	…	$t-2$	$t-1$
1	$r_0\lambda_1$	$r_1\lambda_2$	…	$r_{t-2}\lambda_{t-1}$	$r_{t-1}\lambda_t$
2	$r_0\lambda_2$	$r_1\lambda_3$	…	$r_{t-2}\lambda_t$	
⋮	⋮	⋮	⋮		
$t-1$	$r_0\lambda_{t-1}$	$r_1\lambda_t$			
t	$r_0\lambda_t$				

$$\sum_{s=0}^{t-1} r_s = 1$$

(*arithmetic* separation method)

$$\hat{\lambda}_k = \frac{\sum_{j=1}^{k} Y_{j,\,k-j}}{\sum_{s=0}^{k-1} \hat{r}_s}, \quad k = 1, \ldots, t-1; \quad \hat{\lambda}_t = \sum_{j=1}^{t} Y_{j,\,t-j}; \quad \hat{r}_s = \frac{\sum_{j=1}^{t-s} Y_{j,\,s}}{\sum_{k=s+1}^{t} \hat{\lambda}_k}, \quad s = 0, \ldots, t-1$$

(separation of factors in *arithmetic separation method*: the given sequence of calculations must be strictly followed)

$$\prod_{s=0}^{t-1} r_s = 1$$

(*geometric* separation method)

$$\hat{\lambda}_k = \left[\frac{\prod_{j=1}^{k} Y_{j,\,k-j}}{\prod_{s=0}^{k-1} \hat{r}_s}\right]^{\frac{1}{k}}, \quad k = 1, \ldots, t-1; \quad \hat{\lambda}_t = \left[\prod_{j=1}^{t} Y_{j,\,t-j}\right]^{\frac{1}{t}}$$

$$\hat{r}_s = \left[\frac{\prod_{j=1}^{t-s} Y_{j,\,s}}{\coprod_{k=s+1}^{t} \hat{\lambda}_k}\right]^{\frac{1}{t-s}}, \quad s = 0, \ldots, t-1$$

(separation of factors in *geometric separation method*: the given sequence of calculations must be strictly followed)

$$\hat{Y}_{j,\,s} = \hat{r}_s \hat{\lambda}_{j+s}, \quad j = 1, \ldots, t; \quad s = 0, \ldots, t-1; \quad j + s > t$$

(*separation method*: the estimates $\hat{\lambda}_k$ for $k = t + 1, \ldots, 2t - 1$, which are necessary in order to complete the (non-cumulative) run-off triangle to the rectangle, are obtained by extrapolating the values $\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_t$)

- *Method of claim ratio*: this method of claim reserve estimation applies the claim ratio CR_j in the year j (see Sect. 21.1) to the earned premium EP_j in the year j subtracting the cumulative claim amount $X_{j,\,t-j}$ (i.e. subtracting all past payments for the claims incurred in the year j):

$$\mathrm{CR}_j\mathrm{EP}_j - X_{j,\,t-j}, \quad j = 1, \ldots, t$$

(*method of claim ratio*: the part of the claim reserve, which in the year t covers all future payments on the claims incurred in the year j)

$$\mathrm{CR}_j\mathrm{EP}_j(1 - 1/\hat{c}_{j,\,\infty}), \quad \text{where} \quad \hat{c}_{j,\,\infty} = \hat{c}_{j,\,t-j}\hat{c}_{j,\,t-j+1} \cdot \ldots \cdot \hat{c}_{j,\,t-2}; \quad j = 1, \ldots, t$$

(*Bornhuetter-Ferguson method*: the part of the claim reserve, which in the year t covers all future payments on the claims incurred in the year j; the estimated development coefficients $\hat{c}_{j,\,s}$ can be obtained similarly as in the Chain-Ladder method (see *thereinbefore*))

21.5 Bonus-Malus Systems

- *Bonus-malus system* (see Sect. 22.7): means that the premium adjusts according to claims experience; in practice the bonus-malus systems are typical mainly for the private motor insurance (i.e. both for the hull insurance and for the motor third party liability insurance); the *bonus systems* without malus, which are also called *No Claim Discount systems* (*NCD systems*), are more frequent in practice
- *Bonus*: is a premium discount (guaranteed in the insurance policy) usually according to the number of the past policy years without claims
- *Malus*: is a premium penalty (specified in the insurance policy) usually according to the number and amount of the past policy claims
- *Bonus load* of a bonus system (P' are the total insurance premiums applying bonus, P are the corresponding total insurance premiums theoretically without applying bonus):

$$\left(1 - \frac{P'}{P}\right) 100 \ \ (\%)$$

- *Hunger for bonus*: is a situation, where clients prefer not to report claims in order not to impair their bonus positions
- Characteristics of bonus systems are: (1) number of *bonus levels* corresponding to particular discounts of the so-called *basic premium* (i.e. the premium without bonus); (2) *bonus level at entry*; (3) *decisive period* (i.e. the period since the previous claim reported in the given insurance policy without another claim reported; (4) *system rules* (e.g. which is the reduction of the attained bonus level after a claim reported, which is the new decisive period in such a case, and the like); *mathematical models of bonus-malus systems* (see Sect. 22.7) are based on the theory of *Markov chains* (see Sect. 30.2): e.g. the transition probabilities between particular bonus levels are important

Further Reading

Booth, P., Chadburn, R.., Cooper, D., Haberman, S., James, D.: Modern Actuarial Theory and Practice. Chapman and Hall/CRC, London (1999)

Daykin, C.D., Pentikäinen, T., Pesonen, M.: Practical Risk Theory for Actuaries. Chapman and Hall, London (1994)

Hart, D.G., Buchanan, R.A., Howe, B.A.: The Actuarial Practice of General Insurance. The Institute of Actuaries of Australia, Sydney, NSW (1996)

Mack, T.: Schadenversicherungsmathematik. VVW, Karlsruhe (2002)

Straub, E.: Non-Life Insurance Mathematics. Springer, Berlin (1988)

Taylor, G.C.: Claims Reserving in Non-Life Insurance. North-Holland, Amsterdam (1986)

Teugels, J., Sundt, B. (eds.): Encyclopedia of Actuarial Science. Wiley, New York (2004)

Wolfsdorf, K.: Versicherungsmathematik (Teil 2: Theoretische Grundlagen, Risikotheorie, Sachversicherung). Teubner, Stuttgart (1986)

Chapter 22
Risk Theory in Insurance

Abstract Chapter 22 presents basic formulas of risk theory in the context of insurance: 22.1. Collective Risk Model, 22.2. Aggregate Claim Distribution, 22.3. Copula, 22.4. Credibility Premium, 22.5. Ruin Probability, 22.6. Deductible, 22.7. Calculations for Bonus-Malus Systems.

22.1 Collective Risk Model

- *Risk model* in insurance: is intended to model the total loss amount Z for claims, which incurred during a time period of length T (e.g. during 1 year) in a given insurance portfolio
- *Individual risk model*: deals with risks corresponding to particular (individual) insurance policies:

$$Z = \sum_{i=1}^{k} Y_i$$

  (total claim amount in *individual risk model*: $Y_1, \ldots, Y_k$ is a sequence of claims corresponding to particular insurance policies for an insurance portfolio consisting of k policies; typically, Y_i are independent random variables (see Sect. 26.6))
- *Collective risk model*: assumes that in homogeneous insurance portfolios (or tariff groups, see Sect. 21.1) the claims incurred due to particular insurance events are identically distributed (and mostly also independent) random variables:

$$Z = \sum_{i=1}^{N} X_i$$

  (total claim amount in *collective risk model*: $X_1, \ldots, X_N$ is a sequence of claims (unlike the individual risk model, the ordering of this sequence is arbitrary

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_22,

regardless of the corresponding policies), and N is the number of claims during a given period; typically, X_i are independent and identically distributed random variables which are independent of the random variable N)

- *Models for number of claims*: assume a probability distribution of the random variable N (the number of claims during a given period, see *thereinbefore*) with values 0, 1, 2, ...; the best used probability distributions of N are:

  – *Poisson distribution* (see Sect. 26.4) $N \sim P(\lambda)$: is the discrete probability distribution with one parameter $\lambda > 0$; such an N may model the number of claims for a large number of independent homogeneous policies with a small probability of the claim (so called "distribution of rare events"):

  $$P(N = j) = \mathrm{e}^{-\lambda}\frac{\lambda^j}{j\,!}, \quad j = 0, 1,\ 2,\ \ldots; \quad \mathrm{E}(N) = \lambda = \mathrm{var}(N)$$

  – *negative binomial distribution* (see Sect. 26.4) $N \sim \mathrm{NB}(r, p)$: is the discrete probability distribution with two parameters $r > 0$ and $0 < p < 1$; such an N may model (for $r \in \mathrm{N}$) the number of failures before the rth success in independent trials with probability p of success:

  $$P(N = j) = \binom{r + j - 1}{j} p^r (1 - p)^j, \quad j = 0, 1,\ 2,\ \ldots$$
  $$\mathrm{E}(N) = \frac{r(1 - p)}{p} < \mathrm{var}(N) = \frac{r(1 - p)}{p^2}$$

  – *mixed Poisson distribution*: is the Poisson distribution with random parameter λ (λ is interpreted as a random intensity with distribution function $F(\lambda)$); it is applied to insurance portfolios with heterogeneous risks, where insurance policies with a small risk (or a large risk) have λ with small values (or large values), respectively; in particular, the case of a constant λ with $P(\lambda = \lambda_0) = 1$ implies the distribution $P(\lambda_0)$ of N (see *thereinbefore*), and the case of the gamma distribution (see Sect. 26.5) with parameters $p/(1 - p)$ and r implies the distribution $\mathrm{NB}(r, p)$ of N (see *thereinbefore*):

  $$P(N = j) = \int_0^\infty \mathrm{e}^{-\lambda}\frac{\lambda^j}{j\,!}\,\mathrm{dF}(\lambda), \quad j = 0, 1,\ 2,\ \ldots$$

- *Models for number of claims in K tariff groups*: N denotes the number of claims during a given period in K mutually independent tariff groups (i.e. $N = N_1 + \ldots + N_K$, where N_i is the number of claims during a given period in the ith tariff group):

– if $N_i \sim P(\lambda_i)$, then $N \sim P(\lambda_1 + \cdots + \lambda_K) \approx N(\lambda_1 + \cdots + \lambda_K, \lambda_1 + \cdots + \lambda_K)$
– if $N_i \sim \mathrm{NB}(r_i, p)$, then $N \sim \mathrm{NB}(r_1 + \cdots + r_K, p)$

- *Models for claim amount*: assume a probability distribution of the random variable X (usually the claim amount per one claim during a given period, see *thereinbefore*) with non-negative values; the best used probability distributions of X are:

  – *logarithmic normal distribution* (see Sect. 26.5) $X \sim \mathrm{LN}(\mu, \sigma^2)$: is the continuous distribution with two parameters $-\infty < \mu < \infty$ and $\sigma > 0$; it holds $\ln X \sim \mathrm{N}(\mu, \sigma^2)$; such an X may model the claim amount e.g. in accident, private motor, fire, windstorm and other insurances:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma\, x} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right), \quad x > 0$$

$$\mathrm{E}(X) = \exp\left(\mu + \frac{1}{2}\sigma^2\right)$$

$$\mathrm{var}(X) = \exp\left(2\mu + \sigma^2\right)\left(\exp(\sigma^2) - 1\right)$$

  – *gamma distribution* (see Sect. 26.5) $X \sim \Gamma(a, \lambda)$: is the continuous distribution with two parameters $a > 0$, $\lambda > 0$:

$$f(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1}\, \mathrm{e}^{-\lambda x}, \quad x \geq 0; \quad \mathrm{E}(X) = \frac{a}{\lambda}; \quad \mathrm{var}(X) = \frac{a}{\lambda^2}$$

  – *Weibull distribution* (see Sect. 26.5): is the continuous distribution with two parameters $a > 0$, $\lambda > 0$:

$$f(x) = a\lambda^a x^{a-1} \exp\left(-\lambda^a x^a\right), \quad x > 0$$

$$\mathrm{E}(X) = \frac{1}{\lambda}\Gamma\left(\frac{a+1}{a}\right)$$

$$\mathrm{var}(X) = \frac{1}{\lambda^2}\left(\Gamma\left(\frac{a+2}{a}\right) - \Gamma^2\left(\frac{a+2}{a}\right)\right)$$

  – *exponential distribution* (see Sect. 26.5) $X \sim \mathrm{Exp}(\lambda)$: is the continuous distribution with one parameter $\lambda > 0$; it is a special case of the gamma distribution and the Weibull distribution for $a = 1$ (see *thereinbefore*); the exponential distribution is also used to model the lengths of periods between insurance claims (see Sect. 22.2):

$$f(x) = \lambda\, \mathrm{e}^{-\lambda x}, \quad x \geq 0; \quad \mathrm{E}(X) = \frac{1}{\lambda}; \quad \mathrm{var}(X) = \frac{1}{\lambda^2}$$

– *beta distribution* (see Sect. 26.5): is the continuous distribution with two parameters $p > 0$, $q > 0$; the U-shaped probability density of the beta distribution for $p < 1$, $q < 1$ is used to model claims e.g. in the fire insurance (either very small claims, or on contrary very large ones are typical for this insurance product):

$$f(x) = \frac{1}{B(p,q)} x^{p-1}(1-x)^{q-1}, \;\; 0 < x < 1$$

$$\mathrm{E}(X) = \frac{p}{p+q}; \quad \mathrm{var}(X) = \frac{pq}{(p+q)^2(p+q+1)}$$

– *Pareto distribution* (see Sect. 26.5): is the continuous distribution with two parameters $a > 0$, $b > 0$; due to heavy tails, it is applied in situations with outlying extreme loss amounts, e.g. in the sickness, fire and other insurances:

$$f(x) = \frac{b\, a^b}{x^{b+1}}, \quad x \geq a$$

$$\mathrm{E}(X) = \frac{ab}{b-1} \quad \text{pro} \;\; b > 1$$

$$\mathrm{var}(X) = \frac{a^2 b}{(b-1)^2(b-2)} \quad \text{for} \;\; b > 2$$

22.2 Aggregate Claim Distribution

Assumptions and denotation:

$X_1, X_2, \ldots$	sequence of *iid* non-negative random variables with the distribution function F_X (claim amounts, see Sect. 22.1)
N	random variable independent of X_1, X_2, ... with non-negative integer values (number of claims, see Sect. 22.1)
$Z = X_1 + \ldots + X_N$	*aggregate claim amount* (*total loss*) in the collective risk model (see Sect. 22.1)

- *Compound* distribution: arises by compounding the distribution of the random variable N and the distribution of the random variables X_i (i.e. the distribution of the random variable X due to the identical distribution of X_i); it models the aggregate claim amount (total loss) Z that aggregated during the given period in the given insurance portfolio; in other words, Z is the random sum of random variables (see Sect. 26.12); its distribution can be derived applying e.g. the moment generating functions (see Sect. 26.10):

$$m_Z(x) = \sum_{j=0}^{\infty} P(N=j)\,(M_X(x))^j = M_N\,(\ln M_X(x))$$

(moment generating function (see Sect. 26.10) of the aggregate claim amount Z by means of the moment generating functions M_N and M_X)

$$F_Z(x) = \sum_{j=0}^{\infty} P(N=j) F_X^{*j}(x), \quad x \in \mathrm{R}$$

(distribution function of the aggregate claim amount Z by means of the convolution (see Sect. 26.11) of the distribution function F_X)

$$\mathrm{E}(Z) = \mathrm{E}(N) \times \mathrm{E}(X);\;\; \mathrm{var}(Z) = \mathrm{E}(N)\mathrm{var}(X) + \mathrm{var}(N)\,(\mathrm{E}(X))^2$$

- *Compound Poisson distribution* (see Sect. 26.12) $Z \sim \mathrm{CP}(\lambda, F_X)$: is the distribution of the aggregate claim amount for $N \sim P(\lambda)$:

$$F_Z(x) = \sum_{j=0}^{\infty} \mathrm{e}^{-\lambda}\, \frac{\lambda^j}{j!}\, F_X^{*j}(x)$$

(distribution of the aggregate claim amount Z with *compound Poisson distribution*)

$$\mathrm{E}(Z) = \lambda \mathrm{E}(X); \quad \mathrm{var}(Z) = \lambda \mathrm{E}(X^2); \quad \gamma_1 = \frac{\mathrm{E}(X^3)}{\left(\lambda\left(\mathrm{E}(X^2)\right)^3\right)^{1/2}}; \quad \gamma_2 = \frac{\mathrm{E}(X^4)}{\lambda\left(\mathrm{E}(X^2)\right)^2}$$

$$X + Y \sim \mathrm{CP}(\lambda_X + \lambda_Y, (\lambda_X F_X + \lambda_Y F_Y)/(\lambda_X + \lambda_Y))\,,$$

where $X \sim \mathrm{CP}(\lambda_X, F_X)$ and $Y \sim \mathrm{CP}(\lambda_Y, F_Y)$ are mutually independent

- *Numerical methods* (for the calculation of the distribution of the aggregate claim amount) applied in practice instead of analytic (complicated) formulas:

  (1) *Panjer's recursive formula*: is applicable in the case of X (the distribution of the claim amounts) with natural values only:

Denotation:

$P(X = k) = x(k),\ k \in \mathrm{N}$
$P(Z = j) = z(j),\ j \in \mathrm{N}_0$

$$z(j+1) = \frac{\lambda}{j+1}\sum_{i=0}^{j} z(j-i)x(i+1)(i+1),\ \ j \in \mathrm{N}_0;\quad z(0) = e^{-\lambda}$$

(distribution of Z by means of *Panjer's recursive formula* for $N \sim P(\lambda)$ (see Sect. 22.1))

$$z(j+1) = \frac{1}{j+1}\sum_{i=0}^{j} z(j-i)x(i+1)(1-p)\,((j-i) + r\cdot(i+1)),\ \ j \in \mathrm{N}_0;\quad z(0) = p^r$$

(distribution of Z by means of *Panjer's recursive formula* for $N \sim \mathrm{NB}(r, p)$ (see Sect. 22.1))

Denotation:

$\mu = \mathrm{E}\,(Z)$	mean value (see Sect. 26.3) of the aggregate claim amount Z
$\sigma^2 = \mathrm{var}\,(Z)$	variance (see Sect. 26.3) of the aggregate claim amount Z
$\gamma_1 = \gamma_1\,(Z)$	skewness (see Sect. 26.3) of the aggregate claim amount Z
$\gamma_2 = \gamma_2\,(Z)$	kurtosis (see Sect. 26.3) of the aggregate claim amount Z
$f_{\mathrm{norm}Z}\,(x)$	probability density (see Sect. 26.3) of the normed aggregate claim amount $(Z-\mu)/\sigma$
$\Phi(x),\ \Phi^{(j)}(x)$	distribution function of $N(0, 1)$ (see Sect. 26.5) and its jth derivative
$\varphi(x)$	probability density of $N(0, 1)$ (see Sect. 26.5)

(2) *Approximation by normal distribution*: is denoted as *approximation* NP1 (in contrast to the *power* approximation by normal distribution denoted as NP2, see *thereinafter*):

$$P\left(\frac{Z-\mu}{\sigma} \le x\right) \approx \Phi(x)$$

(distribution of Z: *approximation by normal distribution*)

(3) *Approximation by gamma distribution* (see Sect. 22.1): due to the asymmetry it takes into account also the skewness (see Sect. 26.3) of Z (one uses the gamma distribution in the form $\Gamma(a, 1)$):

$$P\left(\frac{Z-\mu}{\sigma} \le x\right) \approx \frac{1}{\Gamma(a)}\int_0^{a+\sqrt{a}\cdot x} y^{a-1}\mathrm{e}^{-y}\,\mathrm{d}y,\quad \text{where}\quad a = \frac{4}{\gamma_1^2}$$

(distribution of Z: *approximation by gamma distribution*)

(4) *Approximation by logarithmic-normal distribution* (see Sect. 22.1): again due to the asymmetry it takes into account the skewness (see Sect. 26.3) of Z (one uses the logarithmic-normal distribution shifted by a suitable constant):

$$P\left(\frac{Z-\mu}{\sigma} \le x\right) \approx \Phi\ \left(\frac{b}{2} + \frac{1}{b}\ln\left(x\sqrt{\exp(b^2)-1}+1\right)\right),$$

where

$$b = \sqrt{\ln q}; \quad q = \sqrt[3]{d+\sqrt{d^2-1}} + \sqrt[3]{d-\sqrt{d^2-1}} - 1; \quad d = 1 + \frac{\gamma_1^2}{2}$$

(distribution of Z: *approximation by logarithmic-normal distribution*)

(5) *Gram-Charlier approximation*: is based on the expansion of the probability density of Z by means of orthogonal polynomials:

$$P\left(\frac{Z-\mu}{\sigma} \le x\right) \approx \Phi\ (x) - \frac{\gamma_1}{6}\Phi^{(3)}(x) + \frac{\gamma_2-3}{24}\Phi^{(4)}(x)$$

(distribution of Z: *Gram-Charlier approximation*)

(6) *Edgeworth approximation*: is based on Taylor expansion of the moment generating function (see Sect. 26.10) of normed Z:

$$f_{\mathrm{norm}Z}(x) \approx \varphi\ (x) - \frac{\gamma_1}{6}\varphi^{(3)}(x) + \frac{\gamma_2-3}{24}\varphi^{(4)}(x) + \frac{\gamma_1^2}{72}\varphi^{(6)}(x)$$

(distribution of Z: *Edgeworth approximation*; the first three summands correspond to Gram-Charlier approximation)

(7) *Approximation* NP2: is the power approximation by normal distribution taking into account the skewness (see Sect. 26.3) of Z and using Gram-Charlier approximation (see *thereinbefore*); it is the best used approximation in practice:

$$P\left(\frac{Z-\mu}{\sigma} \le x + \frac{\gamma_1}{6}(x^2-1)\right) \approx \Phi\ (x)$$

(distribution of Z: *approximation* NP2)

(8) *Esscher approximation*: is based on the moment generating function (see Sect. 26.10) of Z under the assumption that one knows the form of this function

- *Risk process* $\{T_1, X_1, T_2, X_2, \ldots\}$: is a sequence of mutually independent, non-negative random variables (see stochastic process in Sect. 30.1); one usually interprets X_i as claims incurred at times $W_i = T_1 + T_2 + \cdots + T_i$ (i.e. T_i is the random *claim interoccurrence time* between X_i and X_{i+1}); there are two other processes closely related to the risk process:
  - *claim number process* $\{N_t, t \geq 0\}$: N_t is the number of claims up to time t (i.e. in the time interval $(0, t\rangle$):

$$N_t = \sum_{i=1}^{\infty} I_{[W_i \leq t]}$$

  - *accumulated claims process* $\{Z_t, t \geq 0\}$: Z_t is the aggregate claim (the total loss) up to time t (i.e. in the time interval $(0, t\rangle$):

$$Z_t = \sum_{i=1}^{N_t} X_i$$

- *Poisson risk process* $\{T_1, X_1, T_2, X_2, \ldots\}$ *with intensity* $\lambda > 0$: is the risk process (see *thereinbefore*), where $T_i \sim \mathrm{Exp}(\lambda)$ (i.e. the claim interoccurrence times are identically distributed with the exponential distribution (see Sect. 26.5)) and X_i are identically distributed with the distribution function F_X; then it holds for the claim number process and the accumulated claims process (see *thereinbefore*):

$$\{N_t, t \geq 0\} \sim \text{Poisson process with intensity } \lambda \text{ (see Sect. 30.4)}$$

$$\{Z_t \sim \mathrm{CP}(\lambda t, F_X), t \geq 0\}$$

(process with compound Poisson distribution, see *thereinbefore*)

$$F_{Z_t}(x) = \sum_{j=0}^{\infty} \mathrm{e}^{-\lambda t} \frac{(\lambda t)^j}{j!} F_X^{*j}(x)$$

$$\mathrm{E}(Z_t) = \lambda t \mathrm{E}(X); \quad \mathrm{var}(Z_t) = \lambda t \mathrm{E}(X^2); \quad \gamma_1(Z_t) = \frac{\mathrm{E}(X^3)}{\left(\lambda t \left(\mathrm{E}(X^2)\right)^3\right)^{1/2}}$$

22.3 Copula

- *Copula*: is an instrument which enables to model specific dependencies among random variables mainly in financial and insurance models, e.g. in the context of the capital adequacy of banks (see e.g. the methodology VaR in Sect. 12.3) or the solvency of insurance companies (see Sect. 24.3), where simulations of portfolios with prescribed dependency structures are necessary; for simplicity, only two- dimensional copulas are presented here, which model dependencies between two random variables

- *Two-dimensional copula*: is a function C: $\langle 0, 1\rangle \times \langle 0, 1\rangle \to \langle 0, 1\rangle$ with following properties:

  - $C(u, 0) = C(0, v) = 0$ for all $u, v \in \langle 0, 1\rangle$
  - $C(u, 1) = u$ and $C(1, v) = v$ for all $u, v \in \langle 0, 1\rangle$
  - $C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0$ for all $u_1, u_2, v_1, v_2 \in \langle 0, 1\rangle$ such that $u_1 \leq u_2, v_1 \leq v_2$

  (in general, the n-dimensional copula $C(\mathbf{u})$ is a function $\langle 0, 1\rangle^n \to \langle 0, 1\rangle$, which (1) is equal to zero for all $\mathbf{u} \in \langle 0, 1\rangle^n$, that have at least one coordinate u_i equal to zero, (2) is equal to the coordinate u_i, if all other coordinates of $\mathbf{u}$ are unities, and (3) is n-increasing (in the same sense as the joint distribution function of random variables $X_1, \ldots, X_n$ (see Sect. 26.6))

$$H(x, y) = C\left(F(x), G(y)\right)$$

(*Sklar's Theorem*: for an arbitrary joint distribution function $H(x, y)$ with marginal distribution functions $F(x)$ and $G(y)$ (see Sect. 26.6), there exists a two-dimensional copula $C(u, v)$ that fulfils the given relation (moreover, this copula is unique on the domain {range of F}×{range of G}, if $F(x)$ and $G(y)$ are continuous); conversely, if $C(u, v)$ is a copula and $F(x)$ and $G(y)$ are distribution functions, then the function $H(x, y)$ defined by the given relation is a joint distribution function with marginal distribution functions $F(x)$ and $G(y)$)

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

where φ is a continuous, strictly decreasing function $\langle 0, 1\rangle \to \langle 0, \infty\rangle$ such that $\varphi(1) = 0$ (it may be $\varphi(0) = \infty$))
(*Archimedean copula*: the function $C(u, v)$ defined by the given relation is a copula, if and only if the function φ is convex (i.e. $\varphi[au + (1-a)v] \leq a\varphi(u) + (1-a)\varphi(v)$ for all $a, u, v \in \langle 0, 1\rangle$); the Archimedean copula is a special case of the general copula (see *thereinbefore*), and the function φ is called its *generator*; if using suitable generators in Sklar's theorem (see *thereinbefore*), one may achieve such dependencies among marginal distributions, which are necessary in financial and insurance models)

22.4 Credibility Premium

- *Credibility theory*: is a set of quantitative tools which allow an insurer to perform prospective *experience rating*, i.e. to construct future premiums given past experience; in this context the insurer is forced to answer the question, how *credible* is the policyholder's own experience (i.e. whether the policyholder is really a better or worse risk than the average of the given rating class or tariff group)

- *Credibility premium*: may be constructed in two ways:

  (1) *American credibility*: constructs the credibility premium directly as a combination of the premium based on *own* past data of the insurer and the premium based on *global* past data (e.g. for the whole insurance market); in practice, such a combination is usually constructed ad hoc without a theoretical background:

$$P_{\text{cred}} = Z \cdot P_{\text{own}} + (1 - Z) \cdot P_{\text{glob}}$$

(credibility premium according to *American credibility*: P_{own} is the premium based on own past data of the insurer; P_{glob} is the premium based on global past data (for the whole insurance market); Z is the *credibility coefficient* (i.e. the weight for the own data); the credibility coefficient should have the following properties: (1) $0 \leq Z \leq 1$; (2) Z increases with increasing volume of the own data of the insurer (other things being equal); (3) Z decreases with increasing significance of external data (other things being equal); (4) Z decreases with increasing volatility of the claims paid by the insurer)

$$P\left(|\theta - \hat{\theta}| \;\leq\; k\theta \right) \;\geq\; 1 - \alpha$$

(*full credibility* of estimated parameter θ: means that the error of the estimator $\hat{\theta}$ does not exceed $k\%$ of θ (k is fixed in advance, e.g. 5% of θ) with a sufficient confidence at least $1 - \alpha$ (e.g. 95%); typically, the estimated parameter is an important insurance indicator in this context (e.g. claim ratio, loss frequency, average claim degree, see Sect. 21.1))

Denotation:

n_f number of claims necessary to achieve the full credibility (i.e. such a volume of data that guarantees the full credibility according to the given definition when estimating the corresponding parameter θ, see *thereinbefore*)

n number of claims in the corresponding portfolio of the insurer (i.e. the volume of the own data)

$$Z = \min\left(\sqrt{\frac{n}{n_f}}, 1\right)$$

(*credibility coefficient*: approach recommended in practice)

  (2) *Bayesian credibility*: constructs the credibility premium by means of Bayesian methods (see Bayes Theorem in Sect. 26.2); the Bayesian approach to credibility has been suggested by Bühlmann and Straub (see *thereinafter*)

- *Bühlmann-Straub credibility model* (B-S model): is based on the principle that the individual risk experience may be described by a (hypothetical) parameter values θ associated with individual risks as unobservable realizations of a random variable (or a random vector) Θ:

  Aim: construction of premiums at the time period $t + 1$ for J risk classes (tariff groups) using observations at past t periods

  Assumptions and denotation:

X_{js}	total claims in the risk class j at the period s ($j = 1, \ldots, J$; $s = 1, \ldots, t$)
Θ_j	random variable for the risk class j ($j = 1, \ldots, J$) with risk parameter values θ_j as its realizations
$\mu(\Theta_j) = \mathrm{E}(X_{js}\|\Theta_j)$	premium for the risk class j ($j = 1, \ldots, J$) using the principle of mean value for net premiums (see Sect. 21.2), i.e. the premium is constructed as the corresponding conditional mean value given the risk parameter (see Sect. 26.8); it does not depend on the time index s
$\sigma^2\,(\Theta_j) = \mathrm{var}(X_{js}\|\Theta_j)$	variance for the risk class j ($j = 1, \ldots, J$) given the risk parameter; it does not depend on the time index s
$\mathrm{cov}(X_{js}, X_{jr}\,\|\Theta_j) = 0$	covariance for the risk class j ($j = 1, \ldots, J$) given the risk parameter; it is zero between different time periods, i.e. one assumes conditional uncorrelated behaviour for $s \neq r$
iid Θ_j	independent and identically distributed random variables Θ_j ($j = 1, \ldots, J$), i.e. one assumes that loss (probability) characteristics in particular risk classes are equal with mutually independent behaviour
$\{\Theta_j, X_{j1}, \ldots, X_{jt}\}$	independent random vectors ($j = 1, \ldots, J$), i.e. one assumes that particular risk classes are mutually independent

$$P_j = Z_j\bar{X}_{j\cdot} + (1 - Z_j)\bar{X}_{\cdot\cdot}, \quad \text{where}$$

$$\bar{X}_{j\cdot} = \frac{1}{t}\sum_{s=1}^{t} X_{js}; \quad \bar{X}_{\cdot\cdot} = \frac{1}{tJ}\sum_{s=1}^{t}\sum_{j=1}^{J} X_{js}; \quad Z_j = \frac{ta}{ta + s^2}$$

$$s^2 = \text{est E}(\sigma^2(\Theta_j)) = \frac{1}{J}\sum_{j=1}^{J}\frac{1}{t-1}\sum_{s=1}^{t}\left(X_{js} - \bar{X}_{j\cdot}\right)^2$$

$$a = \text{est var}(\mu(\Theta_j)) = \frac{1}{(J-1)}\sum_{j=1}^{J}\left(\bar{X}_{j\cdot} - \bar{X}_{\cdot\cdot}\right)^2 - \frac{1}{t}s^2$$

(*credibility premium* in B-S model for the risk class j ($j = 1, \ldots, J$): the premium is derived as the best linear unbiased estimate (see Sect. 27.11))

- *General Bühlmann-Straub credibility model*:

$$P_j = Z_j\bar{X}_{jE} + (1 - Z_j)\bar{X}_{ZE}, \quad \text{where}$$

$$E_j. = \sum_{s=1}^{t} E_{js}; \quad E.. = \sum_{s=1}^{t}\sum_{j=1}^{J} E_{js}; \quad Z_j = \frac{E_j.a}{E_j.a + s^2}; \quad Z. = \sum_{j=1}^{J} Z_j$$

$$\bar{X}_{jE} = \sum_{s=1}^{t} \frac{E_{js}}{E_j.}X_{js}; \quad \bar{X}_{EE} = \sum_{j=1}^{J} \frac{E_j.}{E..}\bar{X}_{jE}$$

$$\bar{X}_{ZE} = \sum_{j=1}^{J} \frac{Z_j}{Z.}\bar{X}_{jE} = \sum_{j=1}^{J} \frac{Z_j}{Z.} \sum_{s=1}^{t} \frac{E_{js}}{E_j.}X_{js}$$

$$s^2 = \text{est E}(\sigma^2(\Theta_j)) = \frac{1}{J}\sum_{j=1}^{J} \frac{1}{t-1} \sum_{s=1}^{t} E_{js}(X_{js} - \bar{X}_{jE})^2$$

$$a = \text{est var}(\mu(\Theta_j)) = \left(\sum_{j=1}^{J} E_j.(\bar{X}_{jE} - \bar{X}_{EE})^2 - (J-1)\, s^2\right) \Big/ \left(E.. - \sum_{j=1}^{J} \frac{E_{j.}^2}{E..}\right)$$

(credibility premium in *general* B-S model for the risk class j ($j = 1, \ldots, J$): it takes into account also the risk volume E_{js} in the risk class j at the period s ($j = 1, \ldots, J$; $s = 1, \ldots, t$), i.e. the claim amounts X_{js} are replaced by loss burdens X_{js}/E_{js} per a risk unit in the risk class j at the period s)

- *More general credibility models*: are e.g.
  - *regression credibility models*: the random variable $\mu(\Theta_j) = \text{E}(X_{js}|\Theta_j)$ for the risk class j ($j = 1, \ldots, J$) is the response variable (regressand) in a linear regression model (see Sect. 27.11) with known explanatory variables (regressors) and unknown parameters, which are functions of Θ_j
  - *evolutionary credibility models*: enable to describe changes of loss characteristics in time, e.g. they model an evolution of $\text{E}(X_{js}|\Theta_j)$ at the periods $s = 1, \ldots, t$ for the risk class j ($j = 1, \ldots, J$) as an autoregressive process (see Sect. 30.4 and Sect. 31.4)

22.5 Ruin Probability

- *Ruin theory*: investigates behaviour of the given insurance process (*risk process*, see Sect. 22.2), where the analyzed reserve (*surplus*) of the insurer increases (starting with an initial value) due to premiums collected and decreases due to insurance claims paid

- *Probability of ruin* $\Psi(U)$: is (in this context) the probability that the surplus will sometimes become negative; from the practical point of view, it is closely related to the *solvency of insurer* (see Sect. 24.3)

  Denotation and assumptions:

  U initial surplus (initial reserve) at the time 0 ($U \geq 0$)
  P premium amount (constant in time) for the unit time period
  Z_t total loss (accumulated claims process) up to time t in the Poisson risk process $\{T_1, X_1, T_2, X_2, \ldots\}$ with intensity λ (see Sect. 22.2)

  $$R_t = U + Pt - Z_t$$

  (surplus accumulated up to time t)

  $$\Psi(U) = P\left(\min_{t \geq 0} R_t < 0\right)$$

  (*probability of ruin*)

  $$\Psi(U) \leq e^{-r \cdot U},$$

  where r is so called *Lundberg's coefficient* (*adjustment coefficient*)
  (*Lundberg's inequality* for probability of ruin: it holds for the Poisson risk process (see Sect. 22.2) under the assumption $P > E(X)$)

  $$\int_0^\infty [1 - F_X(x)]\, e^{rx} dx = \frac{P}{\lambda}$$

  (equation for Lundberg's coefficient: r is its *unique positive solution*)

  $$\Psi(U) \approx \frac{P/\lambda - E(X)}{r \int_0^\infty x\, [1 - F_X(x)]\, e^{rx} dx} e^{-rU}$$

  (*Lundberg's approximation* for probability of ruin)

  $$r \approx \frac{2E(P - X)}{\mathrm{var}(X)} = \frac{2\lambda_1 E(X)}{\mathrm{var}(X)}$$

  (approximation for Lundberg's coefficient : the second expression holds, if the principle of mean value is used, i.e. $P = (1 + \lambda_1) \cdot E(X)$, where λ_1 is the security loading (see Sect. 21.2))

$$U \approx -\frac{\ln \varepsilon}{r} \approx -\ln \varepsilon \frac{\mathrm{var}(X)}{2\mathrm{E}(P-X)} = -\ln \varepsilon \frac{\mathrm{var}(X)}{2\lambda_1 \mathrm{E}(X)}$$

(approximation for initial surplus: it is applicable under the condition that the probability of ruin does not exceed a given bound $\varepsilon > 0$)

$$\frac{U}{\mathrm{E}(X)} \approx -\ln \varepsilon \frac{\mathrm{var}(X)}{2\lambda_1 \, (\mathrm{E}(X))^2} = -\frac{\ln \varepsilon}{2\lambda_1} \, (V(X))^2$$

(approximation for the ratio of the initial surplus U to the net premium E(X) under the condition that the probability of ruin does not exceed a given bound $\varepsilon > 0$: the approximation is formulated by means of the coefficient of variation $V(X)$ (see Sect. 26.3) that measures the relative volatility of insurance claims)

22.6 Deductible

- *Deductibles*: mean (see also Sect. 21.3) that the clients participates in the claim settlements, which implies lower premiums than in the case without deductibles; the formulas used in the case of deductibles (see *thereinafter*) correspond to analogous formulas used in the framework of reinsurance (see Sect. 24.2)

Denotation:

X	claim amount (in a given insurance policy) with distribution function $F(x)$
$Z = X_1 + \cdots + X_N$	aggregate claim amount (total loss) in the collective risk model (see Sect. 22.2)
$\tilde{X},\ \tilde{X}_i, \tilde{Z}, \tilde{N}$	insured's deductible
$\hat{X},\ \hat{X}_i, \hat{Z}, \hat{N}$	insurer's participation
μ, σ, γ_1	mean value, standard deviation and skewness of the claim amount ($\mu = \mathrm{E}(X)$, $\sigma = \sigma(X)$, $\gamma_1 = \gamma_1(X)$, see Sect. 26.3)
Φ, φ	distribution function and probability density of $N(0, 1)$ (see Sect. 26.5)

- *Quota deductible*: the insured's deductible is a stipulated proportion $(1 - q)$ of the insurance claim (*quota q* gives the proportion of the insurer's participation):

$$\tilde{N} = \hat{N} = N$$

$$\tilde{X} = (1-q)X; \quad \hat{X}_i = qX$$

$$\mathrm{E}(\hat{X}) = q\mathrm{E}(X); \quad \mathrm{var}(\hat{X}) = q^2 \mathrm{var}(X); \; V(\hat{X}) = \frac{\sigma(\hat{X})}{\mathrm{E}(\hat{X})} = \frac{\sigma(X)}{\mathrm{E}(X)} = V(X)$$

$$\tilde{S} = (1-q)S; \quad \hat{S} = qS$$

- *Excess deductible*: the insured's deductible is the whole claim amount, which does not exceed a stipulated amount a (the so-called *priority*), or it is the amount a, if the claim amount is higher than a:

$$\tilde{N} = N; \quad \mathrm{E}(\hat{N}) = p_a \mathrm{E}(N)$$

$$\mathrm{var}(\hat{N}) = p_a(1-p_a)\mathrm{E}(N) + p_a^2 \mathrm{var}(N), \quad \text{where} \quad p_a = \mathrm{P}(X > a) = 1 - F(a)$$

$$N \sim P(\lambda) \Rightarrow \hat{N} \sim P(p_a \lambda) \text{ (see Sect. 26.4)}$$

$$N \sim B(n,p) \Rightarrow \hat{N} \sim B(n, p_a p) \text{ (see Sect. 26.4)}$$

$$N \sim NB(r,p) \Rightarrow \hat{N} \sim NB\left(r, \frac{p}{p + p_a(1-p)}\right) \text{ (see Sect. 26.4)}$$

$$\tilde{X} = \min(X,\ a); \quad \hat{X}_i = (X-a)^+$$

$$\mathrm{E}(\tilde{X}) = \int_0^a (1-F(x))\ \mathrm{d}x; \quad \mathrm{var}(\tilde{X}) = 2\int_0^a x\,(1-F(x))\ \mathrm{d}x - [\mathrm{E}(\tilde{X})]^2$$

$$\mathrm{E}(\hat{X}) = \int_a^\infty (1-F(x))\ \mathrm{d}x;$$

$$\mathrm{var}(\hat{X}) = 2\left\{\int_a^\infty x\,(1-F(x))\ \mathrm{d}x - a\int_a^\infty (1-F(x))\ \mathrm{d}x\right\} - [\mathrm{E}(\hat{X})]^2$$

$$\mathrm{E}(\hat{X} \mid X > a) = \frac{1}{1-F(a)}\int_a^\infty (1-F(x))\ \mathrm{d}x$$

$$\mathrm{E}(\hat{X}) < \frac{\sigma^2}{a-\mu}; \quad \mathrm{E}(\hat{X}) < \frac{\sigma}{2}\left(\sqrt{1+z^2} - z\right), \quad \text{where } z = \frac{a-\mu}{\sigma}$$

$$\mathrm{E}(\hat{X}) \approx \sigma \begin{cases} \varphi(v)\left(1 + \frac{v \cdot \gamma_1}{6}\right) - z\,(1-\Phi(v)), & \gamma_1 > 0 \\ \varphi(v)\left(1 - \frac{v \cdot \gamma_1}{6}\right) - z\Phi(v), & \gamma_1 < 0 \end{cases},$$

$$\text{where } z = \frac{a-\mu}{\sigma}; \quad v = -\frac{3}{|\gamma_1|} + \frac{1}{|\gamma_1|}\sqrt{\gamma_1^2 + 6\gamma_1 z + 9}$$

(approximation based on NP2, see Sect. 22.2)

$$\mathrm{E}(\hat{X}^2) \approx \sigma^2 \begin{cases} (1-\Phi(v))\left(z^2+2\left(\frac{\gamma_1}{6}\right)^2+1\right) & \\ \quad +\varphi(v)\left(\left(\frac{\gamma_1}{2}-z\right)\left(\frac{v\gamma_1}{6}+1\right)-v\left(\frac{\gamma_1}{6}\right)^2\right), & \gamma_1>0 \\ 1+z^2-\left\{(1-\Phi(v))\left(z^2+2\left(\frac{\gamma_1}{6}\right)^2+1\right)\right. & \\ \quad \left.+\varphi(v)\left(\left(z-\frac{\gamma_1}{2}\right)\left(1-\frac{v\gamma_1}{6}\right)-v\left(\frac{\gamma_1}{6}\right)^2\right)\right\}, & \gamma_1<0 \end{cases}$$

(approximation based on NP2, see *thereinbefore* and Sect. 22.2)

$$\mathrm{E}(\tilde{S}) = \mathrm{E}(N) \times \mathrm{E}(\tilde{X})$$
$$\mathrm{var}(\tilde{S}) = \mathrm{E}(N)\mathrm{var}(\tilde{X}) + \mathrm{var}(N)\left(\mathrm{E}(\tilde{X})\right)^2$$

$$\mathrm{E}(\hat{S}) = \mathrm{E}(N) \times \mathrm{E}(\hat{X})$$
$$\mathrm{var}(\hat{S}) = \mathrm{E}(N)\mathrm{var}(\hat{X}) + \mathrm{var}(N)\left(\mathrm{E}(\hat{X})\right)^2$$

- *Integral deductible*: the insured's deductible is the whole claim amount, which does not exceed a stipulated amount a (*priority*), or it is zero, if the claim amount is higher than a:

$$\mathrm{E}(\tilde{N}) = (1-p_a)\mathrm{E}(N)$$
$$\mathrm{var}(\tilde{N}) = p_a(1-p_a)\mathrm{E}(N) + (1-p_a)^2\,\mathrm{var}(N),$$

where $p_a = \mathrm{P}(X > a) = 1 - F(a)$

$$\mathrm{E}(\hat{N}) = p_a\mathrm{E}(N)$$
$$\mathrm{var}(\hat{N}) = p_a(1-p_a)\mathrm{E}(N) + p_a^2\mathrm{var}(N)$$

$$\mathrm{E}(\tilde{X}) = \int_0^a (1-F(x))\;\mathrm{d}x - a\,(1-F(a))$$
$$\mathrm{var}(\tilde{X}) = 2\int_0^a x\,(1-F(x))\;\mathrm{d}x - a^2\,(1-F(a)) - \left(\mathrm{E}(\tilde{X})\right)^2$$

$$\mathrm{E}(\hat{X}) = \int_a^\infty (1-F(x))\;\mathrm{d}x + a\,(1-F(a))$$
$$\mathrm{var}(\hat{X}) = 2\int_a^\infty x\,(1-F(x))\;\mathrm{d}x + a^2\,(1-F(a)) - \left(\mathrm{E}(\hat{X})\right)^2$$

22.7 Calculations for Bonus-Malus Systems

- *Bonus-malus system* (see Sect. 21.5): is given by (1) number of *bonus levels* corresponding to particular discounts of the so-called *basic premium* (i.e. the premium without bonus); (2) *bonus level at entry*; (3) *decisive period* (i.e. the period from the previous claim reported in the given insurance policy without another claim reported; (4) *system rules* (e.g. which is the reduction of the attained bonus level after a claim reported, which is the new decisive period in such a case, and the like); the calculations for the bonus-malus systems are usually based on the theory of Markov chains (see Sect. 30.2)

Denotation and assumptions:

J	number of bonus (and malus) levels
$p_{ij}(\theta)$	transition probability (see Sect. 30.2) of going from the level i to the level j ($i, j = 1, \ldots, n$) for the insured with risk characteristics given by the parameter value θ (similarly as for the Bayesian credibility, see Sect. 22.4)
$\mathbf{P}(\theta) = \left(p_{ij}(\theta)\right)_{i,j=1}^{J}$	transition matrix (see Sect. 30.2)
$p_i^n(\theta)$	probability that the insured with the risk parameter value θ attains the level i in the year n
$\mathbf{p}^n(\theta) = \left(p_1^n(\theta), \ldots, p_J^n(\theta)\right)$	row vector of probabilities $p_i^n(\theta)$
$\mathbf{p}^1 = (0, \ldots, 1, \ldots, 0)$	row vector coding by means of the unity and zeroes the bonus level at entry
π_0	basic premium (see *thereinbefore*)
$\mathbf{c} = (c_1, \ldots, c_J)'$	column vector containing per cents of the basic premium π_0 (see *thereinbefore*) for particular bonus levels
$w_n(\theta)$	probability that a randomly chosen insured with a risk parameter value θ persists in the given bonus-malus system just n years: the distribution of the policyholders according to the number of insured years may be e.g. geometric with parameter $p \in (0, 1)$ (see Sect. 26.4)
$e(\theta) = \sum_{n=1}^{\infty} n \cdot w_n(\theta)$	mean number of insured years for the insured with the risk parameter value θ
$u(\theta)$	probability density of risk parameter θ; θ is a random variable (see Sect. 22.4), which may have in practice e.g. gamma distribution with parameters $a > 0$, $\lambda > 0$ (see Sect. 26.5)
X	claim amount distribution in the given insurance portfolio under the *iid* assumption for the claims

$$\mathbf{p}^{n+1}(\theta) = \mathbf{p}^{n}(\theta) \cdot \mathbf{P}(\theta), \quad n = 1,\ 2,\ \ldots$$

(recursive calculation, see Sect. 30.2)

$$\mathbf{p}^{n}(\theta) = \mathbf{p}^{1}(\theta) \cdot \mathbf{P}(\theta)^{n-1}, \quad n = 1,\ 2,\ \ldots$$

$$\pi(\theta) = \pi_0 \sum_{n=1}^{\infty} w_n(\theta) \sum_{j=1}^{J} p_j^n(\theta) c_j$$

(average premium over policyholders with risk parameter value θ)

$$\pi(\theta) = \pi_0 (1-p)\, x(\theta)^{\text{entry}},$$

where $x(\theta)^{\text{entry}}$ is the component corresponding to the bonus level at entry in the vector $\mathbf{x}(\theta)$, which is the solution of the equation $(\mathbf{I} - p\,\mathbf{P}(\theta))\,\mathbf{x}(\theta) = \mathbf{c}$
(average premium over policyholders with risk parameter value θ, if $w_n(\theta)$ is the geometric distribution with parameter $p \in (0,\,1)$, see *thereinbefore*)

$$\int_0^{\infty} u(\theta)\pi(\theta)\,\mathrm{d}\theta = \int_0^{\infty} \theta u(\theta)\,\mathrm{d}\theta \cdot \mathrm{E}(X)$$

(equation for calculation of basic premium π_0)

- Analysis of bonus-malus systems usually looks for a possible stationary distribution (see Sect. 30.2) of the given Markov chain, since such a distribution corresponds to the stabilized behaviour of the insurance portfolio (e.g. which part of the insured drivers attains in such a stable mode the highest bonus level, and the like)

Further Reading

Booth, P., Chadburn, R., Cooper, D., Haberman, S., James, D.: Modern Actuarial Theory and Practice. Chapman and Hall/CRC, London (1999)
Daykin, C.D., Pentikäinen, T., Pesonen, M.: Practical Risk Theory for Actuaries. Chapman and Hall, London (1994)
Denuit, M., Dhaene, J., Goovaerts, M., Kaas, R.: Actuarial Theory for Dependent Risks: Measures, Orders and Models. Wiley, New York (2005)
Embrechts, P., Klüppelberg, C., Mikosch, T.: Modelling External Events for Insurance and Finance. Springer, Berlin (1997)
Heilmann, W.-R.: Fundamentals of Risk Theory. VVW, Karlsruhe (1988)
Kaas, R., Goovaerts, M., Dhaene, J., Denuit, M.: Modern Actuarial Risk Theory, Using R. Springer, Berlin (2008)

Klugman, S.A., Panjer, H.H., Willmot, G.E.: Loss Models. Wiley, New York (1998)
Mack, T.: Schadenversicherungsmathematik. VVW, Karlsruhe (2002)
Nelsen, R.B.: An Introduction to Copulas. Springer, New York (1999)
Straub, E.: Non-Life Insurance Mathematics. Springer, Berlin (1988)
Teugels, J., Sundt, B. (eds.): Encyclopedia of Actuarial Science. Wiley, New York (2004)
Wolfsdorf, K.: Versicherungsmathematik (Teil 2: Theoretische Grundlagen, Risikotheorie, Sachversicherung). Teubner, Stuttgart (1986)

Chapter 23
Health Insurance

Abstract Chapter 23 deals briefly with fundamentals of insurance for private medical treatment.

- *Insurance for private medical treatment*: makes use of calculations, which are based on the *equivalence principle*, i.e. the expected present value (calculated at the time of policy issue) of the premiums is equal to the expected present value of the benefits similarly as for the life insurance in Sect. 18.2 (but now we deal with a class of non-life insurance, see Chap. 16)

- Classification according to type of insurance claims: e.g.
  - sickness insurance (costs of therapy usually including prophylactic treatment)
  - hospitalization insurance (excluding costs of stationary therapy included in sickness insurance, i.e. mainly hotel services)
  - daily benefits insurance in the case of sick leave
  - insurance of over-standard treatment (the application of more expensive medicaments and medical materials, and the like)

- Classification of *health insurer's costs of claims* (this classification is more detailed and dynamic than the one above): e.g.
  - costs of outpatient treatment
  - costs of doctor's visits at home
  - costs of medicaments and medical materials
  - costs of surgical operations
  - costs of hospitalization
  - costs due to pregnancy and birth
  - costs of dental treatment and prosthetics

- *Method of average costs*: is the best used method for health insurance calculations (e.g. in Germany):

  Denotation:

l_x number of policyholders aged x
q_x probability of death at age x (see Sect. 17.2)

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_23,

w_x probability of cancellation (lapse) of health insurance at age x (these probabilities may be estimated e.g. by means of mathematical curves)

$K_x{}^j$ *average costs per capita*: are the average annual costs per capita (classified according to sex) in jth class of health insurance at age x (e.g. $K_x{}^{\text{outpatient}}$ are the average annual costs of outpatient treatment of a male aged x)

$$G^j = K^j_{x_0}$$

(*basic costs per capita* (in practice one often puts $x_0 = 28$ or $x_0 = 43$))

$$l_{x+1} = l_x \cdot (1 - q_x - w_x)$$

(*decrement order of health insurance*)

$$k^j_x = \frac{K^j_x}{K^j_{x_0}} = \frac{K^j_x}{G^j}$$

(*profile* (normed costs per capita): are the average costs per capita $K_x{}^j$ relatively to the costs at a chosen age x_0)

$$K^j_x = G^j \cdot k^j_x$$

(*profile application* in practice of health insurance: unlike the average costs per capita, the profile is mostly *stable* in time so that it is not necessary to change it within a lot of years; in a given year, it is sufficient to multiply the profile by the value G^j, which is specific just for this year)

- Commutations functions in health insurance (see Sect. 17.6):

$$D_x = l_x \cdot v^x; \quad N_x = D_x + D_{x+1} + \ldots + D_\omega$$

$$O^j_x = k^j_x \cdot D_x$$

$$U^j_x = O^j_x + O^j_{x+1} + \ldots$$

$$\ddot{a}_x = \frac{N_x}{D_x}$$

$$A^j_x = \frac{K^j_x \cdot l_x + K^j_{x+1} \cdot l_{x+1} \cdot v + \ldots}{l_x} = G^j \cdot \frac{k^j_x \cdot l_x \cdot v^x + k^j_{x+1} \cdot l_{x+1} \cdot v^{x+1} + \ldots}{l_x \cdot v^x}$$
$$= G^j \cdot \frac{O^j_x + O^j_{x+1} + \ldots}{D_x} = G^j \cdot \frac{U^j_x}{D_x}$$

(net single premium in jth class of health insurance for policyholders aged x at entry)

$$P_x^j = \frac{A_x^j}{\ddot{a}_x}$$

(net annual premium in jth class of health insurance for policyholders aged x at entry)

Further Reading

Bohn, K.: Die Mathematik der deutschen privaten Krankversicherung. Verlag Versicherungswirtschaft, Karlsruhe (1980)
Teugels, J., Sundt, B. (eds.): Encyclopedia of Actuarial Science. Wiley, New York (2004)
Wolfsdorf, K.: Versicherungsmathematik (Teil 1:Personenversicherung). Teubner, Stutgart (1997)

Chapter 24
Reinsurance

Abstract Chapter 24 deals with actuarial problems of modern reinsurance: 24.1. Basic Concepts of Reinsurance, 24.2. Types of Reinsurance, 24.3. Solvency, 24.4. Alternative Risk Transfer ART.

24.1 Basic Concepts of Reinsurance

- *Reinsurance* ("insurance of the risk assumed by the insurer"): is the principal mechanism, which insurance companies use to transfer part the risk assumed through their own underwriting activities to reinsurance companies; the *insurer* (called *cedant* or *direct insurer* or *reinsured* in this context) is said to *cede* risk to the *reinsurer* (*cessionaire*); there is usually no contract between any insured from the insurance portfolio of the insurer on one side and the reinsurer on other side
- *Retention* (*deductible*): is the part of the risk that is not ceded and is kept for own account of the insurer
- *Priority*: is the maximum part of the loss incurred, which is covered by the insurer
- *Reinsurance premium*: is the premium paid by the insurer to the reinsurer for the risk accepted
- *Limit* of reinsurer: is the maximum part of the insured risk that can be ceded to the reinsurer; there may be more reinsurers for a given insurer that cover various layers of the insured risk
- *Capacity* of insurer: is the maximum amount of a risk that can be accepted in insurance; one of the aims of reinsurance (see *thereinafter*) is just to enlarge the underwriting capacity of the insurer given by the retentions of this insurer and by the limits of the participating reinsurers
- *Profit commission*: is a contractually agreed commission paid by the reinsurer to the insurer in proportion to reinsurer's profits in the given reinsurance business; it motivates the insurer to be concerned in profitability of the reinsurer
- *Commission*: is a remuneration paid by the reinsurer to the insurer for insurer's costs related to the reinsurance business
- *Reinsurance basis*: concerns mainly life reinsurance, which can be contracted:

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_24,

  - *on risk basis*: is a system of life reinsurance, under which the reinsurer participates in death risk only, and only insofar as it exceeds the premium reserves (the administration of insurance and premium reserves stay with the cedant)
  - *on normal basis*: all essentials of the insurance contract are ceded in a given ratio
  - *on modified basis*: the only distinction from the normal basis consists in the reinsurer's duty to deposit the premium reserves with the insurer
- *Retrocession*: passes on a part of the ceded risk further to another reinsurer (*retrocessionaire*)
- *Coinsurance*: refers to the joint assumption of risk among various insurers, i.e. a number of insurers share a risk
- *Insurance pool*: is a risk-sharing community in the legal form of a non-trading partnership; the members of the pool agree to submit all the risks falling under the terms of the pooling agreement into the pool; in return, they participate according to a pre-defined distribution in the entire volume of business brought into the pool
- *Coinsurance pool*: means that the pool itself is organized as a risk carrier and the pool members act only as intermediaries: they write business for the pool's account and authorize and oblige the pool directly; the coinsurance pool itself operates as an insurer in the form of a legal entity (e.g. nuclear pool)
- *Reinsurance pool*: is an insurance pool, in which the risks are initially written by individual pool members; the other pool members then participate in all risks by way of reinsurance in accordance with a pre-distributed allocation formula; unlike the coinsurance pool, the reinsurance pool is not a visible legal entity
- *Reciprocity*: is the mutual exchange of reinsurance

- Types of reinsurance:
  - *facultative*: is an individual reinsurance negotiated and placed individually with a separate reinsurance contract and policy terms; it involves a case-by-case review and acceptance of risks by the reinsurer; this arrangement is often used for large or unique risks
  - *obligatory*: is a *treaty*, where the insurer is obliged to cede and the reinsurer is obliged to accept (when treaty stipulations regarding the business to be reinsured are met); the treaty, where reinsurances are ceded optionally by the insurer, but simultaneously the reinsurer is obliged to accept all such cessions, is called *facultative-obligatory*
  - *proportional*: is a form of reinsurance in which exposures (risks, sums insured), claims (losses) and premiums of the insurer are shared proportionally by the insurer and the reinsurer on the basis of some predefined formula, such as fixed or variable percentages of these values
  - *non-proportional*: allocates exposure through non-proportional layers, based on actual claims received (therefore it is sometimes called *claim reinsurance*); the reinsurer usually agrees to pay any losses, which exceed a specified priority up to a maximum limit against payment of a specially calculated reinsurance premium

  - *finite* or *financial*: is a combination of risk transfer and risk financing, where the investment aspects, i.e. how much return can be gained on the sum invested, play a special role (also see alternative risk transfer in Sect. 24.4)

- Reinsurance achieves several important goals:
  - it increases *underwriting capacity* (see *thereinbefore*)
  - it makes an insurance portfolio more *homogeneous*, e.g. by protecting the insurer against low-frequency/high-severity events
  - it creates *profit stability* violated due to various risks, e.g. due to
    - *risk of fluctuations in claim ratio* (it is the ratio of claim amounts to corresponding premiums, see Sect. 21.1)
    - *risk of economic, social and technological changes* (their impact can be hardly involved into calculations of premiums)
    - *risk of errors* such as an inappropriate interpretation of underlying statistics
  - it diversifies insurance risks using
    - *territorial* diversification due to distinct development of insurance results in various countries
    - *product* diversification due to distinct development of insurance results for various products in the portfolio
    - *time* diversification due to joint account of the insurer and the reinsurer over a longer time horizon so that the insurer's profits and losses may balance mutually
  - it enables some *financial benefits* mainly due to finite reinsurance (see Sect. 24.4)
  - it guarantees *professional services from reinsurer* when introducing new products, evaluating risks in a skilled way, educating the staff, and the like

24.2 Types of Reinsurance

Denotation:

N, S, X, Z, P	number of claims, sum insured in a policy, claim amount (loss) in a policy, total claim amount in the reinsured portfolio, premium in a policy
N_I, S_I, X_I, Z_I, P_I	as *thereinbefore*, but for own account of the insurer
N_R, S_R, X_R, Z_R, P_R	as *thereinbefore*, but for own account the reinsurer
${}_XF(x) = \mathrm{P}(X \leq x)$ and ${}_Xf(x)$	distribution function and probability density of the claim amount X (see Sect. 26.3)
${}_XF_I(x) = \mathrm{P}(X_I \leq x)$, ${}_Xf_I(x)$	as *thereinbefore*, but for own account of the insurer
${}_XF_R(x) = \mathrm{P}(X_R \leq x)$, ${}_Xf_R(x)$	as *thereinbefore*, but for own account of the reinsurer

$$X = X_I + X_R, \qquad Z = Z_I + Z_R, \qquad P = P_I + P_R$$

- *Proportional reinsurance* (see Sect. 24.1): is usually applied in life insurance; the proportional reinsurance can be as follows:

(1) *Quota share*: the insurer and reinsurer agree to split exposures (risks, sums insured), claim amounts (losses) and premiums as *fixed* percentages of these values; the fixed percentage of the reinsurer share is called *quota*:

$$S_R = q \cdot S; \qquad X_R = q \cdot X; \qquad P_R = q \cdot P, \text{ where } 0 < q < 1$$

(*quota share*: q is the quota of the reinsurer; the partition of the risk between the insurer and the reinsurer is the same for each policy; the advantage consists in the administrative simplicity (everything is divided in the same ratio); the disadvantages are an insufficient homogeneity of the reinsured portfolio and an unnecessary cession for policies with small sums insured which should be no problem for the insurer, even if no reinsurance is applied)

$$P_R(t) = q \cdot (S - {}_tV_{x:\overline{n}|}) \cdot q_{x+t-1} \approx q \cdot P^{\text{riz}}_{x:\overline{n}|}(t)$$

(*reinsurance premium* at time t in the framework of the *quota share reinsurance on the risk basis* (see Sect. 24.1) with the quota of reinsurer q, if one reinsures a life contract with the net premium reserve ${}_tV_{x:\overline{n}|}$ (see Sect. 18.3) and with the sum insured S; q_{x+t-1} is the probability of death at age $x + t - 1$ and its value is chosen by the reinsurer (it can differ from the one used by the insurer, see Sect. 17.2); $P^{\text{riz}}_{x:\overline{n}|}(t)$ is the risk premium at time t of the direct insurance (see Sect. 18.3))

$$N = N_I = N_R$$

$$S = S_I + S_R$$

$$\mathrm{E}(X_I) = (1-q) \cdot \mathrm{E}(X), \qquad \mathrm{E}(X_R) = q \cdot \mathrm{E}(X)$$

$$\mathrm{var}(X_I) = (1-q)^2 \cdot \mathrm{var}(X), \qquad \mathrm{var}(X_R) = q^2 \cdot \mathrm{var}(X)$$

$$\frac{\sigma(X_I)}{\mathrm{E}(X_I)} = \frac{\sigma(X_R)}{\mathrm{E}(X_R)} = \frac{\sigma(X)}{\mathrm{E}(X)}$$

$${}_XF_I(x) = {}_XF(x/(1-q)), \qquad {}_XF_R(x) = {}_XF(x/q)$$

$${}_Xf_I(x) = \frac{{}_Xf(x/(1-q))}{1-q}, \qquad {}_Xf_R(x) = \frac{{}_Xf(x/q)}{q}$$

(2) *Surplus*: the insurer and reinsurer agree to split exposures (risks, sums insured), claim amounts (losses) and premiums as *variable* percentages

of these values; the percentage of the insurer's share in a given policy corresponds to the ratio of insurer's retention s to the sum insured S in this policy (accordingly the percentage for the reinsurer $(S - s)/S = 1 - s/S$ corresponds to reinsurer's surplus $S - s$):

$$S_R = \begin{cases} 0 & \text{for } S \le s, \\ S - s & \text{for } S > s, \end{cases} \quad X_R = \begin{cases} 0 & \text{for } S \le s, \\ \left(1 - \frac{s}{S}\right) \cdot X & \text{for } S > s, \end{cases}$$

$$P_R = \begin{cases} 0 & \text{for } S \le s, \\ \left(1 - \frac{s}{S}\right) \cdot P & \text{for } S > s, \end{cases} \quad \text{where } s > 0$$

(*surplus*: s is the retention of the insurer; unlike the quota share reinsurance (see *thereinbefore*), the partition of the risk between the insurer and the reinsurer can differ among various policies; the advantage is an efficient homogeneity of the reinsured portfolio; the disadvantage is an administrative complexity (the partition of the risk between the insurer and the reinsurer is variable)); see also formulas given *thereinbefore* for the quota share with

$$q = \begin{cases} 0 & \text{for } S \le s \\ 1 - \frac{s}{S} & \text{for } S > s \end{cases}$$

$$S_R = \begin{cases} 0 & \text{for } S \le s, \\ S - s & \text{for } s < S \le s + L, \\ L & \text{for } S > s + L, \end{cases} \quad X_R = \begin{cases} 0 & \text{for } S \le s, \\ \left(1 - \frac{s}{S}\right) \cdot X & \text{for } s < S \le s + L, \\ \frac{L}{S} \cdot X & \text{for } S > s + L, \end{cases}$$

$$P_R = \begin{cases} 0 & \text{for } S \le s \\ \left(1 - \frac{s}{S}\right) \cdot P & \text{for } s < S \le s + L \\ \frac{L}{S} \cdot P & \text{for } S > s + L \end{cases}$$

(surplus with *limit of reinsurer*: the limit of reinsurer L is usually given as multiples of the retention (so called *lines*): e.g. three lines mean that the reinsurer covers the risk in the amount of the treble of the retention)

- *Non-proportional reinsurance* (see Sect. 24.1): is usually applied in nonlife insurance; it is characterized by a distribution of liability between the insurer and the reinsurer on the basis of losses rather than sums insured; as compensation for the cover provided, the reinsurer receives part of the original premium and not the part of the premium corresponding to the sum reinsured as in the proportional reinsurance; the treaty defines a *priority*, up to which the insurer pays all losses; for his part, the reinsurer obliges himself to pay all losses above the priority and not out of a contractually defined *layer* (there are usually several adjacent layers contracted with several reinsurers); in practical reinsurance one applies as the loss amount the so-called *ultimate net loss UNL* (the reinsured loss after deduction of all reimbursements from other reinsurance contracts) and as the premium amount the so-called *gross net premium* *income GNPI* (the premium in the reinsured business after deduction of all lapses and reinsurance premiums for other reinsurance contracts, which reduce the exposure of the given business); the non-proportional reinsurance may be as follows:

(1) *WXL/R reinsurance* (*working excess of loss cover per risk*): is created to reinsure individual risks on loss basis (its aim is to relieve the insurer of losses which surpass the amount retained for own account on any particular risk); if any contract from the reinsured portfolio is affected by a loss with claim exceeding the priority, then the excess incurred is reimbursed by the reinsurer (but only within the reinsurer's layer):

$$X_R = \begin{cases} 0 & \text{for } X \leq a, \\ X - a & \text{for } X > a, \end{cases} \quad \text{where } a > 0$$

(excess in *WXL/R reinsurance*: a is the priority of the insurer; in practice one uses denotations of the type "1,000,000 EUR *xs* 400,000 EUR" (1,000,000 EUR is the layer of the reinsurer and 400,000 EUR is the priority of the insurer))

$$N_I = N$$

$$\mathrm{P}(N_R = n) = \sum_{i=n}^{\infty} \mathrm{P}(N = i) \cdot \binom{i}{n} \cdot p_a^n \cdot (1 - p_a)^{\,i-n},$$

where $p_a = \mathrm{P}(X > a) = 1 - {}_XF(a)$

$$\mathrm{E}(N_R) = p_a \cdot \mathrm{E}(N); \quad \mathrm{var}(N_R) = p_a \cdot (1 - p_a) \cdot \mathrm{E}(N) + p_a^2 \cdot \mathrm{var}(N)$$

$$\mathrm{P}(N_R = n) = e^{-\lambda \cdot p_a} \cdot \frac{(\lambda \cdot p_a)^n}{n!}; \qquad \mathrm{E}(N_R) = \lambda \cdot p_a; \qquad \mathrm{var}(N_R) = \lambda \cdot p_a$$

for Poisson distribution $N \sim P(\lambda)$ (see Sect. 26.4)

$$\mathrm{P}(N_R = n) = \binom{r+n-1}{n} \left(\frac{p}{p + p_a \cdot (1-p)} \right)^r \cdot \left(1 - \frac{p}{p + p_a \cdot (1-p)} \right)^n$$

for negative binomial distribution $N \sim NB(r, p)$ (see Sect. 26.4)

$$ {}_XF_I(x) = \begin{cases} {}_XF(x) & \text{for } x < a; \\ 1 & \text{for } x \geq a; \end{cases} \qquad {}_XF_R(x) = {}_XF(a + x)$$

$$\mathrm{E}(X_I) = \int_0^a x \, \mathrm{d}\, {}_XF(x) + a \cdot \left(1 - {}_XF(a)\right) = \int_0^a \left(1 - {}_XF(x)\right) \mathrm{d}x$$

$$\mathrm{E}(X_R) = \int_a^\infty x\,\mathrm{d}\,{}_XF(x) \; - a \cdot \; (1 - {}_XF(a)) = \int_a^\infty (1 - {}_XF(x))\,\mathrm{d}x = \mathrm{E}(X) - \mathrm{E}(X_I)$$

$$\mathrm{var}(X_I) = 2 \int_0^a x \cdot (1 - {}_XF(x))\,\mathrm{d}x - [\mathrm{E}(X_I)]^2$$

$$\mathrm{var}(X_R) = 2 \left\{ \int_a^\infty x \cdot (1 - {}_XF(x))\,\mathrm{d}x - a \cdot \int_a^\infty (1 - {}_XF(x))\,\mathrm{d}x \right\} - [\mathrm{E}(X_R)]^2$$

$$\mathrm{E}(X_I{}^j) = \int_0^a x^j\,\mathrm{d}\,{}_XF(x) \; + a^j \cdot (1 - {}_XF(a))\,, \quad \text{where } j \in \mathrm{N}$$

$$\mathrm{E}(X_R{}^j) = \sum_{i=1}^{j} \binom{j}{i} \cdot (-a)^{\,j-i} \cdot [\mathrm{E}(X^i) - \mathrm{E}(X_I{}^i)], \quad \text{where } j \in \mathrm{N}$$

Assumptions:

$$Z_I = \sum_{i=1}^{N} X_{iI}; \qquad Z_R = \sum_{i=1}^{N} X_{iR},$$

where X_1, X_2, ... are *iid* random variables independent on the random variable N

$$\mathrm{E}(Z_I) = \mathrm{E}(N) \cdot \mathrm{E}(X_I) = \mathrm{E}(Z) \cdot {}_XF^{(1)}(a)$$

$$\mathrm{var}(Z_I) = \mathrm{E}(N) \cdot \mathrm{var}(X_I) + \mathrm{var}(N) \cdot [\mathrm{E}(X_I)]^2 = \mathrm{var}(Z) \cdot {}_XF^{(2)}(a)$$

where

$${}_XF^{(j)}(a) = \frac{\int_0^a x^j\,\mathrm{d}\,{}_XF(x) + a^j \cdot (1 - {}_XF(a))}{\mathrm{E}(X^j)}, \quad j \in \mathrm{N}$$

$$\mathrm{E}(Z_R) = \mathrm{E}(N) \cdot \mathrm{E}(X_R); \quad \mathrm{var}(Z_R) = \mathrm{E}(N) \cdot \mathrm{var}(X_R) + \mathrm{var}(N) \cdot [\mathrm{E}(X_R)]^2$$

(2) *WXL/E reinsurance* (*working excess of loss cover per event*): offers the insurer protection against losses caused by the same event of a non-catastrophic character; if several contracts from the reinsured portfolio is affected by a loss event with aggregate claim exceeding the priority contracted for all these policies altogether then the excess incurred is reimbursed by the reinsurer (but only within reinsurer's layer):

$$X_R = \begin{cases} 0 & \text{for } \sum_{i=1}^{n} X_i \le a\,, \\ \sum_{i=1}^{n} X_i - a & \text{for } \sum_{i=1}^{n} X_i > a\,, \end{cases} \quad \text{where } a > 0$$

(excess in *WXL/E reinsurance*: a is the priority of the insurer; $X_1, \ldots, X_n$ are claim amounts for a given loss event in n affected policies)

(3) *CatXL reinsurance* (*catastrophe excess of loss cover*): coincides with WXL/E reinsurance (see *thereinbefore*) except for the catastrophic character of the loss event, which causes in this case a substantial accumulation of losses; for natural catastrophes with exposure persisting a longer time (flood, hurricane. and the like), one contracts in addition the so-called *n-hours clause* (for a given loss event the reinsurer covers only such losses, which accumulate within n h); *Umbrella Cover reinsurance* (*All Risks Cover*) concerns an accumulation of losses from a single (catastrophic) event, but over various branches of insurance

(4) *SL reinsurance* (*stop loss reinsurance*): the priority of the insurer is applied to the annual result of the insurer in one branch against a negative deviation due to a marked increase in the number and the cost of losses; it has usually the form of a boundary for the claim ratio (see Sect. 21.1), to which the claim ratio must be rolled back in consequence of the reinsurer's intervention:

$$X_R = \begin{cases} 0 & \text{for } X/P \le p, \\ X - p \cdot P & \text{for } p < X/P \le l, \\ (l-p) \cdot P & \text{for } l < X/P, \end{cases} \quad \text{where } l > 0;\ p > 0$$

(excess in *SL reinsurance*: p is the priority of the insurer; l is the limit of the reinsurer (p and l are applied to the claim ratio X/P with claim amounts X and premiums P))

(5) *LCR*(r) *reinsurance* (*largest claims reinsurance*): works in such a way that the reinsurer reimburses r largest claims ($r \in \mathrm{N}$ is a given number), which incurred during the covered period (usually during a calendar year):

$$X_R = X_{(1)} + X_{(2)} + \ldots + X_{(r)}$$

(*LCR*(r) *reinsurance*: $X_{(1)} \ge X_{(2)} \ge \ldots \ge X_{(r)} \ge \ldots \ge X_{(n)}$ are ordered claims $X_1, X_2, \ldots, X_n$ in a given year (see Sect. 27.3))

(6) *ECOMOR*(r) *reinsurance*: works in such a way that the reinsurer reimburses only such parts of the claims that exceed the rth largest claim ($r \in \mathrm{N}$ is a given number, see *thereinbefore*):

$$X_R = \left(X_{(1)} - X_{(r)}\right) + \ldots + \left(X_{(r-1)} - X_{(r)}\right) = X_{(1)} + \ldots + X_{(r-1)} - (p-1) \cdot X_{(r)}$$

(*ECOMOR*(r) *reinsurance*: $X_{(1)} \ge X_{(2)} \ge \ldots \ge X_{(r)} \ge \ldots \ge X_{(n)}$ are ordered claims $X_1, X_2, \ldots, X_n$ in a given year (see Sect. 27.3))

24.3 Solvency

- *Solvency of insurer*: is insurer's ability to cover the liabilities accepted (i.e. to cover the entitled claims in events incurred)
- *Solvency reporting*: is a methodology for public supervision and regulation of the insurance industry to safeguard policyholders (or claimants in general) from the disastrous consequences of the insolvency of an insurer; this requires arrangements for preventing insurers from becoming bankrupt; such a supervision consists usually in testing the financial position of each insurer at regular intervals, normally annually; if there appears a substantial risk of insolvency urgent remedial measures will be required (including such hard measures as the withdrawal of the company's licence); solvency methodologies concentrate on the capital adequacy of the given insurer, i.e. whether insurer's assets will be sufficient to meet all claims and other obligations when the insurer would have to settle immediately all liabilities (the so-called *winding-up* or *break-up* schedule), or when no new business is written thereafter nor is any business renewed (the so-called *run-off* schedule), or when the insurer's activities will continue in future (the so-called *going-concern* schedule)
- *Ruin theory*: is a part of risk theory (see Sect. 22.5); it looks for the probability of the event that a value of the insurance process (usually a corresponding reserve) falls under an a priori given boundary (the so-called *ruin probability*); this theoretical context can be applied when modeling solvency
- Measurement of solvency:

  (1) *Analysis of basic accounting indicators*:

$$\frac{\text{available solvency margin}}{\text{net premium}}$$

(*solvency ratio*: *available solvency margin ASM* (*free assets FA*) of the insurer is excess of insurer's assets over and above what is needed to match insurer's liabilities (*ASM* involves usually the shareholders' equity, the undistributed (retained) profit, various reserve funds with exception of the technical reserves (i.e. provisions covering insurance liabilities), hidden reserves, and the like, also see *FA* in Sect. 19.5); *net premium* is the premium for own account of the insurer after deduction of the reinsurance premium, see Sect. 24.1)

$$\frac{\text{technical provisions}}{\text{net premium}} \quad (\textit{reserves ratio} : \text{see Sects. 18.3 and 21.4})$$

$$\frac{\text{net premium}}{\text{premium}}$$

(*retention ratio*: is the ratio of the premium after deduction of the reinsurance premium against the premium before deduction of the reinsurance premium)

(2) *Solvency capital requirements SCR* (*target capital*): is a level for *ASM* (see *thereinbefore*) required usually in the sense of going-concern schedule (see *thereinbefore*); as *SCR* is a “soft” level, the *minimum capital requirements MCR* is a “hard” level, under which the crisis management must be applied in the insurance company

(3) *Risk-based capital RBC*: is an approach to *SCR* (see *thereinbefore*), which looks for the required capital adequacy of the insurer by quantifying the corresponding risks; it is an analogy of the capital adequacy approach in finance (see Sect. 12.3: particular classes of assets of a bank are weighted by prescribed risk weights); this approach has been developed by the National Association of Insurance Commissioners (NAIC) as the regulatory system for the US (see *thereinafter*)

(4) *Simulation models*: are based mostly on the ruin theory (see *thereinbefore*) and simulate various scenarios when modeling future development

(5) *Rating assessment*: is essentially the *RBC* approach (see *thereinbefore*), where the risk weights applied depend on an official rating of the given risk category (which can be rating of the issuers of securities in the context of asset risk, rating of the debtors (including reinsurers) in the context of credit risk, and the like)

- Practical approaches to solvency reporting:

(1) *Solvency I* (Europe):

$$ASM \geq k_1 \cdot R_I + k_2 \cdot RC_I$$

(*solvency reporting in life insurance*: available solvency margin *ASM* (see *thereinbefore*) must not decrease under the solvency capital requirements *SCR* (see *thereinbefore*) calculated as a prescribed percentage $k1$ of the premium reserves (see Sect. 18.3) for own account of the insurer R_I and a prescribed percentage k_2 of the risk capital (the *risk capital* is the difference between the sums insured and the premium reserves, see Sect. 18.3) for own account of the insurer RC_I)

$$AMS \geq \max\{m_1 \cdot X_I;\ m_2 \cdot P_I\}$$

(*solvency reporting in non-life insurance*): available solvency margin *ASM* (see *thereinbefore*) must not be lower than the solvency capital requirements *SCR* (see *thereinbefore*) calculated as the maximum value between a prescribed percentage m_1 of the claim amounts for own account of the insurer X_I and from a prescribed percentage m_2 of the premiums for own account of the insurer P_I)

(2) *RBC* (the US):

$$TAC \geq RBC$$

(*solvency reporting*: *total adjusted capital TAC* (an analogy of *AMS*, see *thereinbefore*) must not be lower than the risk-based capital *RBC* (see *thereinafter*))

$$RBC = C_0 + \sqrt{(C_1 + C_3)^2 + C_2^2} + C_4$$

(*risk-based capital in life insurance*: C_i are required capital amounts, which cover the corresponding types of risks: C_0 for investments in insurance affiliates, for off-balance sheet risks and for contingent obligations; C_1 for asset risk, including risk of default, concentration risk and risk of reinsurance non-recoverability; C_2 for insurance risk, arising from high levels of claims, i.e. for risk of technical provisions and written premium risk; C_3 interest rate risk, arising from changes in levels of interest rates; C_4 for business risk, including expense overruns and management incompetence)

$$RBC = R_0 + \sqrt{R_1^2 + R_2^2 + \left(\frac{1}{2}R_3\right)^2 + \left(\frac{1}{2}R_3 + R_4\right)^2 + R_5^2}$$

(*risk-based capital in non-life insurance*: R_i are required capital amounts, which cover the corresponding types of risks: R_0 for investments in insurance affiliates, for off-balance sheet risks and for contingent obligations; R_1 for fixed interest investments; R_2 for equities, real estate investments and other asset risks; R_3 for credit risk including reinsurance non-recoverability; R_4 for loss reserve risk, i.e. for risk of technical provisions excepting the unearned premium reserves, and for reserve growth risk; R_5 for written premium risk and for premium growth risk)

(3) *Solvency II* (world): a new schedule that can make use of insurer's internal models

24.4 Alternative Risk Transfer ART

- *Alternative risk transfer ART*: is a product, channel or solution that transfers risk exposures between the insurance and capital markets to achieve prescribed risk management goals; the ART market is the combined risk management marketplace for innovative insurance and capital market solutions

- *Finite reinsurance*: is an important part of ART that addresses the problem of the limited (i.e. finite) risk in the context of reinsurance (sometimes a more special term *financial reinsurance* is used, see Sect. 24.1); the features of finite reinsurance are as follows:
  - limited acceptance of insurance risk by the reinsurer (finite products stabilize reinsurance costs)

- smoothing of fluctuations in insurance and reinsurance results (e.g. finite products attenuate cyclical trends in the insurance market)
- expansion of insurance and reinsurance capacity (e.g. the securitization of insurance risks makes use of the enormous capacity of capital markets, see *thereinafter*)
- separation of underwriting and timing risk
- experience account (the direct insurer participates in the losses as well in the profits so that the insurer's balance sheet is optimized)
- future investments income is explicitly taken into account in the calculation of the premiums (the finite reinsurance promotes the establishment of long-term relationships between the insurer and the reinsurer)
- reduction of credit risk (ways and means for ART reimbursements are often prepared in advance and in extenso)
- multi-years cover
- multiple, inter-connected triggers
- several lines of business may be covered
- application of various financial and tax aspects

- Some products, vehicles and solutions in the framework of ART:
  (1) *Captive*: is a risk channel, which is used to facilitate a company's own insurance/reinsurance risk financing or risk transfer strategies; a captive is generally formed as a licenced insurance/reinsurance company and may be controlled by a single owner or multiple owners (or sponsors)
  (2) *Risk retention group RRG*: is a retention vehicle (similar to the group captive), where a group assumes and spreads the liability risks of its members via pooling
  (3) *Coinsurance and reinsurance pools* (see Sect. 24.1): are sometimes considered to be forms of ART, since they mobilize sufficient capacities to cover large risks
  (4) *Securitization*: is the process of removing assets, liabilities or cash flows from the balance sheet and conveying them to third parties through tradable securities (the so-called *insurance-linked securities ILS*, see *thereinafter*); the factors that support the securitization of insurance risks are e.g. as follows:
    - *ILS triggers*: are triggers on the insurance-linked securities, where the suspension of interest and/or principal (see *thereinafter*) occurs when actual losses sustained by the issuer in a pre-defined segment of business reach a certain level (the so-called *indemnity triggers*), or when the value of a recognized third-party index reaches a certain threshold (the so-called *index triggers*, which are based on corresponding *loss indices* measured by reputable agencies as a standardized average loss development in the given region for the given type of insurance risk)
    - *risk exchanges* trading with securitized insurance risk, e.g. CATEX in the US (Catastrophe Risk Exchange)
    - *special purpose vehicles SPV*: are organizers of ILS trading

- *Catastrophe bonds* (*CatBonds*): are highly profitable bonds (their coupon rate is usually much higher than the market average, see Sect. 9.2), for which the suspension of coupons and/or principal occurs in the case of a predefined natural catastrophe; e.g. an annual reinsurance treaty, according to which the reinsurer reimburses a sum insured L at the end of the contract year, if the catastrophe has occurred, can be replaced by the issue of a 1-year catastrophe bond with an annual coupon; the following table contains the appropriate cash flows, which comply with requirements of all participating sides: q_{cat} is the probability of the natural catastrophe, i is the annual coupon rate, F is the face value (principal) of the bond (see Sect. 9.2), P_R is the reinsurance premium (see Sect. 24.1):

	Time $t = 0$	Time $t = 1$	
		Occurrence of catastrophe (with probability q_{cat})	Non-occurrence of catastrophe (with probability $1 - q_{\text{cat}}$)
Insurer	$-P_R = -\frac{1}{1+i} \cdot q_{\text{cat}} \cdot L$	L	0
Reinsurer=Issuer (*CatBond*)	$P_R + F$	$-L$	$-L$
Investor (*CatBond*)	$-F = -\frac{1}{1+i} \cdot (1 - q_{\text{cat}}) \cdot L$	0	L

$$q_b = \frac{L - F^* \cdot (1+i)}{L} .$$

(*probability of catastrophe priced by the bond market* (unlike the estimate q_{cat} by the reinsurance market, see *thereinbefore*): F^* is the market price of the given catastrophe bond)

$$P_{R_b} = \frac{1}{1+i} \cdot q_b \cdot L = \frac{L}{1+i} - F^*$$

(*reinsurance premium priced by the bond market*: see *thereinbefore*)

- *Insurance derivatives* (insurance swaps, insurance options, insurance futures, and the like): are derivative instruments, which are similarly as catastrophe bonds related to insurance (e.g. by means of the loss indices, see *thereinbefore*)

Further Reading

Banks, E.: Alternative Risk Transfer. Wiley, Chichester (2004)
Booth, P., Chadburn, R.., Cooper, D., Haberman, S., James, D.: Modern Actuarial Theory and Practice. Chapman and Hall/CRC, London (1999)

Dacey, M.S.: Reinsurance. Chatham, Kent (1992)
Daykin, C.D., Pentikäinen, T., Pesonen, M.: Practical Risk Theory for Actuaries. Chapman and Hall, London (1994)
Dienst, H.-R.: Mathematische Verfahren der Rückversicherung. Verlag Versicherungswirtschaft, Karlsruhe (1988)
Gerathewohl, K. et al.: Rückversicherung – Grundlagen und Praxis. Verlag Versicherungswirtschaft, Karlsruhe (Band I 1976, Band II 1979)
Grossmann, M.: Rückversicherung – eine Einführung. Hochschule St. Gallen, St. Gallen (1990)
Hart, D.G., Buchanan, R.A., Howe, B.A.: The Actuarial Practice of General Insurance. The Institute of Actuaries of Australia, Sydney (1996)
Kiln, R., Kiln, S.: Reinsurance in Practice. Witherby, London (2001)
Liebwein, P.: Klassische und moderne Formen der Rückversicherung. Verlag Versicherungswirtschaft, Karlsruhe (2000)
Pfeiffer, C.: Einführung in die Rückversicherung: Das Standardwerk für Theorie und Praxis. Gabler, Wiesbaden (1994)
Straub, E.: Non-Life Insurance Mathematics. Springer, Berlin (1988)
Teugels, J., Sundt, B. (eds.): Encyclopedia of Actuarial Science. Wiley, New York (2004)
Tiller, J.E., Fagerberg, D.: Life, Health, and Annuity Reinsurance. ACTEX Publications, Winsted and Avon, Connecticut (1990)

Part III
Formulas of Related Disciplines

Chapter 25
Mathematical Compendium

Abstract Chapter 25 reminds basic mathematical skills.

25.1 Powers with Integral Exponents ($a, b \in \mathrm{R}; a, b \neq 0; p, q \in \mathrm{Z}$)

$$a^p b^p = (ab)^p; \qquad \frac{a^p}{b^p} = \left(\frac{a}{b}\right)^p$$

$$a^p a^q = a^{p+q}; \qquad \frac{a^p}{a^q} = a^{p-q}; \qquad \left(a^p\right)^q = a^{p \cdot q}$$

25.2 Roots of Real Numbers ($a, b \in \mathrm{R}; a, b > 0; m, n \in \mathrm{N}; p \in \mathrm{Z}$)

$$\sqrt[n]{ab} = \sqrt[n]{a}\,\sqrt[n]{b}; \qquad \sqrt[n]{\frac{a}{b}} = \frac{\sqrt[n]{a}}{\sqrt[n]{b}}$$

$$\sqrt[n]{a^p} = \left(\sqrt[n]{a}\right)^p; \qquad \sqrt[m]{\sqrt[n]{a}} = \sqrt[m \cdot n]{a}$$

25.3 Powers with Rational Exponents ($a \in \mathrm{R}; a > 0; m, n \in \mathrm{N}$)

$$a^{1/n} = \sqrt[n]{a}; \qquad a^{-1/n} = \frac{1}{\sqrt[n]{a}}; \qquad a^{m/n} = \sqrt[n]{a^m}$$

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_25,

25.4 Powers with Real Exponents ($a, b \in \mathrm{R}; a, b > 0; x, y \in \mathrm{R}$)

$$a^x = \lim_{n\to\infty} a^{r_n}, \quad \text{where } \lim_{n\to\infty} r_n = x, \quad r_n \in Q$$

$$1^x = 1; \quad a^x b^x = (ab)^x; \quad \frac{a^x}{b^x} = \left(\frac{a}{b}\right)^x$$

$$a^x a^y = a^{x+y}; \quad \frac{a^x}{a^y} = a^{x-y}; \quad \left(a^x\right)^y = a^{x\cdot y}$$

25.5 Formulas $a^n \pm b^n$ ($a, b \in \mathrm{R}; n, k \in \mathrm{N}$)

$$a^2 - b^2 = (a+b)(a-b)$$

$$a^3 \pm b^3 = (a \pm b)(a^2 \mp ab + b^2)$$

$$a^n - b^n = (a-b)\left(a^{n-1} + a^{n-2}b + a^{n-3}b^2 + \ldots + ab^{n-2} + b^{n-1}\right)$$

$$a^{2k} - b^{2k} = (a+b)\left(a^{2k-1} - a^{2k-2}b + a^{2k-3}b^2 - \ldots - b^{2k-1}\right)$$

$$a^{2k+1} + b^{2k+1} = (a+b)\left(a^{2k} - a^{2k-1}b + a^{2k-2}b^2 - \ldots + b^{2k}\right)$$

25.6 Logarithms ($x, y \in \mathrm{R}; x, y > 0; a, b, c \in \mathrm{R}; a, b > 0; a, b \neq 1$)

$$y = \log_a x \quad \Leftrightarrow \quad a^y = x$$

$$\log x = \log_{10} x \ \ (\textit{common logarithm})$$

$$\ln x = \log_e x$$

(*natural logarithm*: $\mathrm{e} = \lim_{n\to\infty}(1 + 1/n)^n = 2{,}718\ 28\ldots$ is the base of natural logarithm)

$$a^{\log_a x} = x; \quad \log_a a = 1; \quad \log_a 1 = 0$$

$$\log_a xy = \log_a x + \log_a y; \quad \log_a \frac{x}{y} = \log_a x - \log_a y; \quad \log_a x^c = c \cdot \log_a x$$

$$\log_b x = \frac{\log_a x}{\log_a b}$$

25.7 Factorial and Binomial Coefficients ($k, m, n \in \mathrm{N}_0$; $k \leq m$, $k \leq n$)

$$n\,! = 1 \cdot 2 \cdot \ldots \cdot n; \quad 0\,! = 1 \quad (\textit{factorial})$$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{1 \cdot 2 \cdot \ldots \cdot k} \quad (\textit{binomial coefficient})$$

$$\binom{n}{k} = \binom{n}{n-k}\ ; \quad \binom{n}{0} = \binom{n}{n} = 1\ ; \quad \binom{n}{1} = \binom{n}{n-1} = n$$

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}\ ; \quad \binom{k}{k} + \binom{k+1}{k} + \ldots + \binom{n}{k} = \binom{n+1}{k+1}$$

$$\binom{n}{0}\binom{m}{k} + \binom{n}{1}\binom{m}{k-1} + \ldots + \binom{n}{k}\binom{m}{0} = \binom{n+m}{k}\ ; \quad \sum_{k=0}^{n} \binom{n}{k} = 2^n;$$

$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}$$

- Binomial coefficients may be ordered in *Pascal's triangle*:

					Binomial coefficients						
$n = 0$:						1					
$n = 1$:					1		1				
$n = 2$:				1		2		1			
$n = 3$:			1		3		3		1		
$n = 4$:		1		4		6		4		1	
$n = 5$:	1		5		10		10		5		1
$\vdots$	$\iddots$	$\iddots$	$\iddots$		$\iddots$		$\iddots$		$\iddots$		

- Binomial coefficients and corresponding relations (see *thereinbefore*) remain in force for $n \in \mathrm{R}$

25.8 Binomial Theorem ($a, b \in \mathbb{R}$; $n \in \mathbb{N}$)

$$(a \pm b)^n = \sum_{k=0}^{n} (\pm 1)^k \binom{n}{k} a^{n-k} b^k = a^n \pm \binom{n}{1} a^{n-1} b + \binom{n}{2} a^{n-2} b^2 \pm \ldots + (\pm 1)^n b^n$$

$$(a \pm b)^2 = a^2 \pm 2ab + b^2$$

$$(a \pm b)^3 = a^3 \pm 3a^2 b + 3ab^2 \pm b^3$$

25.9 Sums of Powers of Natural Numbers ($n \in \mathbb{N}$)

$$1 + 2 + \ldots + n = \frac{n(n+1)}{2}$$

$$1^2 + 2^2 + \ldots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

$$1^3 + 2^3 + \ldots + n^3 = \frac{n^2(n+1)^2}{4}$$

$$1^4 + 2^4 + \ldots + n^4 = \frac{n(n+1)(2n+1)(3n^2+3n-1)}{30}$$

25.10 Numerical Series ($a_1, d, q, v \in \mathbb{R}$; $n \in \mathbb{N}$)

$$\sum_{k=0}^{n-1} (a_1 + kd) = a_1 + (a_1 + d) + \ldots + (a_1 + (n-1)d) = na_1 + \frac{n(n-1)}{2} d \quad (\textit{arithmetic series})$$

$$\sum_{k=0}^{n-1} a_1 q^k = a_1 + a_1 q + \ldots + a_1 q^{n-1} = \begin{cases} a_1 \frac{1-q^n}{1-q}, & q \neq 1 \\ a_1 n, & q = 1 \end{cases} \quad (\textit{geometric series})$$

$$\sum_{k=0}^{\infty} q^k = 1 + q + q^2 + \ldots = \frac{1}{1-q}, \ |q| < 1; \quad \sum_{k=1}^{\infty} q^k = q + q^2 + \ldots = \frac{q}{1-q}, \ |q| < 1$$

$$\sum_{k=1}^{n} kq^k = q + 2q^2 + \ldots + nq^n = \begin{cases} q \cdot \frac{nq^{n+1} - (n+1)q^n + 1}{(1-q)^2}, & q \neq 1 \\ \frac{n(n+1)}{2}, & q = 1 \end{cases}$$

$$\sum_{k=1}^{\infty} kq^k = q + 2q^2 + \ldots = \frac{q}{(1-q)^2}, \quad |q| < 1$$

$$\sum_{k=1}^{n} \frac{k}{q^k} = \frac{1}{q} + \frac{2}{q^2} + \ldots + \frac{n}{q^n} = \begin{cases} \frac{1}{q^n} \cdot \frac{q^{n+1}-(n+1)q+n}{(1-q)^2}, & q \neq 1 \\ \frac{n(n+1)}{2}, & q = 1 \end{cases}$$

$$\sum_{k=1}^{\infty} \frac{k}{q^k} = \frac{1}{q} + \frac{2}{q^2} + \ldots = \frac{q}{(1-q)^2}, \quad |q| > 1$$

$$\sum_{k=1}^{n} \left(kq^k\right)^2 = q^2 + \left(2q^2\right)^2 + \ldots \left(nq^n\right)^2 =$$

$$= \begin{cases} \dfrac{-n^2q^{2n+6} + (2n^2+2n-1)q^{2n+4} - (n+1)^2q^{2n+2} + q^4 + q^2}{(1-q^2)^3}, & q \neq 1 \\ \dfrac{n(n+1)(2n+1)}{6}, & q = 1 \end{cases}$$

$$\sum_{k=1}^{\infty} \left(kq^k\right)^2 = q^2 + \left(2q^2\right)^2 + \ldots = \frac{q^2(q^2+1)}{(1-q^2)^3}, \quad |q| < 1$$

$$\sum_{k=1}^{n} kq^{n-k+1} = q^n + 2q^{n-1} + \ldots + nq = \begin{cases} q \cdot \frac{q^{n+1}-(n+1)q+n}{(1-q)^2}, & q \neq 1 \\ \frac{n(n+1)}{2}, & q = 1 \end{cases}$$

$$\sum_{k=1}^{n} \frac{k}{q^{n-k+1}} = \frac{1}{q^n} + \frac{2}{q^{n-1}} + \ldots + \frac{n}{q} = \begin{cases} \frac{1}{q^n} \cdot \frac{nq^{n+1}-(n+1)q^n+1}{(1-q)^2}, & q \neq 1 \\ \frac{n(n+1)}{2}, & q = 1 \end{cases}$$

$$\sum_{k=0}^{n-1} (a_1 + kd) \cdot v^k = a_1 + (a_1 + d) \cdot v + \ldots + (a_1 + (n-1)d) \cdot v^{n-1}$$
$$= \begin{cases} a_1 \frac{1-v^n}{1-v} + dv\frac{(n-1)v^n - nv^{n-1}+1}{(1-v)^2}, & v \neq 1 \\ na_1 + d\frac{(n-1)n}{2}, & v = 1 \end{cases}$$

$$\sum_{k=0}^{n-1} (a_1 + kd) \cdot v^{n-k-1} = a_1 \cdot v^{n-1} + (a_1 + d) \cdot v^{n-2} + \ldots + (a_1 + (n-1)d)$$
$$= \begin{cases} a_1 \frac{1-v^n}{1-v} + d\frac{v^n - nv + n - 1}{(1-v)^2}, & v \neq 1 \\ na_1 + d\frac{(n-1)n}{2}, & v = 1 \end{cases}$$

$$\sum_{k=0}^{n-1} \left(a_1q^k\right) \cdot v^k = a_1 + (a_1q) \cdot v + \ldots + \left(a_1q^{n-1}\right) \cdot v^{n-1} = \begin{cases} a_1 \frac{1-(qv)^n}{1-qv}, & v \neq \dfrac{1}{q} \\ na_1, & v = \dfrac{1}{q} \end{cases}$$

$$\sum_{k=0}^{n-1} \left(a_1q^k\right) \cdot v^{n-k-1} = a_1 \cdot v^{n-1} + (a_1q) \cdot v^{n-2} + \ldots + \left(a_1q^{n-1}\right) = \begin{cases} a_1 \frac{q^n - v^n}{q-v}, & v \neq q \\ na_1q^{n-1}, & v = q \end{cases}$$

25.11 Means ($x_1, \ldots, x_n \in \mathrm{R}; k, n, n_1, \ldots, n_k \in \mathrm{N}; n = \sum_{i=1}^{k} n_i$)

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \quad (\textit{arithmetic mean})$$

$$\bar{x}_G = \sqrt[n]{x_1 x_2 \ \ldots \ x_n} \quad (\textit{geometric mean})$$

$$\bar{x}_H = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} \quad (\textit{harmonic mean})$$

$$\bar{x}_K = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}} \quad (\textit{quadratic mean})$$

$$\bar{x} = \frac{\sum_{i=1}^{k} n_i x_i}{n} \quad (\textit{weighted}\text{ arithmetic mean})$$

$$\bar{x}_G = \sqrt[n]{x_1^{n_1} x_2^{n_2} \ \ldots \ x_k^{n_k}} \quad (\textit{weighted}\text{ geometric mean})$$

$$\bar{x}_H = \frac{n}{\sum_{i=1}^{k} \frac{n_i}{x_i}} \quad (\textit{weighted}\text{ harmonic mean})$$

$$\bar{x}_K = \sqrt{\frac{\sum_{i=1}^{k} n_i x_i^2}{n}} \quad (\textit{weighted}\text{ quadratic mean})$$

$$\bar{x}_H \le \bar{x}_G \le \bar{x} \le \bar{x}_K$$

25.12 Beta and Gamma Function ($x, p, q \in \mathrm{R}; p > 0, q > 0; n \in \mathrm{N}$)

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} \, \mathrm{d}t \quad (\textit{gamma function})$$

$$\Gamma(x+1) = x \cdot \Gamma(x); \quad \Gamma(n) = (n-1)!\,; \quad \Gamma(1) = 1; \ \Gamma(2) = 1; \ \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi};$$
$$\Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\sqrt{\pi}$$

$$B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}\,\mathrm{d}x \quad (\textit{beta function})$$

$$B(p,q) = B(q,p); \quad B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

Further Reading

Rektorys, K. et al.: Survey of Applicable Mathematics. Kluwer, Dordrecht (1994)

Chapter 26
Probability Theory

Abstract Chapter 26 deals with formulas and laws of probability theory: 26.1. Random Events and Probability, 26.2. Conditional Probability and Independent Events, 26.3. Random Variables and Their Basic Characteristics, 26.4. Important Discrete Distributions, 26.5. Important Continuous Distributions, 26.6. Random Vectors and Their Basic Characteristics, 26.7. Transformation of Random Variables, 26.8. Conditional Mean Value, 26.9. Martingales, 26.10. Generating Function, 26.11. Convolutions and Sums of Random Variables, 26.12. Random Sums of Random Variables, 26.13. Some Inequalities, 26.14. Limit Theorems of Probability Theory.

26.1 Random Events and Probability

$$0 \leq P(A) \leq 1 \ (\textit{probability of random event } A)$$

$$P(\emptyset) = 0 \ (\textit{impossible event}); \ P(\Omega) = 1 \ (\textit{certain event})$$

$$P(\overline{A}) = 1 - P(A) \ (\textit{complementary} \text{ random event})$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \ (\textit{union} \text{ of random events})$$

$$P(A \cup B) = P(A) + P(B) \ (\text{union of } \textit{disjoint random events}: \ A \cap B = \emptyset)$$

$$P(A - B) = P(A) - P(A \cap B) \ (\textit{difference} \text{ of random events}: \ A - B = A \cap \overline{B})$$

$$A \subset B \Rightarrow P(A) \leq P(B) \ (\text{random event } A \textit{ implies} \text{ random event } B)$$

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} (-1)^{i-1} \cdot \sum_{1 \leq k_1 < \ldots < k_i \leq n} P(A_{k_1} \cap \ldots \cap A_{k_i})$$

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_26,

$$P\left(\bigcap_{i=1}^{n} A_i\right) \geq 1 - \sum_{i=1}^{n} [1 - P(A_i)] \text{ (\textit{Bonferroni Inequality})}$$

$$P(A) = \frac{\binom{K}{k} \cdot \binom{N-K}{n-k}}{\binom{N}{n}}$$

(probability that there are precisely k defective products in a sample of n products randomly selected from a lot of N products that contains K defective ones)

$$(\Omega, \boldsymbol{B}, P)$$

(*probability space* with elementary events $\omega \in \Omega$, σ-algebra $\boldsymbol{B}$ (see *thereinafter*) and probability P; it is the key concept of axiomatic probability theory)

$\boldsymbol{B}$ from triplet $(\Omega, \boldsymbol{B}, P)$ (see *thereinbefore*), where (*i*) $\Omega \in \boldsymbol{B}$; (*ii*) $\bar{A} \in \boldsymbol{B}$ for $A \in \boldsymbol{B}$; (*iii*) $A_1 \cup A_2 \cup A_3 \cup \ldots \in \boldsymbol{B}$ for $A_1, A_2, A_3, \ldots \in \boldsymbol{B}$ (*σ-algebra* of probability space; it is a non-empty family of all considered random events, which are sets of elementary events from Ω)

P from triplet $(\Omega, \boldsymbol{B}, P)$ (see *thereinbefore*), where (*i*) $P(\Omega) = 1$; (*ii*) $P(A) \geq 0$ for $A \in \boldsymbol{B}$; (*iii*) $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ for mutually disjoint $A_1, A_2, A_3, \ldots \in \boldsymbol{B}$ (*probability* (or strictly speaking *probability measure*) of probability space)

$$\begin{cases} P\left(\bigcap_{k=1}^{\infty} \bigcup_{i=k}^{\infty} A_i\right) = 0 \text{ for } \sum_{i=1}^{\infty} P(A_i) < \infty \\ \qquad\qquad\qquad\qquad = 1 \text{ for } \sum_{i=1}^{\infty} P(A_i) = \infty \text{ with mutually disjoint } A_1, A_2, \ldots \in \Omega \end{cases}$$

(*Borel–Cantelli lemma*: the corresponding probability can be interpreted as $P\{\omega : \omega \in A_i$ for infinite number of $i\}$)

26.2 Conditional Probability and Independent Events

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

(*conditional probability* of event A given event B, where $P(B) > 0$)

$$P(B) = \sum_{i=1}^{n} P(A_i) \cdot P(B|\,A_i),$$

where $A_1, \ldots, A_n$ are disjoint, and their union is certain event (*total probability rule*)

$$P(A_k \mid B) = \frac{P(A_k) \cdot P(B \mid A_k)}{\sum\limits_{i=1}^{n} P(A_i) \cdot P(B \mid A_i)}, \; k = 1, \ldots, n,$$

where $A_1, \ldots, A_n$ are disjoint, and their union is certain event (*Bayes Theorem*)

$$P(A \cap B) = P(A) \cdot P(B) \; (\textit{independent events } A \text{ and } B)$$

$$P(A_{i_1} \cap \ldots \cap A_{i_r}) = P(A_{i_1}) \cdot \ldots \cdot P(A_{i_r}), \; 1 \le i_1 < \ldots < i_r \; (\textit{independent events } A_1, A_2, \ldots)$$

$$P(A|B) = P(A),$$

where A and B are independent events and $P(B) > 0$

26.3 Random Variables and Their Basic Characteristics

$$\{\omega : X(\omega) \le x\} \in \boldsymbol{B} \quad \text{for } x \in \mathrm{R}$$

(*random variable* X: is *measurable* (see *thereinbefore*) real function X: $\Omega \to \mathrm{R}$ on probability space $(\Omega, \boldsymbol{B}, P)$ (see Sect. 26.1); the smallest σ-algebra, for which X is measurable, is called *σ-algebra generated by random variable X* and is denoted $\sigma(X) \subset \boldsymbol{B}$ (analogously one can define σ-algebra generated by more random variables))

$$F(x) = P(X \le x), \quad x \in \mathrm{R}$$

(*distribution function* of random variable X: $F(x)$ is non-decreasing and continuous from the right)

$$\lim_{x \to -\infty} F(x) = 0; \quad \lim_{x \to \infty} F(x) = 1$$

$$P(a < X \le b) = F(b) - F(a); \quad P(X = a) = F(a) - F(a - 0), \quad a, b \in \mathrm{R}; \quad a < b$$

$$F(x) = \sum_{x_j \le x} p_j$$

(distribution function of *discrete* random variable with $p_j = P(X = x_j)$, $\sum\limits_{x_j} p_j = 1$)

$$F(x) = \int_{-\infty}^{x} f(t) \, \mathrm{d}t$$

(distribution function of *continuous* random variable with *probability density* f)

$$P(a < X \leq b) = \int_a^b f(x)\,\mathrm{d}x, \quad a, b \in \mathrm{R}; \quad a < b,$$

where X has the probability density $f(x)$

$$\mathrm{E}(X) = \int_\Omega X\,\mathrm{d}P(\omega)$$

(*mean value* of random variable: general definition of mean value as the integral of X with respect to a probability measure P; the existence of *finite* mean value is equivalent with *integrability* of X with respect to P)

$$\mathrm{E}(X) = \sum_{x_j} x_j \cdot p_j$$

(*mean value* of discrete random variable: "average of this variable", see Sect. 27.5)

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,\mathrm{d}x$$

(*mean value* of continuous random variable: "average of this variable")

$$\mathrm{E}\,(g(X)) = \sum_{x_j} g(x_j)p_j; \quad \mathrm{E}\,(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)\,\mathrm{d}x,$$

where g is a function of random variable X

$$\mathrm{var}(X) = \mathrm{E}\left((X - \mathrm{E}(X))^2\right) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2 \ \ (\textit{variance}, \text{ see Sect. 27.6})$$

$$\sigma(X) = \sqrt{\mathrm{var}(X)}$$

(*standard deviation*: "average deviation from average", see Sect. 27.6)

$$\mathrm{E}\,(|X - \mathrm{E}X|)\ (\textit{mean deviation})$$

$$V(X) = \frac{\sqrt{\mathrm{var}(X)}}{|\mathrm{E}(X)|}$$

(*coefficient of variation*: "relative standard deviation with respect to average", see Sect. 27.6)

$$\gamma_1 = \frac{\mathrm{E}\,(X - \mathrm{E}(X))^3}{\sigma^3}$$

(*skewness*: "distinction in concentration of higher and lower values", see Sect. 27.7)

$$\gamma_2 = \frac{\mathrm{E}\,(X - \mathrm{E}(X))^4}{\sigma^4} - 3$$

(*kurtosis*: "distinction in concentration of inner and outer values", see Sect. 27.7)

$$\mu'_k = \mathrm{E}(X^k),\ k \in N_0\ (k\text{th } \textit{moment}\text{:}\ \ \mu'_0 = 1,\ \ \mu'_1 = \mathrm{E}X)$$

$$\mu_k = \mathrm{E}\,(X - \mathrm{E}(X))^k,\ \ k \in N_0$$

(kth *central moment*: $\mu_0 = 1$, $\mu_1 = 0$, $\mu_2 = \mathrm{var}(X)$)

$$\mu_k = \sum_{j=0}^{k} \binom{k}{j} (-\mu'_1)^j \mu'_{k-j};\quad \mu'_k = \sum_{j=0}^{k} \binom{k}{j} (\mu'_1)^j \mu_{k-j},\ \ k \in N_0$$

$$F(x_p - 0) \le p,\ \ F(x_p) \ge p,\ \ 0 < p < 1$$

(*p-quantile* x_p of random variable: "it separates $100p\%$ lower values from $100\cdot(1-p)\ \%$ higher values", see Sect. 27.4)

$$P(X \le x_p) = p,\ \ 0 < p < 1$$

(*p-quantile* x_p of continuous random variable, see Sect. 27.4)

$x_{0,\,5}$ (*median*: "50% quantile", see Sects. 27.4 and 27.5)

$x_{0,\,25}$ (*lower quartile*, see Sect. 27.4); $x_{0,\,75}$ (*upper quartile*, see Sect. 27.4)

$x_{k/10}, k = 1, \ldots, 9$ (*kth decile*); $x_{k/100}, k = 1, \ldots, 99$ (*kth percentile*, see Sect. 27.4)

$x_{0,\,75} - x_{0,\,25}$ (*interquartile range*, see Sect. 27.6); $x_{0,\,9} - x_{0,\,1}$ (*interdecile range*, see Sect. 27.6)

$$P(X = \hat{x}) \ge P(X = x_j) \text{ for all } x_j$$

(*mode* $\hat{x}$ of discrete random variable with $p_j = P(X = x_j)$, $\sum_{x_j} p_j = 1$, see Sect. 27.5)

$$f(\hat{x}) \ge f(x) \text{ for all } x$$

(*mode* $\hat{x}$ of continuous random variable with probability density $f(x)$, see Sect. 27.5)

$$m(t) = \mathrm{E}\left(e^{t \cdot X}\right), \quad t \in \mathrm{R} \ (\textit{moment generating function})$$

$$\mu_k' = m^{(k)}(0), \quad k \in N_0,$$

where μ_k' is kth moment and $m^{(k)}(\cdot)$ is kth derivative of moment generating function

26.4 Important Discrete Distributions

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, \ n$$

(*binomial distribution* $B(n, p)$ with parameters $n \in \mathrm{N}, p \in (0, 1)$: $\mathrm{E}(X) = np$; $\mathrm{var}(X) = np(1-p)$; $\gamma_1 = \frac{1-2p}{\sqrt{np(1-p)}}$; $\gamma_2 = \frac{1-6p(1-p)}{np(1-p)}$)

$$P(X = x) = \begin{cases} p^r, & x = 0 \\ \frac{(r+x-1)(r+x-2)\ldots r}{x!} p^r (1-p)^x, & x = 1, 2, \ldots \end{cases}$$

(*negative binomial distribution* $NB(n, p)$ with parameters $r > 0$, $p \in (0, 1)$: $\mathrm{E}(X) = \frac{r(1-p)}{p}$; $\mathrm{var}(X) = \frac{r(1-p)}{p^2}$; $\gamma_1 = \frac{2-p}{\sqrt{r(1-p)}}$; $\gamma_2 = \frac{p^2-6p+6}{r(1-p)}$)

$$P(X = x) = p(1-p)^x, \quad x = 0, 1, \ldots$$

(*geometric distribution* with parameter $p \in (0, 1)$: is the negative binomial distribution with parameters $r = 1$ and $p \in (0, 1)$)

$$P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!}, \quad x = 0, 1, \ldots$$

(*Poisson distribution* $P(\lambda)$ with intensity $\lambda > 0$: $\mathrm{E}(X) = \lambda$; $\mathrm{var}(X) = \lambda$; $\gamma_1 = 1/\sqrt{\lambda}$; $\gamma_2 = 1/\lambda$)

$$P(X = x) = \int_{-\infty}^{\infty} \mathrm{e}^{-\vartheta} \frac{\vartheta^x}{x!} \, \mathrm{d}F(\vartheta), \quad x = 0, 1, \ldots$$

(*mixed Poisson distribution* with distribution function $F(\vartheta)$ of random intensity Θ:
– for Θ constant with $P(\Theta = \lambda) = 1$: Poisson distribution with intensity λ
– for Θ with gamma distribution with parameters $p/(1-p)$ and r (see Sect. 26.5): negative binomial distribution with parameters r and p)

$$P(X = x) = \frac{1}{|\log q|} \cdot \frac{p^x}{x}, \quad x = 1, 2, \ldots$$

(*logarithmic distribution* with parameter $p \in (0,\ 1)$, where $q = 1 - p$: $\mathrm{E}(X) = \frac{1}{|\log q|} \frac{p}{q}$; $\mathrm{var}(X) = \frac{1}{(\log q)^2} \frac{p}{q^2} (|\log q| - p)$; $\gamma_1 = \frac{(1+p)(\log q)^2 - 3pq\ |\log q| + 2p^2}{\sqrt{p}(|\log q| - p)^{3/2}}$)

26.5 Important Continuous Distributions

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \mathrm{R}$$

(*normal distribution* $N(\mu,\ \sigma^2)$ with parameters $\mu \in \mathrm{R}$, $\sigma^2 > 0$: $\mathrm{E}(X) = \mu$; $\gamma_1 = 0$; $\gamma_2 = 0$; $P(a \le X \le b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$; $P(|X - \mu| \le a) = 2\Phi\left(\frac{a}{\sigma}\right) - 1$;
$P(|X-\mu| > k\sigma) = \begin{cases} 0.317\,3 & \text{for } k = 1 \\ 0.045\,5 & \text{for } k = 2 \\ 0.002\,7 & \text{for } k = 3 \end{cases}$)

$$f(x) = \varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \quad x \in \mathrm{R}$$

(*standard normal distribution* $N(0,\ 1)$: $F(x) = \Phi(x) = \int\limits_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right)\,\mathrm{d}t$;
$\mathrm{E}(X) = 0$; $\mathrm{var}(X) = 1$; $\gamma_1 = 0$; $\gamma_2 = 0$; p-quantiles u_p (i.e. $P(X \le u_p) = p$, $0 < p < 1$) are tabulated)

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right), \quad x > 0$$

(*logarithmic normal distribution* $LN(\mu,\ \sigma^2)$ with parameters $\mu \in \mathrm{R}$, $\sigma^2 > 0$: $X = \exp(Z)$ for $Z \sim N\left(\mu, \sigma^2\right)$; $\mathrm{E}(X) = \exp\left(\mu + \frac{\sigma^2}{2}\right)$; $\mathrm{var}(X) = \exp\left(2\mu + \sigma^2\right) \times \left(\exp(\sigma^2) - 1\right)$; $\gamma_1 = \left(\exp(\sigma^2) + 2\right)\sqrt{\exp(\sigma^2) - 1}$; $\gamma_2 = \exp(4\sigma^2) + 2\exp(3\sigma^2) + 3\exp(2\sigma^2) - 6$)

$$f(x) = \frac{1}{b-a}, \quad a \le x \le b$$

(*uniform distribution* with parameters $a,\ b \in \mathrm{R}$ $(a < b)$: $\mathrm{E}(X) = \frac{a+b}{2}$; $\mathrm{var}(X) = \frac{(b-a)^2}{12}$; $\gamma_1 = 0$; $\gamma_2 = -\frac{6}{5}$)

$$f(x) = \lambda\, \mathrm{e}^{-\lambda \cdot x}, \quad x \geq 0$$

(*exponential distribution* $Exp(\lambda)$ with parameter $\lambda > 0$: $F(x) = 1 - \mathrm{e}^{-\lambda \cdot x}$ for $x \geq 0$; $\mathrm{E}(X) = 1/\lambda$; $\mathrm{var}(X) = 1/\lambda^2$; $\gamma_1 = 2$; $\gamma_2 = 6$; $P(X > t + s\,|X > s) = P(X > t)$ for $s, t > 0$)

$$f(x) = \frac{\lambda}{2}\, \mathrm{e}^{-\lambda \cdot |x-a|}, \quad x \in \mathrm{R}$$

(*double exponential distribution* with parameters $a \in \mathrm{R}$, $\lambda > 0$: $\mathrm{E}(X) = a$; $\mathrm{var}(X) = 2/\lambda^2$; $\gamma_1 = 0$; $\gamma_2 = 3$)

$$f(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1}\, \mathrm{e}^{-\lambda \cdot x}, \quad x \geq 0$$

(*gamma distribution* $\Gamma(a, \lambda)$ with parameters $a > 0$, $\lambda > 0$: $\mathrm{E}(X) = a/\lambda$; $\mathrm{var}(X) = a/\lambda^2$; $\gamma_1 = 2/\sqrt{a}$; $\gamma_2 = 6/a$; *Erlang distribution*: is gamma distribution with $a \in \mathrm{N}$; $\chi^2(n)$ *distribution*: is gamma distribution with $a = n/2$, $\lambda = 1/2$)

$$f(x) = \frac{1}{B(p,q)} x^{p-1}(1-x)^{q-1}, \quad 0 < x < 1$$

(*beta distribution* with parameters $p > 0, q > 0$: $\mathrm{E}(X) = \frac{p}{p+q}$; $\mathrm{var}(X) = \frac{pq}{(p+q)^2(p+q+1)}$)

$$f(x) = a\lambda^a x^{a-1} \exp\left(-\lambda^a x^a\right), \quad x > 0$$

(*Weibull distribution* with parameters $a > 0$, $\lambda > 0$: $\mathrm{E}(X) = \frac{1}{\lambda}\Gamma\left(\frac{a+1}{a}\right)$; $\mathrm{var}(X) = \frac{1}{\lambda^2}\left(\Gamma\left(\frac{a+2}{a}\right) - \Gamma^2\left(\frac{a+2}{a}\right)\right)$; $\mathrm{Exp}(\lambda)$: is Weibull distribution with $a = 1$; *Rayleigh distribution*: is Weibull distribution with $a = 2$)

$$f(x) = \frac{1}{\pi} \cdot \frac{b}{b^2 + (x-a)^2}, \quad x \in \mathrm{R}$$

(*Cauchy distribution* with parameters $a \in \mathrm{R}$, $b > 0$: has no moments (i.e. mean value, variance, etc.))

$$f(x) = \frac{1}{\pi} \cdot \frac{2b}{b^2 + (x-a)^2}, \quad x \geq a$$

(*one-sided Cauchy distribution* with parameters $a \in \mathrm{R}$, $b > 0$: has no moments (i.e. mean value, variance, ...))

$$f(x) = \frac{ba^b}{x^{b+1}}, \quad x \geq a$$

(*Pareto distribution* with parameters $a > 0, b > 0$:
$\mathrm{E}(X) = \dfrac{ab}{b-1}$ for $b > 1$; $\mathrm{var}(X) = \dfrac{a^2 b}{(b-1)^2(b-2)}$ for $b > 2$;
$\gamma_1(X) = \dfrac{2\sqrt{b-2}\,(b+1)}{\sqrt{b}(b-3)}$ for $b > 3$)

$$f(x) = \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} \mathrm{e}^{-x/2}, \quad x > 0$$

(χ^2 *distribution* with $n \in \mathrm{N}$ degrees of freedom $\chi^2(n)$: $\mathrm{E}(X) = n$; $\mathrm{var}(X) = 2n$; $X_1^2 + \ldots + X_n^2$ for independent random variables $X_1, \ldots, X_n$ with distribution $N(0, 1)$ has the distribution $\chi^2(n)$; p-quantiles $\chi_p^2(n)$ (i.e. $P(X \leq \chi_p^2(n)) = p$, $0 < p < 1$) are tabulated)

$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\,\Gamma(n/2)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \quad x \in \mathrm{R}$$

(*t distribution* (*Student's*) with n degrees of freedom $t(n)$:

$\mathrm{E}(X) = 0$ for $n > 1$; $\mathrm{var}(X) = \frac{n}{n-2}$ for $n > 2$; $X/\sqrt{Y/n}$ for independent random variables X and Y with distribution $N(0, 1)$ and $\chi^2(n)$ has the distribution $t(n)$; p-quantiles $t_p(n)$ (i.e. $P(X \leq t_p(n)) = p$, $0 < p < 1$) are tabulated)

$$f(x) = \frac{1}{B(n_1/2, n_2/2)} \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} x^{\frac{n_1}{2} - 1} \left(1 + \frac{n_1}{n_2} x\right)^{-\frac{n_1+n_2}{2}}, \quad x > 0$$

(*F distribution* (*Fisher–Snedecor*) with n_1 and n_2 degrees of freedom $F(n_1, n_2)$:

$\mathrm{E}(X) = \frac{n_2}{n_2-2}$ for $n_2 > 2$; $\mathrm{var}(X) = \frac{2n_2^2(n_1+n_2-2)}{n_1(n_2-2)^2(n_2-4)}$ for $n_2 > 4$; $(X/n_1)/(Y/n_2)$ for independent random variables X and Y with distribution $\chi^2(n_1)$ and $\chi^2(n_2)$ has the distribution $F(n_1, n_2)$; p-quantiles $F_p(n_1, n_2)$ (i.e. $P(X \leq F_p(n_1, n_2)) = p$, $0 < p < 1$) are tabulated)

26.6 Random Vectors and Their Basic Characteristics

$$F(x_1, \ldots, x_n) = P(X_1 \leq x_1, \ldots, X_n \leq x_n), \quad x_1, \ldots, x_n \in \mathrm{R}$$

(*distribution function* of random variables $X_1, \ldots, X_n$: non-decreasing and continuous from the right in each of its arguments x_i)

$$\lim_{x_j \to -\infty} F(x_1, \ldots, x_n) = 0, \ \ j = 1, \ldots, n; \qquad \lim_{x_1 \to \infty, \ldots, x_n \to \infty} F(x_1, \ldots, x_n) = 1$$

$$P(a_1 < X \le b_1, \ldots, a_n < X \le b_n) = \sum_{\delta_1, \ldots, \delta_n} (-1)^{\sum_{i=1}^{n} \delta_i} F(c_1, \ldots, c_n),$$

$$a_i, b_i \in \mathrm{R}; \ \ a_i < b_i \ (i = 1, \ldots, n), \text{ where } c_i = \delta_i a_i + (1 - \delta_i) b_i; \ \delta_i = \pm 1$$

$$P(a_1 < X_1 \le b_1, \ a_2 < X_2 \le b_2) = F(b_1, \ b_2) - F(a_1, \ b_2) - F(a_2, \ b_1) + F(a_1, \ a_2),$$

$$a_i, b_i \in \mathrm{R}; \ \ a_i < b_i \ \ (i = 1, \ 2)$$

$$F(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_n} f(t_1, \ldots, t_n) \, \mathrm{d}t_1 \ldots \mathrm{d}t_n$$

(distribution function of *continuous* random variables with *probability density* $f(x_1, \ldots, x_n)$)

$$P(a_1 < X \le b_1, \ldots, a_n < X \le b_n) = \int_{a_1}^{b_1} \ldots \int_{a_n}^{b_n} f(x_1, \ldots, x_n) \ \mathrm{d}x_1 \ldots \ \mathrm{d}x_n,$$

$$a_i, b_i \in \mathrm{R}; \ a_i < b_i \ (i = 1, \ldots, n)$$

$$F_1(x_1) = \lim_{x_2 \to -\infty} F(x_1, \ x_2); \quad F_2(x_2) = \lim_{x_1 \to -\infty} F(x_1, \ x_2)$$

(*marginal* distribution functions)

$$P(X_1 = x_1) = \sum_{x_2} P(X_1 = x_1, \ X_2 = x_2); \quad P(X_2 = x_2) = \sum_{x_1} P(X_1 = x_1, \ X_2 = x_2)$$

(*marginal* probability functions of discrete random variables)

$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, \ x_2) \, \mathrm{d}x_2; \quad f_2(x_2) = \int_{-\infty}^{\infty} f(x_1, \ x_2) \, \mathrm{d}x_1$$

(*marginal* probability densities of continuous random variables)

$$P(X_1 = x_1 | \ X_2 = x_2) = \begin{cases} P(X_1 = x_1, \ X_2 = x_2)/P(X_2 = x_2), & P(X_2 = x_2) \ne 0 \\ 0, & P(X_2 = x_2) = 0 \end{cases}$$

(*conditional* probability function of discrete random variable)

$$f(x_1 \,|x_2) = \begin{cases} f(x_1,\ x_2)/f_2(x_2), & f_2(x_2) \neq 0 \\ 0, & f_2(x_2) = 0 \end{cases}$$

(*conditional* probability density of continuous random variable)

$$F(x_1, \ldots,\ x_n) = F(x_1) \cdot\ \ldots \cdot F(x_n),\quad x_1, \ldots,\ x_n \in \mathrm{R}$$

(distribution function of *independent* random variables $X_1, \ldots, X_n$)

$$\mathrm{E}(X_1^{k_1} \cdot \ldots \cdot X_n^{k_n}) = \mathrm{E}(X_1^{k_1}) \cdot\ \ldots \cdot \mathrm{E}(X_n^{k_n}),\quad k_1, \ldots,\ k_n \in N_0,$$

where random variables $X_1, \ldots, X_n$ are independent

$$\mathrm{cov}(X, Y) = \mathrm{E}\left((X - \mathrm{E}(X))\,(Y - \mathrm{E}(Y))\right) = \mathrm{E}(XY) - \mathrm{E}(X) \cdot \mathrm{E}(Y)$$

(*covariance*, see Sect. 27.8)

$$\rho(X, Y) = \frac{\mathrm{cov}(X, Y)}{\sqrt{\mathrm{var}(X)\mathrm{var}(Y)}}$$

(*correlation coefficient*: $-1 \leq \rho(X, Y) \leq 1$; $\rho(X, Y) = 0$ for independent X and Y;

$\rho(X, Y) = 1$, if and only if $Y = aX + b$ a.s. for $a > 0$;

$\rho(X, Y) = -1$, if and only if $Y = aX + b$ a.s. for $a < 0$, see Sect. 27.8)

$$\rho(X, Y) = 0 \text{ (or equivalently, } \mathrm{cov}(X, Y) = 0)$$

(*uncorrelated* random variables, see Sect. 27.8)

$$\mathbf{\Sigma_{xx}} = \left(\mathrm{cov}(X_i, X_j)\right)_{\substack{i=1,\ldots,\,n \\ j=1,\ldots,\,n}}$$

(*covariance matrix* of random variables $X_1, \ldots, X_n$, see Sect. 27.8)

$$\mathbf{R_{xx}} = \left(\rho(X_i, X_j)\right)_{\substack{i=1,\ldots,\,n \\ j=1,\ldots,\,n}}$$

(*correlation matrix* of random variables $X_1, \ldots, X_n$, see Sect. 27.8)

$$P(X_1 = x_1, \ldots,\ X_k = x_k) = \frac{n!}{x_1!\ \ldots\ x_k!} p_1^{x_1}\ \ldots\ p_k^{x_k},\ \ x_j = 0, 1, \ldots,\ n;\ \ x_1 + \ldots + x_k = n$$

(*k-variate multinomial distribution* with parameters $n \in \mathrm{N}$, $p_j \in (0,\ 1)$, $p_1 + \ldots + p_k = 1$: $\mathrm{E}(X_j) = np_j$; $\mathrm{var}(X_j) = np_j(1 - p_j)$; $\mathrm{cov}(X_i, X_j) = -np_ip_j$ for $i \neq j$)

$$f(x_1,\dots,x_n) = \frac{1}{(2\pi)^{n/2}\,|\boldsymbol{\Sigma}|^{1/2}} \times \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right),\quad \mathbf{x} = (x_1,\dots,x_n)' \in \mathrm{R}^n$$

(*n-variate normal distribution* $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with parameters $\boldsymbol{\mu} \in \mathrm{R}^n$, $\boldsymbol{\Sigma} > 0$:

$\mathrm{E}(X_i) = \mu_i$; $\mathrm{cov}(X_i, X_j) = \sigma_{ij}$; $(X_1, \dots, X_n)$ has n-variate normal distribution, if and only if the random variable $c_1X_1 + \dots + c_nX_n$ has the normal distribution for arbitrary $c_1, \dots, c_n \in \mathrm{R}$)

$$f(x) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times \exp\left(-\frac{1}{2(1-\rho^2)}\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)\right),$$
$$x_1, x_2 \in \mathrm{R}$$

(*bivariate normal distribution* with parameters $\mu_1, \mu_2 \in \mathrm{R}; \sigma_1^2, \sigma_2^2 > 0; \rho \in (-1, 1)$: $\mathrm{E}(X_1) = \mu_1$; $\mathrm{E}(X_2) = \mu_2$; $\mathrm{var}(X_1) = \sigma_1^2$; $\mathrm{var}(X_2) = \sigma_2^2$; $\mathrm{cov}(X_1, X_2) = \sigma_1\sigma_2\rho$)

26.7 Transformation of Random Variables

$$Z = aX + b: \qquad q(z) = \frac{1}{|a|}f\left(\frac{z-b}{a}\right),\quad z \in \mathrm{R},$$

where X and Z have densities $f(x)$ and $q(z)$; $a \neq 0$

$$Z = X^2: \ q(z) = \frac{1}{2\sqrt{z}}\left(f\left(\sqrt{z}\right) + f\left(-\sqrt{z}\right)\right),\ z > 0,$$

where X and Z have densities $f(x)$ and $q(z)$

$$Z = |X|: \ q(z) = f(z) + f(-z),\quad z \geq 0,$$

where X and Z have densities $f(x)$ and $q(z)$

$$Z = \mathrm{e}^X: \ q(z) = \frac{1}{|z|}f(\ln z),\quad z > 0,$$

where X and Z have densities $f(x)$ and $q(z)$

$$Z = 1/X: \ q(z) = \frac{1}{z^2} f\left(\frac{1}{z}\right), \quad z \neq 0,$$

where X and Z have densities $f(x)$ and $q(z)$

$$Z = \sqrt{X}: \ q(z) = 2zf\left(z^2\right), \quad z \geq 0,$$

where non-negative X and Z have densities $f(x)$ and $q(z)$

$$Z = \ln X: \ q(z) = \mathrm{e}^z f\left(\mathrm{e}^z\right), \quad z \in \mathrm{R},$$

where positive X and real Z have densities $f(x)$ and $q(z)$

$$Z = X + Y: \ q(z) = \int_{-\infty}^{\infty} f(x)g(z-x)\,\mathrm{d}x = \int_{-\infty}^{\infty} f(z-y)g(y)\,\mathrm{d}y, \quad z \in \mathrm{R},$$

where X, Y and Z have densities $f(x)$, $g(y)$ and $q(z)$

$$Z = XY: \ q(z) = \int_{-\infty}^{\infty} \frac{1}{|x|} f(x) g\left(\frac{z}{x}\right) \mathrm{d}x = \int_{-\infty}^{\infty} \frac{1}{|y|} f\left(\frac{z}{y}\right) g(y)\,\mathrm{d}y, \quad z \in \mathrm{R},$$

where X, Y and Z have densities $f(x)$, $g(y)$ and $q(z)$

$$Z = X/Y: \ q(z) = \int_{-\infty}^{\infty} |y|\, f(yz)\, g(y)\,\mathrm{d}y, \quad z \in \mathrm{R},$$

where X, Y and Z have densities $f(x)$, $g(y)$ and $q(z)$

26.8 Conditional Mean Value

$$\mathrm{E}\,(X\,|\boldsymbol{Y})$$

(*conditional mean value* of integrable (see Sect. 26.3) random variable X in probability space $(\Omega, \boldsymbol{B}, P)$ *with respect to σ-algebra* $\boldsymbol{Y}$ ($\boldsymbol{Y} \subset \boldsymbol{B}$): $\mathrm{E}\,(X\,|\boldsymbol{Y})$ is integrable random variable in probability space $(\Omega, \boldsymbol{Y}, P)$ fulfilling

$$\int_A E(X\,|\boldsymbol{Y})\,\mathrm{d}P = \int_A X\,\mathrm{d}P \text{ for all } A \in \boldsymbol{Y})$$

$$\mathrm{E}\,(X\,|Y)$$

(*conditional mean value* of random variable X in probability space $(\Omega, \boldsymbol{B}, P)$ *with respect to random variable* Y: $\mathrm{E}\,(X\,|Y)$ is $\mathrm{E}\,(X\,|\boldsymbol{Y})$ for σ-algebra $\boldsymbol{Y} = \sigma(Y)$ generated by random variable, see Sect. 26.3)

$$\mathrm{E}\,(X\,|\boldsymbol{B}) = \mathrm{E}\,(X\,|X) = X \text{ a.s.};\ \mathrm{E}\,(X\,|\,\{\emptyset,\ \Omega\}) = E(X) \text{ a.s.}$$

$$\mathrm{E}\,(\mathrm{E}\,(X\,|\boldsymbol{Y})|\boldsymbol{Z}) = \mathrm{E}(X\,|\boldsymbol{Z}) \text{ a.s., where } \boldsymbol{Z} \subset \boldsymbol{Y}$$

$$\mathrm{E}\,(a_0 + a_1X_1 + a_2X_2\,|Y) = a_0 + a_1\mathrm{E}\,(X_1\,|Y) + a_2\mathrm{E}\,(X_2\,|Y) \text{ a.s.,} \quad a_0, a_1, a_2 \in R$$

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

$$\mathrm{E}\,(g(Y)\cdot X\,|Y) = g(Y)\cdot \mathrm{E}\,(X\,|Y) \text{ a.s.,}$$

where g is a measurable function of random variable Y

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

where X and Y are independent random variables

$$P\,(A\,|Y)$$

(*conditional probability* of random event $A \in \boldsymbol{B}$ in probability space $(\Omega, \boldsymbol{B}, P)$ *with respect to random variable* Y: $P\,(A\,|Y)$ is $\mathrm{E}\,(I_A\,|Y)$ for indicator I_A of event A (i.e. $I_A\,(\omega) = 1$ for $\omega \in A$ and $I_A\,(\omega) = 0$ for $\omega \notin A$))

$$f = \frac{\mathrm{d}Q}{\mathrm{d}P}, \text{ where } Q(A) = \int_A f\ \mathrm{d}P \text{ for arbitrary event } A \in \boldsymbol{B}$$

(*Radon–Nikodym derivative* of Q with respect to P: if a probability measure Q in space $(\Omega, \boldsymbol{B})$ is *dominated* by a probability measure P in the same space $(\Omega, \boldsymbol{B})$ (i.e. $P(A) = 0$ implies $Q(A) = 0$), then there exists a non-negative random variable f in $(\Omega, \boldsymbol{B})$ (uniquely except for a zero probability set), which has the property given *thereinbefore*)

26.9 Martingales

$$\{(X_t, \Im_t),\ t \geq 0\},$$

where random variables X_t and σ-algebras $\Im_t \subset \boldsymbol{B}$ in $(\Omega, \boldsymbol{B}, P)$ fulfil:

(i) $\Im_t \subset \Im_T$ for $0 \leq t \leq T$ (the so-called *filtration*: $\Im_t$ is the total of information up to time t)
(ii) X_t are measurable with respect to $\Im_t$ (i.e. X_t are *adapted* with respect to $\Im_t$)
(iii) X_t are integrable, i.e. $\mathrm{E}(|X_t|) < \infty$
(iv) $\mathrm{E}\,(X_T\,|\Im_t) = X_t$ for $0 \leq t \leq T$ (one writes simply $\mathrm{E}_t(X_T) = X_t$)

(*martingale*: is a stochastic process (see Sect. 30.1), which has properties (*i*)–(*iv*) for given filtration $\{\Im_t\}$ and probability P (one can also define martingale at discrete time as a stochastic sequence); the key property (*iv*) can be rewritten to the form $E_t(X_T - X_t) = 0$, i.e. the expected future change of martingale is zero, so that one cannot expect that martingale will increase or decrease systematically in future (the best prediction of its future value is its current value): this fact is suitable when modeling financial assets; if there is inequality $E_t(X_T) \geq X_t$ (or $E_t(X_T) \leq X_t$) in (*iv*), then one has *submartingale* (or *supermartingale*, respectively); *continuous* martingale has trajectories (see Sect. 30.1), which are continuous at time; martingale can be *continuous from right* only if its trajectories are continuous with exception of jumps that are continuous from the right; *quadratically integrable* martingale has $E({X_t}^2) < \infty$; trajectories of quadratically integrable continuous martingale in an infinite time interval have infinite variations, finite quadratic variations and zero variations of the higher than second order (e.g. the quadratic variation is the limit value of sum of squares of increments of a given trajectory when one refines more and more the partition of the given time interval)

$$\{(E(Y\,|\Im_t), \Im_t),\ t \geq 0\},$$

where Y is integrable random variable in $(\Omega,\ \boldsymbol{B},\ P)$ (example of martingale (in particular, Y can be constant))

$$\{(Y_0 + Y_1 + \ldots + Y_n,\ \sigma(Y_1, \ldots, Y_n)),\ n \in N_0\},$$

where $\{Y_n\}$ is a sequence of independent random variables with zero mean value in $(\Omega, \boldsymbol{B}, P)$ and $Y_0 = 0$; it is called in this context *process of martingale differences* ($\sigma(Y_1, \ldots, Y_n)$ is σ-algebra generated by random variables $Y_1, \ldots, Y_n$, see Sect. 26.3) (example of martingale)

$$\{(|X_t|\,, \Im_t),\ t \geq 0\},$$

where $\{(X_t, \Im_t),\ t \geq 0\}$ is martingale (example of submartingale)

$$\{(W_t,\ \sigma(W_t)),\ t \geq 0\},$$

where $\{W_t\}$ is Wiener process (see Sect. 30.4) (example of martingale)

$$\left\{\left(W_t^2 - t,\ \sigma(W_t)\right),\ t \geq 0\right\},$$

where $\{W_t\}$ is Wiener process (see Sect. 30.4) (example of martingale)

$$\left\{\left(e^{\alpha \cdot W_t - (\alpha^2/2) \cdot t},\ \sigma(W_t)\right),\ t \geq 0\right\},$$

where $\{W_t\}$ is Wiener process (see Sect. 30.4) and $\alpha \in R$ (example of martingale)

$$\{(N_t - \lambda \cdot t,\ \sigma(N_t)),\ t \geq 0\},$$

where $\{N_t\}$ is Poisson process with intensity λ (see Sect. 30.4) (example of martingale continuous from right)

$$X_t = M_t + I_t$$

(submartingale $\{X_t, \Im_t\}$ continuous from right can be decomposed to a martingale $\{M_t, \Im_t\}$ continuous from right and to a process $\{I_t\}$ measurable with respect to $\{\Im_t\}$ with non-decreasing trajectories)

(*Doob–Meyer decomposition*)

26.10 Generating Function

$$g_N(z) = \mathrm{E}(z^N) = \sum_{j=0}^{\infty} P(N = j) \cdot z^j, \quad |z| \leq 1$$

(*generating function* of *discrete* random variable N with

$$p_j = P(N = j); \quad \sum_j p_j = 1, j = 0, 1, \ldots)$$

$$P(N = j) = \frac{1}{j!}\, g_N^{(j)}(0), \quad j = 0, 1, \ldots$$

$$\mathrm{E}N = g_N'(1);\ \mathrm{var}N = g_N''(1) + g_N'(1) - \left(g_N'(1)\right)^2$$

$$\begin{aligned} g_{N_1,\ldots,\,N_n}(z_1, \ldots, z_n) &= \mathrm{E}(z_1^{N_1}\ \ldots\ z_n^{N_n}) \\ &= \sum_{j_1=0}^{\infty} \ldots \sum_{j_n=0}^{\infty} P(N_1 = j_1, \ldots,\ N_n = j_n) \cdot z_1^{j_1}\ \ldots\ .z_n^{j_n}, \quad |z_1|, \ldots,\ |z_n| \leq 1 \end{aligned}$$

(*generating function* of *discrete* random variables $N_1, \ldots, N_n$)

$$P(N_1 = j_1, \ldots,\ N_n = j_n) = \frac{\partial^{\,j_1 + \ldots + j_n}}{\partial z_1^{j_1}\ \ldots\ \partial z_n^{j_n}} g_{N_1,\ldots,\,N_n}(0, \ldots, 0), \quad j_1, \ldots,\ j_n \in N_0$$

$$\mathrm{E}(N_i) = \frac{\partial}{\partial z_i} g_{N_1,\ldots,\,N_n}(1, \ldots, 1), \quad i = 1, \ldots,\ n$$

$$\mathrm{cov}(N_i, N_j) = \frac{\partial^2}{\partial z_i \partial z_j} g_{N_1,\ldots,\,N_n}(1, \ldots, 1)$$

$$- \frac{\partial}{\partial z_i} g_{N_1,\ldots,N_n}(1,\ldots,1) \cdot \frac{\partial}{\partial z_j} g_{N_1,\ldots,N_n}(1,\ldots,1), \quad i,\ j = 1,\ldots,n$$

$$g_{N_1+\ldots+N_n}(z) = \prod_{i=1}^{n} g_{N_i}(z), \quad |z| \le 1,$$

where random variables $N_1,\ldots,N_n$ are independent

$$m_X(z) = \mathrm{E}\left(\mathrm{e}^{z \cdot X}\right), \quad z \in \mathrm{R} \quad (\textit{moment generating function} \text{ of random variable } X)$$

$$\mathrm{E}\left(X^k\right) = m_X^{(k)}(0), \quad k \in N_0$$

$$m_N(z) = g_N(\mathrm{e}^z); \quad g_N(z) = m_N(\ln z),$$

where N is discrete random variable

$$m_X(z) = 1 + z \int_0^{\infty} (1 - F(x))\, \mathrm{e}^{z \cdot x} \mathrm{d}x, \quad z \in \mathrm{R},$$

where X is non–negative random variable with distribution function $F(x)$

$$m_{X_1,\ldots,X_n}(z_1,\ldots,z_n) = \mathrm{E}\left(\mathrm{e}^{z_1 X_1 + \ldots + z_n X_n}\right), \quad z_1,\ldots,z_n \in \mathrm{R}$$

(*moment generating function* of random variables $X_1,\ldots,X_n$)

$$\mathrm{E}(X_1^{k_1} \cdot \ldots \cdot X_n^{k_n}) = \frac{\partial^{k_1+\ldots+k_n}}{\partial z^{k_1} \ldots \partial z^{k_n}} m_{X_1,\ldots,X_n}(0,\ldots,0), \quad k_1,\ldots,k_n \in \mathrm{N}_0$$

$$\mathrm{cov}(X_i, X_j) = \frac{\partial^2}{\partial z_i \partial z_j} \ln m_{X_1,\ldots,X_n}(0,\ldots,0), \quad i,\ j = 1,\ldots,n$$

$$m_{X_1+\ldots+X_n}(z) = \prod_{i=1}^{n} m_{X_i}(z), \quad z \in \mathrm{R},$$

where random variables $X_1,\ldots,X_n$ are independent

26.11 Convolutions and Sums of Random Variables

$$p(j) = (p_1 * p_2)(j) = \sum_{i=0}^{\infty} p_1(j-i)p_2(i) = \sum_{i=0}^{\infty} p_2(j-i)p_1(i), \quad j = 0, 1, \ldots$$

(*convolution of counting densities* p_1 and p_2 of discrete random variables N_1 and N_2, where

$$p_1(j) = P(N_1 = j), \ \sum_j p_1(j) = 1,$$
$$p_2(j) = P(N_2 = j), \ \sum_j p_2(j) = 1, \ j = 0, 1, \ldots :$$

is the counting density of *sum* of *independent* random variables $N_1 + N_2$)

$$F(x) = (F_1 * F_2)(x) = \int_{-\infty}^{\infty} F_1(x-y)\,\mathrm{d}F_2(y) = \int_{-\infty}^{\infty} F_2(x-y)\,\mathrm{d}F_1(y), \quad x \in \mathrm{R}$$

(*convolution of distribution functions* F_1 and F_2 of random variables X_1 and X_2: is the distribution function of *sum* of *independent* random variables $X_1 + X_2$)

$$f(x) = (f_1 * f_2)(x) = \int_{-\infty}^{\infty} f_1(x-y)\,\mathrm{d}f_2(y) = \int_{-\infty}^{\infty} f_2(x-y)\,\mathrm{d}f_1(y), \quad x \in \mathrm{R}$$

(*convolution of probability densities* f_1 and f_2 of random variables X_1 and X_2: is the probability density of *sum* of *independent* random variables $X_1 + X_2$)

$p^{*n} = p^{*(n-1)} * p; \ F^{*n} = F^{*(n-1)} * F; \quad f^{*n} = f^{*(n-1)} * f, \ \ n = 2, 3, \ldots$
(*convolution powers*: $p^{*1} = p, \ \ F^{*1} = F, \ \ f^{*1} = f$)

$$p^{*n}(0) = (p(0))^n,$$
$$p^{*n}(j) = \frac{1}{j \cdot p(0)} \sum_{i=1}^{j} ((n+1) \cdot i - j) \cdot p(i) \cdot p^{*n}(j-i), \ j = 1, 2, \ldots,$$

where $p(0) > 0$

$$B(n_1, p) * B(n_2, p) = B(n_1 + n_2, p), \quad n_1, \ n_2 \in N; \ p \in (0, 1)$$

$$P(\lambda_1) * P(\lambda_2) = P(\lambda_1 + \lambda_2), \quad \lambda_1, \ \lambda_2 > 0$$

$$N(\mu_1, \sigma_1^2) * N(\mu_2, \sigma_2^2) = N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2), \quad \mu_1, \ \mu_2 \in \mathrm{R}; \quad \sigma_1^2, \ \sigma_2^2 > 0$$

$$\Gamma(a, \lambda_1) * \Gamma(a, \lambda_2) = \Gamma(a, \lambda_1 + \lambda_2), \quad a, \ \lambda_1, \ \lambda_2 > 0$$

$$(\mathrm{Exp}(\lambda))^{*n} = \Gamma(n, \lambda), \quad n \in N; \quad \lambda > 0$$

26.12 Random Sums of Random Variables

$$F_S(x) = \sum_{j=0}^{\infty} P(N = j) \cdot F_X^{*j}(x), \quad x \in \mathrm{R},$$

where $S = X_1 + \ldots + X_N$ is a sum of random number N of independent and identically distributed (*iid*) random variables $X_1, \ldots, X_N$, which have distribution function F_X and are independent of N

$$\mathrm{E}(S) = \mathrm{E}(N) \cdot \mathrm{E}(X), \quad \mathrm{var}(S) = \mathrm{E}(N) \cdot \mathrm{var}(X) + \mathrm{var}(N) \cdot (\mathrm{E}(X))^2 ,$$

where $S = X_1 + \ldots + X_N$ is a sum of random number N of independent and identically distributed random variables $X_1, \ldots, X_N$, which have moments $E(X)$ and $\mathrm{var}(X)$ and are independent of N

$$S \sim CP(\lambda, F_X),$$

where $S = X_1 + \ldots + X_N$ is a sum of random number N of independent and identically distributed random variables $X_1, \ldots, X_N$, which have distribution function $F_X(x)$ and are independent of $N \sim P(\lambda)$
(*compound Poisson distribution* with intensity $\lambda > 0$:

$$F_S = \sum_{j=0}^{\infty} \mathrm{e}^{-\lambda} \tfrac{\lambda^j}{j!} F_X^{*j}(x) \text{ (see Sect. 26.11)}; \ \mathrm{E}(S) = \lambda \cdot \mathrm{E}(X); \mathrm{var}(X) = \lambda \cdot \mathrm{E}\left(X^2\right);$$
$$\gamma_1 = \frac{\mathrm{E}\left(X^3\right)}{\left(\lambda \cdot \left(\mathrm{E}\left(X^2\right)\right)^3\right)^{1/2}}; \quad \gamma_2 = \frac{\mathrm{E}\left(X^4\right)}{\lambda \cdot \left(\mathrm{E}\left(X^2\right)\right)^2}$$

26.13 Some Inequalities

$$P\left(X > \lambda \cdot \mathrm{E}(X)\right) \leq 1/\lambda,$$

where X is positive; $\mathrm{E}(X) < \infty$; $\lambda > 1$ (*Markov Inequality*)

$$P\left(|X - \mathrm{E}(X)| \geq \varepsilon\right) \leq \frac{\mathrm{var}(X)}{\varepsilon^2},$$

where $\mathrm{E}(X^2) < \infty$; $\varepsilon > 0$ (*Chebyshev Inequality*)

$$P\left(X - \mathrm{E}(X) \geq \varepsilon\right) \leq \frac{\mathrm{var}(X)}{\mathrm{var}(X) + \varepsilon^2},$$

where $\mathrm{E}(X^2) < \infty$; $\varepsilon > 0$ (*one-sided Chebyshev Inequality*)

$$\mathrm{E}\left(g(X)\right) \geq g\left(\mathrm{E}(X)\right),$$

where $\mathrm{E}(|X|) < \infty$; g is convex (see Sect. 22.3) function

(*Jensen Inequality*)

$$\mathrm{E}\left(g(X)\right) \leq g\left(\mathrm{E}(X)\right),$$

where $\mathrm{E}(|X|) < \infty$; g is concave function

(*Jensen Inequality*)

$$P\left(\max_{1 \leq k \leq n} \left| \sum_{i=1}^{k} (X_i - \mathrm{E}(X_i)) \right| \geq \varepsilon\right) \leq \frac{\sum\limits_{i=1}^{n} \mathrm{var}(X_i)}{\varepsilon^2},$$

where $X_1, \ldots, X_n$ are independent; $\mathrm{E}({X_i}^2) < \infty$; $\varepsilon > 0$ (*Kolmogorov Inequality*)

$$|F_n(x) - \Phi(x)| \leq A \ \frac{\mathrm{E}\left(|X_1 - \mu|^3\right)}{\sigma^3 \sqrt{n}}, \quad x \in \mathrm{R},$$

where $X_1, \ldots, X_n$ are independent and identically distributed; $\mathrm{E}(X_i) = \mu$; $\mathrm{var}(X_i) = \sigma^2$; $\mathrm{E}(|{X_1}^3|) < \infty$; $F_n(x)$ is distribution function of $\sum_{i=1}^{n} (X_i - \mu)/(\sigma\sqrt{n})$; A is a constant independent of x

(*Berry–Essén Inequality*)

26.14 Limit Theorems of Probability Theory

$$\lim_{n \to \infty} P\left(\, |X_n - X| \geq \varepsilon\right) = 0,$$

where $\varepsilon > 0$ is arbitrary (*convergence in probability* $X_n \xrightarrow{P} X$)

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - p\right| \geq \varepsilon\right) = 0,$$

where independent X_i's have distribution $B(1, p)$, and $\varepsilon > 0$ is arbitrary

(*Weak Law of Large Numbers*: *Bernoulli Theorem*)

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \varepsilon\right) = 0,$$

where independent X_i's are identically distributed with mean value μ, and $\varepsilon > 0$ is arbitrary

(*Weak Law of Large Numbers*: *Khintchine Theorem*)

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}(X_i)\right| \geq \varepsilon\right) = 0,$$

where independent X_i's fulfil $\lim_{n\to\infty} \frac{1}{n}\left(\sum_{i=1}^{n} \mathrm{var}(X_i)\right)^{1/2} = 0$, and $\varepsilon > 0$ is arbitrary

(*Weak Law of Large Numbers*: *Markov Theorem*)

$$P\left(\lim_{n\to\infty} X_n = X\right) = 1$$

(*convergence almost surely* $X_n \to X$ a.s. (or *with probability one*))

$$P\left(\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i = \mu\right) = 1,$$

where independent X_i's are identically distributed with mean value μ

(*Strong Law of Large Numbers*: *Kolmogorov Theorem*)

$$\lim_{n\to\infty} F_n(x) = F(x)$$

(*convergence in distribution* (*weak convergence*) $L(X_n) \to L(X)$: distribution function F_n of random variables X_n converges to distribution function F of random variable X at every point x of continuity of the function F)

$$\lim_{n\to\infty}\int_{-\infty}^{\infty} f(x)\,\mathrm{d}F_n(x) = \int_{-\infty}^{\infty} f(x)\,\mathrm{d}F(x)$$

for arbitrary real continuous bounded function f

(*convergence in distribution* (*weak convergence*) $L(X_n) \to L(X)$: an equivalent definition)

$$L\left(\frac{1}{\sqrt{np(1-p)}}\left(\sum_{i=1}^{n} X_i - np\right)\right) \to N(0,\ 1),$$

where independent X_i's have distribution $B(1, p)$

(*Central Limit Theorem*: *Moivre–Laplace Theorem*)

$$L\left(\frac{1}{\sigma\sqrt{n}}\left(\sum_{i=1}^{n} X_i - n\mu\right)\right) \to N(0,\ 1),$$

where independent X_i's are identically distributed with mean value μ and variance σ^2

(*Central Limit Theorem*: *Lévy–Lindeberg Theorem*)

$$L\left(\frac{1}{\left(\sum_{i=1}^{n}\operatorname{var}(X_i)\right)^{1/2}}\left(\sum_{i=1}^{n} X_i - \sum_{i=1}^{n}\mathrm{E}(X_i)\right)\right) \to N(0,\ 1),$$

where independent X_i's are identically distributed with mean value μ and variance σ^2 and fulfil $\lim\limits_{n\to\infty}\sum\limits_{i=1}^{n}\mathrm{E}\left(|X_i-\mathrm{E}(X_i)|^k\right)\Big/\left(\sum\limits_{i=1}^{n}\operatorname{var}(X_i)\right)^{k/2} = 0$ for a real $k > 2$

(*Central Limit Theorem*: *Lyapunov Theorem*)

$$\lim_{n\to\infty}\mathrm{E}\left((X_n - X)^2\right) = 0$$

(*convergence in the mean square* $X_n \xrightarrow{L_2} X$ for $\mathrm{E}\left(X_n^2\right) < \infty$; $n \in \mathrm{N}$)

$$X_n \xrightarrow{L_2} X \Rightarrow X_n \xrightarrow{P} X \Rightarrow L\left(X_n\right) \to L\left(X\right), \qquad X_n \to X \text{ a.s.} \Rightarrow X_n \xrightarrow{P} X$$

Further Reading

Daykin, C.D., Pentikäinen, T., Pesonen, M.: Practical Risk Theory for Actuaries. Chapman and Hall, London (1994)

Feller, W.: An Introduction to Probability Theory and Its Applications. Wiley, New York (1968)

Heilmann, W.-R.: Fundamentals of Risk Theory. Verlag Versicherungswirtschaft, Karlsruhe (1988)

Johnson, N.L., Kotz, S.: Distributions in Statistics. Discrete Distributions. Wiley, New York (1969)

Johnson, N.L., Kotz, S.: Distributions in Statistics. Continuous Univariate Distributions. Wiley, New York (1970)

Johnson, N.L., Kotz, S.: Distributions in Statistics. Multivariate Distributions. Wiley, New York (1972)

Malliaris, A.G., Brock, W.A.: Stochastic Methods in Economics and Finance. North-Holland, Amsterdam (1982)

Panjer, H.H., Willmot, G.E.: Insurance Risk Models. Society of Actuaries, Schaumburg (1992)

Rektorys, K. et al.: Survey of Applicable Mathematics. Kluwer, Dordrecht (1994)

Chapter 27
Descriptive and Mathematical Statistics

Abstract Chapter 27 deals with basic theory and practical methods of statistical inference: 27.1. Sampling Theory: Simple Random Sample, 27.2. Sampling Theory: Stratified Random Sample, 27.3. Elementary Statistical Treatment, 27.4. Sample Quantiles, 27.5. Measures of Sample Level, 27.6. Measures of Sample Variability, 27.7. Measures of Sample Concentration, 27.8. Measures of Sample Dependence, 27.9. Point and Interval Estimators, 27.10. Hypothesis Testing, 27.11. Regression Analysis, 27.12. Analysis of Variance (ANOVA), 27.13. Multivariate Statistical Analysis.

27.1 Sampling Theory: Simple Random Sample

$$P = 1/\binom{N}{n}$$

(probability that a *sample* S of size $n \in \mathrm{N}$ will be selected from a *population* of size $N \in \mathrm{N}$ ($n \leq N$), when one applies *simple random sampling* (it means that each individual (or statistical unit) has the same probability of being chosen at any stage during the sampling process); simple random sample can be obtained by random sampling *without replacement* when an individual is chosen from the population of N individuals randomly with the same probability $1/N$, then another individual is chosen from the population of $N - 1$ remaining individuals with the same probability$1/(N - 1)$, etc.; in practice, samples are collected usually, when one observes a statistical attribute y_i of individuals in a given population ($i = 1, \ldots, N$))

$$p = n/N$$

(probability that a given individual of the population is selected as an element of the sample)

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_27,

$$\bar{Y} = \frac{1}{N} \sum_{i=1}^{N} y_i \quad (\textit{population mean})$$

$$\bar{y} = \frac{1}{n} \sum_{i \in S} y_i \quad (\textit{sample mean}: \ \text{is unbiased estimate of population mean } \bar{Y})$$

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (y_i - \bar{Y})^2 \quad (\textit{population variance})$$

$$s^2 = \frac{1}{n} \sum_{i \in S} (y_i - \bar{y})^2 \quad (\textit{sample variance})$$

$$\hat{s}^2 = \frac{n(N-1)}{N(n-1)} s^2 \quad \text{is unbiased estimate of population variance } \sigma^2$$

$$\bar{y} \mp \sqrt{\frac{N-n}{n \cdot (N-1)}} \cdot \hat{s} \cdot u_{\alpha/2}$$

($(1 - \alpha)100\%$ *confidence interval* for population mean $\bar{Y}$, where $u_{\alpha/2}$ is the quantile of normal distribution for $\alpha \in (0, 1)$: it holds asymptotically under general assumptions)

- *Simple* random sampling: see *thereinbefore*
- Random sampling *with replacement*: when the same individual is chosen again then the corresponding draw is repeated (or the whole sampling process may be repeated from beginning)
- *Systematic* random sampling: involves selection of individuals from an ordered sampling frame, in which every kth element is selected
- *Stratified* random sampling (see Sect. 27.2): stratification is the process of grouping members of the population into relatively homogeneous subgroups (strata) before sampling; the random sampling is then applied within each stratum

27.2 Sampling Theory: Stratified Random Sample

$$\bar{Y} = \frac{1}{N} \sum_{i=1}^{L} N_i \bar{Y}_i$$

(*population mean*, where population of $N \in \mathrm{N}$ individuals is divided into L disjoint subgroups (*strata*) of sizes N_i ($N_1 + \ldots + N_L = 1$) and $\bar{Y}_i$ is the mean in the ith strata)

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{L} n_i \bar{y}_i$$

(*sample mean*, where $\bar{y}_i$ is the sample mean in the ith strata corresponding here to a sample of size $n_i(n_1 + \ldots + n_L = n)$: is unbiased estimate of population mean $\bar{Y}$ only for *representative* (*quota*) *sample* with $n_i/n = N_i/N$, $i = 1, \ldots, L$)

$$\hat{y} = \frac{1}{N} \sum_{i=1}^{L} N_i \bar{y}_i$$

(unbiased estimate of population mean $\bar{Y}$; it has variance

$\mathrm{var}(\hat{y}) = \dfrac{1}{N^2} \displaystyle\sum_{i=1}^{L} N_i^2 \cdot \dfrac{N_i - n_i}{N_i - 1} \cdot \dfrac{\sigma_i^2}{n_i}$,

where $\sigma_i^2 = \dfrac{1}{N_i} \displaystyle\sum_{j=1}^{N_i} \left(y_{i_j} - \bar{y}_i\right)^2$ is variance in the ith strata)

- Two-stage sample: it proceeds as a selection of strata in the first stage and a sampling of individuals in each strata in the second stage

27.3 Elementary Statistical Treatment

$$x_1, x_2, \ldots, x_n$$

(sample *observations* of a statistical attribute $x_i \in \mathrm{R}$, which originate as realization of a random sampling of size $n \in \mathrm{N}$ (in practice they are obtained as results of measurements or experiments); some values among $x_1, x_2, \ldots, x_n$ can repeat several times, e.g. $x_2 = x_5$)

$$x_{(1)} \le x_{(2)} \le \ldots \le x_{(n)} \quad (\textit{ordered}\ \text{sample})$$

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n) \quad (\textit{bivariate}\ \text{sample observations})$$

$$p_j = n_j / n$$

(*relative frequencies* ($j = 1, \ldots, k$, $\sum_{j=1}^{k} n_j = n$): they apply when
– particular values repeat among observations (n_j is the number of value x_j)
– or one sorts the observation to k *class intervals* (*cells*) where the jth class approximated by the value x_j contains just n_j observations;

they may presented in the form of *frequency tables*:

	Frequency		Cumulative frequency	
Values x_j	Absolute	Relative	Absolute	Relative
x_1	n_1	p_1	n_1	p_1
x_2	n_2	p_2	$n_1 + n_2$	$p_1 + p_2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_k	n_k	p_k	$n_1 + n_2 + \ldots + n_k$	$p_1 + p_2 + \ldots + p_k$
$\sum$	n	1		

or graphically by means of various graphical instruments: *histograms*, *graphs* (*area*, *bar*, *column*, *pie*), *boxplots* (presenting minimal value, lower quartile, arithmetic mean, median, upper quartile and maximal value of given observations, see Sects. 27.4 and 27.5), *STEM-and-LEAF*, *correlation tables* (for multivariate data) and others)

27.4 Sample Quantiles

$$\tilde{x}_p, \quad 0 < p < 1$$

(*p-quantile* of sample observations $x_1, x_2, \ldots, x_n$: separates $100p\%$ of low values in ordered sample from $100(1-p)\%$ of high values, see Sect. 26.3)

$\tilde{x}_{0,5}$ (*median*: “sample 50% quantile”, i.e. “middle value of sample”, see Sect. 26.3)

$\tilde{x}_{0,25}$ (*lower quartile*, see Sect. 26.3); $\tilde{x}_{0,75}$ (*upper quartile*, see Sect. 26.3)

$\tilde{x}_{k/10}$, $k = 1, \ldots, 9$ (kth *decile*, see Sect. 26.3); $\tilde{x}_{k/100}$, $k = 1, \ldots, 99$ (kth *percentile*, see Sect. 26.3)

$$F_n(x) = \begin{cases} 0 & \text{for } x < x_{(1)} \\ i/n & \text{for } x_{(i)} \le x < x_{(i+1)}, \end{cases} \quad x \in \mathbf{R};\ x_{(1)} \le x_{(2)} \le \ldots \le x_{(n)} \text{ is ordered sample}$$

$$\quad\quad \begin{cases} 1 & \text{for } x \ge x_{(n)} \end{cases}$$

(*empirical distribution function*: if n increases, then $F_n(x)$ converges with probability one uniformly over all real x to the theoretical distribution function of the probability distribution, from which the observed sample has been obtained (*Glivenko theorem*))

27.5 Measures of Sample Level

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \quad ((\textit{arithmetic} \ \text{or} \ \textit{sample}) \ \textit{mean})$$

$$\bar{x}_G = \sqrt[n]{x_1 \cdot \ldots \cdot x_n}, \quad x_1, \ldots, x_n > 0 \ \ (\textit{geometric mean})$$

$$\bar{x}_H = \frac{n}{1/x_1 + \ldots + 1/x_n}, \quad x_1, \ldots, x_n > 0 \quad (\textit{harmonic mean})$$

$$\bar{x}_Q = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2} \quad (\textit{quadratic mean})$$

$$\bar{x} = \frac{1}{n}\sum_{j=1}^{k} x_j n_j; \quad \bar{x}_G = \sqrt[n]{x_1^{n_1} \cdot \ldots \cdot x_k^{n_k}};$$

$$\bar{x}_H = \frac{n}{n_1/x_1 + \ldots + n_k/x_n}; \quad \bar{x}_K = \sqrt{\frac{1}{n}\sum_{j=1}^{n} x_j^2 n_j},$$

where n_j is the number of the value x_j in the sample, $j = 1, \ldots, k$; $\sum_{j=1}^{k} n_j = n$

$$\bar{x} \approx \frac{1}{n}\sum_{j=1}^{k} a_j n_j; \quad \bar{x}_G \approx \sqrt[n]{a_1^{n_1} \cdot \ldots \cdot a_k^{n_k}};$$

$$\bar{x}_H \approx \frac{n}{n_1/a_1 + \ldots + n_k/a_n}; \quad \bar{x}_K \approx \sqrt{\frac{1}{n}\sum_{j=1}^{n} a_j^2 n_j},$$

where the sample is grouped to k classes (intervals) approximated by values x_j, $j = 1, \ldots, k$; $\sum_{j=1}^{k} n_j = n$

$$\bar{x} = \frac{1}{n}\sum_{j=1}^{k} \bar{x}_j n_j,$$

where the sample is grouped to k classes represented by arithmetic means of particular classes (*class means*) $\bar{x}_j, j = 1, \ldots, k$; $\sum_{j=1}^{k} n_j = n$

$$x_{(1)} \le \bar{x}_H \le \bar{x}_G \le \bar{x} \le \bar{x}_K \le \bar{x}_{(n)}$$

$$\mathbf{\bar{x}} = (\bar{x}_1, \ldots, \bar{x}_m) = \left(\frac{1}{n}\sum_{i=1}^{n} x_{i1}, \ldots, \frac{1}{n}\sum_{i=1}^{n} x_{im}\right) = \frac{1}{n}\mathbf{e}'\,\mathbf{X}$$

(mean of *m-variate* sample $(x_{11}, \ldots, x_{1m}), \ldots, (x_{n1}, \ldots, x_{nm})$, which is ordered to matrix $\mathbf{X}$ of type $n \times m$)

$$\tilde{x}_{0,5} = \begin{cases} x_{(m)} & \text{for } n = 2m - 1 \\ \frac{1}{2}\left(x_{(m)} + x_{(m+1)}\right) & \text{for } n = 2m \end{cases} \quad (\textit{median}\text{, see Sects. 26.3 and 27.4})$$

$$\hat{x}$$

(*mode*: is the "most frequent" value of the given sample, i.e. the value with the highest relative frequency, see Sect. 26.3)

27.6 Measures of Sample Variability

$$s_x^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \frac{1}{n^2} \left(\sum_{i=1}^{n} x_i \right)^2$$
$$= \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \bar{x}^2 \text{ ((\textit{sample}) \textit{variance}, see Sect.26.3)}$$

$$s_x^2 = \frac{1}{n} \sum_{j=1}^{k} (x_j - \bar{x})^2 n_j$$
$$= \frac{1}{n} \sum_{j=1}^{k} x_j^2 n_j - \frac{1}{n^2} \left(\sum_{j=1}^{k} x_j n_j \right)^2,$$

where n_j is the number of the value x_j in the sample, $j = 1, \ldots, k$, $\sum_{j=1}^{k} n_j = n$

$$s_x^2 = \frac{1}{n} \sum_{j=1}^{k} (\bar{x}_j - \bar{x})^2 n_j + \frac{1}{n} \sum_{j=1}^{k} s_{xj}^2 n_j,$$

where the sample is grouped to k classes, the first summand is the variance of class means $\bar{x}_j$ (*among-means variability*) and the second summand is the mean of *class variances* s_{xj}^2 (*inside-class variability*), $j = 1, \ldots, k$; $\sum_{j=1}^{k} n_j = n$

$$s_x = \sqrt{s_x^2} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2} \quad \text{(\textit{standard deviation}, see Sect. 26.3)}$$

$$V_x = \frac{s_x}{|\,\bar{x}\,|} \quad \text{(\textit{coefficient of variation}, see Sect. 26.3)}$$

$$R = x_{\max} - x_{\min} = x_{(n)} - x_{(1)} \quad (\textit{range})$$

$$\tilde{x}_{0,75} - \tilde{x}_{0,25} \quad (\textit{interquartile range}, \text{ see Sect. 26.3});$$

$$\tilde{x}_{0,90} - \tilde{x}_{0,10} \quad (\textit{interdecile range}, \text{ see Sect. 26.3})$$

$$\frac{\tilde{x}_{0,75} - \tilde{x}_{0,25}}{2} \quad (\textit{quartile deviation}); \qquad \frac{\tilde{x}_{0,90} - \tilde{x}_{0,10}}{2} \quad (\textit{decile deviation})$$

$$M = \frac{n^2 - \sum_{j=1}^{k} n_j^2}{n(n-1)} \cdot 100\%$$

(*mutability*: variability of a discrete variable with k possible states (e.g. public inquiry answers) around its mode, where n_j is the observed absolute frequency of $j = 1, \ldots, k; \sum_{j=1}^{k} n_j = n$)

27.7 Measures of Sample Concentration

$$\hat{\gamma}_1 = \frac{1}{n} \cdot \frac{\sum_{i=1}^{n} (x_i - \bar{x})^3}{s_x^3}$$

(*skewness*: describes "distinction in concentration of higher and lower values", see Sect. 26.3; the positive skewness $\gamma_1 > 0$ means a higher concentration of lower values in the sample, i.e. a longer right-hand tail of the probability density (vice versa for the negative skewness $\gamma_1 < 0$))

$$\hat{\gamma}_1 = \frac{n_L - n_U}{n}$$

(*simplified skewness*: n_L is the number of observations in the sample that are smaller than the sample mean $\bar{x}$; n_U is the number of observations that are larger than $\bar{x}$)

$$\hat{\gamma}_2 = \frac{1}{n} \cdot \frac{\sum_{i=1}^{n} (x_i - \bar{x})^4}{s_x^4} - 3$$

(*kurtosis*: describes "distinction in concentration of inner and outer values", see Sect. 26.3; the positive kurtosis $\gamma_2 > 0$ means a higher concentration of inner values in comparison with outer values, i.e. a peaked shape of the probability density; vice versa, the negative kurtosis $\gamma_2 < 0$ means comparable concentrations of inner and outer values, i.e. a flat shape of the probability density)

27.8 Measures of Sample Dependence

$$s_{xy} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

((*sample*) *covariance* of bivariate sample observations $(x_1, y_1), \ldots, (x_n, y_n)$, see Sect. 26.6)

$$r_{xy} = \frac{s_{xy}}{s_x s_y} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}}$$

$$= \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n} \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\left(\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2\right)\left(\sum_{i=1}^{n} y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2\right)}}$$

(*correlation coefficient*: is a measure of linear dependence between two statistical attributes, $-1 \le r_{xy} \le 1$; $r_{xx} = 1$, see Sect. 26.6)

$r_{xy} = 0$ (or equivalently $s_{xy} = 0$) (*uncorrelated* statistical attributes, see Sect. 26.6)

$$\boldsymbol{\Sigma}_{\mathbf{xx}} = \left(s_{x_i x_j}\right)_{\substack{i=1,\ldots,m \\ j=1,\ldots,m}} = \frac{1}{n}\mathbf{X}'\,\mathbf{X} - \bar{x}'\,\bar{x}$$

(*covariance matrix* of m-variate sample $(x_{11}, \ldots, x_{1m}), \ldots, (x_{n1}, \ldots, x_{nm})$, which is ordered to matrix $\mathbf{X}$ of type $n \times m$, see Sects. 26.6 and 27.5)

$$\mathbf{R}_{\mathbf{xx}} = \left(r_{x_i x_j}\right)_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}$$

(*correlation matrix*: has unities in the principal diagonal $r_{x_i x_i} = 1$, see Sect. 26.6)

$$r_{y\mathbf{x}} = \sqrt{\mathbf{R}_{y\mathbf{x}} \mathbf{R}_{\mathbf{xx}}^{-1} \mathbf{R}'_{y\mathbf{x}}},$$

where

$$\mathbf{R}_{\mathbf{xx}} = \left(r_{x_i x_j}\right)_{\substack{i=1,\ldots,m \\ j=1,\ldots,m}}, \quad \mathbf{R}_{y\mathbf{x}} = \left(r_{y x_i}\right)_{i=1,\ldots,m}$$

(*multiple correlation coefficient* between univariate sample $y_1, \ldots, y_n$ and m-variate sample $(x_{11}, \ldots, x_{1m}), \ldots, (x_{n1}, \ldots, x_{nm})$: is a measure of linear dependence between a statistical attribute and a group of statistical attributes)

$$r_{yz.\mathbf{x}} = \frac{r_{yz} - \mathbf{R}_{y\mathbf{x}}\mathbf{R}_{\mathbf{xx}}^{-1}\mathbf{R}'_{z\mathbf{x}}}{\sqrt{1 - \mathbf{R}_{y\mathbf{x}}\mathbf{R}_{\mathbf{xx}}^{-1}\mathbf{R}'_{y\mathbf{x}}}\sqrt{1 - \mathbf{R}_{z\mathbf{x}}\mathbf{R}_{\mathbf{xx}}^{-1}\mathbf{R}'_{z\mathbf{x}}}},$$

where

$$\mathbf{R}_{\mathbf{xx}} = \left(r_{x_i x_j}\right)_{\substack{i=1,\dots,m\\ j=1,\dots,m}}; \quad \mathbf{R}_{y\mathbf{x}} = \left(r_{y x_i}\right)_{i=1,\dots,m}; \quad \mathbf{R}_{z\mathbf{x}} = \left(r_{z x_i}\right)_{i=1,\dots,m}$$

(*partial correlation coefficient* between samples $y_1, \dots, y_n$ and $z_1, \dots, z_n$ given fixed values of m-variate sample $(x_{11}, \dots, x_{1m}), \dots, (x_{n1}, \dots, x_{nm})$: is a measure of linear dependence between two statistical attributes when fixing a given group of other statistical attributes so that they cannot affect this dependence)

27.9 Point and Interval Estimators

$$X_1, X_2, \dots, X_n$$

(*random sample* of size $n \in \mathrm{N}$ from a given probability distribution: is a sequence of independent random variables, which have this distribution; when observing a random sample one obtains the (statistical) sample $x_1, x_2, \dots, x_n$, see Sect. 27.3; the random sample can be *multivariate* one, i.e. a sequence of independent random vectors, see Sect. 26.6)

$$\hat{\theta} = T(X_1, \dots, X_n)$$

(*point estimator* of parameter $\theta \in \mathrm{R}$: is a function called *statistics*, whose probability distribution (affected by θ) concentrates maximally in a given sense around the unknown parameter value θ; the estimated parameter can be multivariate one $\theta \in \mathrm{R}^m$)

$$\hat{\theta} = T(x_1, \dots, x_n)$$

(*point estimate* of parameter $\theta \in \mathrm{R}$ for observed sample $x_1, x_2, \dots, x_n$: is a numerical value in contrast to the random variable $T(X_1, \dots, X_n)$ (see *thereinbefore*))

$$\mathrm{E}\,(T(X_1, \dots, X_n)) = \theta$$

(*unbiased estimator*: if this equality holds only asymptotically for increasing size n of random sample, then the estimator is *asymptotically unbiased*; in the multivariate case, the given equality must hold for all components of the vector $\theta \in \mathrm{R}^m$)

$$\mathrm{var}\,(T(X_1, \dots, X_n)) \leq \mathrm{var}\,(S(X_1, \dots, X_n))$$

(*best unbiased estimator*: is such an unbiased estimator $T(X_1, \dots, X_n)$ of parameter θ that the given inequality holds for all the unbiased estimators $S(X_1, \dots, X_n)$

of θ; in the multivariate case $\theta \in \mathrm{R}^m$, the difference of covariance matrices $\operatorname{var}(S(X_1, \ldots, X_n))$–$\operatorname{var}(T(X_1, \ldots, X_n))$ must be always positive semidefinite)

$$\operatorname{var}(T(X_1, \ldots, X_n)) = \frac{1}{n \cdot I(\theta)}$$

(*efficient estimator* (the fraction on the right-hand side with *Fisher's measure of information* $I(\theta)$ is the so-called *Cramer-Rao lower bound* of variances of unbiased estimators): is the unbiased estimator of θ with minimal variance; if the lower bound is achieved only in limit for increasing size n of the sample, then the corresponding estimator is called *asymptotically efficient*; the definition can be generalized again for the multivariate case $\theta \in \mathrm{R}^m$ using *Fisher's matrix of information*)

$$T(X_1, \ldots, X_n) \xrightarrow{P} \theta$$

(*consistent estimator*: achieves the unknown parameter value θ as the limit in probability (see Sect. 26.14) with increasing size n of the random sample)

$$\prod_{i=1}^{n} f(x_i; \hat{\theta}) \geq \prod_{i=1}^{n} f(x_i; \theta) \text{ for all admissible } \theta$$

(*maximum likelihood estimation* (f is the probability density of the distribution, from which the corresponding sample has been obtained): is the most probable parameter value for the sample observations $x_1, x_2, \ldots, x_n$; the maximum likelihood estimator is (under general assumptions) asymptotically efficient (hence also asymptotically unbiased, see *thereinbefore*))

$$\hat{p} = \frac{\bar{x}}{N}$$

(efficient estimate of $p \in (0, 1)$ for a known $N \in \mathrm{N}$ in distribution $B(N, p)$, see Sect. 26.4)

$$\hat{\mu} = \bar{x}; \quad \hat{\sigma}^2 = \frac{n}{n-1} s_x^2$$

(best unbiased estimate of $\mu \in \mathrm{R}$ and $\sigma^2 > 0$ in distribution $N(\mu, \sigma^2)$, see Sect. 26.5)

$$\hat{\mu} = \bar{x}, \quad \hat{\sigma}^2 = s_x^2$$

(maximum likelihood estimate of $\mu \in \mathrm{R}$ and $\sigma^2 > 0$ in distribution $N(\mu, \sigma^2)$, see Sect. 26.5)

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} \ln x_i, \quad \hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\ln x_i - \hat{\mu})^2$$

(best unbiased estimate of $\mu \in$ R and $\sigma^2 > 0$ in distribution $LN(\mu, \sigma^2)$, see Sect. 26.5)

$$\hat{a} = \frac{n}{n-1}x_{(1)} - \frac{1}{n-1}x_{(n)}, \quad \hat{b} = \frac{n}{n-1}x_{(n)} - \frac{1}{n-1}x_{(1)}$$

(best unbiased estimate of $a, b \in$ R ($a < b$) in uniform distribution, see Sect. 26.5)

$\hat{\lambda} = \frac{1}{\bar{x}}$ (best unbiased estimate of $\lambda > 0$ in distribution Exp (λ) see Sect. 26.5)

$$\hat{\lambda} = \frac{a}{\bar{x}}$$

(efficient estimate of $\lambda > 0$ for a known $a > 0$ in distribution $\Gamma(a, \lambda)$, see Sect. 26.5)

$$T_d(X_1, \ldots, X_n) < \theta < T_h(X_1, \ldots, X_n)$$

($100(1 - \alpha)\%$ *two-sided confidence interval* (*interval estimator*) of parameter $\theta \in$ R: is the interval, in which θ lies with probability $(1 - \alpha)$ $(0 < \alpha < 1)$, i.e. $P(T_d(X_1, \ldots, X_n) < \theta < T_h(X_1, \ldots, X_n)) = 1 - \alpha$)

$$-\infty < \theta < T_h(X_1, \ldots, X_n) \quad \text{or} \quad T_d(X_1, \ldots, X_n) < \theta < \infty$$

($100(1 - \alpha)\%$ *one-sided confidence intervals* of parameter $\theta \in$ R: are one-sided intervals, in which θ lies with probability $(1 - \alpha)$ $(0 < \alpha < 1)$, i.e. $P(-\infty < \theta < T_h(X_1, \ldots, X_n)) = 1 - \alpha$ or $P(T_d(X_1, \ldots, X_n) < \theta < \infty) = 1 - \alpha$)

$$T_d(x_1, \ldots, x_n) < \theta < T_h(x_1, \ldots, x_n);$$
$$-\infty < \theta < T_h(x_1, \ldots, x_n); T_d(x_1, \ldots, x_n) < \theta < \infty$$

(confidence intervals (interval estimation) of parameter $\theta \in$ R for sample observations $x_1, x_2, \ldots, x_n$: are numerical intervals)

$$\bar{x} - t_{1-\alpha/2}(n-1) \cdot s_x/\sqrt{n-1} < \mu < \bar{x} + t_{1-\alpha/2}(n-1) \cdot s_x/\sqrt{n-1}$$

$$n \cdot s_x^2 \Big/ \chi^2_{1-\alpha/2}(n-1) < \sigma^2 < n \cdot s_x^2 \Big/ \chi^2_{\alpha/2}(n-1)$$

$$s_x\sqrt{n \Big/ \chi^2_{1-\alpha/2}(n-1)} < \sigma < s_x\sqrt{n \Big/ \chi^2_{\alpha/2}(n-1)}$$

($100(1 - \alpha)\%$ two-sided confidence intervals of parameters $\mu \in$ R, $\sigma^2 > 0$ and $\sigma > 0$ in distribution $N(\mu, \sigma^2)$, see Sect. 26.5)

$$\frac{d-1}{d+1} < \rho < \frac{h-1}{h+1},$$

where

$$d = \exp\left(2Z - \frac{2u_{1-\alpha/2}}{\sqrt{n-3}}\right); \quad h = \exp\left(2Z + \frac{2u_{1-\alpha/2}}{\sqrt{n-3}}\right); \quad z = \frac{1}{2}\ln\frac{1+r_{xy}}{1-r_{xy}}$$

($100(1-\alpha)\%$ two-sided confidence interval of parameter $\rho \in (-1, 1)$ (see Sect. 26.6) in bivariate normal distribution: z is called *Fisher's z-transformation*)

$$\bar{x} - u_{1-\alpha/2} \cdot \sqrt{p_A(1-p_A)/n} < \pi_A < \bar{x} + u_{1-\alpha/2} \cdot \sqrt{p_A(1-p_A)/n}$$

(approximate $100(1-\alpha)\%$ two-sided confidence interval of probability π_A of occurrence of an event A, which has (in random counting sample of large size n) the relative frequency $p_A = n_A/n$, see Sect. 27.3)

$$n \geq 4\left(u_{1-\alpha/2}\right)^2 \frac{p_A(1-p_A)}{\Delta^2}$$

(minimal approximate size of sample so that the width of two-sided confidence interval of π_A (see *thereinbefore*) is not larger than a preset bound Δ)

$$\bar{x} - u_{1-\alpha/2} \cdot s_x/\sqrt{n} < \mathrm{E}(X) < \bar{x} + u_{1-\alpha/2} \cdot s_x/\sqrt{n}$$

$$s_x^2 - u_{1-\alpha/2} \cdot s_x^2\sqrt{2/n} < \mathrm{var}(X) < s_x^2 + u_{1-\alpha/2} \cdot s_x^2\sqrt{2/n}$$

$$s_x - u_{1-\alpha/2} \cdot s_x/\sqrt{2n} < \sqrt{\mathrm{var}(X)} < s_x + u_{1-\alpha/2} \cdot s_x/\sqrt{2n}$$

$$r_{xy} - u_{1-\alpha/2}(1-r_{xy}^2)/\sqrt{n} < \rho(X,Y) < r_{xy} + u_{1-\alpha/2}(1-r_{xy}^2)/\sqrt{n}$$

(approximate $100(1-\alpha)\%$ two-sided confidence intervals of $\mathrm{E}(X)$, $\mathrm{var}(X)$, $\sqrt{\mathrm{var}(X)}$, $\rho(X, Y)$ in a general random sample of large size n, see Sects. 26.3 and 26.6)

- Nonparametric estimation: is an estimation of a probability distribution characteristic without a parametric specification (e.g. an estimation of median)

27.10 Hypothesis Testing

$$(x_1, \ldots, x_n)' \in W$$

(*critical region* $W \subset \mathrm{R}^n$ of *null hypothesis* H_0 against *alternative hypothesis* H_1 with *significance level* $\alpha (0 < \alpha < 1)$: if the observed values fulfil $(x_1, \ldots, x_n)' \in W$, then one rejects H_0 against H_1; every test of statistical hypothesis takes a risk of a wrong decision with two possibilities of making an error: (1) *type-one error* is committed, if H_0 is rejected when it is true, and (2) *type-two error* is committed, if H_0 is not rejected when it is false; the objective is to find such W that the probability of type-one error does not exceed α (the prescribed *significance level*) and, at the same time, the probability β of type-two error is as small as possible ($1 - \beta$ is called *power of test*))

$$\sqrt{n-1}\ \left|\bar{x}-\mu_0\right|/s_x \geq t_{1-\alpha/2}(n-1)$$

(critical region of two-sided *one-sample t-test* of hypothesis H_0: $\mu = \mu_0$ against alternative H_1: $\mu \neq \mu_0$ ($\mu_0 \in \mathrm{R}$ is a given constant) with significance level α: test is based on a sample from distribution $N(\mu, \sigma^2)$, see Sect. 26.5)

$$\sqrt{n-1}\,(\bar{x}-\mu_0)/s_x \geq t_{1-\alpha}(n-1)$$

(critical region of one-sided *one-sample t-test* of hypothesis H_0: $\mu \leq \mu_0$ against alternative H_1: $\mu > \mu_0$ ($\mu_0 \in \mathrm{R}$ is a given constant) with significance level α: test is based on a sample from distribution $N(\mu, \sigma^2)$, see Sect. 26.5)

$$n \cdot s_x^2/\sigma_0^2 \leq \chi^2_{\alpha/2}(n-1) \text{ or } n \cdot s_x^2/\sigma_0^2 \geq \chi^2_{1-\alpha/2}(n-1)$$

(critical region of two-sided one-sample test of hypothesis H_0: $\sigma^2 = \sigma_0^2$ against alternative H_1: $\sigma^2 \neq \sigma_0^2$ ($\sigma_0^2 > 0$ is a given constant) with significance level α: test is based on a sample from distribution $N(\mu, \sigma^2)$, see Sect. 26.5)

$$n \cdot s_x^2/\sigma_0^2 \geq \chi^2_{1-\alpha}(n-1)$$

(critical region of one-sided one-sample test of hypothesis H_0: $\sigma^2 \leq \sigma_0^2$ against alternative H_1: $\sigma^2 > \sigma_0^2$ ($\sigma_0^2 > 0$ is a given constant) with significance level α: test is based on a sample from distribution $N(\mu, \sigma^2)$, see Sect. 26.5)

$$\frac{|\bar{x}-\bar{y}|}{\sqrt{\dfrac{s_x^2}{n_1-1}+\dfrac{s_y^2}{n_2-1}}} \geq \frac{t_{1-\alpha/2}(n_1-1)\frac{s_x^2}{n_1-1}+t_{1-\alpha/2}(n_2-1)\frac{s_y^2}{n_2-1}}{\dfrac{s_x^2}{n_1-1}+\dfrac{s_y^2}{n_2-1}}$$

(critical region of two-sided *two-sample t-test* of hypothesis H_0: $\mu_1 = \mu_2$ against alternative H_1: $\mu_1 \neq \mu_2$ with significance level α: test is based on two independent samples from distribution $N(\mu_1, \sigma_1^2)$ of size n_1 and from distribution $N(\mu_2, \sigma_2^2)$ of size n_2, see Sect. 26.5)

$$\sqrt{n}\cdot\frac{|\bar{x}-\bar{y}|}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}((x_i-y_i)-(\bar{x}-\bar{y}))^2}}\geq t_{1-\alpha/2}(n-1)$$

(critical region of two-sided *paired t-test* of hypothesis H_0: $\mu_1=\mu_2$ against alternative H_1: $\mu_1\neq\mu_2$ with significance level α: test is based on a sample from bivariate normal distribution with parameters μ_1, μ_2, σ_1^2, σ_2^2, ρ, see Sect. 26.6)

$$\sqrt{n-2}|r_{xy}|\Big/\sqrt{1-r_{xy}^2}\geq t_{1-\alpha/2}(n-2)$$

(critical region of two-sided *paired t-test* of hypothesis H_0: $\rho=0$ against alternative H_1: $\rho\neq 0$ with significance level α: test is based on a sample from bivariate normal distribution with parameter ρ, see Sect. 26.6)

- *Nonparametric tests*: can be used without an assumption on the type of probability distribution, e.g.
  - *sign test*: is test of comparability of paired values of a bivariate sample
  - *one-sample Wilcoxon test*: as for the sign test
  - *two-sample Wilcoxon test*: is test of matching of two unknown probability distributions
  - *Mann-Whitney test*: as for the two-sample Wilcoxon test
  - *Kruskal-Wallis test*: generalizes the two-sample Wilcoxon test for more samples
  - *Friedman test*: as for Kruskal-Wallis test
  - *Spearman test*: is test of independence between paired values of a bivariate sample
  - *Kendall test*: as for Spearman test
  - *test based on iterations under and above median*: test of randomness of a sample
  - *test based on turning points*: test of randomness of a sample
  - *rank tests*: are special cases of the tests *thereinbefore* (e.g. Spearman or Kendall rank tests) when replacing the observations by their ranks (*rank* R_i of value x_i is the number of such $x_1, x_2, \ldots, x_n$, which are smaller or equal to x_i)
- *Goodness of fit tests*: enable to test whether a given sample has been chosen from probability distribution of a given type (*chi-square test*, *Kolmogorov-Smirnov test* and others)

27.11 Regression Analysis

$$Y=f(x_1,\ldots,x_k)+\varepsilon$$

(*regression model*: explains *response variable* (*regressand*) Y (with the character of random variable) by means of (nonrandom) *explanatory variables* (*regressors*) $x_1,\ldots,x_k$, which are arguments of *regression function f*, and by means of a *residual* (*error variable*) ε)

$$Y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_k x_{ik} + \varepsilon_i,$$

where

$$E(\varepsilon_i) = 0; \quad \operatorname{cov}(\varepsilon_i, \varepsilon_j) = \delta_{ij}\sigma^2 \quad (i,j = 1, \ldots, n)$$

(*linear regression model*: the regression function is a linear function of unknown parameters $\beta_0, \beta_1, \ldots, \beta_k \in \mathrm{R}$; Y_i denotes the value of the response variable corresponding to the values $x_{i1}, \ldots, x_{ik}$ of the explanatory variables $x_1, \ldots, x_k$ ($i = 1, \ldots, n$; $n > k + 1$); residuals ε_i fulfil the given assumptions, i.e. they are uncorrelated random variables ("measurements of the model for various i do not affect mutually") with zero mean value ("residuals are random fluctuations without a systematic component") and with constant variance $\sigma^2 > 0$, which is also an unknown parameter of the model ("measurement error for various i remains the same")

$$y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_k x_{ik} + \varepsilon_i, \quad i = 1, \ldots, n$$

(*observed* model of linear regression: unknown parameters $\beta_0, \beta_1, \ldots, \beta_k$ can be estimated by means of observations $y_i, x_{i1}, \ldots, x_{ik}$ ($i = 1, \ldots, n$; $n > k + 1$); the residuals ε_i represent measurement errors, imperfections due to improper choice of the model, etc.)

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

where

$$\mathbf{y} = (y_1, \ldots, y_n)'; \quad \mathbf{X} = \left(x_{ij}\right)_{\substack{i=1,\ldots,n \\ j=0,1,\ldots,k}} \quad (x_{i0} \equiv 1); \quad \boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)'$$

$$E(\boldsymbol{\beta}) = \mathbf{0}; \quad \operatorname{var}(\boldsymbol{\varepsilon}) = \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon\varepsilon}} = \sigma^2\,\mathbf{I}$$

(*matrix form* of the observed linear regression model: the first column of $\mathbf{X}$ is the column vector of unities $\mathbf{e}$; one assumes usually that $\mathbf{X}$ has linearly independent columns, i.e. $h(\mathbf{X}) = k + 1$)

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}, \quad \boldsymbol{\Sigma}_{\mathbf{bb}} = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$$

(*best linear unbiased estimate* (see Sect. 27.9) of parameters $\boldsymbol{\beta}$ and its covariance matrix (*Gauss-Markov theorem*); it may be obtained from the system of *normal equations* $(\mathbf{X}'\mathbf{X})\mathbf{b} = \mathbf{X}'\mathbf{y}$)

$$\mathbf{b} = \underset{\boldsymbol{\beta} \in \mathrm{R}^{k+1}}{\arg\min} \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_{i1} + \ldots + \beta_k x_{ik}))^2 = \underset{\boldsymbol{\beta} \in \mathrm{R}^{k+1}}{\arg\min}\, (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

($\mathbf{b}$ is obtained by the *method of least squares* (*OLS-estimate*, Ordinary Least Squares))

$$\mathbf{b}_t = \mathbf{b}_{t-1} + \mathbf{P}_t\mathbf{x}_t(y_t - \mathbf{x}_t'\mathbf{b}_{t-1}); \quad \mathbf{P}_t = \mathbf{P}_{t-1} - (\mathbf{x}_t'\mathbf{P}_{t-1}\mathbf{x}_t + 1)^{-1}\mathbf{P}_{t-1}\mathbf{x}_t\mathbf{x}_t'\mathbf{P}_{t-1}$$

(*recursive method of least squares*: the estimate $\mathbf{b}_t$ is calculated at time t using the previous estimate $\mathbf{b}_{t-1}$ and the observed values y_t and $\mathbf{x}_t = (x_{t1}, \ldots, x_{tk})$; one must preset initial values $\mathbf{b}_0$ and $\mathbf{P}_0$)

$$\hat{y} = b_0 + b_1x_1 + \ldots + b_kx_k = \mathbf{x}'\,\mathbf{b}$$

(*point prediction* $\hat{y}$: is the value of the response variable Y estimated ("predicted") by means of the estimated model for values $\mathbf{x}' = (1, x_1, \ldots, x_k)$ of the explanatory variables)

$$\hat{\mathbf{y}} = (\hat{y}_1, \ldots, \hat{y}_n)' = \mathbf{Xb} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}$$

(*estimated value* $\hat{y}_i = b_0 + b_1x_{i1} + \ldots + b_kx_{ik}$: is the value of the response variable Y_i calculated by means of the estimated model ($i = 1, \ldots, n$))

$$\hat{\boldsymbol{\varepsilon}} = (\hat{\varepsilon}_1, \ldots, \hat{\varepsilon}_n)' = \mathbf{y} - \mathbf{Xb} = \left(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right)\mathbf{y} = \mathbf{My}$$

(*estimated residual* $\hat{\varepsilon}_i = y_i - \hat{y}_i = y_i - (b_0 + b_1x_{i1} + \ldots + b_kx_{ik})$: is the value of the residual ε_i calculated by means of the estimated model ($i = 1, \ldots, n$); the square matrix $\mathbf{M}$ is symmetric and idempotent)

$$\hat{y}_1 + \ldots + \hat{y}_n = y_1 + \ldots + y_n$$

$$\hat{\varepsilon}_1 + \ldots + \hat{\varepsilon}_n = 0$$

$$\hat{y}_1\hat{\varepsilon}_1 + \ldots + \hat{y}_n\hat{\varepsilon}_n = 0$$

$$x_{1j}\hat{\varepsilon}_1 + \ldots + x_{nj}\hat{\varepsilon}_n = 0, \quad j = 0, 1, \ldots, k$$

$$S_t = \sum_{i=1}^{n} (y_i - \bar{y})^2, \quad \text{where} \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

(*total sum of squares*: describes the variability of the response variable since it holds $s_y^2 = S_t/n$, see Sect. 27.6)

$$S_r = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$$

(*regression sum of squares*: describes such a part of variability of the response variable, which is explained by the model)

$$S_\varepsilon = \sum_{i=1}^{n} \hat{\varepsilon}_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

(*residual sum of squares*: describes such a part of variability of the response variable, which is not explained by the model)

$$S_t = S_r + S_\varepsilon$$

$$R^2 = \frac{S_r}{S_t} = 1 - \frac{S_\varepsilon}{S_t}$$

(*coefficient of determination*: indicate, how the considered linear regression model is capable to explain variability of the response variable, i.e. R^2 assesses the quality of the model for the given data; $0 \le R^2 \le 1$ (it is common to express R^2 in per cent, where 100% means the ideal case)

$$\bar{R}^2 = 1 - \frac{n-1}{n-k-1} \cdot \frac{S_\varepsilon}{S_t}$$

(*adjusted coefficient of determination*: in comparison with R^2, the adjusted coefficient has better properties when estimating the coefficient of multivariate correlation $r_{y\mathbf{x}}$, see Sect. 27.8)

$$s^2 = \tfrac{1}{n-k-1}\, S_\varepsilon = \tfrac{1}{n-k-1} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \quad (\textit{unbiased estimate} \text{ of } \sigma^2)$$

$$\mathbf{S_{bb}} = s^2(\mathbf{X}'\mathbf{X})^{-1} \quad (\textit{unbiased estimate} \text{ of } \textstyle\sum_{\mathbf{bb}} = \sigma^2(\mathbf{X}'\mathbf{X}^{-1}))$$

$$s_{b_j}^2 = (\mathbf{S_{bb}})_{jj} = s^2\left((\mathbf{X}'\mathbf{X})^{-1}\right)_{jj},\ j = 0, 1, \ldots, k \ (\textit{unbiased estimate} \text{ of } \mathrm{var}(b_j))$$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

where $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I})$
(*normal model of linear regression*: $\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$; $\mathbf{b} \sim N(\boldsymbol{\beta}, \sigma^2\,(\mathbf{X}'\mathbf{X})^{-1}$, see Sect. 26.6)

$$\hat{\boldsymbol{\beta}} = \mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}; \quad \sigma^2 = \frac{1}{n} S_\varepsilon = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \frac{n-k-1}{n}\, s^2$$

(*maximum likelihood estimate* of $\boldsymbol{\beta}$ and σ^2, see Sect. 27.9)

$$|b_j| \,/ s_{b_j} \ \ge \ t_{1-\alpha/2}(n-k-1)$$

(*test of significance for parameter* β_j in the normal model of linear regression ($j = 0, 1, \ldots, k$): is the test of hypothesis H_0: $\beta_j = 0$ against alternative H_1: $\beta_j \ne 0$ with significance level α, see Sect. 27.10; the test statistics $|b_j| \,/ s_{b_j}$ is called *t-ratio*)

$$|b_j|/s_{b_j} \geq u_{1-\alpha/2}$$

(asymptotic *test of significance for parameter* β_j in the model of linear regression without assumption of normality ($j = 0, 1, \ldots, k$; n is large): is the test of hypothesis H_0: $\beta_j = 0$ against alternative H_1: $\beta_j \neq 0$ with significance level α, see Sect. 27.10)

$$\frac{n-k-1}{k} \cdot \frac{R^2}{1-R^2} \geq F_{1-\alpha}(k,\ n-k-1)$$

(*test of significance for model* (*test of significance for coefficient of determination*) in the normal model of linear regression: is the test of hypothesis H_0: $(\beta_1, \ldots, \beta_k)' = \mathbf{0}$ against alternative H_1: $(\beta_1, \ldots, \beta_k)' \neq \mathbf{0}$ with significance level α, see Sect. 27.10)

$$\Big(\hat{y} - t_{1-\alpha/2}(n-k-1) \cdot s \cdot \sqrt{1 + \mathbf{x}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}},\ \hat{y} + t_{1-\alpha/2}(n-k-1) \cdot s \cdot \sqrt{1 + \mathbf{x}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}}\Big)$$

($100(1-\alpha)\%$ *prediction interval* in the normal model of linear regression, see Sect. 27.9: one predicts the value of the response variable Y for values $\mathbf{x}' = (1, x_1, \ldots, x_k)$ of the explanatory variables; $\hat{y}$ is the point prediction, see *thereinbefore*)

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i,$$

where

$$\mathrm{E}(\varepsilon_i) = 0;\quad \mathrm{cov}(\varepsilon_i, \varepsilon_j) = \delta_{ij}\sigma^2\ \ (i, j = 1, \ldots, n)$$

(*regression model of line*: the regression function is a line; in particular, $k = 1$)

$$b_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})y_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2};$$

$$b_0 = \bar{y} - b_1\bar{x} = \frac{\left(\sum_{i=1}^{n} y_i - b_1\sum_{i=1}^{n} x_i\right)}{n},$$

(best linear unbiased estimates of β_0 and β_1 in the regression model of line)

$$s^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \bar{y} - b_1(x_i - \bar{x}))^2 = \frac{1}{n-2}\left(\sum_{i=1}^{n} y_i^2 - b_0\sum_{i=1}^{n} y_i - b_1\sum_{i=1}^{n} x_i y_i\right)$$

(unbiased estimate of σ^2 in the regression model of line)

$$s_{b_1}^2 = \frac{s^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}; \quad s_{b_0}^2 = s^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \right)$$

(unbiased estimates of var(b_0) and var(b_1) in the regression model of line)

$$\frac{|b_1|}{s} \sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2} \geq t_{1-\alpha/2}(n-k-1)$$

(*test of linearity* in the normal regression model of line: is the test of hypothesis H_0: $\beta_1 = 0$ against alternative H_1: $\beta_1 \neq 0$ with significance level α, see Sect. 27.10; the test statistics is again the t-ratio)

$$b_0 + b_1 x \ \pm \ t_{1-\alpha/2}(n-2) \cdot s \cdot \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}};$$

($100(1-\alpha)\%$ prediction interval in the normal regression model of line, see Sect. 27.9: one predicts the value of the response variable Y for a value of the explanatory variable x)

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

where

$$\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}; \quad \mathrm{var}(\boldsymbol{\varepsilon}) = \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}} = \sigma^2\boldsymbol{\Omega} \ \ (\sigma^2 > 0; \ \ \boldsymbol{\Omega} > 0)$$

(*generalized* linear regression model: $\boldsymbol{\Omega}$ is a general positive definite matrix with unknown elements, i.e. unlike the classical model, the residuals ε_i need not be mutually uncorrelated with constant variance)

$$\tilde{\mathbf{b}} = (\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{y}; \quad \boldsymbol{\Sigma}_{\tilde{\mathbf{b}}\tilde{\mathbf{b}}} = \sigma^2(\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X})^{-1}$$

(best linear unbiased estimate of parameters $\boldsymbol{\beta}$ (*Aitken estimate*) and its covariance matrix, see Sect. 27.9: the possibilities of its practical application are restricted due to the fact that the matrix $\boldsymbol{\Omega}$ can be estimated only in such special cases, in which it can be described using only a small number of (unknown) parameters (see *thereinafter* the special cases of heteroscedasticity and autocorrelated residuals))

$$\mathbf{b}_w = \underset{\boldsymbol{\beta} \in \mathrm{R}^{k+1}}{\arg\min} \ \sum_{i=1}^{n} w_i \left(y_i - (\beta_0 + \beta_1 x_{i1} + \ldots + \beta_k x_{ik})\right)^2,$$

where $w_1, \ldots, w_n$ are given positive weights

($\mathbf{b}_w$ is obtained by the *weighted* method of least squares (*WLS-estimate*, Weighted Least Squares); it is Aitken estimate with $\mathbf{\Omega} = \mathbf{diag}\{w_1, \ldots, w_n\}$; one can construct it as the OLS-estimate in the transformed model

$$\sqrt{w_i}y_i = \beta_0\sqrt{w_i} + \beta_1\sqrt{w_i}x_{i1} + \ldots + \beta_k\sqrt{w_i}x_{ik} + \varepsilon_i' \quad (i = 1, \ldots, n))$$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

where

$$\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}; \quad \mathrm{var}(\boldsymbol{\varepsilon}) = \mathbf{\Sigma}_{\boldsymbol{\varepsilon\varepsilon}} = \sigma^2\mathbf{diag}\{k_1, \ldots, k_n\} \quad (\sigma^2 > 0; \; k_1, \ldots, k_n > 0)$$

(*heteroscedasticity*: the residuals ε_i are uncorrelated but they have variances $\sigma^2 k_i$ with unknown values k_i; in such a case one can try to remove the heteroscedasticity using transformation to $y_i/x_{ij} = \beta_0/x_{ij} + \beta_1 x_{i1}/x_{ij} + \ldots + \beta_k x_{ik}/x_{ij} + \varepsilon_i'$ $(i = 1, \ldots, n)$, where x_j is a suitably chosen explanatory variable with observed positive values $x_{1j}, \ldots, x_{nj}$)

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

where

$$\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}; \quad \mathrm{var}(\boldsymbol{\varepsilon}) = \mathbf{\Sigma}_{\boldsymbol{\varepsilon\varepsilon}} = \sigma^2 \begin{pmatrix} 1 & \rho & \cdots & \rho^{n-1} \\ \rho & 1 & \cdots & \rho^{n-2} \\ \vdots & \vdots & \vdots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \cdots & 1 \end{pmatrix} \quad (\sigma^2 > 0; \; -1 < \rho < 1)$$

(*autocorrelated residuals*: the residuals ε_i have the constant variance σ^2 but they are correlated (with correlation structure $AR(1)$, see Sects. 30.4 and 31.4); in such a case one can remove the autocorrelation of residuals by transformation to

$$y_i - \hat{\rho}y_{i-1} = \beta_0(1 - \hat{\rho}) + \beta_1(x_{i1} - \hat{\rho}x_{i-1,1}) + \ldots + \beta_k(x_{ik} - \hat{\rho}x_{i-1,k}) + \varepsilon_i'$$

$(i = 2, \ldots, n)$, where $\hat{\rho} = \sum_{i=1}^{n-1}\hat{\varepsilon}_i\hat{\varepsilon}_{i+1}/\sum_{i=1}^{n}\hat{\varepsilon}_i^2$ and $\hat{\varepsilon}_i$ are residuals calculated by the *OLS*-method in the original model)

$$D = \frac{\sum_{i=1}^{n-1}\left(\hat{\varepsilon}_{i+1} - \hat{\varepsilon}_i\right)^2}{\sum_{i=1}^{n}\hat{\varepsilon}_i^2} = \frac{\sum_{i=1}^{n-1}\left(\hat{\varepsilon}_{i+1} - \hat{\varepsilon}_i\right)^2}{S_\varepsilon}$$

(*Durbin-Watson statistics* for the test of hypothesis H_0: $\rho = 0$ (uncorrelated residuals) against alternative H_1: $\rho > 0$ (*positive autocorrelations of residuals*) or H_1: $\rho < 0$ (*negative autocorrelations of residuals*))

$$Y_i = f(\mathbf{x}_i; \boldsymbol{\theta}) + \varepsilon_i, \quad i = 1, \ldots, n$$

(*nonlinear regression model*: the regression function is a nonlinear function of unknown parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)' \in \mathrm{R}^m$; Y_i denotes the value of the response variable corresponding to the values $\mathbf{x}_i = (x_{i1}, \ldots, x_{ik})'$ of the explanatory variables;

in the simplest case, the residuals ε_i fulfil the assumptions of the classical linear regression model)

$$\hat{\boldsymbol{\theta}}_{t+1} = \hat{\boldsymbol{\theta}}_t + \left(\mathbf{F}(\hat{\boldsymbol{\theta}}_t)'\mathbf{F}(\hat{\boldsymbol{\theta}}_t)\right)^{-1} \mathbf{F}(\hat{\boldsymbol{\theta}}_t)' \left(\mathbf{y} - \mathbf{f}(\hat{\boldsymbol{\theta}}_t)\right), \quad t = 0, 1, \ldots,$$

where

$$\mathbf{F}(\hat{\boldsymbol{\theta}}_t) = \left(\frac{\partial f(\mathbf{x}_i, \hat{\boldsymbol{\theta}}_t)}{\partial \theta_j}\right)_{\substack{i=1,\ldots,n \\ j=1,\ldots,m}}; \quad \mathbf{f}(\hat{\boldsymbol{\theta}}_t) = \left(f(\mathbf{x}_i, \hat{\boldsymbol{\theta}}_t)\right)_{i=1,\ldots,n}; \quad \mathbf{y} = (y_i)_{i=1,\ldots,n}$$

(*Gauss-Newton method*: provides an iterative estimation of parameters $\boldsymbol{\theta}$ using a suitable stop rule)

$$s^2 = \frac{1}{n-m} \sum_{i=1}^{n} \left(y_i - f(\mathbf{x}_i, \hat{\boldsymbol{\theta}})\right)^2; \mathbf{S}_{\hat{\boldsymbol{\theta}}\hat{\boldsymbol{\theta}}} = s^2 \left(\mathbf{F}(\hat{\boldsymbol{\theta}})'\mathbf{F}(\hat{\boldsymbol{\theta}})\right)^{-1}$$

(estimates of σ^2 and $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}\hat{\boldsymbol{\theta}}}$ in Gauss-Newton method)

27.12 Analysis of Variance (ANOVA)

$$x_{pi} = \mu + \alpha_i + \varepsilon_{pi}, \quad i = 1, \ldots, I; \quad p = 1, \ldots, n_i$$

(*model of one-way classification* according to a single factor A: the observed sample of size n is subdivided according to the factor A into I groups; the ith group contains observations $x_{1i}, \ldots, x_{n_i,i}$; ε_{pi} are independent random variables with distribution $N(0, \sigma^2)$; μ, α_i, σ^2 are unknown parameters ($i = 1, \ldots, I$; $n_1 + \ldots + n_I = n$))

$$\frac{n-I}{I-1} \cdot \frac{S_A}{S_\varepsilon} = \frac{n-I}{I-1} \cdot \frac{S_A}{S_t - S_A} \geq F_{1-\alpha}(I-1, n-I),$$

where

$$S_t = \sum_{i=1}^{I} \sum_{p=1}^{n_i} (x_{pi} - \bar{x})^2; \quad S_A = \sum_{i=1}^{I} n_i(\bar{x}_i - \bar{x})^2; \quad S_\varepsilon = \sum_{i=1}^{I} \sum_{p=1}^{n_i} (x_{pi} - \bar{x}_i)^2$$

(critical region for *test of hypothesis* H_0: $\alpha_1 = \ldots = \alpha_I = 0$ with significance level α that effects of the factor A in the one-way classification are not significant, see Sect. 26.5)

$$(\bar{x}_r - \bar{x}_s)^2 \geq \frac{n_r + n_s}{n_r n_s} \cdot \frac{I-1}{n-I} \cdot S_\varepsilon \cdot F_{1-\alpha}(I-1, n-I),$$

(rejection of equality of the rth and sth group ($r, s = 1, \ldots, I$; $r \neq s$) with significance level α in the one-way classification, provided that H_0 has been rejected, see Sect. 26.5)

$$x_{pij} = \mu + \alpha_i + \beta_j + \varepsilon_{pij}, \quad i = 1, \ldots, I; \ j = 1, \ldots, J; \ p = 1, \ldots, P$$

(*model of two-way classification* according to factors A and B: an observed sample of size n is subdivided according to the factors A and B into $I{\cdot}J$ groups; the group corresponding to the values i of A and j of B contains observations $x_{1ij}, \ldots, x_{Pij}$; ε_{pij} are independent random variables with distribution $N(0, \sigma^2)$; $\mu, \alpha_i, \beta_j, \sigma^2$ are unknown parameters ($i = 1, \ldots, I; j = 1, \ldots, J; I{\cdot}J{\cdot}P = n$))

$$\frac{n-I-J+1}{I-1} \cdot \frac{S_A}{S_\varepsilon} \geq F_{1-\alpha}(I-1, n-I-J+1) \quad \text{or}$$

$$\frac{n-I-J+1}{J-1} \cdot \frac{S_B}{S_\eta} \geq F_{1-\alpha}(J-1, n-I-J+1),$$

where

$$S_t = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{p=1}^{P}(x_{pij} - \bar{x})^2; \quad S_A = J \cdot P\sum_{i=1}^{I}(\bar{x}_i - \bar{x})^2$$

$$S_B = I \cdot P\sum_{j=1}^{J}(\bar{x}_j - \bar{x})^2; \quad S_\varepsilon = S_t - S_A - S_B$$

(critical region for *test of hypothesis* H_0: $\alpha_1 = \ldots = \alpha_I = 0$ or H_0: $\beta_1 = \ldots = \beta_J = 0$ with significance level α that effects of the factor A or B in the two-way classification are not significant, respectively, where unlike the one-way classification, one respects the effects of the second factor; in addition, the *two-way classification with interactions* tests also dependence between effects of both factors)

27.13 Multivariate Statistical Analysis

$$\mathbf{z}_i = \mathbf{X}\mathbf{c}_i, \quad i = 1, \ldots, m$$

(ith *principal component* of m-variate sample $(x_{11}, \ldots, x_{1m}), \ldots, (x_{n1}, \ldots, x_{nm})$ arranged to a matrix $\mathbf{X}$ of type $n \times m$, where $\mathbf{c}_1, \ldots, \mathbf{c}_m$ are the orthonormal eigenvectors corresponding to the ordered eigenvalues $\lambda_1 \geq \ldots \geq \lambda_m$ of the covariance matrix $\mathbf{\Sigma}_{\mathbf{xx}}$, see Sect. 27.8: in particular, the principal components are mutually uncorrelated and explain the variability of the original sample as much as possible in their sequence $\mathbf{z}_1, \mathbf{z}_2, \ldots$; therefore several principal components $\mathbf{z}_1, \ldots, \mathbf{z}_k$ ($k << m$) are capable to replace in a sufficient way the original large sample $\mathbf{x}_1, \ldots, \mathbf{x}_m$: they explain $100(\lambda_1 + \ldots + \lambda_k)/(\lambda_1 + \ldots + \lambda_m)$ % of the original variability (or original information))

$$d_j = \boldsymbol{\mu}_j'\mathbf{\Sigma}^{-1}\mathbf{x} - \frac{1}{2}\boldsymbol{\mu}_j'\mathbf{\Sigma}^{-1}\boldsymbol{\mu}_j + \ln p_j, \quad j = 1, \ldots, r$$

(*linear discriminant analysis*: decomposes an m-variate sample into $r \geq 2$ discriminant classes, where the statistical units of jth class ($j = 1, \ldots, r$) have m-variate normal distribution with known mean vector $\boldsymbol{\mu}_j$ and known covariance matrix $\boldsymbol{\Sigma}$ (see Sect. 27.8) and p_j is the probability that a randomly chosen unit belongs to jth class ($\boldsymbol{\mu}_j$, $\boldsymbol{\Sigma}$ and p_j are estimated in advance by means of those statistical units, for which their allocation to discriminant classes is given a priori); a statistical unit $\mathbf{x} = (x_1, \ldots, x_m)'$ is then allocated to hth class, if it fulfils $d_h = \max(d_1, \ldots, d_r)$)

$$(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)'\boldsymbol{\Sigma}^{-1}\mathbf{x} > a,$$

where

$$a = \frac{1}{2}\boldsymbol{\mu}_1'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_1 - \frac{1}{2}\boldsymbol{\mu}_2'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_2 + \ln p_2 - np_1$$

(linear discriminant analysis with *two* discriminant classes): if a statistical unit $\mathbf{x} = (x_1, \ldots, x_m)'$ fulfils the given inequality, then it is allocated to the first class; in the opposite case it is allocated to the second class)

- *Quadratic discriminant analysis*: generalizes the linear one since its decision function is quadratic in $\mathbf{x}$, and particular discriminant classes can have different covariance matrices $\boldsymbol{\Sigma}_j$

Further Reading

Draper, N.R., Smith, H.: Applied Regression Analysis. Wiley, New York (1998)
Jolliffe, I.T.: Principal Component Analysis. Springer, New York (2002)
Lehmann, E.L., Romano, J.P.: Testing Statistical Hypotheses. Springer, New York (2005)
Mardia, K.V., Kent, J.T., Bibby, J.M.: Multivariate Analysis. Academic, London (1979)
Rektorys, K. et al.: Survey of Applicable Mathematics. Kluwer, Dordrecht (1994)
Rice, J.A.: Mathematical Statistics and Data Analysis. Duxbury Press, Boston, MA (1995)
Wasserman, L.: All of Nonparametric Statistics. Springer, New York (2006)
Wilks, S.S.: Mathematical Statistics. Wiley, New York (1962)

Chapter 28
Econometrics

Abstract Chapter 28 presents basic procedures of modern econometrics: 28.1. Multicollinearity, 28.2. A Priori Restrictions, 28.3. Qualitative Regressors, 28.4. Probit and Logit Models, 28.5. Random Regressors and Instrumental Variable Estimation, 28.6. Simultaneous Equation Models and 2SLS-Estimator.

28.1 Multicollinearity

$$\left|\mathbf{X}'\mathbf{X}\right| \approx 0$$

(*multicollinearity* in a model of linear regression $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ (see Sect. 27.11): in economic practice, the assumption $h(\mathbf{X}) = k + 1$ may be "nearly" violated due to correlations among economic variables in positions of regressors in the matrix $\mathbf{X}$ of type $n \times (k + 1)$ (the matrix $\mathbf{X}'\mathbf{X}$ is *ill-conditioned*); the OLS-estimate $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ with covariance matrix $\boldsymbol{\Sigma}_{\mathbf{bb}} = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$ has no practical sense; some practical recommendations how to face multicollinearity are as follows:
– to omit (or replace) the regressors inducing multicollinearity
– to transform regressors (e.g. to centre by subtracting arithmetic mean or estimated trend, to standardize dividing by standard deviation, to apply first differences instead of original regressors)
– to make use of an a priori restrictions (see Sect. 28.2)
– to apply the method of principal components to regressors (see *thereinafter*))

$$\left|r_{x_i x_j}\right| \geq 0.8$$

(*multicollinearity criterion* due to correlations between pairs of regressors)

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_28,

$$r_i \;\geq\; R$$

(*multicollinearity criterion* due to correlations between particular regressors and remaining regressors: $r_i = r_{x_i,(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)}$ is the multiple correlation coefficient (see Sect. 27.8) and R^2 is the coefficient of determination (see Sect. 27.11))

$$-\,(n-1-(2\,k+5)/6)\ln|R_{xx}| \geq \chi^2_{1-\alpha}\,(k(k-1)/2)$$

(*statistical test of multicollinearity*: the critical region of *Farrar-Glauber test* with significance level α and null hypothesis that none of correlation coefficients for particular pairs of regressors is significantly different from zero)

$$\frac{n-k}{k-1}\frac{r_i^2}{1-r_i^2} \geq F_{1-\alpha/2}(k-1, n-k),$$

where

$$r_i = r_{x_i,(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)}$$

with

$$\frac{r_i^2}{1-r_i^2} = r^{ii} - 1 \quad \text{and} \quad R_{xx}^{-1} = \left(r^{ij}\right) \text{(see Sect. 27.8)}$$

(*statistical test of multicollinearity*: the critical region of the test with significance level α and null hypothesis that a regressor x_i is uncorrelated with the remaining regressors)

$$\sqrt{n-k}\frac{\left|r_{(i,j)}\right|}{\sqrt{1-r^2_{(i,j)}}} \geq t_{1-\alpha/2}(n-k),$$

where

$$r_{(i,j)} = -\frac{r^{ij}}{\sqrt{r^{ii}r^{jj}}} \text{ with } r_{(i,j)} = r_{x_i x_j \cdot (x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_{j-1},x_{j+1},\ldots,x_n)} \text{(see Sect. 27.8)}$$

(*statistical test of multicollinearity*: the critical region of the test with significance level α and null hypothesis that regressors x_i and x_j are uncorrelated when one eliminates the influence of the remaining regressors)

$$\mathbf{y} - \bar{y}\mathbf{e} = \mathbf{Z}_m\boldsymbol{\gamma} + \boldsymbol{\varepsilon}$$

(*multicollinearity eliminated by method of principal components*: the original model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ is transformed in a suitable way, where the matrix $\mathbf{Z}_m$ is created by m principal components $\mathbf{z}_1, \ldots, \mathbf{z}_m$ of centered regressors $\mathbf{x}_1, \ldots, \mathbf{x}_k$ (the response variable y is also centered); the number m of (ordered) principal components ($m << k$) may be chosen according to the usual rule (see Sect. 27.13); finally, the

OLS-estimate $\mathbf{b}$ of parameters $\boldsymbol{\beta}$ in the original model is obtained by the corresponding reverse transformation of the OLS-estimate $\mathbf{c}$ of parameters $\boldsymbol{\gamma}$ in the transformed model)

28.2 A Priori Restrictions

$$\left.\begin{array}{l}\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ \mathbf{R}\boldsymbol{\beta} = \mathbf{r}\end{array}\right\},$$

where $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ is a linear regression model (see Sect. 27.11); $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$ are *a priori restrictions* for parameters $\boldsymbol{\beta}$, which have form of *linear equalities* ($\mathbf{R}$ is a known matrix $p \times (k+1)$, $\mathbf{r}$ is a known vector $p \times 1$, $h(\mathbf{X}) = p < k+1$)

$$\mathbf{b}^* = \mathbf{b} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\left(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right)^{-1}(\mathbf{r} - \mathbf{R}\mathbf{b}), \text{where } \mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

(*OLS-estimate* (i.e. by the method of least squares) of parameters $\boldsymbol{\beta}$ in the linear regression model with *a priori restrictions* for parameters $\boldsymbol{\beta}$ in the form of *linear equalities*; the estimate $\mathbf{b}^*$ is unbiased with var($\mathbf{b}$) – var($\mathbf{b}^*$) ≥ 0 (see Sect. 27.9); in the case of general restrictions, one must apply numerical optimization procedures of estimation)

$$\left.\begin{array}{l}\mathbf{y} = x_1\beta_1 + \boldsymbol{\varepsilon} \\ \beta_d \leq \beta_1 \leq \beta_h\end{array}\right\},$$

where $\mathbf{y} = x_1\beta_1 + \boldsymbol{\varepsilon}$ is the *regression model of line without intercept* (see Sect. 27.11); $\beta_d \leq \beta_1 \leq \beta_h$ is the *a priori restriction for parameter* β_1 (i.e. β_d and β_h are known bounds)

$$b_1^* = \begin{cases}\beta_d \text{ for } b_1 < \beta_d \\ b_1 \text{ for } \beta_d \leq b_1 \leq \beta_h \\ \beta_h \text{ for } b_1 > b_h\end{cases}$$

(*OLS-estimate* of the parameter β_1 in the regression model of line without intercept with the given *a priori restriction for parameter* β_1 (see *thereinbefore*))

28.3 Qualitative Regressors

$$y = \beta_0 + \beta_1 v_1 + \beta_2 v_2 + \beta_3 v_3 + \varepsilon$$

(*qualitative regressors* x_1 with three states A, B, C and x_2 with two states D, E in the linear regression model (a special example): qualitative regressors are "coded" by means of *dummy variables* (*dummies*, *zero-one variables*) v_0, v_1, v_2, v_3; for each

x_1	x_2	v_0	v_1	v_2	v_3
A	E	1	0	0	0
A	F	1	0	0	1
B	E	1	1	0	0
B	F	1	1	0	1
C	E	1	0	1	0
C	F	1	0	1	1

observed value y_i ($i = 1, \ldots, n$) one applies the code corresponding to such states of both regressors, for which this value has been observed (e.g. salaries according to three age classes A, B, C and according to gender D (male), E (female))

$$y = \beta_0 + f(x_1) + ... + f(x_k) + \varepsilon,$$

where $f(x_i)$ is a symbolic expression of the qualitative regressor x_i with k_i states by means of a linear combination of $k_i - 1$ dummies

(*qualitative regressors* $x_1, \ldots, x_k$ in the linear regression model (a general case))

$$y = \beta_0 + \beta_1 v_1 + \beta_2 v_2 + \beta_3 v_3 + \beta_4 v_4 + \beta_5 v_5 + \varepsilon$$

x_1	x_2	v_0	v_1	v_2	v_3	v_4	v_5
A	E	1	0	0	0	0	0
A	F	1	0	0	1	0	0
B	E	1	1	0	0	0	0
B	F	1	1	0	1	1	0
C	E	1	0	1	0	0	0
C	F	1	0	1	1	0	1

(*qualitative regressors* x_1 with three states A, B, C and x_2 with two states D, E in the linear regression model with *interactions* between regressors (a special example): unlike the previous example, the influence of x_1 on y and influence of x_2 on y interact)

28.4 Probit and Logit Models

$$P(y_t = 1 \mid \mathbf{x}_{t\cdot}, \boldsymbol{\beta}) = \Phi(\mathbf{x}_{t\cdot}\boldsymbol{\beta}), \quad P(y_t = 0 \mid \mathbf{x}_{t\cdot}, \boldsymbol{\beta}) = 1 - \Phi(\mathbf{x}_{t\cdot}\boldsymbol{\beta}),$$

where y is a two-state qualitative variable with value 1 (e.g. a family owns a car) or 0 (a family owns no car); Φ is the distribution function of $N(0, 1)$ (see Sect. 26.5) (*probit model*: unlike the qualitative regressors (see Sect. 28.3), now the response variable y has the qualitative (binary) character; moreover, the predicted value $\Phi(\hat{\beta}_0 + \hat{\beta}_1 x_1 + \ldots + \hat{\beta}_k x_k)$ (or $1 - \Phi(\hat{\beta}_0 + \hat{\beta}_1 x_1 + \ldots + \hat{\beta}_k x_k)$) with estimated parameters $\boldsymbol{\beta}$ is interpreted as the probability that the given values of regressors induce the value 1 (or 0) of the response variable, respectively)

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta} \in \mathrm{R}^{k+1}}{\arg\max} \prod_{t=1}^{T} \left(\Phi(\mathbf{x}_{t\cdot}\boldsymbol{\beta})\right)^{y_t} \left(1 - \Phi(\mathbf{x}_{t\cdot}\boldsymbol{\beta})\right)^{1-y_t},$$

(*maximum likelihood estimate* of parameters in *probit model* (see Sect. 27.9))

$$\mathrm{P}(y_t = 1 \mid \mathbf{x}_{t\cdot}, \boldsymbol{\beta}) = \frac{\mathrm{e}^{\mathbf{x}_{t\cdot}\boldsymbol{\beta}}}{1 + \mathrm{e}^{\mathbf{x}_{t\cdot}\boldsymbol{\beta}}}, \quad \mathrm{P}(y_t = 0 \mid \mathbf{x}_{t\cdot}, \boldsymbol{\beta}) = \frac{1}{1 + \mathrm{e}^{\mathbf{x}_{t\cdot}\boldsymbol{\beta}}},$$

where y is a two-state qualitative variable with value 1 or 0
(*logit model*: unlike the probit model, the logit model uses the distribution function of the logistic distribution $\mathrm{e}^{\mathbf{x}_{t\cdot}\boldsymbol{\beta}}/(1+\mathrm{e}^{\mathbf{x}_{t\cdot}\boldsymbol{\beta}})$; the maximum likelihood estimation looks analogously as in the probit model)

28.5 Random Regressors and Instrumental Variable Estimation

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

where $\mathbf{y} = (y_1, \ldots, y_n)'$;
$\mathbf{X} = (x_{ij})_{\substack{i=1,\ldots,n \\ j=0,1,\ldots,k}}$ $(x_{i0} \equiv 1)$; $\mathbf{X}$ is a matrix of $(k + 1)$-variate (row) random vectors

$\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)'$; $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$; $\mathrm{var}(\boldsymbol{\varepsilon}) = \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon\varepsilon}} = \sigma^2\mathbf{I}$

(*linear regression model with random regressors*: in practice, $\mathbf{X}$ has frequently the form of $(k + 1)$-variate stochastic process (i.e. the index i is interpreted as the time index t, see Sect. 30.1) with a regular matrix of the second moments, which is constant in time; special cases corresponding to various types of correlations between $\mathbf{X}$ and $\boldsymbol{\varepsilon}$ can be considered)

$$\mathbf{b}^* = (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{y}; \quad \mathbf{S}_{\mathbf{b}^*\mathbf{b}^*} = s^{2*}(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1} \text{ for } s^{2*} = \frac{\boldsymbol{\varepsilon}^{*\prime}\boldsymbol{\varepsilon}^*}{n-k-1};$$

$$\boldsymbol{\varepsilon}^* = \mathbf{y} - \mathbf{X}\mathbf{b}^*,$$

where $\mathbf{Z}$ is a matrix $n \times (k + 1)$ of observed *instrumental variables* $z_0, z_1, \ldots, z_k$, which are simultaneously uncorrelated with the residual ε and significantly correlated with the original regressors $x_0, x_1, \ldots, x_k$

($\mathbf{b}^*$ is obtained by the *method of instrumental variables* (*IV-estimate*) including estimation of its covariance matrix and parameter σ^2: IV-estimate is consistent (see Sect. 27.9); in practice, the instrumental variables may be found using practical considerations)

$$y_t = \alpha + \beta\lambda x_{t-1} + \beta\lambda^2 x_{t-2} + \beta\lambda^3 x_{t-3} + \ldots + \varepsilon_t, \quad t = 0, \ldots, T \quad (0 < \lambda < 1)$$

(*distributed lag model*: is a special case of the linear regression model with random regressors, which have the form of time lags of variables x)

$$y_t = \alpha(1-\lambda) + \lambda y_{t-1} + \beta x_t + (\varepsilon_t - \lambda\varepsilon_{t-1}), \quad t = 1, \ldots, T$$

(*Koyck's transformation*: consists in subtracting the equation for time t–1 multiplied by the parameter λ from the equation for time t in the distributed lag model; there are recommended procedures, how to estimate the model after such a transformation)

28.6 Simultaneous Equation Models and 2SLS-Estimator

$$\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_m \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 & \ldots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{X}_m \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_m \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_m \end{pmatrix},$$

$$\mathrm{var}(\boldsymbol{\varepsilon}) = \boldsymbol{\Omega} = \begin{pmatrix} \sigma_{11}\mathbf{I} & \sigma_{12}\mathbf{I} & \ldots & \sigma_{1m}\mathbf{I} \\ \sigma_{21}\mathbf{I} & \sigma_{22}\mathbf{I} & \ldots & \sigma_{2m}\mathbf{I} \\ \vdots & \vdots & & \vdots \\ \sigma_{m1}\mathbf{I} & \sigma_{m2}\mathbf{I} & \ldots & \sigma_{mm}\mathbf{I} \end{pmatrix},$$

where for $\boldsymbol{\varepsilon}_{t\,\cdot} = (\varepsilon_{t1}, \ldots, \varepsilon_{tm})$:

$$\mathrm{E}(\boldsymbol{\varepsilon}_{t\,\cdot}) = \mathbf{0}; \quad E(\boldsymbol{\varepsilon}'_{s\,\cdot}\boldsymbol{\varepsilon}_{t\,\cdot}) = \delta_{st}\boldsymbol{\Sigma}; \quad \boldsymbol{\Sigma} = \left(\sigma_{ij}\right)_{\substack{i=1,\ldots,m \\ j=1,\ldots,m}} > 0$$

(*SUR-model* (seemingly unrelated regressions): is a special case of *multiple linear regression model*, where one considers m linear regression equations (see Sect. 27.11) as a system (one has n observations for each equation); residuals at various time periods are uncorrelated but they are correlated simultaneously at the same time with a fixed correlation structure (across particular equations); the estimation of SUR-system may be organized as Aitken estimate in a generalized linear regression model (see Sect. 27.11))

$$\left.\begin{aligned} &\gamma_{11}y_{t1} + \ldots + \gamma_{m1}y_{tm} + \beta_{11}x_{t1} + \ldots + \beta_{k1}x_{tk} + \varepsilon_{t1} = 0 \\ &\vdots \\ &\gamma_{1m}y_{t1} + \ldots + \gamma_{mm}y_{tm} + \beta_{1m}x_{t1} + \ldots + \beta_{km}x_{tk} + \varepsilon_{tm} = 0 \end{aligned}\right\} \quad t = 1, \ldots, n,$$

where

$$\boldsymbol{\varepsilon}_{t\cdot} = (\varepsilon_{t1}, \dots, \varepsilon_{tm}); \quad \mathrm{E}(\boldsymbol{\varepsilon}_{t\cdot}) = \mathbf{0}; \quad E(\boldsymbol{\varepsilon}'_{s\cdot}\boldsymbol{\varepsilon}_{t\cdot}) = \delta_{st}\boldsymbol{\Sigma}; \quad \boldsymbol{\Sigma} = (\sigma_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,m}} > 0$$
$$\mathbf{x}_{t\cdot} = (x_{t1}, \dots, x_{tk}); \quad \mathrm{E}(\mathbf{x}'_{t\cdot}\mathbf{x}_{t\cdot}) = \boldsymbol{\Sigma}_{\mathbf{xx}} \text{ regular for all } t; \quad \mathrm{E}(\mathbf{x}'_{t\cdot}\boldsymbol{\varepsilon}_{t\cdot}) = \mathbf{0}$$

(*structural form of simultaneous equation model* (*SEM*): is a system of m equations, which describe (using unknown parameters $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$) relations among m *endogenous variables* $y_1, \dots, y_m$ (the endogenous variables are outputs of the system: their number is the same as the number of equations) and k *exogenous variables* $x_1, \dots, x_k$ (the exogenous variables are inputs into the system at the given time); one has n observations for each variable; in addition, the exogenous variables can be either *strictly exogenous* ones (originating outside the system) or *predetermined* ones (originating within the system, but in a past time))

$$\mathbf{Y}\boldsymbol{\Gamma} + \mathbf{X}\mathbf{B} + \mathbf{E} = \mathbf{0},$$

where $\mathbf{Y}$ is $n \times m$ matrix of observed endogenous variables; $\mathbf{X}$ is $n \times k$ matrix of observed exogenous variables; $\mathbf{E}$ is $n \times m$ matrix of residuals; $\boldsymbol{\Gamma}$ is $m \times m$ matrix and $\mathbf{B}$ is $k \times m$ matrix of unknown parameters

(*matrix structural form of simultaneous equation model*)

$$\mathbf{Y} = -\mathbf{X}\mathbf{B}\boldsymbol{\Gamma}^{-1} - \mathbf{E}\boldsymbol{\Gamma}^{-1} = \mathbf{X}\boldsymbol{\Pi} + \mathbf{V},$$

where $\mathbf{V}$ is $n \times m$ matrix of residuals; $\boldsymbol{\Pi}$ is $k \times m$ matrix of unknown parameters

(*matrix reduced form of simultaneous equation model*: describes the explicit dependence of endogenous variables on exogenous ones (one assumes regularity of the matrix $\boldsymbol{\Gamma}$)

$$\mathbf{P} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

(*OLS-estimate of parameters* $\boldsymbol{\Pi}$ *in reduced form of simultaneous equation model*: the estimate $\mathbf{P}$ is consistent (under general assumptions, see Sect. 27.9))

$$\begin{pmatrix} \hat{\mathbf{Y}}'_j\hat{\mathbf{Y}}_j & \hat{\mathbf{Y}}'_j\mathbf{X}_j \\ \mathbf{X}'_j\hat{\mathbf{Y}}_j & \mathbf{X}'_j\mathbf{X}_j \end{pmatrix} \begin{pmatrix} \mathbf{c}_{\cdot j} \\ \mathbf{b}_{\cdot j} \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{Y}}'_j\mathbf{y} \\ \mathbf{X}'_j\mathbf{y} \end{pmatrix}$$

or equivalently

$$\begin{pmatrix} \mathbf{Y}'_j\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}_j & \mathbf{Y}'_j\mathbf{X}_j \\ \mathbf{X}'_j\mathbf{Y}_j & \mathbf{X}'_j\mathbf{X}_j \end{pmatrix} \begin{pmatrix} \mathbf{c}_{\cdot j} \\ \mathbf{b}_{\cdot j} \end{pmatrix} = \begin{pmatrix} \mathbf{Y}'_j\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\ \mathbf{X}'_j\mathbf{y} \end{pmatrix}$$

where $\mathbf{y}_{\cdot j} = \mathbf{Y}_j \boldsymbol{\gamma}_{\cdot j} + \mathbf{X}_j \boldsymbol{\beta}_{\cdot j} + \boldsymbol{\varepsilon}_{\cdot j}$ is jth equation (of the structural form) to be estimated ($\mathbf{y}_j$ a $\mathbf{Y}_j$ are endogenous variables and $\mathbf{X}_j$ are exogenous variables in the jth equation); $\mathbf{Y}_j = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}_j$

(estimates $\mathbf{c}_j$ and $\mathbf{b}_j$ of parameters $\boldsymbol{\gamma}_j$ and $\boldsymbol{\beta}_j$ are obtained by the *method of two-stage least squares* (2-*SLS-estimate*): 2-SLS-estimate is consistent (under general assumptions, see Sect. 27.9); 2-SLS-estimate can be interpreted as the IV-estimate (see Sect. 28.5) constructed by means of the instrumental variables $\mathbf{Y}_j$, which are obtained as estimated endogenous variables obtained by means of OLS-estimation in the reduced form)

Further Reading

Campbell, J.Y., Lo, A.W., MacKinlay, A.C.: The Econometrics of Financial Markets. Princeton University Press, Princeton, NJ (1997)

Davidson, J.: Econometric Theory. Blackwell, Oxford (2000)

Greene, W.H.: Econometric Analysis. Prentice Hall, Englewood Cliffs, NJ (2003)

Heij, C. et al.: Econometric Methods with Applications in Business and Economics. Oxford University Press, Oxford (2004)

Judge, G.G. et al.: Introduction to the Theory and Practice of Econometrics. Wiley, New York (1988)

Plasman, J.: Modern Linear and Nonlinear Econometrics. Springer, Heidelberg (2006)

Chapter 29
Index Numbers

Abstract Chapter 29 provides a basic information on index numbers (mainly price indices) and stock exchange indicators: 29.1. Indices as Instruments of Comparison, 29.2. Indices in Practice, 29.3. Stock Exchange Indicators.

29.1 Indices as Instruments of Comparison

- *Index numbers* (or simply *indices*): serve as instruments of comparison of the same attribute (mostly price or quantity of goods or services at a given time period and in a given region) over time (time series) or over regions (cross-section)

$$p = \frac{Q}{q}$$

(common denotation of *indicators*, various changes of which may be described by means of indices: p is a *price* (more generally, an *intensity* indicator); q is a *quantity* (more generally, an *extension* indicator); Q is a *value*)

$$I_p; \quad I_q; \quad I_Q, \text{ respectively}$$

(*index of price change*; *index of quantity change*; *index of value change*, respectively: the corresponding change occurs during a *time* interval between two time periods $s < t$ denoted usually as 0 (*base period*) and 1 (*current period*); when comparing over regions, one considers changes between two places; a more specific symbol for such indices is I_{st} (or e.g. $I_{p;st}$))

$$(i) \;\; I_{tt} = 1; \quad (ii) \;\; I_{st} = I_{ts} \text{ (change of time)}; \quad (iii) \;\; \prod_{t=0}^{T-1} I_{t,\,t+1} = I_{0T} \text{ (chain rule)}$$

(axiomatic *properties of indices*: the indices frequent in practice (see *thereinafter*) do not fulfil some of these properties; $I_{01}, I_{02}, \ldots, I_{0T}$ are called *base indices*; $I_{01}, I_{12}, \ldots, I_{T-1,T}$ are called *chain indices*)

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_29,

$$I_p = \frac{p_1}{p_0}; \quad I_q = \frac{q_1}{q_0}; \quad I_Q = \frac{Q_1}{Q_0}, \quad (\textit{individual indices})$$

29.2 Indices in Practice

- *Compound indices* (or simply *indices*): are designed to compare how prices, quantities (or other attributes) change (over time or place), when they are taken as a whole for many items; they are usually weighted by means of individual indices

$$I_{\Sigma q} = \frac{\sum_{i=1}^n q_{1i}}{\sum_{i=1}^n q_{0i}}; \quad I_{\Sigma Q} = \frac{\sum_{i=1}^n Q_{1i}}{\sum_{i=1}^n Q_{0i}}; \quad I_{\bar{p}} = \frac{\dfrac{\sum_{i=1}^n Q_{1i}}{\sum_{i=1}^n q_{1i}}}{\dfrac{\sum_{i=1}^n Q_{0i}}{\sum_{i=1}^n q_{0i}}} = \frac{\dfrac{\sum_{i=1}^n p_{1i}q_{1i}}{\sum_{i=1}^n q_{1i}}}{\dfrac{\sum_{i=1}^n p_{0i}q_{0i}}{\sum_{i=1}^n q_{0i}}}$$

$$= \frac{\dfrac{\sum_{i=1}^n Q_{1i}}{\sum_{i=1}^n Q_{1i}/p_{1i}}}{\dfrac{\sum_{i=1}^n Q_{0i}}{\sum_{i=1}^n Q_{0i}/p_{0i}}},$$

where p_{0i} denotes the price of the ith item at time 0 (i =1, ..., n); q_{1i} denotes quantity of the ith item at time 1 ($i = 1, \ldots, n$), and the like

(*aggregated indices*: the $I_{\bar{p}}$ describes a change of the average price of the whole)

$$I_p^L = \frac{\sum_{i=1}^n I_{p_i} p_{0i} q_{0i}}{\sum_{i=1}^n p_{0i} q_{0i}} = \frac{\sum_{i=1}^n p_{1i} q_{0i}}{\sum_{i=1}^n p_{0i} q_{0i}}; \quad I_q^L = \frac{\sum_{i=1}^n I_{q_i} p_{0i} q_{0i}}{\sum_{i=1}^n p_{0i} q_{0i}} = \frac{\sum_{i=1}^n p_{0i} q_{1i}}{\sum_{i=1}^n p_{0i} q_{0i}}$$

(*price* and *quantity Laspeyres index*, respectively: one preserves the base quantity (or the base price, respectively) from the base period to the current one; in general, Laspeyres index does not fulfil the properties (*ii*) and (*iii*), see Sect. 29.1)

$$I_p^L = \sum\nolimits_{i=1}^n w_i \frac{p_{1i}}{p_{0i}}; \quad I_q^L = \sum\nolimits_{i=1}^n w_i \frac{q_{1i}}{q_{0i}},$$

where

$$w_i = \frac{p_{0i} q_{0i}}{\sum\limits_{j=1}^n p_{0j} q_{0j}}$$

(*weighted* form of the price and quantity Laspeyres index, respectively)

$$I_p^P = \frac{\sum_{i=1}^n p_{1i}q_{1i}}{\sum_{i=1}^n \frac{p_{1i}q_{1i}}{I_{p_i}}} = \frac{\sum_{i=1}^n p_{1i}q_{1i}}{\sum_{i=1}^n p_{0i}q_{1i}}; \quad I_q^P = \frac{\sum_{i=1}^n p_{1i}q_{1i}}{\sum_{i=1}^n \frac{p_{1i}q_{1i}}{I_{q_i}}} = \frac{\sum_{i=1}^n p_{1i}q_{1i}}{\sum_{i=1}^n p_{1i}q_{0i}};$$

(*price* and *quantity Paasche index*: one preserves the current quantity (or the current price, respectively) from the base period to the current one; Paasche index can be written the weighted form analogously as Laspeyres index (see *thereinbefore*); Paasche index does not fulfil the properties (*ii*) and (*iii*), see Sect. 29.1)

$$I_p^{LW} = \frac{\sum_{i=1}^n p_{1i}q_i}{\sum_{i=1}^n p_{0i}q_i}; \quad I_q^{LW} = \frac{\sum_{i=1}^n p_i q_{1i}}{\sum_{i=1}^n p_i q_{0i}}$$

(*price* and *quantity Lowe index*, respectively: one preserves a constant "hypothetic" quantity q_i (or a constant "hypothetic" price p_i, respectively) from the base period to the current one; Lowe index fulfils all conditions (*i*), (*ii*) and (*iii*), see Sect. 29.1)

$$I_p^F = \sqrt{I_p^L I_p^P} = \left(\frac{\sum_{i=1}^n p_{1i}q_{0i}}{\sum_{i=1}^n p_{0i}q_{0i}} \frac{\sum_{i=1}^n p_{1i}q_{1i}}{\sum_{i=1}^n p_{0i}q_{1i}} \right)^{1/2}$$
$$I_q^F = \sqrt{I_q^L I_q^P} = \left(\frac{\sum_{i=1}^n p_{0i}q_{1i}}{\sum_{i=1}^n p_{0i}q_{0i}} \frac{\sum_{i=1}^n p_{1i}q_{1i}}{\sum_{i=1}^n p_{1i}q_{0i}} \right)^{1/2}$$

(*price* and *quantity Fisher index*: is the geometric mean of the price (or the quantity, respectively) Laspeyres and Paasche indices; Fisher index fulfils all conditions (*i*), (*ii*) and (*iii*), see Sect. 29.1)

$$I_p^{\text{EM}} = \frac{\sum_{i=1}^n p_{1i}(q_{0i} + q_{1i})}{\sum_{i=1}^n p_{0i}(q_{0i} + q_{1i})}; \quad I_q^{\text{EM}} = \frac{\sum_{i=1}^n (p_{0i} + p_{1i})q_{1i}}{\sum_{i=1}^n (p_{0i} + p_{1i})q_{0i}}$$

(*price* and *quantity Edgeworth-Marshall index*: one applies the arithmetic mean of quantities q_{0i} and q_{1i} (or the arithmetic mean of prices p_{0i} and p_{1i}, respectively) from the base period to the current one)

$$\frac{I_p^P}{I_p^L} = 1 + r_{I_p I_q} V_{I_p} V_{I_q},$$

where V_{I_p} and V_{I_q} is the variation coefficients (see Sect. 27.6) of individual indices I_{p_i} and I_{q_i}, respectively (see Sect. 29.1); $r_{I_p I_q}$ is the correlation coefficient (see Sect. 27.8) between individual indices I_{p_i} and I_{q_i}

(*Bortkiewicz decomposition* for comparison of the price Laspeyres and Paasche index: in practice, one has usually $r_{I_p I_q} < 0$ so that $I_p^L > I_p^P$; Bortkiewicz decomposition can be applied to explain the difference between two weighted means calculated using the same values)

- *Consumer price indices*: are an important example of price indices regularly published by the statistical offices for each national economy, since one usually measures the inflation by means of them; they are denoted by specific abbreviations, e.g.
  - *CPI* (Consumer Price Index) in U.S.
  - *RPI* (Retail Price Index) in UK, and the like

29.3 Stock Exchange Indicators

- *Stock exchange indicators* (*stock index*, *market indicators*) are constructed in a similar way as the price Laspeyres indices using prices of assets on a given market (mostly stocks, see Sect. 9.3)

$$I_t = C_{t_0} \sum_{i=1}^{n} w_{ti}\, p_{ti},$$

where

$$C_{t_0} = \frac{I_{t_0}}{I'_{t_0}}; \quad w_{ti} \geq 0; \quad \sum_{i=1}^{n} w_{ti} = 1$$

(construction of *stock exchange indicator* at time t: is a weighted mean of prices p_{ti} of corresponding assets at time t (comp. with the weighted form of price Laspeyres index in Sect. 29.2); the weight w_{ti} corresponds mostly to the relative turnover of the ith asset at time t on the market (the capital, financial, commodity one); C_{t_0} is the *continuity factor*, which enables to adjust a possible jump at time t_0 due to methodology changes, mergers, replacements of old representatives by new ones, and the like (the continuity factor is also applied just in the beginning of construction of the indicator to start with an initial value $I_0 = 100$ or 1,000, and the like))

$$\mathrm{DJIA}_t = \frac{1}{D_t} \sum_{i=1}^{30} p_{ti},$$

where e.g. $D_{1928} = 30$; $D_{1991} = 0.559$

(*Dow Jones Index DJIA* (Dow Jones Industrial Average): is one of the oldest and most renowned stock exchange indicators recorded since 1897; the *divisor* D_t has the function of continuity factor as well (see *thereinbefore*); DJIA was initially an ordinary arithmetic mean of thirty stock titles with $D_t = 30$)

Further Reading

Chance, W.A.: A note on the origins of index numbers. The Review of Economics and Statistics 48, 108–110 (1966)
Diewert, W.E.: Index numbers. In: Diewert, W.E., Nakamura, A.O. (eds.) Essays in Index Number Theory, Chapter 5. Elsevier, Amsterdam (1993)
Dupacova, J., Hurt, J., Stepan, J.: Stochastic Modeling in Economics and Finance. Kluwer, Dordrecht (2002)
Sears, R.S., Trennenpohl, G.L.: Investment Management. The Dryden Press, Hinsdale, IL (1993)

Chapter 30
Stochastic Processes

Abstract Chapter 30 deals with methodology of stochastic processes (see also time series in Chap. 31): 30.1. Classification and Basic Characteristics of Stochastic Processes, 30.2. Markov Chains, 30.3. Markov Processes, 30.4. Important Stochastic Processes, 30.5. Spectral Properties of Stochastic Processes.

30.1 Classification and Basic Characteristics of Stochastic Processes

$$\{X_t, t \in T\}, \quad \text{where } T \subset R$$

(univariate *stochastic* (or *random*) *process*: is a family of random variables defined on the same probability space (Ω, $\boldsymbol{B}$, P) (see Sect. 26.1) and indexed by means of *time* indices t from a set T; according to the form of T one distinguishes:
- *stochastic process in continuous time*: T is an interval on the real line, e.g. $T = \langle 0, \infty)$ (i.e. $\{X_t, t \geq 0\}$)
- *stochastic process in discrete time* (*time series*): T consists of discrete real values, e.g. $T = \mathrm{N}_0$ (i.e. $\{X_0, X_1, \ldots\}$) or $T = \mathrm{Z}$ (i.e. $\{\ldots, X_{-1}, X_0, X_1, \ldots\}$)

According to the *state space S* of values of the random variables X_t one distinguishes:
- *stochastic process with discrete states*: e.g. *counting process* $X_t \in \mathrm{N}_0$, which registers a number of defined events in time
- *stochastic process with continuous states*: e.g. $X_t \in \langle 0, \infty)$
- *real* (or *complex*) *stochastic process*: the corresponding state space is the real line (or the complex plane, respectively)
- *multivariate stochastic process*: X_t are n-variate random vectors $\mathbf{X}_t$ (see Sect. 26.6)

$$\{X_t(\omega), \ t \in T\} \text{ for fixed } \omega \in \Omega$$

(*trajectory* (realization) of stochastic process: is the deterministic function of time, which is observed in the given sample)

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_30,

$$F_{t_1,\dots,t_n}(x_1,\dots,x_n) = P(X_{t_1} \le x_1,\dots,X_{t_n} \le x_n), \quad n \in \mathrm{N};$$
$$t_1,\dots,t_n \in T;\ x_1,\dots,x_n \in \mathrm{R}$$

(*distribution function* of stochastic process (see Sect. 26.6))

$$\left.\begin{array}{r} F_{t_{i_1},\dots,t_{i_n}}(x_{i_1},\dots,x_{i_n}) = F_{t_1,\dots,t_n}(x_1,\dots,x_n) \text{ for arbitrary} \\ \text{permutation } i_1,\dots,i_n \text{ of } 1,\dots,n \\ \lim\limits_{x_{n+1}\to\infty} F_{t_1,\dots,t_n,t_{n+1}}(x_1,\dots,x_n,x_{n+1}) = F_{t_1,\dots,t_n}(x_1,\dots,x_n) \end{array}\right\}$$

(*consistent system of distribution functions*: fulfils the properties given above for all $n \in \mathrm{N}$; $t_1,\dots,t_n \in T$; $x_1,\dots,x_n \in \mathrm{R}$; a system of distribution functions is consistent if and only if it corresponds to a stochastic process (*Kolmogorov theorem*))

$\mu_t = \mathrm{E}(X_t),\ t \in T,$ where $\mathrm{E}\,|X_t| < \infty,\ t \in T$ (*mean value* of stochastic process)

$$R(s,t) = \mathrm{cov}(X_s, X_t) = \mathrm{E}\left((X_s - \mu_s)(X_t - \mu_t)\right), \quad s,t \in T,$$

where $\mathrm{E}\left(X_t^2\right) < \infty;\ t \in T$
(*autocovariance function* of stochastic process: in particular, $R(t, t) = \mathrm{var}(X_t)$ is *variance* of stochastic process)

$$(i)\ |R(s,t)| \le \sqrt{R(s,s)R(t,t)}; \quad (ii)\ R(s,t) = R(t,s); \quad (iii)\ \sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j R(t_i,t_j) \ge 0,$$
$$n \in \mathrm{N}; c_1,\dots,c_n \in \mathrm{R}; t_1,\dots,t_n \in T$$

(*properties* of autocovariance function: (*ii*) is *symmetry* and (*iii*) is *non-negative definiteness*)

$$B(s,t) = \rho(X_s, X_t) = \frac{R(s,t)}{\sqrt{R(s,s)\cdot R(t,t)}}, \quad s,\ t \in T, \quad \text{where } \mathrm{E}\left(X_t^2\right) < \infty;\ t \in T$$

(*autocorrelation function* of stochastic process:$|B(s, t)| \le 1$; in particular, it holds $B(t, t) = 1$)

$$F_{t_1,\dots,t_n}(x_1,\dots,x_n) = F_{t_1+h,\dots,t_n+h}(x_1,\dots,x_n)$$

for all $n \in \mathrm{N}$; $x_1,\dots,x_n \in \mathrm{R}$; $t_1,\dots,t_n, t_1+h,\dots,t_n+h \in T$

(*strict stationarity* of stochastic process: means that finite probability distributions of the process are invariant with respect to shifts in time)

$$\left.\begin{array}{l} \mu_t = \mu \\ R(s,t) = R(t-s) \end{array}\right\} \quad s,t \in T;$$

in particular,

$$R(0) = \text{var}(X_t); \quad R(t) = R(-t); \quad B(t) = \frac{R(t)}{R(0)}$$

(*stationarity* (or more exactly, *weak stationarity*) of stochastic process: means that its mean value and autocovariance function are invariant with respect to shifts in time)

$$\mathbf{R}(s,t) = \big(R_{ij}(s,t)\big)_{\substack{i=1,\ldots,n\\ j=1,\ldots,n}} = \boldsymbol{\Sigma}_{\mathbf{x}_s\mathbf{x}_t}, \quad s,t \in T$$

(autocovariance function of n-variate stochastic process $\mathbf{X}_t = (X_{t1}, \ldots, X_{tn})'$: $R_{ij}(t)$ is *mutual covariance function* of processes $\{X_{ti}\}$ and $\{X_{tj}\}$; in the case of stationarity, it holds $\mathbf{R}(s, t) = \mathbf{R}(t - s)$; $R_{ij}(-t) = R_{ji}(t)$)

$$\mathbf{B}(s,t) = \big(B_{ij}(s,t)\big)_{\substack{i=1,\ldots,n\\ j=1,\ldots,n}} = \mathbf{R}_{\mathbf{x}_s\mathbf{x}_t} = \mathbf{D}(s,s)^{-1/2}\mathbf{R}(s,t)\mathbf{D}(t,t)^{-1/2}, \quad s,t \in T,$$

where $\mathbf{D}(t, t) = \mathbf{diag}\{R_{11}(t), \ldots, R_{nn}(t)\} = \mathbf{diag}\{\text{var}(X_{t1}), \ldots, \text{var}(X_{tn})\}$

(autocorrelation function of n-variate stochastic process $\mathbf{X}_t = (X_{t1}, \ldots, X_{tn})'$; in the case of stationarity, it holds $\mathbf{B}(s, t) = \mathbf{B}(t - s)$; $B_{ij}(-t) = B_{ji}(t)$)

30.2 Markov Chains

$$P(X_{n+1} = j \mid X_n = i,\ X_{n-1} = i_{n-1}, \ldots,\ X_0 = i_0) = P(X_{n+1} = j \mid X_n = i)$$

for all $n \in \mathrm{N}_0$; $i, j, i_0, \ldots, i_{n-1} \in S = \mathrm{N}_0 = \{0, 1, \ldots\}$ (one assumes existence of the given conditional probabilities, see Sect. 26.2)

(*Markov property*: the stochastic process $\{X_0, X_1, \ldots\}$ in discrete time, with discrete states $i, j, \ldots$ (i.e. with the discrete state space $S = \mathrm{N}_0$) and with the Markov property is called *Markov chain*; all that matters in determining the next state X_{n+1} of the process is its current state X_n, and it does not matter how the process got to X_n; the probability of X_{n+1} being in state j, given that X_n is in state i, is called *transition probability*)

$$p_{ij} = P(X_{n+1} = j \mid X_n = i) \quad \text{for all } n \in \mathrm{N}_0;\ i, j \in \mathrm{N}_0$$

(*homogeneous* Markov chain: its transition probabilities do not depend on the time n)

$$p_{ij}(n) = P(X_{k+n} = j \mid X_k = i), \quad \text{where } k, n \in \mathrm{N}_0;\ i, j \in \mathrm{N}_0$$

(*n-step* transition probabilities: $p_{ij}(0) = \delta_{ij}$, $p_{ij}(1) = p_{ij}$)

$$p_i = P(X_0 = i), \quad \text{where } i \in \mathrm{N}_0$$

(*initial probability distribution* of Markov chain: $\mathbf{p} = (p_0, p_1, \ldots)'$)

$$p_i(n) = P(X_n = i), \quad \text{where } n \in \mathrm{N}_0;\ i \in \mathrm{N}_0$$

(*probability distribution* of Markov chain at time n: $\mathbf{p}(n) = (p_0(n), p_1(n), \ldots)'$; $\mathbf{p}(0) = \mathbf{p}$)

$$\mathbf{P} = \left(p_{ij}\right); \quad \mathbf{P}(n) = \left(p_{ij}(n)\right), \quad \text{where } n \in \mathrm{N}_0$$

(*transition matrix*: is a *random matrix*, i.e. a matrix of non-negative elements with row sums equal to unities; $\mathbf{P}(0) = \mathbf{I}$; $\mathbf{P}(1) = \mathbf{P}$)

$$p_{ij}(n_1 + n_2) = \sum\nolimits_{k=0}^{\infty} p_{ik}(n_1) \cdot p_{kj}(n_2),$$

i.e. in the matrix form $\mathbf{P}(n_1 + n_2) = \mathbf{P}(n_1) \cdot \mathbf{P}(n_2)$
(*Chapman-Kolmogorov equality*)

$$\mathbf{P}(n) = \mathbf{P}^n;\ p_i(n) = \sum\nolimits_{j=0}^{\infty} p_j(0) \cdot p_{ji}(n), \text{ i.e. in the matrix form } \mathbf{p}(n)' = \mathbf{p}' \cdot \mathbf{P}^n$$

(properties of transition probabilities)

$$\hat{p}_{ij} = \frac{n_{ij}}{\sum_{k=0}^{m} n_{ik}},$$

where n_{ij} is the number of observed transitions from state i to state j

(maximum likelihood estimate (see Sect. 27.9) of transition probabilities in a homogeneous Markov chain with finite state space $S = \{0, 1, \ldots, m\}$)

$$\tau_j = \inf\{n > 0:\ X_n = j\} \quad (\textit{hitting time} \text{ of state } j)$$

$$P(\tau_j < \infty \mid X_0 = j) = 1$$

(*recurrent* (or *persistent*) state j of homogeneous Markov chain: the chain starting in j will return to j after a finite number of steps with probability one)

$$P(\tau_j = \infty \mid X_0 = j) > 0$$

(*transient* state j of homogeneous Markov chain: given that the chain starts in j, there is a non-zero probability that it will never return to j)

$$\mathrm{E}(\tau_j \mid X_0 = j) < \infty; \quad \mathrm{E}(\tau_j \mid X_0 = j) = \infty, \text{ respectively}$$

(*non-null* and *null recurrent* state j of homogeneous Markov chain, respectively)

$$p_{jj}(k \cdot d_j) > 0 \quad \text{for all}\, k \in \mathrm{N},$$

where $d_j \in \mathrm{N}$ is the smallest number with this property
(*periodic* state j with *period* d_j of homogeneous Markov chain: if such a number d_j does not exist or is equal to one, then j is *aperiodic*)

$$\sum\nolimits_{n=0}^{\infty} p_{jj}(n) = \infty, \quad \text{if and only if the state}\, j\, \text{is recurrent}$$

$$\lim_{n \to \infty} p_{ij}(n) = 0, \text{where the state}\, j \text{ is null recurrent or transient}$$

$$\lim_{n \to \infty} p_{ij}(n) = \frac{P(\tau_j < \infty \mid X_0 = i)}{\mathrm{E}(\tau_j \mid X_0 = j)},$$

where the state j is non-null recurrent and aperiodic

$$\lim_{k \to \infty} p_{jj}(k \cdot d_j) = \frac{d_j}{\mathrm{E}(\tau_j \mid X_0 = j)},$$

where the state j is non-null recurrent and periodic with period d_j

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} p_{ij}(k) = \frac{P(\tau_j < \infty \mid X_0 = i)}{\mathrm{E}(\tau_j \mid X_0 = j)},$$

where the state j is non-null recurrent and periodic

$$p_{ij}(n) > 0 \quad \text{for a number}\, n \in \mathrm{N}_0 \text{ (the state}\, j\, \text{is}\, \textit{accessible}\, \text{from}\, i)$$

$$p_{ij} = 0 \quad \text{for all}\, i \in C\, \text{and}\, j \notin C$$

(*closed* class of states $C \subset S$: the probability of leaving such a class is zero:
- *irreducible* Markov chain contains no closed class of states (with exception of the class S of all states), i.e. one can get to any state from any state in it; otherwise the chain is *reducible*
- all states of the irreducible Markov chain are of the same type (i.e. transient, or null recurrent, or non-null recurrent, and simultaneously aperiodic, or periodic with the same period)
- *absorbing* state j forms a closed class $\{j\}$ with a single state only, i.e. $p_{jj} = 1$)

$$S = T \cup C_1 \cup C_2 \cup \ldots$$

(*decomposition* of state space S of Markov chain: T is the class of transient states and $C_1, C_2, \ldots$ are disjoint closed irreducible classes of recurrent states)

$$\mathbf{P} = \begin{pmatrix} \mathbf{P}_1 & \mathbf{0} \\ \mathbf{A} & \mathbf{B} \end{pmatrix}$$

(the transition matrix of a *finite-state* Markov chain can be permuted to this form ($\mathbf{P}_1$ and $\mathbf{B}$ are square submatrices), if and only if the chain is reducible)

- Finite-state Markov chain has following properties:
  - it is not possible that all states are transient
  - there are no null recurrent states
  - if the chain is irreducible, then all states are non-null recurrent

$\pi_j = \sum_{k=0}^{\infty} \pi_k \cdot p_{kj}$, i.e. in the matrix form $\boldsymbol{\pi}' = \boldsymbol{\pi}' \cdot \mathbf{P}$

(*stationary distribution* of homogeneous Markov chain: such a distribution:
– does not exist in irreducible chain with transient states, or with null recurrent states
– exists uniquely in irreducible chain with non-null recurrent states
– exists in finite-state irreducible chain)

$$\mathbf{p}(n) = \boldsymbol{\pi},$$

where a chain with stationary distribution $\boldsymbol{\pi}$ has the initial distribution $\mathbf{p}(0) = \boldsymbol{\pi}$ (homogeneous Markov chain with such an initial distribution, which is stationary, is the strictly stationary stochastic process (see Sect. 30.1))

$$\lim_{n\to\infty} p_{ij}(n) = \pi_j > 0,$$

i.e. in the matrix form: $\lim_{n\to\infty} \mathbf{P}^n = \boldsymbol{\Pi} = \begin{pmatrix} \boldsymbol{\pi}' \\ \boldsymbol{\pi}' \\ \vdots \end{pmatrix}$; $\lim_{n\to\infty} p_j(n) = \pi_j > 0$,

where an irreducible chain has non-null recurrent and aperiodic states

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_{ij}(k) = \pi_j > 0; \quad \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_j(k) = \pi_j > 0,$$

where an irreducible chain has non-null recurrent and periodic states

$$\mathbf{Z} = \left(z_{ij}\right), \quad \text{where } i, j \in \mathrm{N}_0$$

(*matrix of transition rewards* in homogeneous *Markov reward chain*: z_{ij} for $i, j \in \mathrm{N}_0$ is the reward (the profit or loss) due to the transition from the state i to the state j within a time unit)

$$q_i = \sum_{j=0}^{\infty} z_{ij} \cdot p_{ij}$$

(*expected reward* due to a transition from the state $i \in \mathrm{N}_0$ within a time unit)

$$\mathbf{v}(n) = \mathbf{q} + \mathbf{P} \cdot \mathbf{v}(n-1) = \sum_{k=0}^{n-1} \mathbf{P}^k \mathbf{q} \approx (n-1)\boldsymbol{\Pi}\mathbf{q} + (\mathbf{I} - (\mathbf{P} - \boldsymbol{\Pi}))^{-1}\mathbf{q},$$

$n \in \mathrm{N}$; $\mathbf{q} = (q_0, \ldots, q_m)' = \mathbf{v}(1)$; $\boldsymbol{\Pi}$ is matrix with the same rows $\boldsymbol{\pi}'$ (see *thereinbefore*)

(vector of *expected rewards* $\mathbf{v}(n) = (v_0(n), \ldots, v_m(n))'$ within n time units in an irreducible finite-state homogeneous Markov reward chain with non-null recurrent and aperiodic states $S = \{0, 1, \ldots, m\}$)

$$\mathbf{v}(n) = \mathbf{q} + \beta \cdot \mathbf{P} \cdot \mathbf{v}(n-1) = \sum_{k=0}^{n-1} \beta^k \mathbf{P}^k \mathbf{q}$$

$$\lim_{n\to\infty} \mathbf{v}(n) = (\mathbf{I} - \beta \cdot \mathbf{P})^{-1}\mathbf{q}, \quad 0 < \beta < 1 \text{ is discount factor}$$

(vector of *discounted* expected rewards: $\beta^k z_{ij}$ is the discounted reward due to the transition from the state i at time k to the state j at time $(k+1)$ discounted to the initial time 0)

30.3 Markov Processes

$$P(X_{s+t} = j \mid X_s = i, X_{t_n} = i_n, \ldots, X_{t_1} = i_1) = P(X_{s+t} = j \mid X_s = i)$$

for all $n \in \mathrm{N}$; $i, j, i_1, \ldots, i_n \in S = \mathrm{N}_0 = \{0, 1, \ldots\}$; $0 \leq t_1 < \ldots < t_n < s \leq s + t$ (one assumes existence of the given conditional probabilities, see Sect. 26.2)

(*Markov property*: the stochastic process $\{X_t\}$ in continuous time, with discrete states $i, j, \ldots$ (i.e. with the discrete state space $S = \mathrm{N}_0$) and with the Markov property is called *Markov process with discrete states*; all that matters in determining the future state X_{s+t} of the process is its current state X_s, and it does not matter how the process got to X_s; the probability of X_{s+t} being in state j, given that X_s is in state i, is called *transition probability*)

$$p_{ij}(t) = P(X_{s+t} = j \mid X_s = i) \quad \text{for all } s \geq 0;\ i, j \in \mathrm{N}_0$$

(*homogeneous* Markov process with discrete states: its transition probabilities do not depend on the time s (they depend only on the time distance t between corresponding time periods); $p_{ij}(0) = \delta_{ij}$)

$$p_i = P(X_0 = i), \quad \text{where } i \in \mathrm{N}_0$$

(*initial probability distribution* of Markov process with discrete states: $\mathbf{p} = (p_0, p_1, \ldots)'$)

$$p_i(t) = P(X_t = i), \quad \text{where } t \geq 0;\ i \in \mathrm{N}_0$$

(*probability distribution* of Markov process with discrete states at time t:
$\mathbf{p}(t) = (p_0(t), p_1(t), \ldots)'$; $\mathbf{p}(0) = \mathbf{p}$)

$$\mathbf{P}(t) = \big(p_{ij}(t)\big), \quad \text{where } t \geq 0$$

(*transition matrix*: is a *random matrix*, i.e. a matrix of non-negative elements with row sums equal to unities; $\mathbf{P}(0) = \mathbf{I}$)

$$p_{ij}(t_1 + t_2) = \sum_{k=0}^{\infty} p_{ik}(t_1) \cdot p_{kj}(t_2),$$

i.e. in the matrix form $\mathbf{P}(t_1 + t_2) = \mathbf{P}(t_1) \cdot \mathbf{P}(t_2)$
(*Chapman-Kolmogorov equality*)

$$p_i(t) = \sum_{j=0}^{\infty} p_j(0) \cdot p_{ji}(t), \quad \text{i.e. in the matrix form } \mathbf{p}(t)' = \mathbf{p}' \cdot \mathbf{P}(t)$$

(property of transition probabilities)

$$q_{ii} = \lim_{h \to 0+} \frac{p_{ii}(h) - 1}{h}; \quad q_{ij} = \lim_{h \to 0+} \frac{p_{ij}(h)}{h}, \quad \text{where } i \neq j$$

(*transition intensities*: fulfil $p_{ii}(h) = 1 + q_{ii} \cdot h + o(h)$; $p_{ij}(h) = q_{ij} \cdot h + o(h)$, where $i \neq j$)

$\mathbf{Q} = \big(q_{ij}\big)$ (*intensity matrix* of homogeneous Markov process with discrete states)

$$p'_{ij}(t) = \sum_{k=0}^{\infty} p_{ik}(t) \cdot q_{kj},$$

where $t > 0$; $i, j \in \mathrm{N}_0$, i.e. in the matrix form $\mathbf{P}'(t) = \mathbf{P}(t) \cdot \mathbf{Q}$
(*prospective Kolmogorov differential equations* for transition probabilities: hold in the homogeneous Markov process with discrete states under general assumptions; the derivatives $p'_{ij}(t)$ are explained by means of all the transition probabilities from the state i)

$$p'_{ij}(t) = \sum_{k=0}^{\infty} q_{ik} \cdot p_{kj}(t),$$

where $t > 0$; $i, j \in \mathrm{N}_0$, i.e. in the matrix form $\mathbf{P}'(t) = \mathbf{Q} \cdot \mathbf{P}(t)$
(*retrospective Kolmogorov differential equations* for transition probabilities: the derivatives $p'_{ij}(t)$ are explained by means of all the transition probabilities to the state j)

$$p'_i(t) = \sum_{k=0}^{\infty} p_k(t) \cdot q_{ki}, \quad \text{where } t > 0;\ i \in \mathrm{N}_0$$

(Kolmogorov differential equations for probability distribution of the process)

$$\pi_j = \sum_{k=0}^{\infty} \pi_k \cdot p_{kj}(t), \quad \text{i.e. in the matrix form } \boldsymbol{\pi}' = \boldsymbol{\pi}' \cdot \mathbf{P}(t), \quad \text{where } t \geq 0$$

(*stationary distribution* of homogeneous Markov process with discrete states)

$$\lim_{t\to\infty} p_{ij}(t) = \lim_{t\to\infty} p_j(t) = \pi_j; \sum_{k=0}^{\infty} \pi_k \cdot q_{ki} = 0 \;\; (\text{tj.}\boldsymbol{\pi}'\cdot\mathbf{Q} = \mathbf{0}') \;\; \text{for all } i, j \in \mathrm{N}_0$$

(properties of stationary distribution: hold under general assumptions)

$$\mathbf{p}(t) = \boldsymbol{\pi},$$

where a process with stationary distribution $\boldsymbol{\pi}$ has the initial distribution $\mathbf{p}(0) = \boldsymbol{\pi}$

(homogeneous Markov process with discrete states and with such an initial distribution, that it is the stationary distribution, is the strictly stationary stochastic process (see Sect. 30.1))

$$P\left(X_{s+t} \leq x \,\middle|\, (X_\tau)_{\tau\leq s}\right) = P(X_{s+t} \leq x \,|X_s)$$

for all $x \in \mathrm{R}$; $0 \leq s \leq s + t$ (one assumes existence of the given conditional probabilities, see Sect. 26.2); $P(\cdot|X_t)$ is the conditional probability given σ-algebra generated by the random variable X_t (see Sect. 26.3)

(*Markov property*: the stochastic process $\{X_t, t \geq 0\}$ in continuous time, with continuous states (i.e. with the state space $S = \mathrm{R}$) and with the Markov property is called *Markov process with continuous states*; all that matters in determining the future state X_{s+t} of the process is its current state X_s, and it does not matter how the process got to X_s; the probabilities in the Markov property are *transition probabilities* from time s to time $s + t$ for events of the type $X_{s+t} \leq x$, but one can also consider more general events of the type $X_{s+t} \in B$ for arbitrary Borel sets B on the real line)

30.4 Important Stochastic Processes

$$\{Y_t,\, t \in N\}, \text{ where } P(Y_t = 1) = P(Y_t = -1) = 1/2$$

(*binary process*: its trajectory is a record of results when tossing an ideal coin)

$\{X_t,\, t \in \mathrm{N}_0\}$, where $X_0 = 0$; $X_t = \sum_{i=1}^{t} Y_i$, $t \in \mathrm{N}$; $\{Y_t,\, t \in \mathrm{N}\}$ is the binary process

(*symmetric random walk on line*: is a homogeneous Markov chain in discrete time with the state space $S = \mathrm{Z} = \{\ldots, -1, 0, 1, \ldots\}$, where

$$p_{ij} = \begin{cases} 1/2 & \text{for } j = i \pm 1 \\ 0 & \text{otherwise} \end{cases}; \quad p_i = \begin{cases} 1 & \text{for } i = 0 \\ 0 & \text{for } i \neq 0 \end{cases} \quad \text{for all } i, j \in \mathrm{Z};$$

X_t at time $t \in \mathrm{N}_0$ describes the position of a particle that moves (from the starting position in the origin) along integer points on the line in each step with the same probabilities in both directions)

$$\{X_t,\ t\ \in\ \mathrm{N}_0\},$$

where $X_0 = 1;\ X_{t+1} = \sum_{j=1}^{X_t} Z_{tj},\ \ t \in \mathrm{N}_0$; Z_{tj} are *iid* random variables with values in N_0

(*branching process* (*Galton-Watson process*): is a homogeneous Markov chain in discrete time with the state space $S = \mathrm{N}_0 = \{0, 1, \ldots\}$; X_t at time $t \in \mathrm{N}_0$ describes the number of members of tth generation (the initial 0th generation has only one member), where the jth member of tth generation gives rise to a random number Z_{tj} of members (descendants) of $(t + 1)$th generation)

$$\{N_t,\ t\ \geq\ 0\},\ \text{where}$$

$$\begin{cases} (i)\quad N_0 = 0 \\ (ii)\quad N_{t_2} - N_{t_1}, \ldots,\ N_{t_n} - N_{t_{n-1}}\quad \text{are independent for arbitrary } 0 \leq t_1 < \ldots < t_n \\ (iii)\ N_t - N_s \sim P\,(\lambda \cdot (t - s))\quad \text{for arbitrary } 0 \leq s < t \end{cases}$$

(*Poisson process with intensity* $\lambda > 0$ (see also Sect. 15.2): is a homogeneous Markov process in continuous time with the state space $S = \mathrm{N}_0 = \{0, 1, \ldots\}$, where

$$p_{ij}(h) = \begin{cases} \lambda \cdot h + o(h) & \text{for } j = i + 1 \\ 1 - \lambda \cdot h + o(h) & \text{for } j = i \\ o(h) & \text{for } j > i + 1 \\ 0 & \text{for } j < i \end{cases};$$

N_t at time $t \geq 0$ describes the number of occurrences of an observed event (e.g. insurance claims) in the time interval $\langle 0, t\rangle$ (in the interval of a small length h the given event occurs just once with probability $\lambda \cdot h + o(h)$ (which is proportional approximately to the length h) and more than once with probability $o(h)$); in particular, it is

$$p_i(t) = P(N_t = i) = e^{-\lambda\,\cdot\,t}\frac{(\lambda \cdot t)^i}{i\,!},\quad t\ \geq\ 0;\ i\ \in\ \mathrm{N}_0$$

(see Poisson distribution in Sect. 26.4 with mean value $\lambda \cdot t$); the times between particular occurrences of the given event are *iid* random variables with distribution $Exp(\lambda)$ (see exponential distribution in Sect. 26.5 with mean value $1/\lambda$); an efficient estimate of the process intensity is $\hat{\lambda} = n/T$, where n is the observed number of occurrences of the given event within the time T)

$$\{X_t,\ t\ \geq\ 0\},$$

(*Yule process* (*linear birth process*) with parameter $\lambda > 0$: is a homogeneous Markov process in continuous time with the state space $S = \mathrm{N}_0 = \{0, 1, \ldots\}$, where

$$p_{ij}(h) = \begin{cases} i \cdot \lambda \cdot h + o(h) & \text{for } j = i + 1 \\ 1 - i \cdot \lambda \cdot h + o(h) & \text{for } j = i \\ o(h) & \text{for } j > i + 1 \\ 0 & \text{for } j < i \end{cases};$$

X_t at time $t \geq 0$ describes a population size with given initial size $X_0 = k_0$ (in the interval of a small length h each individual gives to rise to just one individual with probability $\lambda \cdot h + o(h)$ and to more individuals with probability $o(h)$ independently of behaviour of other individuals); in particular, it is

$$p_i(t) = \begin{cases} \binom{i-1}{i-k_0} e^{-k_0 \cdot \lambda \cdot t} \left(1 - e^{-\lambda \cdot t}\right)^{i-k_0} & \text{for } i \geq k_0 \\ 0 & \text{for } i < k_0 \end{cases};$$

similarly one defines *general birth process* (it has a general term λ_i instead of the linear term $i \cdot \lambda$) and *linear* or *general birth-and-death process* (here the individuals originate and die away with given probabilities))

$\{\varepsilon_t,\ t \in Z\}$, where $\mathrm{E}(\varepsilon_t) = 0$; $\mathrm{cov}(\varepsilon_s;\ \varepsilon_t) = \delta_{st} \cdot \sigma^2\ (\sigma^2 > 0)$, $s, t \in \mathrm{Z}$ (*white noise* (*WN*): is a (weakly) stationary stochastic process (see Sect. 30.1) in discrete time with state space $S = \mathrm{R}$ (ε_t are mutually uncorrelated with zero mean value and constant variance $\sigma^2 > 0$); moreover, if ε_t are *iid*, then the white noise is strictly stationary; autocovariance function $R(t)$ (see Sect. 30.1) of the white noise is

$$R(t) = \begin{cases} \sigma^2 & \text{for } t = 0 \\ 0 & \text{for } t \neq 0 \end{cases}$$

$$X_t = \varepsilon_t + \theta_1\, \varepsilon_{t-1} + \ldots + \theta_q\, \varepsilon_{t-q},\ t \in \mathrm{Z},$$

where $\{\varepsilon_t\}$ is white noise with variance σ^2; $\theta_q \neq 0$

(*moving average process* of order q denoted as $MA(q)$: is a (weakly) stationary stochastic process in discrete time with the state space $S = \mathrm{R}$, zero mean value and autocovariance function (see Sect. 30.1)

$$R(t) = \begin{cases} \sigma^2 \sum_{i=0}^{q-|t|} \theta_{i+|t|}\, \theta_i & \text{for } |t| = 0,\ 1, \ldots,\ q \\ 0 & \text{for } |t| > q \end{cases};$$

if the polynomial $\lambda^q + \theta_1 \lambda^{q-1} + \ldots + \theta_q$ has all roots inside the unit circle in the complex plane, then the process $MA(q)$ is *invertible*, i.e. predictable using the autoregressive form $X_t = \pi_1\, X_{t-1} + \pi_2\, X_{t-2} + \ldots + \varepsilon_t$ (see Sect. 31.4))

$$X_t = \varphi_1 X_{t-1} + \ldots + \varphi_p X_{t-p} + \varepsilon_t,\ t \in Z,$$

where $\{\varepsilon_t\}$ is white noise with variance σ^2; $\varphi_p \neq 0$; the polynomial $\lambda^p - \varphi_1 \lambda^{p-1} - \ldots - \varphi_p$ has all roots inside the unit circle in the complex plane (*autoregressive process* of order p denoted as $AR(p)$: is a (weakly) stationary stochastic process in discrete time with the state space $S = \mathrm{R}$, zero mean value and autocovariance function fulfilling *Yule-Walker equations*

$$R(t) = \begin{cases} \varphi_1 R(t-1) + \ \dots \ + \varphi_p R(t-p) & \text{for } t = 1, \ 2, \ \dots \\ \varphi_1 R(1) + \ \dots \ + \varphi_p R(p) + \sigma^2 & \text{for } t = 0 \end{cases};$$

in particular, $AR(1)$ has $R(t) = \frac{\sigma^2}{1-\varphi_1^2}\varphi_1^{|t|}, \ \ t \in \mathrm{Z}$ (see Sect. 31.4))

$$X_t = \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, \ t \in \mathrm{Z},$$

where $\{\varepsilon_t\}$ is white noise with variance σ^2; $\varphi_p \neq 0$; $\theta_q \neq 0$; the polynomial $\lambda^p - \varphi_1 \lambda^{p-1} - \dots - \varphi_p$ has all roots inside the unit circle in the complex plane

(*mixed process* of orders p and q denoted as $ARMA(p, q)$: is a (weakly) stationary stochastic process in discrete time with state space $S = \mathrm{R}$, zero mean value and autocovariance function fulfilling $R(t) = \varphi_1 R(t-1) + \dots + \varphi_p R(t-p)$ for $t = q+1, q+2, \dots$; if the polynomial $\lambda^q + \theta_1 \lambda^{q-1} + \dots + \theta_q$ has all roots inside the unit circle in the complex plane, then the process $ARMA(p, q)$ is *invertible*, i.e. predictable using the autoregressive form $X_t = \pi_1 X_{t-1} + \pi_2 X_{t-2} + \dots + \varepsilon_t$ (see Sect. 31.4))

$$X_t = \sum_{i=1}^{K} \left(A_i \cos(\lambda_i t) + B_i \sin(\lambda_i t)\right), \quad t \in \mathrm{R},$$

where $\mathrm{E}(A_i) = \mathrm{E}(B_i) = 0$; $\mathrm{var}(A_i) = \mathrm{var}(B_i) = {\sigma_i}^2 > 0$; $\mathrm{E}(A_i A_j) = \mathrm{E}(B_i B_j) = 0$ for $i \neq j$; $\mathrm{E}(A_i B_j) = 0$; $\lambda_i \in \mathrm{R}$

(*harmonic process*: is a (weakly) stationary stochastic process in continuous time with the state space $S = \mathrm{R}$, zero mean value and autocovariance function of the form

$$R(t) = \sum_{i=1}^{K} \sigma_i^2 \cos(\lambda_i t), \quad t \in \mathrm{R}\Big)$$

$\{W_t, t \geq 0\}$, where

$$\begin{cases} (i) \quad W_0 = 0 \\ (ii) \quad \text{trajectories are continuous in time} \\ (iii) \ W_{t_2} - W_{t_1}, \dots, \ W_{t_n} - W_{t_{n-1}} \ \text{are independent for arbitrary } 0 \leq t_1 < \dots < t_n \\ (iv) \ \ W_t - W_s \sim N(0, \ t-s) \quad \text{for arbitrary } \ 0 \leq s < t \end{cases}$$

(*Wiener process* (*Brownian motion*, see also Sect. 15.1): is a homogeneous Markov process in continuous time with the state space $S = \mathrm{R}$; increments $W_{t+h} - W_t$ have normal distribution $N(0, h)$ (see Sect. 26.5); W_t is applied when modeling movements of physical particles, interest rates, prices of some assets (it is the base of majority of financial models), and the like)

30.5 Spectral Properties of Stochastic Processes

$$X_t = \int_{-\pi}^{\pi} \cos(\lambda t)\, \mathrm{d}U_\lambda + \int_{-\pi}^{\pi} \sin(\lambda t)\, \mathrm{dV}_\lambda, \ t \in \mathrm{Z} \ \text{ or}$$

$$X_t = \int_{-\infty}^{\infty} \cos(\lambda t)\, \mathrm{d}U_\lambda + \int_{-\infty}^{\infty} \sin(\lambda t)\, \mathrm{dV}_\lambda, \ t \in \mathrm{R},$$

where $\{U_\lambda\}$ and $\{V_\lambda\}$ are stochastic processes with *independent increments*, i.e. $\mathrm{E}(\mathrm{d}U_\lambda) = \mathrm{E}(\mathrm{d}V_\lambda) = 0$; $\mathrm{E}(\mathrm{d}U_\lambda \cdot \mathrm{d}U_\kappa) = \mathrm{E}(\mathrm{d}V_\lambda \cdot \mathrm{d}V_\kappa) = 0$ for $\lambda \neq \kappa$; $\mathrm{E}(\mathrm{d}U_\lambda \cdot \mathrm{d}V_\kappa) = 0$
(*spectral decomposition of stationary process* in discrete or continuous time, respectively: "each stationary process $\{X_t\}$ is a mixture of periodic components with various frequencies, where amplitudes of particular components are independent random increments $\mathrm{d}U_\lambda$ and $\mathrm{d}V_\lambda$")

$$R(t) = \int_{-\pi}^{\pi} \cos(\lambda t)\, \mathrm{d}F(\lambda), \quad t \in \mathrm{Z} \qquad \text{or } R(t) = \int_{-\infty}^{\infty} \cos(\lambda t)\, \mathrm{d}F(\lambda), \quad t \in \mathrm{R},$$

where *spectral distribution function* $F(\lambda)$ is non-decreasing; continuous from the right; $F(-\pi) = 0$ or $F(-\infty) = 0$; $F(\pi) = R(0)$ or $F(\infty) = R(0)$, respectively

(*spectral decomposition of autocovariance function* of stationary process in discrete or continuous time, respectively: the function $F(\lambda)$ of given properties exists uniquely for arbitrary stationary process)

$$\mathrm{var}(X_t) = R(0) = \int_{-\pi}^{\pi} \mathrm{d}F(\lambda), \quad t \in \mathrm{Z} \quad \text{or} \quad \mathrm{var}(X_t) = R(0) = \int_{-\infty}^{\infty} \mathrm{d}F(\lambda), \quad t \in \mathrm{R},$$

where $\{X_t\}$ is a stationary process in discrete or continuous time, respectively

$$F(\lambda) = \int_{-\pi}^{\lambda} f(x)\, \mathrm{d}x, \quad -\pi \le \lambda \le \pi; \quad F(\lambda) = \int_{-\infty}^{\lambda} f(x)\, \mathrm{d}x$$

(*spectral density* $f(\lambda)$: exists for absolutely continuous spectral distribution functions; $f(\lambda)$ can be chosen as even, i.e. $f(\lambda) = f(-\lambda)$; $\mathrm{d}F(\lambda) = f(\lambda)\mathrm{d}\lambda$ represents the intensity (expressed by the variance) of the periodic component with frequency λ in the spectral decomposition of the given stationary process (see *thereinbefore*)

$$R(t) = \int_{-\pi}^{\pi} \cos(\lambda t) f(\lambda)\, d\lambda = 2 \int_{0}^{\pi} \cos(\lambda t) f(\lambda)\, d\lambda, \quad t \in \mathrm{Z} \quad \text{or}$$

$$R(t) = \int_{-\infty}^{\infty} \cos(\lambda t) f(\lambda)\, d\lambda = 2 \int_{0}^{\infty} \cos(\lambda t) f(\lambda)\, d\lambda, \quad t \in \mathrm{R}$$

(spectral decomposition of autocovariance function of stationary process in discrete or continuous time, respectively, when there exists the spectral density)

$$f(\lambda) = \frac{1}{2\pi} \sum_{t=-\infty}^{\infty} R(t)\cos(\lambda t) = \frac{1}{2\pi}\left(R(0) + 2\sum_{t=1}^{\infty} R(t)\cos(\lambda t)\right), \quad -\pi \le \lambda \le \pi,$$

where $\sum_{t=0}^{\infty} |R(t)| < \infty$

or $f(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R(t)\cos(\lambda t)\, dt = \frac{1}{\pi} \int_{0}^{\infty} R(t)\cos(\lambda t)\, dt, \quad \lambda \in \mathrm{R},$

where $\int_{0}^{\infty} |R(t)|\, dt < \infty$

(*inverse formula*: expresses the spectral density by means of the autocovariance function; it can be interpreted as the Fourier expansion of the spectral density, where Fourier coefficients are the autocovariances)

$f(\lambda) = \frac{\sigma^2}{2\pi}, \quad -\pi \le \lambda \le \pi$ (spectral density of *white noise*, see Sect. 30.4.)

$$f(\lambda) = \frac{\sigma^2}{2\pi} \cdot \frac{\left|e^{iq\lambda} + \theta_1 e^{i(q-1)\lambda} + \ldots + \theta_q\right|^2}{\left|e^{ip\lambda} - \varphi_1 e^{i(p-1)\lambda} - \ldots - \varphi_p\right|^2}, \quad -\pi \le \lambda \le \pi$$

(spectral density of process *ARMA*(p, q), see Sect. 30.4; in particular, *AR*(1) has the spectral density of the form $f(\lambda) = \frac{\sigma^2}{2\pi} \cdot \frac{1}{1+\varphi_1^2 - 2\varphi_1 \cos\lambda}, \quad -\pi \le \lambda \le \pi$)

- *Harmonic process* (see Sect. 30.4) has no spectral density: its spectral distribution function is piecewise constant with jumps of the size $\sigma_i^2/2$ at the points $\pm\lambda_1, \ldots, \pm\lambda_K$

$$f_{ij}(\lambda) = \frac{1}{2\pi} \sum_{t=-\infty}^{\infty} R_{ij}(t)\, e^{-i\lambda t}, \quad -\pi \le \lambda \le \pi, \quad \text{where} \quad \sum_{t=-\infty}^{\infty} \left|R_{ij}(t)\right| < \infty$$

(*mutual spectral density* of processes $\{X_{ti}\}$ and $\{X_{tj}\}$ in the framework of n-variate stochastic process $\mathbf{X}_t = (X_{t1}, \ldots, X_{tn})'$ in discrete time (the construction in continuous time is analogous); an alternative sufficient condition for the existence of $f_{ij}(\lambda)$ is the existence of the spectral densities $f_i(\lambda)$ and $f_j(\lambda)$ of processes $\{X_{ti}\}$ and $\{X_{tj}\}$)

$$f_{ij}(\lambda) = c(\lambda) - \mathrm{i} \cdot q(\lambda), \quad -\pi \leq \lambda \leq \pi$$

(*cospectrum* $c(\lambda)$ and *quadratic spectrum* $q(\lambda)$: are the real and imaginary part of $f_{ij}(\lambda)$ except for the sign)

$$C_{ij}(\lambda) = \frac{|f_{ij}(\lambda)|}{\left(f_i(\pi)f_j(\lambda)\right)^{1/2}}, \quad -\pi \leq \lambda \leq \pi$$

(*coefficient of coherence*: is a measure of dependence between processes $\{X_{ti}\}$ and $\{X_{tj}\}$, when the dependence concerns the periodic components of $\{X_{ti}\}$ and $\{X_{tj}\}$ with frequency λ ($0 \leq C_{ij}(\lambda) \leq 1$, where the values near to one indicate the highest dependence))

$$\Phi_{ij}(\lambda) = \operatorname{arctg}\frac{q(\lambda)}{c(\lambda)}, \quad -\pi \leq \lambda \leq \pi$$

(*coefficient of phase shift*: describes the phase delay of the periodic component with frequency λ of the process $\{X_{tj}\}$ against the periodic component with the same frequency λ of the process $\{X_{ti}\}$)

$$R_{ij}(\lambda) = C_{ij}(\lambda) \cdot \sqrt{\frac{f_i(\lambda)}{f_j(\lambda)}}, \quad -\pi \leq \lambda \leq \pi$$

(*coefficient of gain*: is the regression coefficient, when the regression of the process $\{X_{ti}\}$ on the process $\{X_{tj}\}$ is performed for the periodic components with frequency λ, which can be written in a schematic way as $X_{ti}(\lambda) = R_{ij}(\lambda) \cdot X_{tj}(\lambda)$)

$$Z_t = \sum_{k=-\infty}^{\infty} \delta_k X_{t-k}, \quad t \in \mathrm{Z}; \quad Z_t = \int_{-\infty}^{\infty} \delta(\tau) X_{t-\tau}\, \mathrm{d}\tau, \quad t \in \mathrm{R}$$

(linear *filter*: the stochastic process $\{Z_t\}$ originates by filtering the process $\{X_t\}$ applying the filter δ)

$$\psi(\lambda) = \sum_{k=-\infty}^{\infty} \delta_k \mathrm{e}^{-\mathrm{i}\lambda k}, \quad -\pi \leq \lambda \leq \pi \ \text{ or } \ \psi(\lambda) = \int_{-\infty}^{\infty} \delta(\tau)\mathrm{e}^{-\mathrm{i}\lambda\tau}\, \mathrm{d}\tau, \quad \lambda \in \mathrm{R}$$

(*transfer function* of filter: the spectral densities $f_X(\lambda)$ and $f_Z(\lambda)$ of processes $\{X_t\}$ and $\{Z_t\}$ fulfil $f_Z(\lambda) = |\psi(\lambda)|^2 \cdot f_X(\lambda)$ under general assumptions; the transfer function of a low-pass filter attains small values for higher frequencies, and the like)

Further Reading

Brockwell, P.J., Davis, R.A.: Time Series: Theory and Methods. Springer, New York (1987)

Cramer, H., Leadbetter, M.R.: Stationary and Related Stochastic Processes. Wiley, New York (1967)

Feller, W.: An Introduction to Probability Theory and Its Applications. Wiley, New York (1968)

Fuller, W.A.: Introduction to Statistical Time Series. Wiley, New York (1976)

Hamilton, J.D.: Time Series Analysis. Princeton University Press, Princeton, NJ (1994)

Leon-Garcia, A.: Probability and Random Processes. Addison-Wesley, Reading, MA (1989)

Malliaris, A.G., Brock, W.A.: Stochastic Methods in Economics and Finance. North-Holland, Amsterdam (1982)

Neftci, S.N.: Mathematics of Financial Derivatives. Academic, London (2000)

Priestley, M.B.: Spectral Analysis and Time Series. Academic, London (1981)

Rektorys, K. et al.: Survey of Applicable Mathematics. Kluwer, Dordrecht (1994)

Chapter 31
Statistical Analysis of Time Series

Abstract Chapter 31 contains formulas relevant for time series analysis: 31.1. Predictions in Time Series, 31.2. Decomposition of (Economic) Time Series, 31.3. Estimation of Correlation and Spectral Characteristics, 31.4. Linear Time Series, 31.5 Nonlinear and Financial Time Series, 31.6 Multivariate Time Series, 31.7. Kalman Filter.

31.1 Predictions in Time Series

$$e_T = x_T - \hat{x}_T$$

(*prediction error*: $\hat{x}_T$ is a (point) prediction of the value x_T in an observed time series, i.e. a statistical estimate of the future (predicted) value x_T $(T > n)$ constructed by means of observations $\{x_1, x_2, \ldots, x_n\}$; the prediction error can be calculated explicitly only as late as the observation x_T is known)

$$\left(\hat{x}_T^D(\alpha),\ \hat{x}_T^H(\alpha)\right)$$

$(100(1 - \alpha)\%$ *interval prediction* (*confidence prediction interval*): is an interval estimated statistically by means of known observations $\{x_1, x_2, \ldots, x_n\}$ such that the future (predicted) value X_T $(T > n)$ lies in it with probability $100(1 - \alpha)$ (e.g. for $\alpha = 0.05$, it is the 95 % interval prediction))

$$\mathrm{SSE} = \sum\nolimits_{t=n+1}^{n+h} \left(x_t - \hat{x}_t\right)^2 = \sum\nolimits_{t=n+1}^{n+h} \mathrm{e}_t^2 \quad (\textit{sum of squared errors})$$

$$\mathrm{MSE} = (1/h) \sum\nolimits_{t=n+1}^{n+h} \left(x_t - \hat{x}_t\right)^2 = (1/h) \sum\nolimits_{t=n+1}^{n+h} e_t^2 \quad (\textit{mean squared error})$$

$$\frac{\left(\bar{\hat{x}} - \bar{x}\right)^2}{\mathrm{MSE}} + \frac{(s_{\hat{x}} - s_x)^2}{\mathrm{MSE}} + \frac{2(1 - r_{\hat{x}x})s_{\hat{x}}s_x}{\mathrm{MSE}} = 100\%$$

(decomposition of *MSE* to *proportional bias*, *proportional variance* and *proportional covariance*, respectively: $\bar{\hat{x}}$, $\bar{x}$, $s_{\hat{x}}$, s_x, $r_{\hat{x}x}$ are the corresponding sample

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0_31,

means, sample standard deviations and sample correlation coefficient of the values $\hat{x}$ and x)

$$\mathrm{MAE} = (1/h) \sum_{t=n+1}^{n+h} \left|x_t - \hat{x}_t\right| = (1/h) \sum_{t=n+1}^{n+h} |e_t|$$

(*mean absolute error*: is less sensitive against large errors than *MSE*)

$$\mathrm{RMSE} = \sqrt{\frac{1}{h} \sum_{t=n+1}^{n+h} \left(x_t - \hat{x}_t\right)^2} \; (\textit{root mean squared error})$$

$$\mathrm{MAPE} = \frac{100}{h} \sum_{t=n+1}^{n+h} \left|\frac{x_t - \hat{x}_t}{x_t}\right|$$

(*mean absolute percentage error*: is measured in %)

$$\mathrm{AMAPE} = \frac{100}{h} \sum_{t=n+1}^{n+h} \left|\frac{x_t - \hat{x}_t}{(x_t + \hat{x}_t)/2}\right|$$

(*adjusted MAPE*: is symmetric with respect to x and $\hat{x}$)

$$U = \frac{\sqrt{\sum_{t=n+1}^{n+h} (x_t - \hat{x}_t)^2}}{\sqrt{\sum_{t=n+1}^{n+h} \hat{x}_t^2} + \sqrt{\sum_{t=n+1}^{n+h} x_t^2}} \quad (\textit{Theil's U-statistics})$$

$$\frac{100}{h} \sum_{t=n+1}^{n+h} z_t, \text{ where } z_t = \begin{cases} 1 & \text{for } x_t \cdot \hat{x}_t > 0 \\ 0 & \text{others} \end{cases}$$

(*percentage of correct predictions of signs + or −*)

$$\frac{100}{h} \sum_{t=n+1}^{n+h} z_t, \text{ where } z_t = \begin{cases} 1 & \text{for } (x_t - x_{t-1}) \cdot (\hat{x}_t - x_{t-1}) > 0 \\ 0 & \text{others} \end{cases}$$

(*percentage of correct predictions of growth or decrease*)

31.2 Decomposition of (Economic) Time Series

$$X_t = Tr_t + C_t + I_t + \varepsilon_t, \quad t \in Z$$

(*additive decomposition* of stochastic process $\{X_t\}$: Tr_t is the *trend component* (*trend*), which reflects long-run movements in the process; C_t is the *cyclical component* (*cycle*), which refers to recurring up and down movements around trend levels (with periodicity usually longer than one year: e.g. cycles economic, demographic, climatic and others); I_t is the *seasonal component* (*seasonal variations*), i.e. periodic patterns, which complete themselves within the period of a calendar year and are then repeated on a yearly basis (it is caused by factors such as weather and customs); ε_t is the *residual* (*random*, *irregular*) *component* (see Sect. 27.11), which are erratic fluctuations around *systematic components* Tr_t, C_t and I_t in the process; the classical decomposition analysis assumes that $\{\varepsilon_t\}$ is the *white noise* (see Sect. 30.4), i.e. in particular, $\text{var}(\varepsilon_t) = \sigma^2 > 0$ for all t)

$$X_t = Tr_t \cdot C_t \cdot I_t \cdot \varepsilon_t, \quad t \in Z$$

(*multiplicative decomposition* of stochastic process $\{X_t\}$: can be transferred by the logarithmic transformation to the additive decomposition (however, there may be problems to estimate particular components after such a transformation))

$$\{x_t,\ t \in \{1, 2, \ldots, n\}\} = \{x_1, x_2, \ldots, x_n\}, n \in \mathrm{N};\ x_t \in \mathrm{R}$$

(observed *time series* corresponding to stochastic process with systematic components: can be used to estimate systematic components and consequently to smooth and predict such a time series)

$$\hat{x}_t = \hat{T}r_t + \hat{C}_t + \hat{I}_t, \quad \hat{x}_t = \hat{T}r_t \cdot \hat{C}_t \cdot \hat{I}_t\text{, respectively}$$

(*smoothed* time series ($t \le n$) or *prediction* in time series ($t > n$): one makes use of estimated or predicted values of the systematic components in the additive and multiplicative decomposition, respectively (see Sect. 27.11))

$X_t = \beta_0 + \varepsilon_t, \quad t = 1, \ldots, n \quad$ (*constant trend*)

– the parameter estimation: $b_0 = \bar{x} = (1/n) \sum_{t=1}^{n} x_t$
– the point prediction (see Sect. 31.1): $\hat{x}_T = b_0$
– the $100(1 - \alpha)$ prediction (see Sects. 26.5 and 31.1):

$$\left(b_0 \ \pm \ t_{1-\alpha/2}(n-1) \cdot \sqrt{\frac{\sum_{t=1}^{n} (x_t - \bar{x})^2}{n-1}} \cdot \sqrt{1 + \frac{1}{n}} \right)$$

$X_t = \beta_0 + \beta_1\, t + \varepsilon_t, \quad t = 1, \ldots, n$ (*linear trend*)

– is characterized by constant first differences $Tr_{t+1} - Tr_t = \beta_1$
– the parameter estimation:

$$b_1 = \frac{\sum_{t=1}^{n} tx_t - ((n+1)/2)\ \sum_{t=1}^{n} x_t}{n(n^2-1)/12}; \quad b_0 = \bar{x} - \frac{n+1}{2} b_1$$

– the point prediction (see Sect. 31.1): $\hat{x}_T = b_0 + b_1 T$

– the $100(1-\alpha)$ interval prediction (see Sects. 26.5 and 31.1):

$$\left(b_0 + b_1 T \ \pm \ t_{1-\alpha/2}(n-2) \cdot \sqrt{\frac{\sum_{t=1}^{n} (x_t - b_0 - b_1 t)^2}{n-2}} \cdot \sqrt{1 + \frac{1}{n} + \frac{(T-(n+1)/2)^2}{n(n^2-1)/12}} \right)$$

$X_t = \beta_0 + \beta_1\, t + \beta_2\, t^2 + \varepsilon_t, \ \ t = 1, \dots, n$ (*quadratic trend*)

– is characterized by constant second differences $(Tr_{t+2} - Tr_{t+1}) - (Tr_{t+1} - Tr_t) = \beta_2$

– the parameter estimation can be obtained from the system of equations:

$$\begin{aligned} b_0 n \quad &+ b_1 \sum t \quad + b_2 \sum t^2 = \sum x_t \\ b_0 \sum t \quad &+ b_1 \sum t^2 \quad + b_2 \sum t^3 = \sum t x_t \\ b_0 \sum t^2 \quad &+ b_1 \sum t^3 \quad + b_2 \sum t^4 = \sum t^2 x_t \end{aligned}$$

– the point prediction (see Sect. 31.1): $\hat{x}_T = b_0 + b_1 T + b_2 T^2$

– the $100(1-\alpha)$ interval prediction (see Sects. 26.5 and 31.1) with

$$\mathbf{X}' = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & n \\ 1 & 4 & \dots & n^2 \end{pmatrix}:$$

$$\left(b_0 + b_1 T + b_2 T^2 \ \pm \ t_{1-\varepsilon/2}(n-3) \cdot \sqrt{\frac{\sum_{t=1}^{n} \left(x_t - b_0 - b_1 t - b_2 t^2\right)^2}{n-3}} \cdot \sqrt{1 + (1,\ T,\ T^2)(\mathbf{X}'\mathbf{X})^{-1} \begin{pmatrix} 1 \\ T \\ T^2 \end{pmatrix}} \right)$$

$X_t = \alpha\beta^t + \varepsilon_t, \ \ t = 1, \dots, n$, where $\beta > 0$ (*exponential trend*)

– the trend is modeled by the *exponential curve*: it is increasing for $\alpha > 0$ and $\beta > 1$ (or it is decreasing for $\alpha > 0$ and $0 < \beta < 1$)

– is characterized by the constant *coefficient of growth* $Tr_{t+1}/Tr_t = \beta$

– the parameter estimation can be obtained from the system of equations:

$$\begin{aligned}\left(\sum x_t^2\right)\ln\alpha \ + \left(\sum tx_t^2\right)\ln\beta \ &= \sum x_t^2 \ln x_t \\ \left(\sum tx_t^2\right)\ln\alpha \ + \left(\sum t^2x_t^2\right)\ln\beta \ &= \sum tx_t^2 \ln x_t\end{aligned}$$

$X_t = \gamma + \alpha\beta^t + \varepsilon_t, \ \ t = 1, \ldots, \ n$, where $\beta > 0$ (*modified exponential trend*)

– the trend is modeled by the *modified exponential curve with asymptotic limit* (for $\alpha < 0, 0 < \beta < 1, \gamma > 0$)
– is characterized by the constant ratio of adjacent first differences: $(Tr_{t+2} - Tr_{t+1})/(Tr_{t+1} - Tr_t) = \beta$
– the approximate parameter estimation for $n = 3m$ can be obtained as follows ($\mathbf{\Sigma}_1$, $\mathbf{\Sigma}_2$, $\mathbf{\Sigma}_3$ denote the sums of the values in the first, second and last third of the given time series, respectively):

$$b = \left(\frac{\sum_3 x_t - \sum_2 x_t}{\sum_2 x_t - \sum_1 x_t}\right)^{1/m}; \ \ a = \frac{b-1}{b(b^m-1)^2}\left(\sum\nolimits_2 x_t - \sum\nolimits_1 x_t\right);$$

$$c = \tfrac{1}{m}\left(\sum\nolimits_1 x_t - \frac{ab(b^m-1)}{b-1}\right)$$

$X_t = \frac{\gamma}{1+\alpha\beta^t} + \varepsilon_t, \ \ t = 1, \ldots, \ n,$ where $\beta > 0, \gamma > 0$ (*logistic trend*)

– the trend is modeled by the *S-curve symmetric around inflex point*
– is characterized by the constant ratio of adjacent first differences of inverted values: $(1/Tr_{t+2} - 1/Tr_{t+1})/(1/Tr_{t+1} - 1/Tr_t) = \beta$
– the approximate parameter estimation can be obtained by estimating the following linear regression model (see Sect. 27.11):

$$\frac{x_t - x_{t-1}}{x_t} = -\ln\beta + \frac{\ln\beta}{\gamma}x_t + \eta_t, \ \ t = 1, \ldots, \ n$$

– the parameter α can be estimated by *Rhodes formula*:

$$\ln\alpha = -\frac{(n+1)\ln\beta}{2} + \frac{1}{n}\sum_{t=1}^{n} \ln\left(\frac{\beta}{x_t} - 1\right)$$

$X_t = \exp\left(\gamma + \alpha\beta^t\right) \ + \varepsilon_t, \ \ t = 1, \ldots, \ n$, where $\beta > 0$ (*Gompertz trend*)

– the trend is modeled by the *S-curve asymmetric around inflex point*
– is characterized by the constant ratio of adjacent first differences of log values: $(\ln Tr_{t+2} - \ln Tr_{t+1})/(\ln Tr_{t+1} - \ln Tr_t) = \beta$
– the parameter estimation can be obtained same way as in the case of the modified exponential trend (see *thereinbefore*) for the time series $\{\ln x_t\}$

$$\hat{x}_t = \sum\nolimits_{i=-m}^{m} w_i x_{t+i} = w_{-m}x_{t-m} + \ldots \ + w_0 x_t + \ldots \ + w_m x_{t+m}$$

(*moving average* of length $2m + 1$ with weights $\{w_i\}$: is used most often in the case of $X_t = Tr_t + \varepsilon_t$; moreover, if locally in the neighbourhood of time t the trend is a polynomial of degree r, namely $\beta_0(t) + \beta_1(t)\tau + \ldots + \beta_r(t)\tau^r$ for $\tau = -m, \ldots, 0, \ldots, m$, then the corresponding moving average is called the moving average of length $2m + 1$ and *degree r* (it has always symmetric weights $\{w_{-m}, \ldots, 0, \ldots, w_m\}$ summing up to one, which are the same for even r and odd $r + 1$)

$$\hat{x}_t = \frac{1}{35}\left(17 \sum\nolimits_{\tau=-2}^{2} x_{t+\tau} - 5 \sum\nolimits_{\tau=-2}^{2} \tau^2 x_{t+\tau}\right)$$
$$= \frac{1}{35}\left(-3x_{t-2} + 12x_{t-1} + 17x_t + 12x_{t+1} - 3x_{t+2}\right),\ t = 3, 4, \ldots, n-2$$

(an example of the moving average of length 5 and degree 2 or 3: the weights for various lengths and degrees are tabulated or produced by software systems; for the first smoothed values $\hat{x}_1,\ \hat{x}_2$ and for the last smoothed values $\hat{x}_{n-1},\ \hat{x}_n$ and for the predicted values $\hat{x}_{n+k}(n)$ with $k > 0$ one must apply special weights, which are again tabulated, see *thereinafter*)

$$\hat{x}_{n+1}(n) = \frac{1}{10}\left(-4x_{n-4} - x_{n-3} + 2x_{n-2} + 5x_{n-1} + 8x_n\right)$$

(an example of the *prediction* moving average of length 5 and degree 1 producing a one-period-ahead prediction (at time n for time $n + 1$, see *thereinbefore*): the weights for various lengths and degrees are tabulated or produced by software systems (however, they are asymmetric and distinct for even r and odd $r + 1$))

$$\hat{x}_t = \frac{1}{2m+1}\left(x_{t-m} + \ldots + x_t + \ldots + x_{t+m}\right)$$

(*arithmetic* moving average of length $2m + 1$)

$$\hat{x}_t = \frac{1}{4m}\left(x_{t-m} + 2x_{t-m+1} + \ldots + 2x_{t+m-1} + x_{t+m}\right)$$

(*centred* moving average over a season of length $2m$: e.g. for monthly observations one puts $2m = 12$ and

$$\hat{x}_{\text{July 08}} = (1/24)(\hat{x}_{\text{January 08}} + 2\hat{x}_{\text{February 08}} + \ldots + 2\hat{x}_{\text{December 08}} + \hat{x}_{\text{January 09}})$$

$$\left.\begin{aligned} \hat{x}_t &= \alpha x_t + (1-\alpha)\hat{x}_{t-1} \\ \hat{x}_{t+k}(t) &= \hat{x}_t \quad \text{for } k > 0 \end{aligned}\right\}$$

(*simple exponential smoothing*: is used most often in the case $X_t = \beta_0(t) + \varepsilon_t$ (i.e. for a constant trend, which is flexible in time); $0 < \alpha < 1$ is a fixed *smoothing constant* (the level of smoothing increases for α approaching zero, and the recommended values of α are $0 < \alpha \le 0.3$); in general, the principle of *exponential smoothing* includes *recursive* smoothing and predicting methods, in which the values observed up to the present period get weights, which decrease exponentially with the age of observations (e.g. an equivalent non-recursive representation of the simple exponential smoothing is $\hat{x}_t = \alpha \sum_{i=0}^{\infty} (1-\alpha)^i x_{t-i})$); in practice one must choose the smoothing constant and the initial value $\hat{x}_0$ (e.g. in the simple exponential smoothing, the initial value can be chosen as the arithmetic average of several (usually six) first observations of the given time series and the smoothing constant α may be found by minimizing *SSE* (see Sect. 31.1) over a grid of possible values of this constant))

$$\left.\begin{aligned} \hat{x}_t &= 2S_t - S_t^{[2]} \\ \hat{x}_{t+k}(t) &= \left(2 + \tfrac{\alpha \cdot k}{1-\alpha}\right) S_t - \left(1 + \tfrac{\alpha \cdot k}{1-\alpha}\right) S_t^{[2]} \quad \text{for } k > 0 \\ S_t &= \alpha x_t + (1-\alpha) S_{t-1} \\ S_t^{[2]} &= \alpha S_t + (1-\varepsilon) S_{t-1}^{[2]} \end{aligned}\right\}$$

(*double* exponential smoothing (*Brown's method*): is used most often in the case $X_t = \beta_0(t) + \beta_1(t) \cdot t + \varepsilon_t$ (i.e. for a linear trend, which is flexible in time); $0 < \alpha < 1$ is a fixed smoothing constant (the recommended values of α are $0 < \alpha \leq 0.3$); S_t and $S_t{}^{[2]}$ are auxiliary *smoothing statistics*; the recommended initial values are $S_0 = b_0(0) - ((1-\alpha)/\alpha) \cdot b_1(0)$, $S_0{}^{[2]} = b_0(0) - (2(1-\alpha)/\alpha) \cdot b_1(0)$, where $b_0(0)$ and $b_1(0)$ are estimated parameters β_0 and β_1 applying the linear trend regression (see *thereinbefore*) to several first observations of the time series; the smoothing constant α can be found again e.g. by minimizing *SSE*)

$$\left.\begin{aligned} L_t &= \alpha x_t + (1-\alpha)(L_{t-1} + T_{t-1}) \\ T_t &= \gamma (L_t - L_{t-1}) + (1-\gamma) T_{t-1} \\ \hat{x}_t &= L_t \\ \hat{x}_{t+k}(t) &= L_t + T_t \cdot k \quad \text{for } k > 0 \end{aligned}\right\}$$

(*Holt's method*: is used most often in the case $X_t = \beta_0(t) + \beta_1(t) \cdot t + \varepsilon_t$ (i.e. for a linear trend, which is flexible in time); $0 < \alpha, \gamma < 1$ are fixed smoothing constants (the recommended values are $0 < \alpha, \gamma \leq 0.2$); L_t and T_t are auxiliary values representing the level and the slope of the linear trend; it is a generalization of the double exponential smoothing (see *thereinbefore*) to the case of two smoothing constants)

$$x_t = Tr_t + I_t + \varepsilon_t \text{ or } x_t = Tr_t \cdot I_t \cdot \varepsilon_t$$

(additive or multiplicative *seasonal decomposition* with length of season s (e.g. it is $s = 12$ for monthly observations): the seasonal component I_t is often called *seasonal index* (in the additive case it is measured in the same units as x_t, in the multiplicative case it is a percentage)

$$\sum_{i=1}^{s} I_{i+s \cdot j} = 0 \text{ for all } j = 0, 1, \ldots$$

(*normalization of seasonal indices* for additive seasonal decomposition with length of season s: it guarantees the uniqueness of decomposition)

$$\sum_{i=1}^{s} I_{i+s\cdot j} = p \text{ or } \prod_{i=1}^{s} I_{i+s\cdot j} = 1 \text{ for all} j = 0, 1, \ \ldots$$

(*normalization of seasonal indices* for multiplicative seasonal decomposition with length of season s: it guarantees the uniqueness of decomposition; the latter normalization corresponds after the logarithmic transformation to the normalization for the additive seasonal decomposition (see *thereinbefore*))

$$x_t = \beta_0 + \beta_1 t + \alpha_2 v_{t_2} + \alpha_3 v_{t_3} + \alpha_4 v_{t_4} + \varepsilon_t$$

(an example of the additive seasonal decomposition with linear trend and length of season $s = 4$ (i.e. for quarterly observations): the seasonality is modeled by means of a qualitative seasonal regressor using three dummies v_1, v_2, v_3 (see Sect. 28.3)):

t	v_{t1}	v_{t2}	v_{t3}
1	0	0	0
2	1	0	0
3	0	1	0
4	0	0	1
5	0	0	0
6	1	0	0
7	0	1	0
8	0	0	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$

$$x_t = \beta_0 + \beta_1 \cdot t + \beta_2 \cdot t \cdot \sin\left(\frac{2\pi t}{s}\right) + \beta_3 \cdot t \cdot \cos\left(\frac{2\pi t}{s}\right) + \varepsilon_t$$

(an example of the multiplicative seasonal decomposition with linear trend and length of season s: the seasonality is modeled by means of goniometric functions)

$$\left.\begin{aligned} L_t &= \alpha(x_t - I_{t-p}) + (1-\alpha)(L_{t-1} + T_{t-1}) \\ T_t &= \gamma(L_t - L_{t-1}) + (1-\gamma)T_{t-1} \\ I_t &= \delta(x_t - L_t) + (1-\delta)I_{t-s} \\ \hat{x}_t &= L_t + I_t \\ \hat{x}_{t+k}(t) &= L_t + T_t \cdot k + I_{t+k-s} \quad \text{for } k = 1, \ldots, s \\ &= L_t + T_t \cdot k + I_{t+k-2s} \ \ \text{for } k = s+1, \ldots, 2s \\ &\vdots \end{aligned}\right\}$$

(additive *Holt-Winters' method* with season of length p: $0 < \alpha, \gamma, \delta < 1$ are fixed smoothing constants; it is a generalization of Holt's method (see *thereinbefore*) for additive seasonal decomposition with three smoothing constants)

$$\left.\begin{aligned}
L_t &= \alpha\,(x_t/I_{t-s}) + (1-\alpha)(L_{t-1}+T_{t-1})\\
T_t &= \gamma(L_t - L_{t-1}) + (1-\gamma)T_{t-1}\\
I_t &= \delta\,(x_t/L_t) + (1-\delta)I_{t-s}\\
\hat{x}_t &= L_t \cdot I_t\\
\hat{x}_{t+k}(t) &= (L_t + T_t\cdot k)\cdot I_{t+k-s} \quad \text{for } k=1,\ldots,s\\
&= (L_t + T_t\cdot k)\cdot I_{t+k-2s} \quad \text{for } k=s+1,\ldots,2s\\
&\vdots
\end{aligned}\right\}$$

(multiplicative *Holt-Winters' method* with season of length p: $0 < \alpha, \gamma, \delta < 1$ are fixed smoothing constants; it is a generalization of Holt's method (see *thereinbefore*) for multiplicative seasonal decomposition with three smoothing constants)

$$\frac{|\,k-(n-1)/2\,|}{\sqrt{(n+1)/12}} \;\geq\; u_{1-\alpha/2},$$

where k is the number of positive differences $x_{t+1} - x_t$ (i.e. the number of the so-called *points of growth*) in time series $x_1, \ldots, x_n$

(critical region for test of hypothesis H_0: $x_t \sim iid$ at level of significance α (see Sect. 27.10): it is one of the simplest *tests of randomness* applied to the detrended and deseasonalised time series (e.g. the additive trend or seasonal adjustment is obtained by subtracting the estimated components Tr_t or I_t) in order to appreciate the quality of such an adjustment (is the residual really a random component?); the test is derived asymptotically (i.e. it requires a larger n in practice)

$$\frac{|\,r-2(n-2)/3\,|}{\sqrt{(16n-29)/90}} \;\geq\; u_{1-\alpha/2},$$

where r is the number of upper and lower turning points in the time series $x_1, \ldots, x_n$

(the *upper turning point* x_t is the local maximum $x_{t-1} < x_t > x_{t+1}$ and the *lower turning point* x_t is the local minimum $x_{t-1} > x_t < x_{t+1}$)
(critical region for test of hypothesis H_0: $x_t \sim iid$ at level of significance α (see Sect. 27.10): it is another *test of randomness* (see *thereinbefore*); the test is derived asymptotically (i.e. it requires a larger n in practice) and is applied, when there is a suspicion that the seasonal adjustment is not sufficient)

31.3 Estimation of Correlation and Spectral Characteristics

$$\{x_t,\ t \in T = \{1,2,\ldots,n\}\} = \{x_1, x_2, \ldots, x_n\}, n \in \mathrm{N}; x_t \in \mathrm{R}$$

(observed *time series* of stationary (see Sect. 30.1) stochastic process: can be used to estimate correlation and spectral characteristics)

$$\hat{\mu} = \bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t \quad \text{(estimated mean of stationary time series)}$$

$$\hat{R}(t) = \frac{1}{n}\sum_{i=1}^{n-\,t}(x_i - \bar{x})(x_{i+t} - \bar{x}), \quad t = 0,\ 1,\ldots,n-1$$

(estimated autocovariance function of stationary time series (see Sect. 30.1): is preferred in practice to the estimate of the form

$$\hat{R}(t) = \frac{1}{n-t}\sum_{i=1}^{n-t}(x_i - \bar{x})(x_{i+t} - \bar{x})\Bigg)$$

$$\hat{B}(t) = \frac{R(t)}{R(0)}, \quad t = 0, 1, \ldots,\ n-1$$

(estimated autocorrelation function of stationary time series (see Sect. 30.1))

$$\sigma\left(\hat{B}(t)\right) = \sqrt{\mathrm{var}\left(\hat{B}(t)\right)} \sim \sqrt{\frac{1}{n}\left(1 + 2\sum\nolimits_{k=1}^{k_0}\left(\hat{B}(k)\right)^2\right)},\ t > k_0,$$

where $B(t) = 0$ for $t > k_0$

(*Bartlett's approximation*: approximates the standard deviation of estimated autocorrelations in a stationary time series beyond the so-called *truncation point* (i.e. for zero theoretical autocorrelations $B(t) = 0$); it is applied to identify this point in the sequence of estimated autocorrelations (i.e. in the so-called *correlogram*), which has implications when identifying a suitable model for a given time series)

$$B(t|t) = \rho(X_s, X_{s+t}|\ X_{s+1}, \ldots,\ X_{s+t-1}),\ t \in N_0 \text{ for arbitrary } s \in Z$$

(*partial autocorrelation function* of stationary stochastic process $\{X_t\}$ in discrete time: is the partial correlation coefficient between X_s and X_{s+t} with fixed values $X_{s+1}, \ldots, X_{s+t-1}$ (see Sect. 27.8); $B(0|0) = 1$; $B(1|1) = B(1)$; the estimated partial autocorrelations are usually constructed recursively, e.g. by means of *Durbin-Levinson algorithm*, which makes use of the fact that the estimate of $B(k|k)$ can be obtained as the OLS-estimate of the parameter φ_k in the $AR(k)$ model $X_t = \varphi_1 X_{t-1} + \ldots + \varphi_k X_{t-k} + \varepsilon_t$)

$$\sigma\left(\hat{B}(t\mid t)\right) = \sqrt{var\left(\hat{B}(t\mid t)\right)} \sim \sqrt{1/n}, \quad t > k_0, \text{ where } B(t\mid t) = 0 \text{ for } t > k_0$$

(*Quenouille's approximation*: approximates the standard deviation of estimated partial autocorrelations in a stationary time series beyond the *truncation point* (i.e. for zero theoretical partial autocorrelations $B(t|t) = 0$); it is applied similarly as Bartlett's approximation to identify this point in the sequence of estimated partial autocorrelations (i.e. in the so-called *partial correlogram*), which has again implications when identifying a suitable model for a given time series)

$$I(\lambda) = \frac{1}{2\pi n}\left(\left(\sum_{t=1}^{n} x_t \cos(\lambda t)\right)^2 + \left(\sum_{t=1}^{n} x_t \sin(\lambda t)\right)^2\right)$$
$$= \frac{1}{2\pi}\left(\hat{R}(0) + 2\sum_{t=1}^{n-1} \hat{R}(t)\cos(\lambda t)\right), \quad -\pi \le \lambda \le \pi$$

(*periodogram* of stationary time series: is an important instrument for *spectral analysis* (in particular, when estimating spectral density, see Sect. 30.5); it is a principal of test statistics for the so-called *tests of periodicity* when identifying and verifying periodic components in time series, e.g. *Fisher's* test of periodicity)

$$\hat{f}(\lambda_0) = \int_{-\pi}^{\pi} s(\lambda - \lambda_0) \cdot I(\lambda)\, \mathrm{d}\lambda, \quad \text{or}$$

$$\hat{f}(\lambda_0) = w_0\hat{R}(0) + 2\sum_{t=1}^{n-1} w_t\hat{R}(t)\cos(\lambda_0 t), \ -\pi \le \lambda_0 \le \pi$$

(estimated spectral density of stationary time series: by smoothing the periodogram ($s(\cdot)$ is a function chosen in a suitable way), or by weighing estimated autocovariances (w_t are weights chosen in a suitable way))

$$w_t = \begin{cases} \dfrac{1}{2\pi}\left(1 - \dfrac{6t^2}{m^2}\left(1 - \dfrac{t}{m}\right)\right) & \text{for } t = 0, 1, \ldots, \dfrac{m}{2} \\ \dfrac{1}{\pi}\left(1 - \dfrac{t}{m}\right)^3 & \text{for } t = \dfrac{m+1}{2}, \ldots, m \\ 0 & \text{for } t > m \end{cases}$$

(*Parzen's estimator* of spectral density: w_t are weights of estimated autocovariances in $\hat{f}(\lambda_0)$ (see *thereinbefore*); choice of $m \in \mathrm{N}$ is recommended between $n/6$ and $n/5$)

31.4 Linear Time Series

$$Bx_t = x_{t-1}; B^j x_t = B(B^{j-1})x_t = x_{t-j}, j \in \mathrm{N} \quad (\textit{lag operator})$$

$$\varphi(B)x_1 = \theta(B)\varepsilon_t,$$

where $\varphi(B) = 1 - \varphi_1 B - \ldots - \varphi_p B^p$ and $\theta(B) = 1 + \theta_1 B + \ldots + \theta_q B^q$

(mixed process *ARMA*(p, q) (see Sect. 30.4) written by means of the lag operator)

$$X_t = \varepsilon_t + \psi_1\varepsilon_{t-1} + \psi_2\varepsilon_{t-2} + \ldots = \psi(B)\varepsilon_t,\ t \in \mathrm{Z},$$

where $\{\varepsilon_t\}$ is a white noise with variance σ^2 (see Sect. 30.4); $\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \ldots$

(*linear process*: is a theoretical background when modeling linear time series; if the power series $\psi(z) = 1 + \psi_1\, z + \psi_2\, z^2 + \ldots$ converges for $|z| \leq 1$ (i.e. inside and on the unit circle in the complex plane), then the linear process exists (in sense of the convergence in the mean square for the corresponding infinite sum, see Sect. 26.14) and is (weakly) stationary (see Sect. 30.1) with zero mean value; the stationary processes *MA*, *AR*, *ARMA* (see Sect. 30.4) are special cases of the linear process and, for time series in practice, they are usually constructed in a systematic way by means of the so-called *Box-Jenkins methodology*, which does make use of the linear models of time series, namely in three steps: (1) identification of model (e.g. for a time series $x_1, \ldots, x_n$ one identifies a model *AR*(1)); (2) estimation of model (e.g. the estimated model $x_t = \hat{\varphi}_1 x_{t-1} + \varepsilon_t$ is $x_t = 0.68x_{t-1} + \varepsilon_t$ with $\hat{\sigma} = 11.24$); (3) diagnostic checking (e.g. the model from the step (2) is verified with the 95% certainty))

$$\hat{X}_{t+k}(t) = \psi_k\, \varepsilon_t + \psi_{k+1}\, \varepsilon_{t-1} + \ldots,\ k \in \mathrm{N}$$

(*prediction* in linear invertible (see *thereinafter*) process for time $t + k$ at time t: it is a linear prediction (i.e. a linear function of the process), which is optimal in sense of the criterion *MSE* (see Sect. 31.1))

$$\mathrm{MSE} = \mathrm{var}(\psi_{t+k} + \psi_1\varepsilon_{t+k-1} + \ldots + \psi_{k-1}\varepsilon_{t+1}) = (1 + \psi_1^2 + \ldots + \psi_{k-1}^2)\sigma^2$$

(*MSE* (see Sect. 31.1) of prediction in linear invertible process for time $t + k$ at time t)

$$X_t - \pi_1 X_{t-1} - \pi_2 X_{t-2} - \ldots = \pi(B)X_t = \varepsilon_t,\ t \in \mathrm{Z},$$

where $\{\varepsilon_t\}$ is a white noise with variance σ^2; $\pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \ldots$

(*invertible* form of linear process: is a theoretical base when constructing predictions in practice; if the power series $\pi(z) = 1 - \pi_1\, z - \pi_2\, z^2 - \ldots$ converges for $|z| \leq 1$ (i.e. inside and on the unit circle in the complex plane), then the linear process is invertible)

$X_t = \varepsilon_t + \theta_1\varepsilon_{t-1}$, $t \in \mathrm{Z}$, where $\{\varepsilon_t\}$ is a white noise with variance σ^2; $\theta_1 \neq 0$

(*moving average process* of the first order *MA*(1) (see Sect. 30.4): is a (weakly) stationary stochastic process with zero mean value and autocorrelation and partial autocorrelation functions (see Sect. 30.1) of the form

$$B(t) = \begin{cases} \frac{\theta_1}{1+\theta_1^2} & \text{for } t = 1 \\ 0 & \text{for } t > 1 \end{cases};$$

$$B(t|t) = \frac{(-1)^{t-1}\theta_1^t(1-\theta_1^2)}{1-\theta_1^{2(t+1)}} \quad \text{for } t = 1,\ 2,\ \ldots;$$

if $|\theta_1| < 1$, then the process $MA(1)$ is invertible)

$$X_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2},\ t \in \mathrm{Z},$$

where $\{\varepsilon_t\}$ is a white noise with variance σ^2; $\theta_2 \neq 0$

(*moving average process* of the second order $MA(2)$ (see Sect. 30.4): is a (weakly) stationary stochastic process with zero mean value and autocorrelation function (see Sect. 30.1) of the form

$$B(t) = \begin{cases} \dfrac{\theta_1(1+\theta_2)}{1+\theta_1^2+\theta_2^2} & \text{for } t = 1 \\ \dfrac{\theta_2}{1+\theta_1^2+\theta_2^2} & \text{for } t = 2 \\ 0 & \text{for } t > 2 \end{cases};$$

if $\theta_1 + \theta_2 > -1$, $-\theta_1 + \theta_2 > -1$, $-1 < \theta_2 < -1$, then the process $MA(2)$ is invertible)

$$X_t = \varphi_1 X_{t-1} + \varepsilon_t,\ \ t \in \mathrm{Z},$$

where $\{\varepsilon_t\}$ is a white noise with variance σ^2; $\varphi_1 \neq 0$; $|\varphi_1| < 1$

(*autoregressive process* of the first order $AR(1)$ (see Sect. 30.4): is a (weakly) stationary invertible stochastic process with zero mean value and autocorrelation and partial autocorrelation functions (see Sect. 30.1) of the form

$$B(t) = \varphi_1^t,\ \ t = 1, 2, \ldots;\ B(t \mid t) = \begin{cases} \varphi_1 & \text{pro } t = 1 \\ 0 & \text{pro } t > 1 \end{cases} \Bigg)$$

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \varepsilon_t,\ t \in \mathrm{Z},$$

where $\{\varepsilon_t\}$ is a white noise with variance σ^2; $\varphi_2 \neq 0$; $\varphi_1 + \varphi_2 < 1$; $-\varphi_1 + \varphi_2 < 1$; $-1 < \varphi_2 < -1$

(*autoregressive process* of the second order $AR(2)$ (see Sect. 30.4): is a (weakly) stationary invertible stochastic process with zero mean value and autocorrelation function (see Sect. 30.1) of the form (G_1 and G_2 are distinct roots of the equation $1 - \varphi_1 B - \varphi_2 B^2 = 0$)

$$B(t) = \frac{G_1^{-1}(1 - G_2^{-2})G_1^{-\,t} - G_2^{-1}(1 - G_1^{-2})G_2^{-\,t}}{(G_1^{-1} - G_2^{-1})(1 + G_1^{-1}G_2^{-1})}, \quad t = 1, 2, \ldots \Big)$$

$$X_t = \varphi_1 X_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1},\ t \in \mathrm{Z},$$

where $\{\varepsilon_t\}$ is a white noise with variance σ^2; $\varphi_1 \neq 0$; $\theta_1 \neq 0$; $|\varphi_1| < 1$

(*mixed process* of the first orders *ARMA*(1, 1) (see Sect. 30.4): is a (weakly) stationary invertible stochastic process with zero mean value and autocorrelation function (see Sect. 30.1) of the form

$$B(t) = \varphi_1^{t-1}\frac{(1 + \varphi_1\theta_1)(\varphi_1 + \theta_1)}{1 + 2\varphi_1\theta_1 + \theta_1^2}, \quad t = 1, 2, \ldots;$$

if $|\theta_1| < 1$, then the process *ARMA*(1, 1) is invertible)

$$\mathrm{AIC}(k,\ l) = \ln \hat{\sigma}_{k,\,l}^2 + \frac{2(k + l)}{n},$$

where $\hat{\sigma}_{k,\,l}^2$ is the estimated variance of the estimated white noise when modeling a given time series $x_1, \ldots, x_n$ by means of the model *ARMA*(k, l)

(*Akaike's criterion* identifying orders of model *ARMA*: the estimated orders $\hat{p}, \hat{q}$ of the model *ARMA* are obtained as arguments k, l minimizing *AIC*(k, l); one can apply other criteria, e.g. *BIC*, and the like)

$$Q = n \sum\nolimits_{t=1}^{K} \left(\hat{B}(t; \hat{\varepsilon}_t)\right)^2,$$

where $\hat{\varepsilon}_t = x_t - \hat{\varphi}_1 x_{t-1} - \ldots - \hat{\varphi}_p x_{t-p} - \hat{\theta}_1 \hat{\varepsilon}_{t-1} - \ldots - \hat{\theta}_q \hat{\varepsilon}_{t-q}$ is the (recursively) estimated white noise in the estimated model *ARMA*(p, q), and $\hat{B}(t; \hat{\varepsilon}_t)$ is the estimated autocorrelation function (see Sect. 31.3) of this estimated white noise

(*portmanteau statistics* verifying constructed model *ARMA*(p, q))

$$Q > \chi_{1-\alpha}^2(K - p - q)$$

(critical region at significance level α (see Sects. 26.5 and 27.10) of *portmanteau test* verifying hypothesis that a given time series is compatible with the constructed model *ARMA*(p, q): K is a chosen natural number with recommended value $K \sim \sqrt{n}$)

$$\Delta^k x_t = \binom{k}{0} x_t - \binom{k}{1} x_{t-1} + \binom{k}{2} x_{t-2} - \ldots + (-1)^k \binom{k}{k} x_{t-\,k},\ k \in N;$$
$$t = \ k + 1, \ldots, n$$

(kth *difference* of time series $x_1, \ldots, x_n$: in particular, the first difference is simply $\Delta x_t = x_t - x_{t-1}$ and the second difference is $\Delta^2 x_t = \Delta(\Delta x_t) = x_t - 2x_{t-1} + x_{t-2}$)

$$\Delta^k = (1 - B)^k,\ k \in \mathrm{N}$$

(*difference operator*: e.g. $\Delta^2 x_t = (1 - B)^2 x_t = x_t - 2Bx_t + B^2 x_t = x_t - 2x_{t-1} + x_{t-2}$)

$$\{X_t\} \sim I(d),\ d \in N_0$$

(*integrated process* of order d: is a stochastic process, for which d is the minimal differencing order such that the differenced process is (weakly) stationary; in particular, the symbol $I(0)$ denotes a (weakly) stationary process)

$$\varphi(B)\Delta^d X_t = \theta(B)\varepsilon_t,\ d \in \mathrm{N},$$

where $\{\varepsilon_t\}$ is a white noise with variance σ^2; $\varphi_p \neq 0$; $\theta_q \neq 0$; the autoregressive polynomial $\lambda^p - \varphi_1\lambda^{p-1} - \ldots - \varphi_p$ has all roots inside the unit circle in the complex plane

(*integrated mixed process* of orders p, d, q (*ARIMA*(p, d, q)): in general, it is a *nonstationary* stochastic process in discrete time with the state space $S = \mathrm{R}$; one can treat it as the model *ARMA*(p, q) for the dth *differences* $\{\Delta^d X_t\}$ of the process $\{X_t\}$; it is suitable for nonstationary time series, which can be transferred into stationary ones by differencing them (e.g. in the case of a time series fluctuating around a constant level, which changes in jumps, one achieves stationarity by means of the first differences); Box-Jenkins methodology makes use of the *ARIMA* models when modeling nonstationary time series)

$$\Delta X_t = \varepsilon_t,\ \text{where}\ \{\varepsilon_t\}\ \text{is a white noise with variance}\ \sigma^2$$

(integrated mixed process *ARIMA*(0, 1, 0)): is a *random walk* with the continuous state space (see Sect. 30.4); the attribute "integrated" is obvious, when one writes $x_t = \varepsilon_t + \varepsilon_{t-1} + \ldots\ldots + \varepsilon_{t-k} + x_{t-k-1}$)

$$H_0 : \varphi_1 = 1\ \text{for model}\ X_t = \varphi_1 X_{t-1} + \varepsilon_t$$

(*Dickey-Fuller tests of unit roots*: answer the important question, whether differencing a given time series is really necessary (e.g.: is the suitable model for $x_1, \ldots, x_n$ a stationary process *AR*(1) or a nonstationary process *ARIMA*(0, 1, 0)?); such tests use statistics of the type $(\hat{\varphi}_1 - 1)/\hat{\sigma}(\hat{\varphi}_1)$ with tabulated critical values and have been augmented for general models *ARIMA*(p, d, q))

$$\varphi(B)\Phi(B^{12})\Delta^d \Delta_{12}^D\, x_t = \theta(B)\Theta(B^{12})\varepsilon_t,$$

where $\{\varepsilon_t\}$ is a white noise with variance σ^2; $\Phi(B^{12}) = 1 - \Phi_1 B^{12} - \ldots - \Phi_P B^{12P}$; $\Theta(B^{12}) = 1 + \Theta_1 B^{12} + \ldots + \Theta_Q B^{12Q}$; $\Delta_{12} = 1 - B^{12}$

(*multiplicative seasonal process* of order p, d, q, P, D, Q with length of season $s = 12$ (*SARIMA*$(p,\ d,\ q) \times (P,\ D,\ Q)_{12}$): is a *seasonal* stochastic process with monthly observations (the model is analogous for other lengths of season); one applies the same model $\Phi(B^{12})\Delta_{12}^{D}\, x_t = \Theta(B^{12})\eta_t$ for all time series of the type $\{\ldots, x_{Jan04}, x_{Jan05}, \ldots\}$, $\{\ldots, x_{Feb04}, x_{Feb05}, \ldots\}, \ldots$, where the series $\{\eta_t\}$, which joins multiplicatively the series for particular months, is *ARIMA*$(p,\ d,\ q)$ of the form $\varphi(B)\ \Delta^d\ \eta_t = \theta(B)\varepsilon_t$; Box-Jenkins methodology makes use just of these *SARIMA* models when modeling seasonal time series)

$$(1 - B)(1 - B^{12})x_t = (1 + \theta_1 B)(1 + \Theta_1 B^{12})\varepsilon_t,$$

$$\text{i.e.}\ x_t = x_{t-1} + x_{t-12} - x_{t-13} + \varepsilon_t + \theta_1\, \varepsilon_{t-1} + \Theta_1\, \varepsilon_{t-12} + \theta_1 \cdot \Theta_1\, \varepsilon_{t-13}$$

(example of *SARIMA*$(0, 1, 1) \times (0, 1, 1)_{12}$)

$$\Delta^d\, X_t = \varepsilon_t,\ d \in \mathrm{R},\ \text{where}\ \{\varepsilon_t\}\ \text{is a white noise with variance}\ \sigma^2$$

(*fractionally integrated process* (*FI*): $\Delta^d = (1 - B)^d$ for $d \notin \mathrm{Z}$ is *fractional difference* obtained by means of (infinite) binomial extension of this expression; the process *FI* is (weakly) stationary, if and only if $d < 1/2$; the process *FI* is invertible, if and only if $d > -1/2$ (then $X_t - \pi_1\, X_{t-1} - \pi_2\, X_{t-2} - \ldots = \varepsilon_t$, where $\pi_k = \Gamma(k - d)/[\Gamma(-d) \cdot \cdot\ \Gamma(k{+}1)]$, see Sect. 25.12); if $0 < d < ½$, then *FI* is stationary, but with $\sum_{t=1}^{\infty} |R(t)| = \infty$ (i.e. values of such a process may be strongly correlated, even if they are very remote at time) so that one calls it *long-memory process* or *persistent process* (with applications in hydrology or for finance time series); more generally, one can use *fractionally integrated mixed process* of orders p, d, q (*ARFIMA*(p, d, q))

31.5 Nonlinear and Financial Time Series

$$\{x_t^{(\lambda)},\ t \in \{1,\ 2, \ldots, n\}\},\ \text{where}\ \lambda \geq 0\ \text{is the parameter of transformation}$$

(*transformed* time series: transformations of time series may have some positive effects, namely (1) homogeneity of volatility; (2) symmetry of distribution, which is skewed before transformation; (3) linearization of the time series so that a linear model can be used after transformation (see Sect. 31.4))

$$x_t^{(\lambda)} = \begin{cases} \dfrac{x_t^{(\lambda)} - 1}{\lambda} & \text{for}\ \lambda \neq 0 \\ \ln x_t & \text{for}\ \lambda = 0 \end{cases}$$

(*Box-Cox transformation*: the value of parameter λ suitable for the given time series (with positive values), which should be transformed, is usually looked for as ML estimate (see Sect. 27.9))

$$x_t^{(\lambda)} = \left\{ \begin{array}{ll} 1/\sqrt{x_t} & \text{for } \lambda = -1/2 \\ \ln x_t & \text{for } \lambda = 0 \\ \sqrt{x_t} & \text{for } \lambda = \quad 1/2 \\ x_t & \text{for } \lambda = \qquad 1 \end{array} \right)$$

(*Jenkins transformation*: confines itself to choice among four possible values; moreover, such a choice may be based on graphical methods)

$$f(X_t, X_{t-1}, \ldots; \varepsilon_1, \varepsilon_{t-1}, \ldots) = 0,$$

where f is a nonlinear function; $\{\varepsilon_t\}$ is a white noise with variance σ^2

(*nonlinear process* in discrete time (*nonlinear time series*): in comparison with the linear process (see Sect. 31.4), it is more suitable to model (1) some real processes, e.g. physical phenomena of the following type: dependence of frequency on amplitude, limit cycles, resonance jumps, and the like; (2) *financial time series*, e.g. special features of development of interest rates, exchange rates and financial indices of the following type: *high frequency data* (e.g. hourly) with irregular record due to continual trading, *volatility clustering* (i.e. tendency for volatility in financial markets to appear in bunches), *conditional heteroscedasticity* (i.e. variable variance conditionally on information on previous time series values), *leptokurtosis* (i.e. tendency for financial asset returns to have distributions with fat tails and excess peakedness at the mean), uncorrelated observations, but highly correlated squared observations (since correlation coefficient is suitable for "linear world" only), distinct response of volatility to large positive and large negative previous values and others)

$$X_t = \sum\nolimits_{i=1}^{q} \sum\nolimits_{j=1}^{p} \beta_{ij} \varepsilon_{t-i} X_{t-j} + \varepsilon_t,$$

where (β_{ij}) is a matrix $(q \times p)$ of parameters; $\{\varepsilon_t\}$ is a white noise with variance σ^2

(*bilinear process* of orders p and q: is based on a bilinear form of previous values of the given process and the white noise; according to the form of matrix (β_{ij}) the bilinear processes are classified to *diagonal* ones with a diagonal matrix, *subdiagonal* ones with a subdiagonal matrix (it has zero elements above the principal diagonal) and *superdiagonal* ones with superdiagonal matrix (it has zero elements under the principal diagonal)

$$X_t = \beta\, \varepsilon_{t-1} X_{t-1} + \varepsilon_t, \text{ where } \{\varepsilon_t\} \text{ is a white noise with variance } \sigma^2$$

(an example of diagonal bilinear process of orders 1 and 1: for $\lambda^2\sigma^2 < 1$ the process is (weakly) stationary with $R(0) = \mathrm{var}(X_t) = \sigma^2(1 + \lambda^2 + \lambda^4)/(1 - \lambda^2)$, $R(1) = \lambda^2\sigma^2$ and $R(t) = 0$ for $t = 2, 3, \ldots$)

$X_t = \beta\, \varepsilon_{t-3} X_{t-2} + \varepsilon_t$, where $\{\varepsilon_t\}$ is a white noise with variance σ^2

(an example of subdiagonal bilinear process of orders 2 and 3: for $\lambda^2\sigma^2 < 1$ the process is (weakly) stationary with $R(0) = \mathrm{var}(X_t) = \sigma^2/(1 - \lambda^2)$ and $R(t) = 0$ for $t = 1, 2, \ldots$)

$$X_t = \sum_{i=1}^{p} \left(\varphi_i + \pi_i \exp(-\gamma\, X_{t-1}^2) \right) + \varepsilon_t,$$

where φ_i, π_i, γ are parameters; $\{\varepsilon_t\}$ is a white noise with variance σ^2

(*exponential autoregressive process* of order p: it models dependence of frequency on amplitude (see *thereinbefore*))

$$X_t = \varepsilon_t + \sum_{j=1}^{q} \left(\theta_j^{(1)} \varepsilon_{t-j}^{+} + \theta_j^{(2)} \varepsilon_{t-j}^{-} \right),$$

where $\theta_j^{(1)}$ and $\theta_j^{(2)}$ are parameters; $\{\varepsilon_t\}$ is a white noise with variance σ^2; $\varepsilon_t^{+} = \max(0,\ \varepsilon_t)$; $\varepsilon_t^{-} = \min(0,\ \varepsilon_t)$

(*asymmetric moving average process* of order q: it behaves asymmetrically for positive and negative values of the white noise; for $\theta_j^{(1)} = \theta_j^{(2)}$ $(j = 1, \ldots, q)$ it becomes the "symmetric" process $MA(q)$, see Sect. 30.4)

$$X_t = \varphi_0^{(k)} + \sum_{i=1}^{p_k} \varphi_i^{(k)} X_{t-i} + \varepsilon_t^{(k)} \quad \text{for}\, x_{t-d} \in T_k (k = 1, \ldots, m),$$

where $d \in \mathrm{N}$, $p_k \in \mathrm{N}_0$, ${\varphi_i}^{(k)} \in \mathrm{R}$, ${\sigma_k}^2 > 0$ are parameters; $\{{\varepsilon_t}^{(k)}\}$ is a white noise with variance ${\sigma_k}^2$; $\{T_k\}$ is a decomposition of the real line into m disjoint subsets (mostly intervals)

(*threshold process SETAR*$(m;\ p_1,\ \ldots,\ p_m)_d$ (Self-Exciting Threshold Autoregressive): shows distinct behaviour presented by distinct autoregressive models according to the location of its previous values among fixed thresholds; it belongs to a more general category of *processes with switching regimes determined by observable process values* (in contrast to Markov-Switching processes *MSW* with random switching of regimes, see thereinafter))

$$\begin{aligned} X_t = {} & \left(\varphi_0^{(1)} + \sum\nolimits_{i=1}^{p_1} \varphi_i^{(1)} X_{t-i}\right) \cdot \left(1 - \frac{1}{1 + \exp\{-\delta \cdot (X_{t-d} - c)\}}\right) \\ & + \left(\varphi_0^{(2)} + \sum\nolimits_{i=1}^{p_2} \varphi_i^{(2)} X_{t-i}\right) \cdot \frac{1}{1 + \exp\{-\delta \cdot (X_{t-d} - c)\}} + \varepsilon_t \end{aligned}$$

where $d \in \mathrm{N}$, $c \in \mathrm{R}$, $\delta > 0$, $p_1, p_2 \in \mathrm{N}_0$, ${\varphi_i}^{(1)}$, ${\varphi_i}^{(2)} \in \mathrm{R}$, $\sigma^2 > 0$ are parameters; $\{\varepsilon_t\}$ is a white noise with variance σ^2

(*process STAR*$(2; p_1, p_2)_d$ (Smooth Transition Autoregressive): differs from the one-threshold model *SETAR*$(2; p_1, p_2)_d$ due to a continuous transition function between the interval under the threshold and the interval above it; instead of the logistic function one can use other transition functions)

$$X_t = \varphi_0^{(s_t)} + \sum_{i=1}^{p_{s_t}} \varphi_i^{(s_t)} X_{t-i} + \varepsilon_t^{(s_t)} \quad \text{for } s_t = 1, \ldots, m,$$

where $p_k \in \mathrm{N}_0$, $\varphi_i^{(k)} \in \mathrm{R}$, $\sigma_k^2 > 0$ ($k = 1, \ldots, m$) are parameters; $\{\varepsilon_t^{(k)}\}$ is a white noise with variance σ_k^2; $P(s_t = j |\ s_{t-1} = i) = p_{ij}$, $i, j = 1, \ldots, m$

(an example of *MSW process* (Markov-Switching, see thereinbefore): shows distinct behavior presented e.g. by distinct autoregressive models, where switching among particular regimes occurs in dependence on the current value of a homogeneous Markov chain $\{s_t\}$ with m states and transition probabilities p_{ij} (see Sect. 30.2); in practice one applies mostly the case of $m = 2$ and $p_1 = p_2$ (the so-called Hamilton's two-state *MSW* process); it belongs to a more general category of *processes with switching regimes determined by unobservable process values*)

$$\sigma_{t+1\,|\,t}^2 = \mathrm{v}ar(X_{t+1}\,|X_t,\ X_{t-1}, \ldots)$$

(conditional variance of stochastic process: is the key concept of nonlinear time series models denoted as *volatility models*, which are mostly consistent with characteristic properties of financial time series (see thereinbefore); when the conditional variance is not constant in time, then such a situation is called *conditional heteroscedasticity*)

$$\left.\begin{aligned} X_t &= \sigma_{t\,|t-1} \times \varepsilon_t \\ \sigma_{t\,|t-1}^2 &= \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j\, \sigma_{t-j\,|t-j-1}^2 = \alpha_0 + \alpha(B) X_t^2 + \beta(B) \sigma_{t\,|t-1}^2 \end{aligned}\right\},$$

where $p, q \in \mathrm{N}$; $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$ are parameters; $\alpha\ (B) = \alpha_1\ B + \ldots + \alpha_p B^p$; $\beta\ (B) = \beta_1\ B + \ldots + \beta_q\ B^q$; random variables $\{\varepsilon_t\}$ are *iid* (0, 1) (or even *iid* $N(0, 1)$)

(*conditional heteroscedasticity process GARCH*$(p,\ q)$ (General Autoregressive Conditional Heteroscedasticity): if $\alpha(1) + \beta(1) < 1$, then the process $\{X_t\}$ is (weakly) stationary (see Sect. 30.1) with the constant variance $\mathrm{var}(X_t) = \alpha_0/(1 - \alpha\ (1) - \beta\ (1))$, the correlation structure of the white noise and the leptokurtic distribution (i.e. fat tails and excess peakedness at the mean, see *thereinbefore*); the process $\{X_t^2\}$ is *ARMA*$(m,\ q)$ where $m = \max\{p,\ q\}$; there are various modifications of the process *GARCH*: *IGARCH* (Integrated *GARCH*, e.g. *IGARCH*(1, 1) is *GARCH*(1, 1) with $\alpha_1 + \beta_1 = 1$, which implies persistency of its volatility), *EGARCH* (Exponential *GARCH* uses logarithms of conditional variances, which enables to reflect the asymmetric effect when modeling volatility), *GARCH-M* (*GARCH*-in-Mean respects dependence between the level and volatility), and others)

$$\left.\begin{array}{l} X_t = \sigma_{t|t-1} \times \varepsilon_t \\ \sigma^2_{t|t-1} = \alpha_0 + \alpha_1 X^2_{t-1} + \beta_1 \sigma^2_{t-1|t-2} \end{array}\right\},$$

where $\alpha_0 > 0$, $\alpha_1 > 0$, $\beta_1 > 0$ are parameters; random variables $\{\varepsilon_t\}$ are *iid* (0, 1) (or even *iid N*(0, 1))

(*conditional heteroscedasticity process GARCH*(1, 1): if $\alpha_1 + \beta_1 < 1$, then the process $\{X_t\}$ is (weakly) stationary (see Sect. 30.1) with the constant variance $\mathrm{var}(X_t) = \alpha_0/(1 - \alpha_1 - \beta_1)$ and the correlation structure of the white noise; if moreover $1 - (\alpha_1 + \beta_1)^2 - 2{\alpha_1}^2 > 0$, then

– kurtosis $\gamma_2 = \dfrac{6\alpha_1^2}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2}$ is higher than for normal distribution

– $\{{X_t}^2\}$ is *ARMA*(1, 1) with autocorrelation function (see Sect. 30.1:

$$B(1) = \alpha_1 + \frac{\alpha_1^2 \beta_1}{1 - 2\alpha_1\beta_1 - \beta_1^2}, \quad B(t) = (\alpha_1 + \beta_1)^{t-1} B(1), \quad t = 2,\ 3, \ldots)$$

$$\left.\begin{array}{l} X_t = \sigma_{t\,|\,t-1} \times \ \varepsilon_t \\ \sigma^2_{t\,|\,t-1} = \alpha_0 + \sum\limits_{i=1}^{p} \alpha_i X^2_{t-i} = \varepsilon_0 + \alpha(B) X_t^2 \end{array}\right\},$$

where $p \in \mathrm{N}$; $\alpha_0 > 0$, $\alpha_i \geq 0$ are parameters; $\alpha\ (B) = \alpha_1\ B + \ldots + \alpha_p\ B^p$; random variables $\{\varepsilon_t\}$ are *iid* (0, 1) (or even *iid N*(0, 1))

(*conditional heteroscedasticity process ARCH*(*p*) (Autoregressive Conditional Heteroscedasticity): if $\alpha(1) < 1$, then the process $\{X_t\}$ is (weakly) stationary (see Sect. 30.1 with the constant variance $\mathrm{var}(X_t) = \alpha_0 /(1 - \alpha(1))$, the correlation structure of white noise and the leptokurtic distribution; the processes *ARCH* in contrast to *GARCH* may not show different responses of volatility to high positive and to high negative previous values of the process)

$$\left.\begin{array}{l} X_t = \sigma_{t\,|\,t-1} \times \ \varepsilon_t \\ \sigma^2_{t\,|\,t-1} = \alpha_0 + \alpha_1 X^2_{t-1} \end{array}\right\},$$

where $\alpha_0 > 0$, $\alpha_1 > 0$ are parameters; random variables $\{\varepsilon_t\}$ are *iid* (0, 1) (or even *iid N*(0, 1))

(*conditional heteroscedasticity process ARCH*(1): if $\alpha_1 < 1$, then the process $\{X_t\}$ is (weakly) stationary (see Sect. 30.1 with the constant variance $\mathrm{var}(X_t) = \alpha_0 / (1 - \alpha_1)$ and the correlation structure of the white noise; if moreover $1 - 3\alpha_1^2 > 0$, then

– kurtosis $\gamma_2 = \dfrac{6\alpha_1^2}{1-3\alpha_1^2}$ is higher than for normal distribution (see Sect. 26.3)
– $\{X_t{}^2\}$ is $AR(1)$ with the autocorrelation function (see Sect. 30.1):

$$B(t) = \alpha_1^t, \quad t = 0, 1, \ldots)$$

31.6 Multivariate Time Series

$$\{\boldsymbol{\varepsilon}_t,\, t \in \mathrm{Z}\},\ \text{where}\, \mathrm{E}(\boldsymbol{\varepsilon}_t) = 0;\, \Sigma_{\boldsymbol{\varepsilon}_s,\boldsymbol{\varepsilon}_t} = \mathrm{E}(\boldsymbol{\varepsilon}_s\boldsymbol{\varepsilon}_t') = \delta_{st}\cdot\ \Sigma(\Sigma > 0),\ s,\, t \in\ \mathrm{Z}$$

(*multivariate white noise*: the random vectors $\{\boldsymbol{\varepsilon}_t\}$ have zero mean values and are serially uncorrelated with constant positively definite variance matrix $\boldsymbol{\Sigma}$)

$$\mathbf{X}_\mathrm{t} = \boldsymbol{\Phi}_1\mathbf{X}_{t-1} + \cdots + \boldsymbol{\Phi}_p\mathbf{X}_{t-p} + \boldsymbol{\varepsilon}_t,\, t \in \mathrm{Z},$$

where $\{\boldsymbol{\varepsilon}_t\}$ is a white noise with variance matrix $\boldsymbol{\Sigma}$; $\boldsymbol{\Phi}_p \neq 0$; all roots of the equation $\det(\mathbf{I}\lambda^p - \boldsymbol{\Phi}_1\lambda^{p-1} - \ldots - \boldsymbol{\Phi}_p) = 0$ lie inside the unit circle in the complex plane

(*multivariate autoregressive process* of order p ($VAR(p)$, Vector Autoregressive): is a (weakly) stationary stochastic process in discrete time with the state space $S = \mathrm{R}^n$, zero mean values and autocovariance function fulfilling *Yule-Walker equations* of the form

$$\mathbf{R}(t) = \begin{cases} \boldsymbol{\Phi}_1\mathbf{R}(t-1) + \ldots\ + \boldsymbol{\Phi}_p\mathbf{R}(t-p) & \text{for } t = 1,\ 2, \ldots \\ \boldsymbol{\Phi}_1\mathbf{R}'(1) + \ldots\ + \boldsymbol{\Phi}_p\mathbf{R}'(p) + \boldsymbol{\Sigma} & \text{for } t = 0 \end{cases};$$

one defines analogously *multivariate moving average process* of order q ($VMA(q)$) and *multivariate mixed process* of orders p and q ($VARMA(p, q)$; the construction of these models in practice is similar as for univariate time series models in the framework of Box-Jenkins methodology (see Sect. 31.4))

$$\mathrm{MSE}\left\{\hat{X}_{t+h} \mid I_t\right\}\ <\ \mathrm{MSE}\left\{\hat{X}_{t+h} \mid I_t \backslash \{Y_\tau\}_{\tau\,\leq\, t}\right\}\ ,$$

where *MSE* on the left hand side of this inequality is the mean squared error of prediction (see Sect. 31.1) for X_{t+h} in the series $\{X_t\}$ based on information I_t up to time t and *MSE* on the right hand side is the mean squared error of prediction for X_{t+h} in the series $\{X_t\}$ based on information I_t up to time t excluding information contained in the past and present of another stochastic process $\{Y_t\}$

(*Granger-causality*: the stochastic process $\{Y_t\}$ Granger-causes the stochastic process $\{X_t\}$; one considers mostly $I_t = \{X_\tau,\, Y_\tau\}_{\tau\,\leq\, t}$; the definition of Granger-causality can be generalized to multivariate stochastic processes $\{\mathbf{X}_t\}$ and $\{\mathbf{Y}_t\}$ and modified in various ways; in particular, the application of this concept is simple in the framework of the vector autoregressive models *VAR* (see *thereinafter*))

$$x_{1t} = \varphi_{11}x_{1,t-1} + \varphi_{12}x_{2,t-1} + \varepsilon_{1t}$$
$$x_{2t} = \varphi_{21}x_{1,t-1} + \varphi_{22}x_{2,t-1} + \varepsilon_{2t}$$

(Granger-causality in bivariate *VAR*(1):
– if $\varphi_{12} \neq 0$: x_2 Granger-causes x_1
– if $\varphi_{21} \neq 0$: x_1 Granger-causes x_2
– if $\varphi_{12} \neq 0$ and $\varphi_{21} \neq 0$: there is *unidirectional relationship from* x_2 *to* x_1
– if $\varphi_{12} = 0$ and $\varphi_{21} \neq 0$: there is *unidirectional relationship from* x_1 *to* x_2
– if $\varphi_{12} \neq 0$ and $\varphi_{21} \neq 0$: there is *feedback* between x_1 and x_2
– if $\varphi_{12} = 0$ and $\varphi_{21} = 0$: x_1 and x_2 are *Granger-independent*)

$$\{\mathbf{a}'\,\mathbf{X}_t\} \sim I(d-b);\ \mathbf{a} \in \mathrm{R}^n,\ d, b \in \mathrm{N},$$

where $\{X_{ti}\} \sim I(d)$ for all components $i = 1, \ldots, n$ (see Sect. 31.4)

(*cointegrated* n-variate process $\{\mathbf{X}_t\} \sim CI(d,\, b)$: if there exists at least one vector $\mathbf{a} \neq \mathbf{0}$ (the so-called *cointegration vector*), then one can by means of a nontrivial linear combination reduce the order d of integrated stochastic processes; there exist at most $0 \leq m < n$ cointegration vectors (the so-called *cointegration rank*); the most frequent case is $d = b$, when one completely eliminates from a group of n stochastic processes by means of a linear combination the long-run tendencies and trends (such a combination is stationary $I(0)$) so that a state of *long-run equilibrium* is achieved)

$$\{X_{t1} - aX_{t2}\} \sim I(0),\ a \neq 0, \text{ where } \{X_{ti}\} \sim I(1) \text{ for } i = 1,\, 2 \text{ (see Sect. 31.4)}$$

(cointegrated processes $\{X_{t1}, X_{t2}\} \sim CI(1,\, 1)$: the simplest case of cointegration)

$$f_{ij}(\lambda),\ C_{ij}(\lambda),\ \Phi_{ij}(\lambda),\ R_{ij}(\lambda)$$

(*mutual spectral density*, *coefficient of coherence*, *coefficient of phase shift*, *coefficient of gain* for processes $\{X_{ti}\}$ and $\{X_{tj}\}$ (see Sect. 30.5))

31.7 Kalman Filter

$$\left.\begin{aligned} \boldsymbol{\theta}_t &= \mathbf{G}_t\boldsymbol{\theta}_{t-1} + \mathbf{w}_t \\ \mathbf{x}_t &= \mathbf{F}_t\boldsymbol{\theta}_t + \mathbf{v}_t \end{aligned}\right\},$$

where $\boldsymbol{\theta}_t$ $(n \times 1)$ is *state vector*; $\mathbf{x}_t$ $(m \times 1)$ is *vector of observations*;
$\mathbf{G}_t$ $(n \times n)$, $\mathbf{F}_t$ $(m \times n)$, $\mathbf{W}_t$ $(n \times n)$, $\mathbf{V}_t$ $(m \times m)$ are matrices known at time t;
$\mathbf{w}_t$ $(n \times 1)$, $\mathbf{v}_t$ $(m \times 1)$ are *residual vectors* fulfilling

$$\mathrm{E}(\mathbf{w}_t) = \mathbf{0},\ \mathrm{E}(\mathbf{v}_t) = \mathbf{0},\ \mathrm{E}(\mathbf{w}_s\,\mathbf{w}_t') = \delta_{st}\mathbf{W}_t,\ \mathrm{E}(\mathbf{v}_s\mathbf{v}_t') = \delta_{st}\mathbf{V}_t,\ \mathrm{E}(\mathbf{w}_s\mathbf{v}_t') = 0$$

(*dynamic linear model DLM*: is a theoretical background of (linear) Kalman filter (in discrete time); the state of a given system (e.g. a trajectory of targetable missile, a parametric model of stochastic process, and the like) is described by means of the state vector $\boldsymbol{\theta}_t$ (e.g. the coordinates of the missile, the vector of all parameters of the statistical model, and the like), which develops in time according to the first equation of *DML* (the so-called *state equation*) under a given *initial condition*, but it is observed indirectly via the vector of observations $\mathbf{x}_t$ according to the second equation of *DML* (the so-called *observation equation*))

$$\left.\begin{aligned} \hat{\boldsymbol{\theta}}_t^t &= \mathrm{E}\,(\boldsymbol{\theta}_t\,|\mathbf{x}_1,\ldots,\mathbf{x}_t) \\ \mathbf{P}_t^t &= \mathrm{E}\left((\hat{\boldsymbol{\theta}}_t^t - \boldsymbol{\theta}_t)(\hat{\boldsymbol{\theta}}_t^t - \boldsymbol{\theta}_t)'\right) \end{aligned}\right\};$$

$$\left.\begin{aligned} \hat{\boldsymbol{\theta}}_t^{t-1} &= \mathrm{E}\,(\boldsymbol{\theta}_t\,|\mathbf{x}_1,\ldots,\mathbf{x}_{t-1}) \\ \mathbf{P}_t^{t-1} &= \mathrm{E}\left((\hat{\boldsymbol{\theta}}_t^{t-1} - \boldsymbol{\theta}_t)(\hat{\boldsymbol{\theta}}_t^{t-1} - \boldsymbol{\theta}_t)'\right) \end{aligned}\right\},\ \text{respectively}$$

(estimated state vector $\boldsymbol{\theta}_t$ at time t and *error matrix* of this estimator; predicted state vector $\boldsymbol{\theta}_t$ at time $t-1$ and *error matrix* of this predictor, respectively)

$$\left.\begin{aligned} \hat{\boldsymbol{\theta}}_t^{t-1} &= \mathbf{G}_t\hat{\boldsymbol{\theta}}_{t-1}^{t-1} \\ \mathbf{P}_t^{t-1} &= \mathbf{G}_t\mathbf{P}_{t-1}^{t-1}\mathbf{G}_t' + \mathbf{W}_t \end{aligned}\right\}\ \text{and}$$

$$\left.\begin{aligned} \hat{\boldsymbol{\theta}}_t^t &= \hat{\boldsymbol{\theta}}_t^{t-1} + \mathbf{P}_t^{t-1}\mathbf{F}_t'\left(\mathbf{F}_t\mathbf{P}_t^{t-1}\mathbf{F}_t' + \mathbf{V}_t\right)^{-1}\left(\mathbf{x}_t - \mathbf{F}_t\hat{\boldsymbol{\theta}}_t^{t-1}\right) \\ \mathbf{P}_t^t &= \mathbf{P}_t^{t-1} - \mathbf{P}_t^{t-1}\mathbf{F}_t'\left(\mathbf{F}_t\mathbf{P}_t^{t-1}\mathbf{F}_t' + \mathbf{V}_t\right)^{-1}\mathbf{F}_t\mathbf{P}_t^{t-1} \end{aligned}\right\}$$

((linear) *Kalman filter* (in discrete time): is a system of recursive formulas for such an estimator of the state vector $\boldsymbol{\theta}_t$ in *DLM*, which is at time t a *linear* function of all past and current observed values $\{\mathbf{x}_1, \ldots, \mathbf{x}_t\}$ and is the *best* one among all linear estimators (i.e. the best linear estimator, see Sect. 27.9); moreover, if $\mathbf{w}_t \sim N(\mathbf{0}, \mathbf{W}_t)$ and $\mathbf{v}_t \sim N(\mathbf{0}, \mathbf{V}_t)$, then this estimator is the best one among all (not only linear) estimators of the state vector; the recursive step from time $t-1$ to time t consists of two substeps (first a construction of prediction for time t at time $t-1$ and then a construction of estimation at time t): however, one can join these substeps to a single explicit formula; *initial values* $\mathbf{x}_0$ and $\mathbf{P}_0$ must be chosen in order to start numerical calculations; a simple example of Kalman filter applications is the method of recursive least squares *RLS* (see Sect. 27.11); there are various generalizations of Kalman filter (for continuous time, nonlinearity, correlated residual vectors and others))

$$\hat{\mathbf{x}}_{t+k}(t) = \hat{\mathbf{x}}_{t+k}^{t} = \mathbf{F}_{t+k}\hat{\boldsymbol{\theta}}_{t+k}^{t} \quad \text{for}\, k \in \mathrm{N}_0, \ \text{where}\ \hat{\boldsymbol{\theta}}_{t+k}^{t} = \mathbf{G}_{t+k}\mathbf{G}_{t+k-1}\ \ldots\ \mathbf{G}_{t+1}\hat{\boldsymbol{\theta}}_{t}^{t}$$

(*prediction in Kalman filter* (for k steps ahead))

$$\left.\begin{aligned} &\boldsymbol{\varphi}_t = \boldsymbol{\varphi}_{t-1} \\ &x_t = \mathbf{X}_t'\boldsymbol{\varphi}_t + \varepsilon_t = (x_{t-1}, \ldots, x_{t-p})\boldsymbol{\varphi}_t + \varepsilon_t \end{aligned}\right\}$$

(*DLM* for autoregressive process $AR(p)$ with parameters, which are fixed at time)

$$\left.\begin{aligned} \hat{\boldsymbol{\varphi}}_t &= \hat{\boldsymbol{\varphi}}_{t-1} + \frac{\mathbf{P}_{t-1}\mathbf{X}_t}{\mathbf{X}_t'\mathbf{P}_{t-1}\mathbf{X}_t + 1}(x_t - \mathbf{X}_t'\hat{\boldsymbol{\varphi}}_{t-1}) \\ \mathbf{P}_t &= \mathbf{P}_{t-1} - \frac{\mathbf{P}_{t-1}\mathbf{X}_t\mathbf{X}_t'\mathbf{P}_{t-1}}{\mathbf{X}_t'\mathbf{P}_{t-1}\mathbf{X}_t + 1} \\ \hat{\sigma}_t^2 &= \frac{1}{t-p}\left((t-p-1)\cdot\hat{\sigma}_{t-1}^2 + \frac{(x_t - \mathbf{X}_t'\hat{\boldsymbol{\varphi}}_{t-1})^2}{\mathbf{X}_t'\mathbf{P}_{t-1}\mathbf{X}_t + 1}\right) \end{aligned}\right\}$$

(recursive estimation of parameters of autoregressive process $AR(p)$ by means of Kalman filter (see *thereinbefore*); the initial values can be chosen e.g. as $\hat{\boldsymbol{\varphi}}_0 = \mathbf{0}$, $\mathbf{P}_0 = c\cdot\mathbf{I}$, where c is a preset high positive number and $\mathbf{I}$ is the unit matrix $p \times p$)

$$\left.\begin{aligned} P_t &= \frac{P_{t-1}}{P_{t-1}x_{t-1}^2 + 1} \\ \hat{\varphi}_t &= \hat{\varphi}_{t-1} + P_t x_{t-1}(x_t - \hat{\varphi}_{t-1}x_{t-1}) \\ \hat{\sigma}_t^2 &= \frac{1}{t-1}\left((t-2)\cdot\hat{\sigma}_{t-1}^2 + \frac{(x_t - \hat{\varphi}_{t-1}x_{t-1})^2}{P_{t-1}x_{t-1}^2 + 1}\right) \end{aligned}\right\}$$

(in particular, recursive estimation of parameters of autoregressive process $AR(1)$ by means of Kalman filter)

$$\left.\begin{aligned} &\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} \\ &x_t = \mathbf{X}_t'\boldsymbol{\beta}_t + \varepsilon_t = (x_{t-1}, \ldots, x_{t-p}, \varepsilon_{t-1}, \ldots,\ \varepsilon_{t-q})\,\boldsymbol{\beta}_t + \varepsilon_t \end{aligned}\right\}$$

(*DLM* for mixed process $ARMA(p, q)$ with parameters, which are fixed at time)

$$\left.\begin{aligned}
\hat{\boldsymbol{\beta}}_t &= \hat{\boldsymbol{\beta}}_{t-1} + \frac{\mathbf{P}_{t-1}\hat{\mathbf{X}}_t}{\hat{\mathbf{X}}_t'\mathbf{P}_{t-1}\hat{\mathbf{X}}_t + 1}(x_t - \hat{\mathbf{X}}_t'\hat{\boldsymbol{\beta}}_{t-1}) \\
\mathbf{P}_t &= \mathbf{P}_{t-1} - \frac{\mathbf{P}_{t-1}\hat{\mathbf{X}}_t\hat{\mathbf{X}}_t'\mathbf{P}_{t-1}}{\hat{\mathbf{X}}_t'\mathbf{P}_{t-1}\hat{\mathbf{X}}_t + 1} \\
\hat{\varepsilon}_t &= x_t - \hat{\mathbf{X}}_t'\hat{\boldsymbol{\beta}}_t; \qquad \hat{\mathbf{X}}_t = (x_{t-1}, \ldots, x_{t-p}, \hat{\varepsilon}_{t-1}, \ldots, \hat{\varepsilon}_{t-q})' \\
\hat{\sigma}_t^2 &= \frac{1}{t-p-q}\left((t-p-q-1)\cdot\hat{\sigma}_{t-1}^2 + \frac{(x_t - \hat{\mathbf{X}}_t'\hat{\boldsymbol{\beta}}_{t-1})^2}{\hat{\mathbf{X}}_t'\mathbf{P}_{t-1}\hat{\mathbf{X}}_t + 1}\right)
\end{aligned}\right\}$$

(recursive estimation of parameters of mixed process *ARMA*(p, q) by means of Kalman filter (see *thereinbefore*))

Further Reading

Abraham, B., Ledolter, J.: Statistical Methods for Forecasting. Wiley, New York (1983)

Bowerman, B.L., O'Connell, R.T.: Time Series and Forecasting. Duxbury Press, North Scituate, MA (1979)

Box, G.E.P., Jenkins, G.M.: Time Series Analysis, Forecasting and Control. Holden-Day, San Francisco, CA (1970)

Brockwell, P.J., Davis, R.A.: Time Series: Theory and Methods. Springer, New York (1987)

Brockwell, P.J., Davis, R.A.: Introduction to Time Series Analysis. Springer, New York (1996)

Brown, R.G.: Smoothing, Forecasting and Prediction of Discrete Time Series. Prentice-Hall, Englewood Cliffs, NJ (1963)

Chatfield, C.: The Analysis of Time Series: Theory and Practice. Chapman and Hall, London (1975)

Fuller, W.A.: Introduction to Statistical Time Series. Wiley, New York (1976)

Granger, C.W.J., Newbold, P.: Forecasting Economic Time Series. Academic, San Diego, CA (1986)

Hamilton, J.D.: Time Series Analysis. Princeton University Press, Princeton, NJ (1994)

Harvey, A.C.: The Econometric Analysis of Time Series. Allan, Oxford (1981)

Kendall, M.: Time-Series. Griffin, London (1976)

Taylor, S.: Modelling Financial Time Series. Wiley, Chichester (1986)

Tsay, R.S.: Analysis of Financial Time Series. Wiley, New York (2002)

Index

T. Cipra, *Financial and Insurance Formulas*, DOI 10.1007/978-3-7908-2593-0,

D

U