UNIVERSITÀ DEGLI STUDI DI UDINE
hic sunt futura
www.uniud.it

Corso di dottorato di ricerca in:
"Scienze Manageriali e Attuariali"

Ciclo 32 °

Heterogeneity and uncertainty in a multistate framework

Dottoranda: Daniela Y. Tabakova

Supervisor: Prof. Gianni Bosi

Co-supervisor: Prof. Ermanno Pitacco

Anno 2020

Abstract

This research further develops the scheme proposed in the paper Pollard (1970). The scheme is based on a two-state model for the analysis of one-year mortality, but the results are also valid for the probabilities related to other types of insurance events such as disablement and accidents. Pollard (1970) proposed a scheme involving calculation of expected value, and the variance of the number of deaths within a given population, under different settings starting from the simplest binomial case, through more general cases where uncertainty is allowed for and more risk classes are considered. In all the settings, the individual events are independent or conditionally independent. Among the main findings presented in the Pollard's paper, we mention the impact of

- splitting the populations into homogeneous classes;
- uncertainty in the assessment of the probability of death

in terms of variance of the number of deaths in the population.
The purpose of this study is to extend the Pollard's original scheme into time-discrete models with more states (active-invalid-dead) together with further investigation into multi-year time horizon. Additionally, hypotheses for real-valued individual frailty are assumed in the models. As a baseline probabilistic structure, we have adopted a traditional three-state model in a Markov context.
We focus on an insurance portfolio. Our outputs of interest are based on the probability distributions of the annual payouts for term insurance policies providing lump sum benefits both in case of death and in case of permanent disability. The analysis of the probability distributions allows us to assess the risk profile of the insurance portfolio, and thus to suggest appropriate actions in terms of premiums and capital allocation. In this regards, we adopt the percentile principle.

Acknowledgements

I would like to express my gratitude to Professor Ermanno Pitacco, for his contributions to our collaboration that has built the main idea of the thesis, along with his essential dedication to the entire project. In particular, special thanks to him for the daily stimulating deliberation, advice and interesting discussion on various heterogeneity and uncertainty aspects, and on their practical implementations in the life insurance area.

I wish to thank the two referees for comments and useful suggestions.

Contents

List of Figures

List of Tables

Introduction

The topic of the present thesis can be placed at the intersection of several scientific and/or technical areas:

- structure of life and health insurance products;
- multistate models, and, in particular, multistate Markov processes;
- heterogeneity modelling, in particular heterogeneity due to unobservable risk factors;
- uncertainty in the parameters of the stochastic models, and assessment of the relevant impact.

In the following Sections we first provide a literature overview, focussing on scientific and technical contributions to the above areas. Then, we single out the main original contributions provided by the present thesis.

Literature overview

Life and health insurance products are addressed in many classical textbooks. In particular, Black and Skipper (1999) provides an extensive and detailed description of a number of insurance products in the area of insurances of the person, from both an economic and a technical perspective.

Mathematical and technical features of life (and, in some cases, health) insurance products are dealt with by all the classical actuarial textbooks. For example: Bowers *et al.* (1997), Dickson *et al.* (2013), Gerber (1995), Olivieri and Pitacco (2015).

The book by Pitacco (2014) is specifically devoted to the basic actuarial models for pricing and reserving for health insurance products (sickness benefits, disability covers, income protection products).

Biometric models (as components of the technical bases for pricing and reserving in life insurance) are presented by all the above textbooks. Here, we only quote the paper by Heligman and Pollard (1980), where new parametric models (or "laws") are proposed, which can represent the age-pattern of mortality over the whole life span. In particular, the first Heligman-Pollard law is adopted in the definition of the technical basis in Chap. 4, and in the numerical examples in Chap. 5 of this thesis.

The use of *Multistate models and, in particular, of Markov structures* to represent the biometric features of insurance products in the area of insurance of the person can be dated back to the seminal contribution by Hoem (1969). The paper by Hoem (1969) places life and other contingencies within the framework of a general, unified, probabilistic theory relying on the Markov assumption; a time-continuous approach is adopted, and formulae and theorems for actuarial values, premiums and reserves are then given. Later contributions to life insurance and related fields, based on multistate models, are given by Amsler (1988), Hoem (1988) and Waters (1984). Conversely, the use of semi-Markov processes in actuarial science and demography is discussed by Hoem (1972). Moving to actuarial textbooks which present the actuarial structure of life and health insurance contracts on multistate models, we quote Haberman and Pitacco (1999), Norberg (2002) and, as specifically regards long-term care insurance, Denuit *et al.* (2019).

As noted by Pitacco (2019), heterogeneity of a population in respect of mortality (and disability) is due to differences among the individuals, which are caused by various risk factors. Some risk factors are observable while others are unobservable. The set of observable risk factors clearly depends on the type of population addressed. It follows that the scientific and technical literature dealing with *heterogeneity modelling* is manifold. Allowing for observable risk factors is a traditional topic in life and health insurance (see, for example, Cummins *et al.* (1983)).

Conversely, the long term and multi-year characteristics of the life insurance contracts and of many health insurance contracts imply difficulties in expressing the impact of unobservable heterogeneity on the individual risk profile. The early contributions to this topic must be credited to Beard (1959) (in the actuarial context) and Vaupel *et al.* (1979) (in the demographic context). The main ideas underlying these contributions are presented in Chap. 1 of this thesis. For a detailed discussion the reader is referred to Pitacco (2019) and bibliographic suggestions therein.

Among the features which heavily affect the risk profile of an insurance portfolio, *uncertainty in the parameters of the stochastic models* should be carefully considered. This topic has raised great interest in the insurance field; see, for example, Cairns

(2000). Parameter uncertainty has a significant impact on the risk profile of a life annuity portfolio or a pension plan; see Pitacco *et al.* (2009) and references therein. As regards uncertainty in the assessment of probabilities (of death in particular) an interesting model has been proposed by Olivieri and Pitacco (2009), while a simple quantification is proposed by Olivieri and Pitacco (2015). The latter is here adopted to express uncertainty in the assessment of probabilities of disablement.

Main contributions of the thesis

Traditional actuarial mathematics has, for a long time, modeled life and health insurance problems according to a quite deterministic approach. Although probabilities are of course used to quantify life contingencies (death, disablement, recovery, etc.), deterministic assessments (typically expected values) have only been provided and used by actuaries.

However, scenarios are affected by a high degree of randomness (concerning biometric issues, financial aspects, etc.), and hence appropriate assessments of results of interest (cash flows, profits, etc.) call for stochastic settings. Stochastic settings are anyway required by current legislations, e.g. the European legislation on solvency requirement.

The main contribution of the present thesis consists in implementing a stochastic approach to the assessment of the annual payouts of a portfolio of life insurance policies also providing lump sum benefits in the case of permanent disability. The "shift" from deterministic to stochastic approach has been realized as follows.

1. Referring to a two-state, one-year setting, interesting generalizations of the classical binomial model have been proposed by Pollard (1970), where the presence of both observable heterogeneity and uncertainty is allowed. In line with generalizaion proposed by Valente (2017) we have defined the following general settings:

   (a) by adding appropriate probability distributions to quantify the uncertainty;

   (b) by allowing for unobservable heterogeneity.

2. The natural setting for the actuarial structure of multi-year life and health insurance policies is provided, since the seminal contributions by Hoem (1969, 1972, 1988) and Amsler (1988), by multistate models with a Markov (or semi-Markov) probabilistic structure. For our model, we have assumed, as the "baseline" setting, a three-state model with a Markov structure. This structure has been adopted, in absence of uncertainty, for homogeneous portfolios as well as for portfolios with

observable heterogeneity. The presence of either uncertainty or unobservable heterogeneity call for more general settings: the resulting structure we have defined consists of Markov processes conditional on the outcomes of the unknown quantities which express uncertainty or unobservable heterogeneity.

3. We have applied the model, structured as described under 1(a), 1(b) and 2, to a portfolio of single-premium, multi-year policies providing death and permanent disability benefits. In line with the need of stochastic assessments, we have calculated, via stochastic simulation, the (simulated) distributions of the payouts for benefits paid in case of death in the active state, in case of death in the disability state, and in case of disablement. The results achieved provides a clear picture of the risk profile of the portfolio throughout time.

The remainder of this thesis is organized as follows. In Chap. 1 we first describe the diverse types of heterogeneity which can affect a life and health insurance portfolio, and some relevant modelling issues; we then provide some insights into the assessment of uncertainty in model parameters. The settings defined by Pollard (1970) and the generalizations we propose are addressed in Chap. 2 (see points 1(a) and 1(b) above).

Starting from the classical multistate model with a Markov structure, described in Chap. 3, a generalized portfolio model embedding possible uncertainty and unobservable heterogeneity has been built-up and described in Chap. 4. Then, in Chap. 5 the generalized model is implemented to perform assessment of the portfolio risk profile; a number of numerical results are presented and discussed.

Finally, suggestions for future research work are provided in the concluding Chap. 6.

Chapter 1

Heterogeneity and uncertainty

1.1 Observable and unobservable risk factors

Evaluations related to life and health insurance products (for pricing, reserving, assessing risk profiles) should account for heterogeneity of the insureds, with respect to probability of disablement, mortality of active (healthy) people, mortality of disabled people, probability of recovery.

Heterogeneity is a consequence of diverse features of the insured risk. In particular, as regards disability and recovery:

- a specific risk exposure, with respect to the probability of disablement;
- the type and the severity of disability, which can affect the probability of recovery.

Generally speaking, heterogeneity is caused by differences among the individuals, which can be explained by a number of risk factors. We can recognize:

- observable risk factors;
- unobservable risk factors.

Some of the observable risk factors are taken into account in the underwriting process at the policy issue, and result, in particular, in the premium rate charged to the policyholder.

Among the observable risk factors, most commonly used when underwriting policies that provide death and disability benefits, we find the following ones:

- occupation risk factors, i.e. related to the specific insured's occupation (working conditions, physical exposures, psychological aspect of work, time pressure, high workload, etc.);

- insured's age at the policy issue;
- insured's gender (which constitutes a risk factor but is not admitted as a rating factor by the current European Union legislation).

According to the observable risk factors, the (potential) insureds' population can be split into a (finite) set of (more or less) homogeneous groups. Indeed, it may be noticed that whatever the type of population concerned, a residual heterogeneity remains among the individuals in each more or less homogeneous group. This residual heterogeneity is due to unobservable risk factors. Mortality is, for example, affected by:

- genetic characteristics;
- lifestyle;
- attitude towards health.

Of course, these risk factors are not embedded into the underwriting process for assessing individual risk exposure of the potential insured persons.

Conversely, the impact of observable risk factors can be assessed and taken into account in the underwriting process. Probabilities of death or disablement for people in poorer or better conditions compared to the "standard" (e.g. average) ones are commonly expressed in relation to the average probabilities. This allows us to work with only one set of probabilities, which will be adjusted in the case we deal with substandard (poorer conditions) or preferred (better conditions) risks.

For instance, additive or multiplicative adjustments to the average mortality rates are well known and widely used in actuarial practice and commonly applied in the pricing procedures (for details, as regards probabilities of death, the reader is referred to Olivieri and Pitacco (2015)).

Hence, when accounting for heterogeneity due to observable risk factors, *individual valuation models* are adopted for adjusting probabilities (of death, disablement, recovery) according to insured's individual features. On the contrary, when allowing for heterogeneity due to unobservable risk factors we need to resort to *collective valuation models*.

Collective models are widely adopted in both the risk theory and the non-life insurance mathematics. The well known Poisson-Gamma model, leading to the negative binomial distribution, constitutes one of the most important examples (see the classical textbook by Bülmann (1970)).

On the contrary, collective models are not very popular in the life and health insurance mathematics and technique. Actually, the long term and multi-year characteristics

of the life insurance contracts and of many health insurance contracts, and hence the dependence of a number of quantitative elements (e.g., the annual expected costs, the policy reserve, etc.) on the insured's attained age, imply difficulties in expressing individual features that cannot be observed. i.e. the unobservable heterogeneity, and (at least partially) justify disregarding those aspects in the life insurance practice (for a detailed discussion, the reader can refer, for example, to Olivieri (2006) and references therein).

Nevertheless, heterogeneity as a result of unobservable risk factors (as well as the relevant impact on the risk profile of insurance portfolios consisting of multi-year policies) can be quantified by adopting the concept of individual "frailty".

1.2 Frailty modelling

1.2.1 General aspects

The first contribution to frailty modelling, in the framework of mortality analysis, dates back to end of the Fifties, and is due to Beard (1959). The proposed approach is based on a non-negative real-valued variable, indeed called frailty, whose level expresses the impact of unobservable risk factors on the individual age-pattern of mortality. The basic conclusion is that people with a higher level of frailty tend to die earlier compared to the others. Although the paper by Beard (1959) constitutes the earliest contribution to frailty, the first formal representation was provided by Vaupel *et al.* (1979), where two hypotheses are assumed:

- all the unobservable risk factors for any individuals are summarized by a non-negative unknown variable (the frailty);
- the value of the individual frailty does not change throughout time, and remains unknown.

More generally, frailty models can be adopted to represent unobservable characteristics in diverse types of risk exposure. So, in the framework of life and health insurance, we can include individual frailty to quantify in stochastic terms contingencies like:

- death (separately of active and disabled insureds, if applicable);
- disablement;
- recovery.

Frailty models and, more generally, heterogeneity models can be classified by adopting diverse criteria. As regards heterogeneity in mortality (of both active and disabled people, as well as in general population), a detailed review is provided by Pitacco (2019) (see also the extensive reference list). Here we only focus on the following classifications, which can help in understanding the features of the model we are proposing.

In relation to the assessment of heterogeneity due to unobservable risk factors, the two following basic approaches can be recognized.

1. A discrete approach: heterogeneity is expressed by a (linear) combination of appropriate functions, where each function expresses mortality (or disability, or recovery) in a homogeneous group inside the heterogeneous population. Because of the characteristics of heterogeneity, the weights in the combination are unknown (and hence constitute random variables). In terms of frailty, each homogeneous group can be labeled by a frailty level, whereas the frailty level changes moving from one group to another. See also Sect. 1.2.2.

2. A continuous approach, based on a non-negative quantity, the frailty, whose level expresses the risk factors affecting the individual mortality, or disability. The individual frailty is unknown (and hence constitutes a random variable). As regard the relation between the frailty and the individual age, we find:

   (a) constant frailty models (fixed-frailty approaches), in which the value of the individual frailty is unknown, but does not change throughout the lifetime;

   (b) variable frailty models, in which the individual frailty level is age-dependent, stochastically changing over the whole life span.

The choice of a particular model obviously depends on what set of unobservable risk factors must be summarized by the model itself.

1.2.2 A discrete approach to heterogeneity

As suggested by Pitacco (2019), most of the models in the framework of the discrete approach can be formalized as follows.

Refer, e.g., to the mortality in a population. For brevity, independent variables (such as age or time) are not shown in the following notation of the mortality functions. Let f denote a biometric function; for example:

- the instantaneous force of mortality (or mortality intensity, or hazard function) μ;
- the annual probabilities of dying q;

- the survival function, i.e. the expected number ℓ of survivors in a cohort, or the probability of survival S;
- the life expectancy (e.g. at the birth) e.

For a population split into r groups, the function f is then expressed as a mixture of the functions $f^{(i)}$, $i = 1, \ldots, r$, pertaining to the various groups:

$$f = w_1 f^{(1)} + w_2 f^{(2)} + \cdots + w_m f^{(r)} \tag{1.1}$$

Several specific models can be placed in the framework described by Eq. (1.1). For example:

1. the functions $f^{(i)}$, $i = 1, \ldots, r$, can be suggested by various observable risk factors (e.g. individual health status, individual occupation, geographical area, etc.), and may be either known or unknown (at least as regards the relevant parameters) depending on information available about the age-pattern of mortality inside each group;
2. the weights w_i, $i = 1, \ldots, r$, may be either known or unknown, depending on information available about the (relative) group sizes;
3. the individual age-pattern of mortality over lifetime may be either fixed, i.e. each individual remains lifelong in a given group, or variable, i.e. each individual can move from one group to another one; the age-pattern is fixed in respect of the heterogeneity factor.

We in particular note that both the setting 1 with unknown functions, and 2 with unknown weights can capture the impact of unobservable heterogeneity.

1.2.3 A continuous approach to frailty: one-year probabilities

When one-year probabilities are involved, a simple approach to modelling individual frailty consists in considering the relevant probabilities as random quantities, following an appropriate distribution.

An appropriate tool is provided by the two-parameter beta distribution, Beta(α, β), whose support is the real interval $[0, 1]$. Thanks to the two parameters, the beta distribution has a high degree of flexibility which allows to capture a broad range of situations.

In particular, expected value and variance of a generic random variable X beta-distributed are respectively given by:

$$\mathbb{E}(X) = \frac{\alpha}{\alpha + \beta} \tag{1.2a}$$

$$\mathrm{Var}(X) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} \tag{1.2b}$$

It may be interesting to set a (reasonable) upper bound smaller than 1 (in particular when probabilities of death or disablement are concerned). To do this, the four-parameter beta distribution can be used, usually denoted as $\mathrm{Beta}(\alpha, \beta, a, c)$, where a and c respectively define the lower bound and the upper bound of the support. In particular, we can set $a = 0$ and, of course $0 < c < 1$.

It can easily be proved that if $X \sim \mathrm{Beta}(\alpha, \beta)$, then the transformation $Y = X(c - a) + a$ has a distribution $\mathrm{Beta}(\alpha, \beta, a, c)$, with expected value and variance respectively given by:

$$\mathbb{E}(Y) = \frac{\alpha c + \beta a}{\alpha + \beta} \tag{1.3a}$$

$$\mathrm{Var}(Y) = \frac{\alpha\beta(c - a)^2}{(\alpha + \beta)^2(\alpha + \beta + 1)} \tag{1.3b}$$

The parameters of the beta distribution can be either constant, that is, independent of the attained age, or variable, and hence depending on the attained age. This way, we can implement either a fixed-frailty or a variable-frailty model (see point 2 in Sect. 1.2.1).

1.3 Uncertainty

As noted in the Introduction, uncertainty in modelling constitutes one of the features which heavily affect the risk profile of an insurance portfolio. Uncertainty may regard diverse aspects of the modelling process. It is usual to distinguish, in particular, between model uncertainty and parameter uncertainty.

For example, referring to biometric modelling in life insurance, we note that:

- model uncertainty is related to the choice of an appropriate model to express the age-pattern of mortality (for example: Gompertz law, Makeham law, Perks law, Weibull law, etc.);

- after chosing a model, parameter uncertainty is related to the parameter values, and can be caused by lack of data in the estimation process.

In probabilistic and actuarial terms, uncertainty risk is the consequence of uncertainty in the choice of the model and/or the relevant parameters. It should be stressed that uncertainty risk is an "aggregate risk", as it simultaneously affects all the individuals in a population (or in an insurance portfolio). As well known, aggregate risks are not diversifiable via pooling, and hence call for transfers outside the usual insurance/reinsurance framework.

Expressing uncertainty risk and relevant impacts is, of course, strictly related to the type of model and its parametrization. In the problems we are dealing with, possible uncertainty will be accounted for in relation to one-year probabilities, as specified in Sect. 2.2.1.

A simple and effective way for expressing uncertainty in the assessment of a probability consists in considering the probability itself as a random quantity, following an appropriate distribution. Again, an appropriate tool is provided by the family of beta distributions (see Sect. 1.2.3): thanks to the parameters, the flexibility of the beta distributions allows us to capture a broad range of uncertainty situations.

Chapter 2

Two-state models: one-year time horizon

2.1 The Pollard scheme: general aspects

The analysis of mortality rates q_x and of expected values of number of deaths has always been of great interest to actuaries. Conversely, little attention has been paid by practitioners to possible uncertainty in the mortality rates and to relevant impact on the number of deaths.

Consider the fact that the mortality rate at any age x, in a given population, summarizes the mortality due to different causes. For some of these causes, for example mortality from pneumonia, we can expect that the population value of q_x varies over time with the climate conditions. Clearly, this causes uncertainty in (future) mortality rates.

Pollard (1970) proposed a scheme involving calculation of the expected value, and the variance of the number of deaths D in a year, within a population of n individuals, under different hypotheses concerning mortality in the population itself. As noted by Pitacco (2019), the Pollard scheme can be described as follows, moving from the simplest to more general settings.

1. The population consists of one group of independent risks, and the same probability of dying q affects all the individuals; hence, the number of events follows a binomial distribution.

2. The population is split into r groups of independent risks, each one with given size n_i and given probability q_i. This setting expresses a situation of "known" (that is,

observable) heterogeneity. It can be proved that the variance of the total number of events is lower than in case 1.

3. The population consists of one group, and an unknown, i.e. random, probability Q, with given expected value and variance, affects all the individuals; the risks are assumed independent conditionally on any possible outcome of Q. The variance of the total number of events is higher than in case 1.

4. Combining features of setting 2 and setting 3 leads to a variance of the total number of events which, compared to setting 1, is lowered because features of 2 and increased because features of 3.

5. Thanks to observable risk factors, the population is split into r groups of independent risks, each group with random size N_i but given probability q_i. This setting expresses a situation of "unknown" heterogeneity, discrete-valued in particular. The variance of the total number of events, compared to setting 1, is lowered thanks to the splitting into groups but increased because of their random sizes.

6. The population is split into r groups, each one with random size N_i and random probability Q_i. Also this setting expresses a situation of "unknown" heterogeneity. Settings 1 to 5 can be recognized as particular cases of this general setting.

It is worth noting that all the above settings can be traced back to the scheme defined by Eq. (1.1), referred, for example, to one-year mortality.

Further, we note that in the original Pollard's scheme no probability distribution has been assumed, neither to quantify the uncertainty in the assessment of the probabilities (settings 3, 4 and 6), nor for the group random sizes (settings 5 and 6). We propose specific assumptions, as described in Sect. 2.2. In particular, to express uncertainty in the assessment of probabilities, we propose the four-parameter beta distribution $\text{Beta}(\alpha, \beta, a, c)$, where, as noted in Sect. 1.3, the parameters a and c respectively define the lower bound and the upper bound of the support of the distribution.

2.2 The Pollard scheme: formalization and generalizations

For the details regarding the original scheme, the reader can refer to the paper by Pollard (1970). In the following Sections, starting from a formalization of the scheme

itself (see Settings 1 to 6 in Sect. 2.2.1), we propose two generalizations (see Settings 7 and 8 in Sect. 2.2.2). Numerical examples are presented in Sect. 2.2.3.

Following the original paper, we focus on the analysis of mortality; however, all the results keep their validity for the probabilities of disablement, and more generally, for probabilities related to whatever type of insured event.

2.2.1 Homogeneous groups

The following settings have been considered by Pollard (1970). Probabilistic assumptions are here proposed to completely define the relevant stochastic structures.

2.2.1.1 Setting 1

Let n persons of a given age have the same known probability of dying q. Assume that individual deaths are independent events. Hence, the number D of deaths in the population follows a binomial distribution $\mathrm{Bin}(n, q)$. Therefore we find:

$$\mathbb{E}(D) = n\,q \tag{2.1a}$$

$$\mathrm{Var}(D) = n\,p\,q, \quad \text{where } p = 1 - q. \tag{2.1b}$$

2.2.1.2 Setting 2

We refer to a population of n persons, split into r sub-groups, each one with known size n_i, $1, 2, \ldots, r$. Within the groups, the probability of dying within the year is q_i; we assume that the probabilities are known, and $q_1 < q_2 < \cdots < q_r$. Again, we assume that individual deaths are independent events.

The number of deaths D is given by $D_1 + D_2 + \cdots + D_r$, where $D_1, D_2, \ldots D_r$ are independent random variables.

We obtain:

$$\mathbb{E}(D) = \sum_{i=1}^{r} \mathbb{E}(D_i) = \sum_{i=1}^{r} n_i q_i = nq, \tag{2.2a}$$

$$\text{where } q = \frac{1}{n} \sum_{i=1}^{r} n_i q_i \qquad \text{is the average rate of mortality.}$$

$$\begin{aligned}
\mathrm{Var}(D) &= \sum_{i=1}^{r} \mathrm{Var}(D_i) = \sum_{i=1}^{r} n_i q_i (1 - q_i) = nq - \sum_{i=1}^{r} n_i q_i^2 \\
&= nq - \sum_{i=1}^{r} n_i (q_i - q)^2 + q^2 \sum_{i=1}^{r} n_i - 2q \sum_{i=1}^{r} n_i q_i \\
&= nq - \sum_{i=1}^{r} n_i (q_i - q)^2 + nq^2 - 2nq^2 \\
&= nq - \sum_{i=1}^{r} n_i (q_i - q)^2 - nq^2 = npq - \sum_{i=1}^{r} n_i (q_i - q)^2.
\end{aligned} \tag{2.2b}$$

Splitting the population into r groups is due to observable heterogeneity, which reduces the variance of the number of deaths D. The fact is obvious considering the extreme case in which a part of the population is certain to die ($q_1 = 1$) and the remaining is certain to survive ($q_2 = 0$); in this case, the outcome is certain and the variance of D equal to zero. More in general, in the case of observable heterogeneity the variance is lower than in the case of homogeneous population (that is, Setting 1).

2.2.1.3 Setting 3

Assume that the n persons in the population have the same unknown probability of dying Q, where Q is a random variable. Assume that Q follows a four-parameter beta distribution $\mathrm{Beta}(\alpha, \beta, a, c)$. Individual deaths are assumed independent events conditional on any possible outcome of Q.

Hence, we obtain:

$$\mathbb{E}(D|Q) = nQ \Rightarrow \mathbb{E}(D) = \mathbb{E}\left[\mathbb{E}(D|Q)\right] = \mathbb{E}(nQ) = n\mathbb{E}(Q) \tag{2.3a}$$

$$\left(\mathbb{E}(D)\right)^2 = n^2\left(\mathbb{E}(Q)\right)^2 \tag{2.3b}$$

$$\begin{aligned}\mathbb{E}(D^2|Q) &= \mathrm{Var}(D|Q) + \left(\mathbb{E}(D|Q)\right)^2 = nQ(1-Q) + n^2Q^2 = nQ - nQ^2 + n^2Q^2 \\ &= n(n-1)Q^2 + nQ\end{aligned} \tag{2.3c}$$

$$\begin{aligned}\mathbb{E}(D^2) &= \mathbb{E}\left[\mathbb{E}(D^2|Q)\right] = \mathbb{E}\left[n(n-1)Q^2 + nQ\right] = n(n-1)\mathbb{E}(Q^2) + n\mathbb{E}(Q) \\ &= n(n-1)\left[\mathrm{Var}(Q) + \left(\mathbb{E}(Q)\right)^2\right] + n\mathbb{E}(Q)\end{aligned} \tag{2.3d}$$

$$\begin{aligned}\mathrm{Var}(D) &= \mathbb{E}(D^2) - \left(\mathbb{E}(D)\right)^2 = n(n-1)\left[\mathrm{Var}(Q) + \left(\mathbb{E}(Q)\right)^2\right] + n\mathbb{E}(Q) - n^2\left(\mathbb{E}(Q)\right)^2 \\ &= n(n-1)\mathrm{Var}(Q) + \cancel{n^2\left(\mathbb{E}(Q)\right)^2} - n\left(\mathbb{E}(Q)\right)^2 + n\mathbb{E}(Q) \cancel{- n^2\left(\mathbb{E}(Q)\right)^2} \\ &= n\mathbb{E}(Q)(1-\mathbb{E}(Q)) + n(n-1)\mathrm{Var}(Q) \\ &= n\frac{\alpha c + \beta a}{\alpha+\beta}\left(1 - \frac{\alpha c + \beta a}{\alpha+\beta}\right) + n(n-1)\frac{\alpha\beta(c-a)^2}{(\alpha+\beta)^2(\alpha+\beta+1)}\end{aligned} \tag{2.3e}$$

We note that, in case Q has beta distribution with considerable high variance, the number of deaths D can have significant random deviations, because of the positive relationship between $\mathrm{Var}(D)$ and $\mathrm{Var}(Q)$.

In this Setting, considering the particular case of two-parameter distribution $\mathrm{Beta}(\alpha, \beta)$, we can obtain interesting results in terms of the probability distribution of D and the coefficient of variation $\mathrm{CV}(D)$. The random variable D follows the beta-binomial distribution, which is a generalization of the binomial distribution, in which the parameter is not certain, but follows a beta distribution with parameters α, β. We find:

$$\mathbb{E}(D) = n\mathbb{E}(Q) = n\frac{\alpha}{\alpha+\beta} \tag{2.4a}$$

$$\begin{aligned}\mathrm{Var}(D) &= n\frac{\alpha}{\alpha+\beta}\left(1 - \frac{\alpha}{\alpha+\beta}\right) + n(n-1)\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \\ &= n\frac{\alpha}{\alpha+\beta}\frac{\beta}{\alpha+\beta} + n^2\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} - n\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \\ &= n\frac{\alpha\beta(\alpha+\beta+1) - \alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} + n^2\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \\ &= n\frac{\alpha\beta(\alpha+\beta)}{(\alpha+\beta)^2(\alpha+\beta+1)} + n^2\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \\ &= n\frac{\alpha\beta}{(\alpha+\beta)(\alpha+\beta+1)} + n^2\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}\end{aligned} \tag{2.4b}$$

Moreover:

$$\mathrm{CV}(D) = \sqrt{\frac{n\frac{\alpha\beta}{(\alpha+\beta)(\alpha+\beta+1)} + n^2\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}}{n^2(\mathbb{E}(Q))^2}} = \sqrt{\frac{\frac{1}{n}\frac{\alpha\beta}{(\alpha+\beta)(\alpha+\beta+1)} + \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}}{(\mathbb{E}(Q))^2}} \quad (2.5)$$

The first term of the numerator under square root depends on the size n and represents random fluctuations, whereas the second one does not depend on n and represents the systematic deviations. We note that:

$$\lim_{n\to\infty} \mathrm{CV}(D) = \sqrt{\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}\frac{1}{(\mathbb{E}(Q))^2}} \quad (2.6)$$

This result has a significant impact on the risk profile of an insurance portfolio: systematic deviations constitute a risk component which is not diversifiable by pooling, and hence calls for transfers outside the usual insurance/reinsurance framework.

2.2.1.4 Setting 4

We refer to a population of n persons, split into r groups each one with given size n_i, $i = 1, 2, \ldots, r$. All the individuals in the group i have the same unknown probability of dying Q_i. Thus, the Q_i are random variables for which we assume the beta distribution $\mathrm{Beta}(\alpha_i, \beta_i, a_i, c_i)$, such that $\mathbb{E}(Q_1) < \mathbb{E}(Q_2) < \cdots < \mathbb{E}(Q_r)$. The number of deaths D in the population is given by $D_1 + D_2 + \cdots + D_r$, where $D_1, D_2, \ldots D_r$ are independent conditional on any outcome of the Q_i.
We obtain:

$$\mathbb{E}(D_i|Q_i) = n_i Q_i \Rightarrow \mathbb{E}(D_i) = n_i\mathbb{E}(Q_i) \Rightarrow (\mathbb{E}(D_i))^2 = n_i^2\,(\mathbb{E}(Q_i))^2 \quad (2.7a)$$

$$\mathbb{E}(D_i^2|Q_i) = \mathrm{Var}(D_i|Q_i) + (\mathbb{E}(D_i|Q_i))^2 = n_i Q_i(1-Q_i) + n_i^2 Q_i^2 \quad (2.7b)$$

$$\begin{aligned}\mathbb{E}(D_i^2) &= \mathbb{E}\left[\mathbb{E}(D_i^2|Q_i)\right] = n_i(n_i-1)\mathbb{E}(Q_i^2) + n_i\mathbb{E}(Q_i) \\ &= n_i(n_i-1)\left[\mathrm{Var}(Q_i) + (\mathbb{E}(Q_i))^2\right] + n_i\mathbb{E}(Q_i)\end{aligned} \quad (2.7c)$$

$$\begin{aligned}
\mathrm{Var}(D) &= \sum_{i=1}^{r} \mathrm{Var}(D_i) = \sum_{i=1}^{r} \left[\mathbb{E}(D_i^2) - (\mathbb{E}(D_i))^2\right] \\
&= \sum_{i=1}^{r} \left[n_i(n_i-1)\left[\mathrm{Var}(Q_i) + (\mathbb{E}(Q_i))^2\right] + n_i\mathbb{E}(Q_i) - n_i^2\left(\mathbb{E}(Q_i)\right)^2\right] \\
&= \sum_{i=1}^{r} \left[n_i\mathbb{E}(Q_i)(1-\mathbb{E}(Q_i)) + n_i(n_i-1)\mathrm{Var}(Q_i)\right] = \\
&= npq - \sum_{i=1}^{r} n_i \left(\mathbb{E}(Q_i) - q\right)^2 + \sum_{i=1}^{r} n_i(n_i-1)\frac{\alpha\beta(c-a)^2}{(\alpha+\beta)^2(\alpha+\beta+1)} \\
&= npq - \sum_{i=1}^{r} n_i \left(\frac{\alpha_i c_i + \beta_i a_i}{\alpha_i+\beta_i} - q\right)^2 + \sum_{i=1}^{r} n_i(n_i-1)\frac{\alpha_i\beta_i(c_i-a_i)^2}{(\alpha_i+\beta_i)^2(\alpha_i+\beta_i+1)}
\end{aligned} \tag{2.7d}$$

where $q = \dfrac{1}{n}\sum_{i=1}^{r} n_i\mathbb{E}(Q_i) = \dfrac{1}{n}\sum_{i=1}^{r} n_i\dfrac{\alpha_i c_i + \beta_i a_i}{\alpha_i+\beta_i}$ is the average rate of mortality.

Thus, the observed heterogeneity reduces the random fluctuations in the number of deaths D, which however will be increased by the variance of the random variables Q_i.

2.2.1.5 Setting 5

We refer to a population of n persons, split into r groups each one with random size N_i, $i = 1, 2, \ldots, r$, where the random vector $\underline{N} = (N_1, N_2, \ldots, N_r)$ has multinomial distribution with parameters $(n, f_1, f_2, \ldots, f_r)$, such that $n = N_1 + N_2 + ... + N_r$ and with $f_1 > f_2 > \cdots > f_r$, $f_1 + f_2 + \cdots + f_r = 1$. Thus, a higher probability of belonging to groups with lower mortality is assumed. The probabilities f_i are assigned, taken from the observation of other groups with similar characteristics. The probability of dying is q_i, given, for all the individuals in the group i, with $q_1 < q_2 < \cdots < q_r$. The number of deaths D is equal to $D_1 + D_2 + \cdots + D_r$.

We obtain:

$$\mathbb{E}(D_i|N_i) = N_i q_i \Rightarrow \mathbb{E}(D_i) = \mathbb{E}(N_i q_i) = \mathbb{E}(N_i) q_i \Rightarrow (\mathbb{E}(D_i))^2 = (\mathbb{E}(N_i))^2 q_i^2 \tag{2.8a}$$

$$\mathbb{E}(D_i^2|N_i) = \mathrm{Var}(D_i|N_i) + (\mathbb{E}(D_i|N_i))^2 = N_i q_i(1-q_i) + N_i^2 q_i^2 \tag{2.8b}$$

$$\begin{aligned}\mathbb{E}(D_i^2) &= \mathbb{E}\left[\mathbb{E}(D_i^2|N_i)\right] = \mathbb{E}(N_i q_i) - \mathbb{E}(N_i q_i^2) + \mathbb{E}(N_i^2 q_i^2) \\ &= \mathbb{E}(N_i) q_i - \mathbb{E}(N_i) q_i^2 + q_i^2 \left[(\mathbb{E}(N_i))^2 + \mathrm{Var}(N_i)\right]\end{aligned} \tag{2.8c}$$

$$\begin{aligned}\mathrm{Var}(D) &= \sum_{i=1}^{r} \mathrm{Var}(D_i) = \sum_{i=1}^{r} \left[\mathbb{E}(D_i^2) - (\mathbb{E}(D_i))^2\right] \\ &= \sum_{i=1}^{r} \left[\mathbb{E}(N_i) q_i - \mathbb{E}(N_i) q_i^2 + q_i^2 \left[(\mathbb{E}(N_i))^2 + \mathrm{Var}(N_i)\right] - (\mathbb{E}(N_i))^2 q_i^2\right] \\ &= \sum_{i=1}^{r} \left[\mathbb{E}(N_i) q_i(1-q_i) + q_i^2 \mathrm{Var}(N_i)\right] \\ &= npq - \sum_{i=1}^{r} \mathbb{E}(N_i) \left(q_i - q\right)^2 + \sum_{i=1}^{r} q_i^2 \mathrm{Var}(N_i) \\ &= npq - \sum_{i=1}^{r} n f_i \left(q_i - q\right)^2 + \sum_{i=1}^{r} q_i^2 n f_i (1-f_i)\end{aligned} \tag{2.8d}$$

where $q = \dfrac{1}{n} \displaystyle\sum_{i=1}^{r} \mathbb{E}(N_i) q_i = \frac{1}{n} \sum_{i=1}^{r} n f_i q_i = \sum_{i=1}^{r} f_i q_i$ is the average rate of mortality.

Thus, the heterogeneity reduces the random fluctuations in the number of deaths D, which however will be increased by the variance of the random variables N_i. We note that this setting captures the features of a situation in which the heterogeneity is due to some observable factors, but actually cannot be observed at an individual level, and this causes the randomness of the group sizes.

2.2.1.6 Setting 6

We refer to a population of n persons all aged x, split into r groups each one with random size N_i, $i = 1, 2, \ldots, r$, where the random vector $\underline{N} = (N_1, N_2, \ldots, N_r)$ has multinomial distribution with parameters $(n, f_1, f_2, \ldots, f_r)$, such that $n = N_1 + N_2 + ... + N_r$ and with $f_1 > f_2 > \cdots > f_r$, $f_1 + f_2 + \cdots + f_r = 1$. Thus, a higher probability of belonging to groups with lower mortality is assumed. The probabilities f_i are assigned, taken from the observation of other groups with similar characteristics. The probability of dying is Q_i for all the individuals in the group i; Q_i are random variables with distribution $\mathrm{Beta}(\alpha_i, \beta_i, a_i, c_i)$, such that $\mathbb{E}(Q_1) < \mathbb{E}(Q_2) < \cdots < \mathbb{E}(Q_r)$. The number of deaths D is equal to $D_1 + D_2 + \cdots + D_r$.

We obtain:

$$\mathbb{E}(D_i|N_i,Q_i) = N_iQ_i \Rightarrow \mathbb{E}(D_i) = \mathbb{E}(N_iQ_i) = \mathbb{E}(N_i)\mathbb{E}(Q_i) \tag{2.9a}$$

with the hypothesis of the stochastic independence for $N_i$and $Q_i \Rightarrow$

$$(\mathbb{E}(D_i))^2 = (\mathbb{E}(N_i))^2\,(\mathbb{E}(Q_i))^2 \tag{2.9b}$$

$$\mathbb{E}(D_i^2|N_i,Q_i) = \mathrm{Var}(D_i|N_i,Q_i) + (\mathbb{E}(D_i|N_i,Q_i))^2$$
$$= N_iQ_i(1-Q_i) + N_i^2Q_i^2 = N_iQ_i - N_iQ_i^2 + N_i^2Q_i^2 \tag{2.9c}$$

$$\mathbb{E}(D_i^2) = \mathbb{E}\left[\mathbb{E}(D_i^2|N_i,Q_i)\right]$$
$$= \mathbb{E}(N_i)\mathbb{E}(Q_i) - \mathbb{E}(N_i)\left[(\mathbb{E}(Q_i))^2 + \mathrm{Var}(Q_i)\right] + \left[(\mathbb{E}(N_i))^2 + \mathrm{Var}(N_i)\right]\left[(\mathbb{E}(Q_i))^2 + \mathrm{Var}(Q_i)\right] \tag{2.9d}$$

$$\mathrm{Var}(D) = \sum_{i=1}^{r}\mathrm{Var}(D_i) = \sum_{i=1}^{r}\left[\mathbb{E}(D_i^2) - (\mathbb{E}(D_i))^2\right]$$
$$= \sum_{i=1}^{r}\left\{\mathbb{E}(N_i)\mathbb{E}(Q_i) - \mathbb{E}(N_i)\left[(\mathbb{E}(Q_i))^2 + \mathrm{Var}(Q_i)\right] + \left[(\mathbb{E}(N_i))^2 + \mathrm{Var}(N_i)\right]\left[(\mathbb{E}(Q_i))^2 + \mathrm{Var}(Q_i)\right]\right\}$$
$$- \sum_{i=1}^{r}(\mathbb{E}(N_i))^2\,(\mathbb{E}(Q_i))^2$$
$$= \sum_{i=1}^{r}\left\{\mathbb{E}(N_i)\mathbb{E}(Q_i) - \mathbb{E}(N_i)(\mathbb{E}(Q_i))^2 - \mathbb{E}(N_i)\mathrm{Var}(Q_i) + (\mathbb{E}(N_i))^2\mathrm{Var}(Q_i) + (\mathbb{E}(Q_i))^2\mathrm{Var}(N_i)\right\} +$$
$$+ \sum_{i=1}^{r}\mathrm{Var}(Q_i)\mathrm{Var}(N_i)$$
$$= \sum_{i=1}^{r}\left\{\mathbb{E}(N_i)\mathbb{E}(Q_i)\,(1-\mathbb{E}(Q_i)) + \mathbb{E}(N_i)(\mathbb{E}(N_i)-1)\mathrm{Var}(Q_i) + (\mathbb{E}(Q_i))^2\mathrm{Var}(N_i)\right\} +$$
$$+ \sum_{i=1}^{r}\mathrm{Var}(Q_i)\mathrm{Var}(N_i)$$
$$= npq - \sum_{i=1}^{r}\mathbb{E}(N_i)\,(\mathbb{E}(Q_i)-q)^2 + \sum_{i=1}^{r}\mathbb{E}(N_i)\,(\mathbb{E}(N_i)-1))\,\mathrm{Var}(Q_i) + \sum_{i=1}^{r}(\mathbb{E}(Q_i))^2\mathrm{Var}(N_i) +$$
$$+ \sum_{i=1}^{r}\mathrm{Var}(N_i)\mathrm{Var}(Q_i)$$

$$\mathrm{Var}(D) = npq - \sum_{i=1}^{r}nf_i\left(\frac{\alpha_ic_i+\beta_ia_i}{\alpha_i+\beta_i} - q\right)^2 + \sum_{i=1}^{r}nf_i(nf_i-1)\frac{\alpha_i\beta_i(c_i-a_i)^2}{(\alpha_i+\beta_i)^2(\alpha_i+\beta_i+1)} +$$
$$+ \sum_{i=1}^{r}nf_i(1-f_i)\left(\frac{\alpha_ic_i+\beta_ia_i}{\alpha_i+\beta_i}\right)^2 + \sum_{i=1}^{r}nf_i(1-f_i)\frac{\alpha_i\beta_i(c_i-a_i)^2}{(\alpha_i+\beta_i)^2(\alpha_i+\beta_i+1)} \tag{2.9e}$$

$$\text{where}\quad q = \frac{1}{n}\sum_{i=1}^{r}\mathbb{E}(N_i)\mathbb{E}(Q_i) = \frac{1}{n}\sum_{i=1}^{r}nf_i\frac{\alpha_ic_i+\beta_ia_i}{\alpha_i+\beta_i} = \sum_{i=1}^{r}f_i\frac{\alpha_ic_i+\beta_ia_i}{\alpha_i+\beta_i}$$

is the average expected rate of mortality.
We note that this setting is a generalization of the previous five settings and gathers together all the observations so far made. Actually, the heterogeneity reduces the random fluctuations in the number of deaths D, which however is increased by the variance of the random variables N_i and Q_i both separately as well as a combined effect.

2.2.2 Unobservable heterogeneity

The settings so far considered assume either homogeneity inside the population, or inside each of the groups. For example, according to settings 3 to 6 the same mortality rate, either known or as a random variable, is assigned either to the whole population or to all the individuals belonging to the same group.

As already noted, some degree of heterogeneity remains because of unobservable factors. The impact of residual heterogeneity can be analyzed and assessed via the concept of individual frailty. Of course, to this purpose more general settings must be defined.

2.2.2.1 Setting 7

We refer to a heterogeneous population which consists of n persons. The event death of the individual j, $j = 1, 2, \ldots, n$, is expressed by the Bernoulli random variable $D_j = |E_j|$ with random parameter Q_j. Because of unobservable heterogeneity, we assume for all the individuals the same beta distribution, that is $Q_j \sim \text{Beta}(\alpha, \beta, a, c)$. We assume that the events $D_j = |E_j|$ are independent conditional on any possible outcome of the random quantities Q_j.

Through an appropriate simulation approach, the expected value, the variance and the coefficient of variation of the total number of deaths in the population, $D = \sum_{j=1}^{n} D_j$, can be calculated.

2.2.2.2 Setting 8

We refer to a population which consists of n persons, split into r groups, each one with random size N_i. We assume that the random vector $\underline{N} = (N_1, N_2, \ldots, N_r)$ has multinomial distribution with parameters $(n, f_1, f_2, \ldots, f_r)$, such that $n = N_1 + N_2 + ... + N_r$ and with $f_1 > f_2 > \cdots > f_r$, $f_1 + f_2 + \cdots + f_r = 1$. Thus, a higher probability of belonging to groups with lower mortality is assumed. The probabilities f_i are assigned, taken from the observation of other groups with similar characteristics. The event death

of the individual j belonging to the group i is expressed by the Bernoulli random variable $D_{ij} = |E_{ij}|$ with random parameter Q_{ij}, $i = 1, \ldots, r$, $Q_{ij} \sim \text{Beta}(\alpha_i, \beta_i, a_i, c_i)$ with parameters such that $\mathbb{E}(Q_{1j}) < \mathbb{E}(Q_{2j}) < \ldots < \mathbb{E}(Q_{rj})$. We note that, thanks to observable heterogeneity each group has distinct parameters, but because of unobservable heterogeneity for all the individuals belonging to a given group the same parameters are assumed.

Through an appropriate simulation approach the expected value, the variance and the coefficient of variation of the total number of deaths $D = \sum_{i=1}^{r} \sum_{j=1}^{N_i} D_{ij}$ can be calculated.

2.2.3 Numerical examples

Although the focus of the present thesis is the stochastic analysis of multi-year insurance policies in a multistate framework, some simple numerical examples in a one-year, two-state context can provide a first insight into the impact of uncertainty and/or unobservable heterogeneity, in particular by assessing the variability in the results of interest in terms of variance and coefficient of variation (also called the risk index).

The following data have been assumed in the numerical examples.

- Total number of individuals $n = 10000$.
- Possible groupings: $r = 3, 6, 12$ (if not otherwise stated, $r = 6$ is assumed).
- Deterministic settings (i.e. no uncertainty, no unobservable heterogeneity):
  - in case of no grouping $q = 0.0073$;
  - in case of grouping $q_1 < q_2 < \cdots < q_r$ with $\overline{q} = 0.0073$.
- Uncertainty and/or frailty
  - in case of no grouping: random $Q \sim \text{Beta}(1.6, 20.31, 0, 0.1)$;
  - in case of grouping: different Beta distributions, with increasing expected values.
- Results of interest:
  - expected value
  - variance
  - coefficient of variation

of the total number of deaths D.

The impact of heterogeneity due to observable risk factors, which allows us to split the population into homogeneous groups, is shown in Table 2.1. Grouping results in a lower variance and hence a lower coefficient of variation.

	Setting 1	Setting 2
$\mathbb{E}(D)$	73	73
$\mathrm{Var}(D)$	72.47	71.81
$\mathrm{CV}(D)$	0.11662	0.11608

TABLE 2.1: Effect of observable risk factors

The impact of uncertainty in a homogeneous population is displayed in Table 2.2. We note a huge increment in the variance and hence in the coefficient of variation. It is worth recalling that uncertainty implies the presence of an undiversifiable risk component, and hence calls for appropriate risk management strategies.

	Setting 1	Setting 3
$\mathbb{E}(D)$	73	73
$\mathrm{Var}(D)$	72.47	3026.92
$\mathrm{CV}(D)$	0.11662	0.75366

TABLE 2.2: Effect of uncertainty

An interesting result is shown in Table 2.3: the effect of uncertainty on the variance of D can be significantly reduced via grouping thanks to observable risk factors.

	Setting 3	Setting 4
$\mathbb{E}(D)$	73	73
$\mathrm{Var}(D)$	3026.92	406.68
$\mathrm{CV}(D)$	0.75366	0.27625

TABLE 2.3: Reducing the effect of uncertainty via grouping

A slight increase in the variance is the effect of group random sizes, as shown by the results in Table 2.4, whereas a much more significant effect is caused by uncertainty in grouping, as displayed in Table 2.5. However, as clearly appears from the results in Table 2.6, the above effect decreases as the number of group increases.

The presence of frailty leads to huge variance, however grouping helps in reducing the variance, as shown by results in Table 2.7.

	Setting 2	Setting 5
$\mathbb{E}(D)$	73	73
$\mathrm{Var}(D)$	71.81	72.90
$\mathrm{CV}(D)$	0.11608	0.11696

TABLE 2.4: Effect of group random sizes

	Setting 5	Setting 6
$\mathbb{E}(D)$	73	73
$\mathrm{Var}(D)$	72.90	408.03
$\mathrm{CV}(D)$	0.11696	0.27671

TABLE 2.5: Effect of uncertainty in grouping

	Setting 6		
	3 Groups	6 Groups	12 Groups
$\mathbb{E}(D)$	73	73	73
$\mathrm{Var}(D)$	592.35	408.03	221.92
$\mathrm{CV}(D)$	0.3334	0.27671	0.20407

TABLE 2.6: Effect of uncertainty in grouping depending on the number of groups

	Setting 7	Setting 8
$\mathbb{E}(D)$	73	73
$\mathrm{Var}(D)$	3095.53	406.72
$\mathrm{CV}(D)$	0.76132	0.27698

TABLE 2.7: Effect of frailty and uncertainty in grouping

Chapter 3

Three-state models: multi-year time horizon

3.1 Multistate models: general aspects

As noted by Pitacco (2014), in the context of insurances of the person (and hence life insurance and health insurance in particular), the evolution of a risk (that is, an insured individual) can be represented in terms of the presence of the risk itself, at every point of time, in a certain "state", belonging to a given set of states, or "state space" (e.g., active, disabled, etc.). A "transition" from one state to another one represents an event (e.g. death, or disablement, or recovery of the individual).

A multistate model consists of a set of states and a set of possible transitions between states. In graphical terms, a multistate model can be represented by a directed graph (or digraph), whose nodes represent the states, whereas the arcs represent possible transitions between states.

The graph only describes what may be the evolution of an individual risk over time. Then, two important items must be added. First,

- a stochastic structure must be defined, in order to quantify in probabilistic terms the possible evolutions of the risk over time.

Second,

- when the multistate model is adopted to represent an insurance cover, we also have to associate possible cash flows (premiums and benefits) to the presence in some states and/or to transitions between states.

3.2 Markov assumption

Let $\mathcal{L}$ denote the state space, consisting of N elements:

$$\mathcal{L} = \{1, 2, \ldots, N\} \tag{3.1}$$

Let $\mathcal{T}$ denote the set of direct transitions between states. $\mathcal{T}$ is a subset of the set of pairs (i, j):

$$\mathcal{T} \subseteq \{(i,j)|\ i \neq j; \quad i,j \in \mathcal{L}\} \tag{3.2}$$

Referring to an insurance policy, we assume that the initial state (that is, the state at the issue date) is 1. All the states $j \in \mathcal{L}$ are assumed to be accessible from the state 1 by direct or indirect transitions.

A probabilistic structure must be defined. Let $S(t)$ denote the random state occupied by the individual risk at time t. At the policy issue date $S(0) = 1$. Then, $\{S(t); t \geq 0\}$ is a stochastic process, taking values in the finite set $\mathcal{L}$. The variable t is often called seniority: it represents the time spent by the policy since its issue date. Any possible outcome $\{s(t)\}$ of the process $S(t)$ is called a sample path.

If, for any n, for any set of times $(0 \leq)\ t_0 < t_1 < \cdots < t_{n-1} < t_n < u$ and the corresponding set of states $i_0, i_1, \ldots, i_{n-1}, i_n, j$ in $\mathcal{L}$ with

$$\Pr\{S(t_0) = i_0 \wedge \cdots \wedge S(t_{n-1}) = i_{n-1} \wedge S(t_n) = i_n \wedge S(u) = j\} > 0 \tag{3.3}$$

we assume:

$$\begin{aligned} &\Pr\{S(u) = j | S(t_0) = i_0 \wedge \ldots \wedge S(t_{n-1}) = i_{n-1} \wedge S(t_n) = i_n\} \\ &= \Pr\{S(u) = j | S(t_n) = i_n\} \end{aligned} \tag{3.4}$$

then, $\{S(t); t \geq 0\}$ is a Markov process.

Thus, it is assumed that the conditional probability on the left side of the above equation only depends on the most recent information $S(t_n) = i_n$, while it is independent of the path before t_n. It follows, in particular, that the above probability does not depend on the time elapsed since the most recent entry into the current state i_n. Conversely, information on time spent in the current state since the most recent entry can be relevant in some applications, for example when the state disability is referred to, and the probabilities relate to events such as death or recovery. In these cases the assumption underlying the Markov models constitutes a controversial issue. An alternative modelling framework is provided by semi-Markov models (see, for example,

Haberman and Pitacco (1999)). These models are formally more complex, but, over all, numerical implementation difficulties might arise because of lack of data.
We note that, in the actuarial context, practical applications of the Markov assumption can be based, for example, on the process $\{S(t)\}$ with t belonging to discrete set, e.g. taking just integer values.

3.3 A three-state model

In this Section an implementation of the Markov model described above is proposed, to represent the features of a specific insurance policy. A time-discrete framework is assumed, with the year as the time unit.

3.3.1 Application to an insurance cover

We consider an insurance cover providing lump sum benefits in case of death and in case of disablement causing a permanent disability.

The state space is

$$\mathcal{L} = \{a, i, d\} \tag{3.5}$$

where:

a = active (or healthy);

i = disabled (or invalid);

d = dead.

Benefits are assumed to be paid at the end of the year in which a relevant event occurs. The benefit amounts are as follows:

B_1 if the individual dies in the year, being in the active state, that is, if the transition $a \to d$ occurs;

B_2 if a disablement occurs and the individual is alive at the end of the year, that is, if the transition $a \to i$ occurs;

B_3 if a disablement occurs and the individual dies before the end of the year, that is, if the transitions $a \to i \to d$ occur.

Because of the benefit structure we have defined, the set of transitions is as follows:

$$\mathcal{T} = \{(a, i), (a, d), (i, d)\} \tag{3.6}$$

We note that, because of the benefit structure, we disregard recovery. The graph representing the set of states and the set of transition is given by Fig. 3.1.

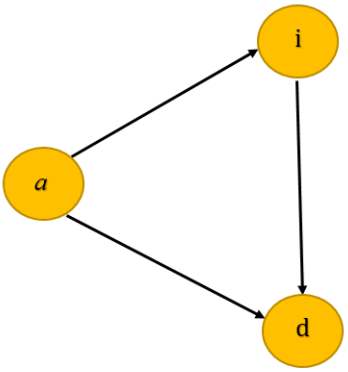

FIGURE 3.1: Three-state model

The above setting can represent a simple insurance package. Assuming $B_3 > B_1$, we indeed recognize the structure of a term insurance policy with a basic death benefit B_1, supplementary benefit in case of death because of accident given by $B_3 - B_1$, and benefit B_2 in the case of disablement leading to permanent disability, provided the individual is alive at the end of the year.

3.3.2 The basic Markov structure

We describe the probabilistic structure required by the three-state model defined above. We adopt the Hamza notation (commonly used in the actuarial framework), and refer to the individual attained age x instead of referring to the policy past-duration t.

Table 3.1 shows transitions either allowed or not allowed by our model, consistently with the following assumptions:

- no recovery;
- no more than one transition in a year (apart from possible death).

For an active individual age x, we define the following probabilities:

- $p_x^{aa} = \Pr\{S(x+1) = a | S(x) = a\}$ = probability of being active at age $x+1$;
- $q_x^{aa} = \Pr\{S(x+1) = d | S(x) = a\}$ = probability of dying within one year, the death occurring in state active (a);
- $p_x^{ai} = \Pr\{S(x+1) = i | S(x) = a\}$ = probability of being disabled at age $x+1$;

TABLE 3.1: Transition between states in a one-year period

State at age x	Transition	State at age $x+1$	Allowed by the model ?
a	$\rightarrow$	i	Yes
i	$\rightarrow$	a	No
a	$\rightarrow$ i $\rightarrow$	a	No
a	$\rightarrow$	d	Yes
i	$\rightarrow$	d	Yes
a	$\rightarrow$ i $\rightarrow$	d	Yes

- $q_x^{ai} = \Pr\{S(x+1) = d | S(x) = a\}$ = probability of dying within one year, the death occurring in state disability (i);
- p_x^a = probability of being alive at age $x+1$;
- q_x^a = probability of dying within one year;
- w_x = probability of becoming disabled within one year.

The following relations hold:

$$p_x^{aa} + p_x^{ai} = p_x^a; \tag{3.7}$$

$$q_x^{aa} + q_x^{ai} = q_x^a; \tag{3.8}$$

$$p_x^a + q_x^a = 1; \tag{3.9}$$

$$p_x^{ai} + q_x^{ai} = w_x. \tag{3.10}$$

We note that, given the definition of the benefit structure, disablement and death constitute two exit causes, so that at the beginning of each year an individual belonging to the insurance portfolio must be in the state active. Hence, probabilities related to individuals in state invalid (at the beginning of the year) are not relevant.

We note that the probability q_x^{ai} refers to an event consisting of two transitions, that is, $a \rightarrow i$ and $i \rightarrow d$. The following approximation is frequently adopted in the actuarial practice:

$$q_x^{ai} \approx w_x \frac{q_x^i}{2} \tag{3.11}$$

where q_x^i generically denotes the mortality of disabled people. The hypotheses underlying approximation (3.11) are as follows (see, for example, Pitacco (2014)):

- uniform distribution of the first transition time within the year;

- the probability that the second transition occurs within the second half of the year is equal to one half that a transition of the same type occurs within the year.

The following approximation, which will be useful in the following, can then be derived:

$$w_x\left(1-\frac{q_x^i}{2}\right)\approx p_x^{ai} \tag{3.12}$$

The three-state Markov model we have described constitutes the baseline setting to define the portfolio stochastic structure. A first (trivial) generalization is required if the insurance portfolio is split, thanks to observable risk factors, into subportfolios, each one consisting of homogeneous risks. The generalization consists in assigning different probabilities to individuals belonging to different subportfolios.

We note that, given the benefit structure as defined in Sect. 3.3.1, our three state model can be interpreted, from a different perspective, as a model with three causes of "decrement", all the causes implying that the individual risk leaves the portfolio. Causes of decrement are as follows:

1. transition $a \to d$, implying benefit B_1;
2. transition $a \to i$, implying benefit B_2;
3. transitions $a \to i \to d$, implying benefit B_3.

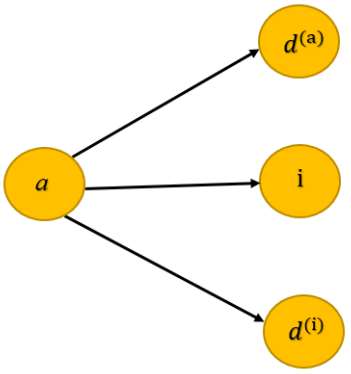

FIGURE 3.2: A model with three decrement causes

3.3.3 Allowing for frailty or uncertainty

The Markov model defined in Sect. 3.3.2 requires, as input data a set of probabilities which quantify, in stochastic terms, the possible paths of an individual risk over time,

i.e. probability of disablement, probability of death in the active state, probability of death in the disability state, possibly differentiated according to the group to which the risk is assigned according to observable risk factors.

However, lack of information might imply a (more or less significant) margin of vagueness in assigning the above probabilities. In particular, this may be caused by:

- heterogeneity due to unobservable risk factors, which implies individual frailty;
- uncertainty in stating the probability for all the individuals belonging to the portfolio (or to a subportfolio).

Our modeling choice consists in representing the vagueness, in both the above situations, by considering the probabilities as random quantities and assigning to these quantities four-parameter beta distributions (see Sects. 1.2.3 and 1.3).

Numerical values of the probabilities, affected either by uncertainty or by individual frailty, will be determined via stochastic simulation. In particular:

- probabilities affected by individual frailty will be determined by simulation for each individual belonging to the portfolio, according to the relevant beta parameters, possibly specific for each subportfolio;
- probabilities affected by uncertainty will be determined by simulation for the whole portfolio, or for each subportfolio, according to the relevant beta parameters.

Hence, stochastic simulation yields completely defined stochastic structures for the three-state model.

A three-state model (and, in general, a multistate model), thanks to probabilistic assumptions and numerical values of the probabilities involved, represents in stochastic terms the evolution of an individual risk over time. Hence, one or more assumptions and one or more set of probabilities are needed according to the portfolio structure. For example:

- just one set of (given) probabilities is needed in the case of
  - homogeneous portfolio;
  - no uncertainty in the assessment of the probabilities;
- one set of (given) probabilities for each homogeneous group (subportfolio) is needed in the case of
  - heterogeneous portfolio if heterogeneity is due to observable risk factors;

  - no uncertainty in the assessment of the probabilities;

- just one set of (given) probabilities and a probability distribution for a probability considered a random quantity is needed in the case of

  - homogeneous portfolio;
  - uncertainty in the assessment of one probability.

Other combinations are of course feasible. Indeed, each portfolio structure calls for a specific mix of assumptions, as we will see in Sect. 4.1.

Chapter 4

The insurance portfolio

4.1 Portfolio structures

In this Section we define, in stochastic terms, a set of portfolio structures, each one labeled as "case". The following features have been considered in order to define the various cases:

- possible presence of observable heterogeneity, and consequent definition of homogeneous groups, or subportfolios;
- possible uncertainty in assigning probabilities;
- possible presence of individual frailty.

Generally speaking, uncertainty and frailty might affect all the probabilities involved in the calculations. Here, we only focus on the probabilities of disablement (see below). More details will be provided in Chap. 5, when defining a specific insurance portfolio.

In all the cases, we assume that the portfolio initially consists of n individuals, all aged x_0. Each individual risk is covered by a m-year insurance policy.

The age-patterns of mortality are described by the first Heligman-Pollard law which, in terms of the mortality odds $\frac{q_x}{1-q_x}$ is given by the following expression:

$$\frac{q_x}{1-q_x} = A^{(x+B)^C} + D\,\mathrm{e}^{-E(\ln x - \ln F)^2} + G\,H^x \tag{4.1}$$

We note that:

- the first term on the right-hand side represents the infant mortality;
- the second term represents the young-adult mortalit;

- the third term represents adult and old mortality.

A given set of parameters is assumed for the mortality of active people expressed by the Heligman-Pollard law, that is for the probabilities q_x^{aa}, whereas for the mortality of disabled people we assume:

$$q_x^i = (1+\mu)\, q_x^{aa} \tag{4.2}$$

that is, a multiplicative model for substandard mortality.

Conversely, probability of disablement is assumed constant over the whole age range involved. More precisely, we have adopted the approximation (3.12), assuming a constant probability of disablement w (that is, independent of the attained age), so that we obtain:

$$p_x^{ai} \approx w\Big(1 - \frac{q_x^i}{2}\Big) \tag{4.3}$$

In the case of frailty or uncertainty, the probability of disablement is a random quantity, denoted with W. From the outcome of W, the (approximate) value of p_x^{ai} can be derived via (4.3).

4.1.1 Case 1

The insurance portfolio consists of one homogeneous group; all individuals have the same known probability of disablement, w.

Summarizing:

- homogeneity;
- no uncertainty.

TABLE 4.1: Case 1

n individuals
1 group
w given

4.1.2 Case 2

A heterogeneous portfolio, thanks to observable heterogeneity, is arranged in r subportfolios, with given sizes $n_1, n_2, \ldots, n_r$, such that $\sum_{j=1}^{r} n_j = n$. Each subportfolio

includes individuals with the same probability of disablement. Probability of disablements are $w_1 < w_2 < \cdots < w_r$. We define the average probability of disablement as follows:

$$\overline{w} = \sum_{j=1}^{r} \frac{n_j}{n} \, w_j \tag{4.4}$$

and assume that $\overline{w} = w$, that is the probability of disablement in Case 1.

Summarizing:

- observable heterogeneity;
- deterministic group sizes;
- no uncertainty.

TABLE 4.2: Case 2

n individuals
r groups
$n_1, n_2, \ldots, n_r$ given
$w_1, w_2, \ldots, w_r$ given

4.1.3 Case 3

The insurance portfolio consists of one homogeneous group; all individuals have the same unknown probability of disablement, W. Hence, W is a random variable with probability distribution Beta(α, β, a, c).

Summarizing:

- homogeneity;
- uncertainty.

TABLE 4.3: Case 3

n individuals
1 group
W random; Beta(α, β, a, c)

4.1.4 Case 4

The insurance portfolio consists of one heterogeneous group; all individuals have unknown probability of disablement. Let $W^{(i)}$ denote the random variable expressing the probability of disablement for the individual i, $i = 1, 2, \ldots, n$. All the $W^{(i)}$ have the same probability distribution Beta(α, β, a, c). This case represents the situation of individual frailty.

Summarizing:

- unobservable heterogeneity;
- continuous frailty modelling.

TABLE 4.4: Case 4

n individuals
1 group
$W^{(i)}$, $i = 1, 2, \ldots, n$ random; Beta(α, β, a, c)

4.1.5 Case 5

A heterogeneous portfolio, thanks to observable risk factors, is arranged in r subportfolios. However, the sizes of the subportfolios, $N_1, N_2, \ldots, N_r$, are random because of unobservable individual heterogeneity, and the vector

$$\underline{N} = (N_1, N_2, \ldots, N_r)$$

has multinomial distribution with parameters $(n, f_1, f_2, \ldots, f_r)$, such that $n = N_1 + N_2 + \cdots + N_r$ and $f_1 + f_2 + \cdots + f_r = 1$. The probabilities of disablement within each subportfolio w_j are given, and such that $w_1 < w_2 < \cdots < w_r$. This multinomial scheme can represent in a discrete framework the individual frailty, thus providing an approximation of the continuous frailty model as defined by Case 4.

Summarizing:

- unobservable heterogeneity;
- discrete frailty modelling.

TABLE 4.5: Case 5

n individuals
r groups
$N_1, N_2, \ldots, N_r$ random
$\underline{N} = (N_1, N_2, \ldots, N_r)$; multinomial $(n, f_1, f_2, \ldots, f_r)$
w_j, $j = 1, 2, \ldots, r$ given

4.1.6 Case 6

A heterogeneous portfolio, thanks to observable risk factors, is arranged in r subportfolios, each with known size, that is, $n_1, n_2, \ldots, n_r$. All individuals in the subportfolio j have the same unknown probability of disablement, W_j. Hence, W_j is a random variable with probability distribution Beta$(\alpha_j, \beta_j, a_j, c_j)$, $j = 1 \ldots, r$.

Summarizing:

- observable heterogeneity;
- uncertainty.

TABLE 4.6: Case 6

n individuals
r groups
$n_1, n_2, \ldots, n_r$ given
W_j random; Beta$(\alpha_j, \beta_j, a_j, c_j)$, $j = 1, \ldots, r$

4.2 Outflows

We briefly recall the benefit structure, and define the annual random payouts generated by the benefit payment.

4.2.1 Benefits

The benefit structure has already been defined in Sect. 3.3.1. To ease the reading of the following Sections, we recall the benefit definitions:

- B_1 if the transition $a \to d$ occurs;
- B_2 if the transition $a \to i$ occurs;

- B_3 if the transitions $a \to i \to d$ occur.

The benefit amounts are the same for all the individuals belonging to the portfolio. We also recall that benefits are assumed to be paid at the end of the year in which a relevant event (that is, a transition) occurs.

4.2.2 Annual payouts

The above benefits define the insurer's annual payouts. We use the following notation:

- $X_1(t)$ = annual payout at time t for benefits B_1;
- $X_2(t)$ = annual payout at time t for benefits B_2;
- $X_3(t)$ = annual payout at time t for benefits B_3;
- $X(t) = X_1(t) + X_2(t) + X_3(t)$ = annual total payout at time t.

where $t = 1, 2, \ldots, m$, with m the common term of all the individual policies.

All the above payouts are random quantities, their outcomes depending on the transitions between states in each year. Our main target is to quantify the randomness in terms of probability distributions of the payouts for $t = 1, 2, \ldots, m$.

4.3 Facing the annual payouts

Random payouts constitute the insurer's liability, which must be met by appropriate assets. Hence, a calculation principle is needed, in order to summarize a sequence of random amounts in terms of deterministic quantities.

Usually, a share of the total amount meeting the liability, i.e. the premiums, is provided by the policyholders, while the remaining share, i.e. the capital allocated, is provided by the insurer.

4.3.1 Premiums

Premium calculation is a two-fold problem:

- a premium calculation principle must be chosen;
- a premium payment arrangement must be defined;

the two aspects being related each other.

As regards the calculation principle, the traditional actuarial approach relies on the equivalence principle, according to which, for each individual risk, the expected present value (shortly, the actuarial value) of the premiums must be equal to the expected present value of the benefits.

The equivalence principle does not explicitly account for the riskiness inherent in the insurance transaction. However, riskiness can implicitly be allowed for by using an appropriate technical basis, in particular probabilities which lead to an overestimation of the insurer's liabilities (with respect to their value according to a realistic basis). This way, an implicit safety loading is included in the premium amounts.

According to an alternative approach, explicit safety loadings can be added to the premiums (calculated by adopting a realistic technical basis), via a specific loading formula.

Various arrangements for the premium payment throughout the policy duration can be chosen. In particular:

1. single premium, paid at policy issue;
2. periodic premiums, paid e.g. at every policy anniversary; for example:
   (a) level (or constant) premiums;
   (b) natural premiums.

The natural premiums are defined as the annual expected costs to the insurer. It follows that arrangement 2(b) implies, according to the equivalence principle, technical equilibrium in each policy year. Conversely, arrangements 1 and 2(a) imply a time-mismatching between premiums and expected costs of benefits, so that a reserve (the so called mathematical reserve) must be set up.

The natural premium arrangement is particularly interesting in the context of group insurance, when the employer acts as the sponsor of the scheme and then pays the premiums. In this case, the annual total premium amount paid by the employer is usually given by the sum of the individual natural premiums, and hence can vary according to the composition of the insured group.

4.3.2 Capital allocation

Capital allocation aims at insurer's solvency. A detailed set of rules defines, at European level, the capital requirements which must be assessed for each insurance company either via a standard formula or via an internal model.

Internal models aim to capture the specific risk profile of an insurance company, and then to allocate the appropriate shareholders' capital. Before being adopted by the insurance company, internal models must be validated by the local supervisor authority.

The assets provided by the premium collection plus the assets backing the shareholders' capital must face, according to some specified principle, the insurer's liabilities. Of course, a time horizon must be stated. In what follows, we will refer to a one-year time horizon; actually, this choice is in line with both the current solvency logic and the natural premium arrangement that we will adopt.

4.3.3 The percentile principle

In what follows (and, in particular, in Sect. 5.4), we will determine, for each year, the amounts of assets facing, according to the percentile principle (that is, according to VaR-like logic), the random value of the payout for benefits falling due in that year.

Hence, we will not distinguish between assets financed by the premium collection and assets backing the shareholders' capital. Of course, the source of assets is absolutely relevant from a corporate perspective. In particular, a huge amount of shareholders' capital implies a cost to the insurer's and hence a poor value creation or even a value destruction.

In formal terms, according to the percentile principle, we have to find, for $t = 1, 2, \ldots, m$, the amount $A(t)$ such that:

$$\Pr\{X(t) > A(t)\} = \epsilon \tag{4.5}$$

where ϵ denotes an assigned (small) probability.

More in detail, we can state the following requirements. Find, for $t = 1, 2, \ldots, m$ and $h = 1, 2, 3$, the amounts $A_h(t)$ such that:

$$\Pr\{X_h(t) > A_h(t)\} = \epsilon_h \tag{4.6}$$

From an operational point of view, only the requirements defined by condition (4.5) are relevant. Nevertheless, requirements defined by conditions (4.6) can be interesting as they provide information about the specific impact of each type of benefit on the total requirement. From a product design perspective, a high impact might suggest a redesign of the insurance product and even the removal of a benefit, or at least the reduction of the related amount.

We note, however, that the sum of the requirements, each assessed according to (4.6), that is $A_1(t) + A_2(t) + A_3(t)$, is usually different from $A(t)$ assessed according to (4.5), because of possible interactions/correlations. We also note that investment issues are not addressed; thus the term "assets" only refers to the amount of money available during the year to cover payouts.

Chapter 5

Stochastic analysis. Numerical results

5.1 Calculation procedures

As clearly emerges from Sect. 4.3.3, probability distributions of the random variables $X_1(t), X_2(t), X_3(t)$ (and then $X(t)$), for $t = 1, 2, \ldots, m$, must be determined. The complexity of the relations between the stochastic structures in input and the results of interest calls for calculations performed via simulation. Hence, simulated distributions of the variables $X_1(t), X_2(t), X_3(t)$ will be calculated.

5.1.1 An overall scheme

Whatever the type of portfolio and the relevant stochastic structure, the simulation procedures yields paths of the stochastic processes $\{X_h(t); t = 1, 2, \ldots\}$, $h = 1, 2, 3$. An example is shown in Fig. 5.1. The portfolio size is $n = 10\,000$.

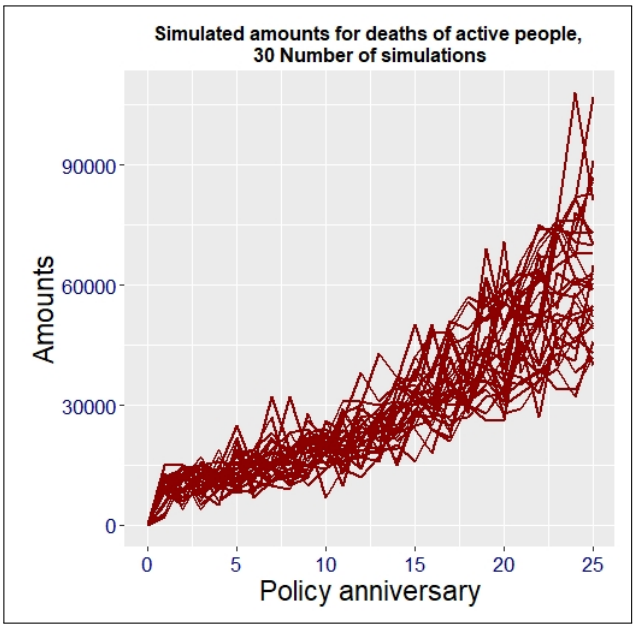

FIGURE 5.1: Simulated annual payout for active death benefits, $n_{sim} = 30$

Appropriate counters allow us to construct the simulated distributions. An example is displayed in Fig. 5.2, which refers to the random variable $X_1(t)$.

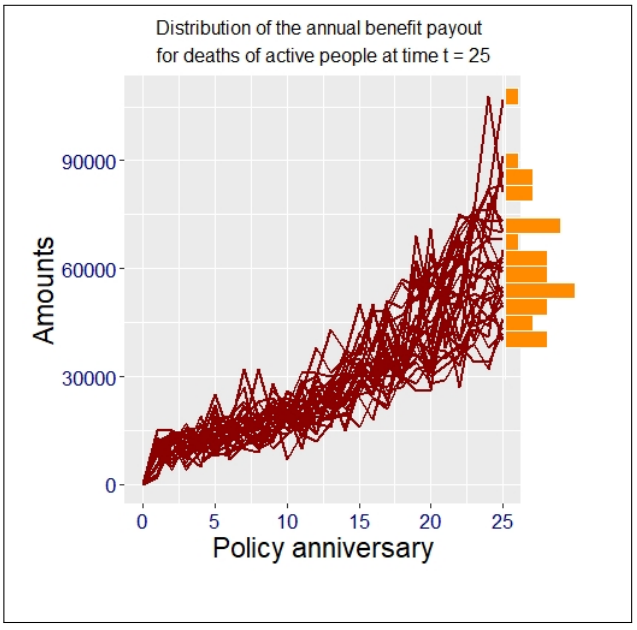

FIGURE 5.2: From simulated annual payout to the simulated distribution at maturity ($t = 25$); $n_{sim} = 30$.

The structure of the simulation procedure depends, of course, on the type of portfolio and the relevant probabilistic structure. A first insight into the diverse procedure structures is provided by Fig. 5.3.

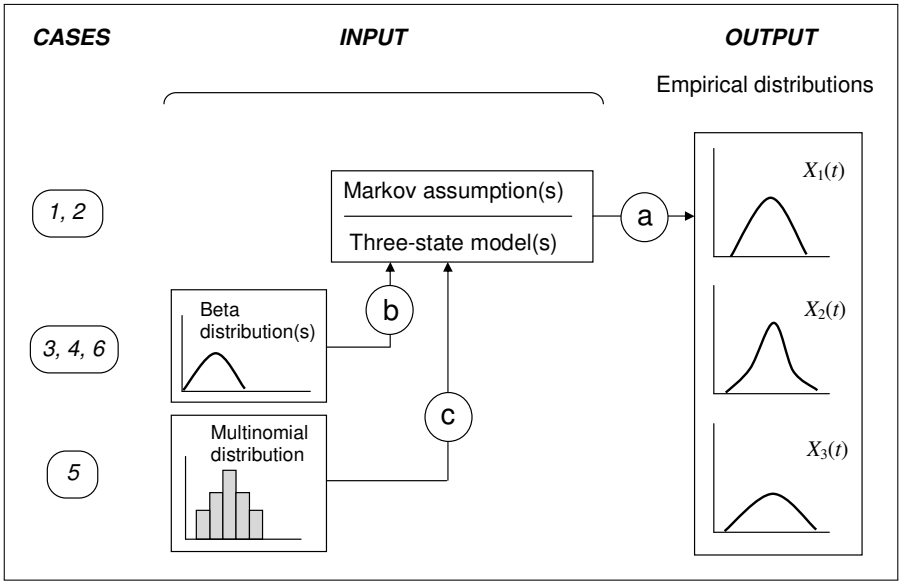

FIGURE 5.3: Calculation procedures to obtain simulated (or empirical) distributions

5.1.2 Specific procedures: Cases 1 and 2

Portfolio structures labeled as Cases 1 and 2 require the simplest procedures. In the absence of both individual frailty and uncertainty, the probabilities involved in the three-state model are completely defined. In detail, one set of probabilities is given for the Case 1, more sets are required for the Case 2, that is, if the heterogeneous portfolio is split into homogeneous subportfolios. In these cases the simulation procedure is directly applied to each individual belonging to the portfolio; see step ⓐ in Fig. 5.3.

For the generic individual aged $x = x_0 + t$, active at time t and hence belonging to the portfolio, one and only one of the following events may occur:

- being active at time $t + 1$, with probability p_x^{aa};
- being disabled at time $t + 1$, with probability p_x^{ai};
- dying before time $t + 1$, death occurring in the state active, with probability q_x^{aa};
- dying before time $t+1$, death occurring in the state disabled, with probability q_x^{ai};

where x denotes the attained age. Of course $p_x^{aa} + p_x^{ai} + q_x^{aa} + q_x^{ai} = 1$. We recall that $p_x^{ai} + q_x^{ai} = w_x$.

The construction, via simulation, of the individual path is based on the (pseudo-) random generation of the event occurring between age x and $x+1$. To this purpose, the outcome of a $(0,1)$-uniform (pseudo-) random number U is generated and then:

$$\begin{aligned} U > p_x^{ai} + q_x^{ai} + q_x^{aa} &\Leftrightarrow \text{being active at time } t+1; \\ q_x^{ai} + q_x^{aa} < U \leq p_x^{ai} + q_x^{ai} + q_x^{aa} &\Leftrightarrow \text{being disabled at time } t+1; \\ q_x^{ai} < U \leq q_x^{ai} + q_x^{aa} &\Leftrightarrow \text{dying from state active before time } t+1; \\ U \leq q_x^{ai} &\Leftrightarrow \text{dying from state disabled before time } t+1. \end{aligned}$$

We note what follows.

- Given, as input data, the value of w_x and the mortality of disabled lives q_x^i, the value of p_x^{ai} can be approximately calculated via (4.3); then q_x^{ai} can be calculated (see above).
- If the portfolio is split into homogeneous subportfolios (that is, in Case 2), the calculation procedure described above is applied with subportfolio-specific values of w_x.

5.1.3 Specific procedures: Cases 3, 4 and 6

Uncertainty or frailty feature Cases 3, 4, and 6. Hence, the simulation procedure must start with the generation of (pseudo-)random numbers beta-distributed; see step ⓑ in Fig. 5.3. In detail:

- the presence of uncertainty in Case 3 calls for the simulation of the random variable W, whose outcome is applied to all the risks in the portfolio;
- the presence of individual frailty in Case 4 requires the simulation of the random variables $W^{(i)}$, $i = 1, 2, \ldots, n$, all with the same beta distribution; then, the outcomes represent the individual frailty levels;
- Case 6 combines observable heterogeneity, which allows us to split the portfolio into r homogeneous subportfolios, with sizes $n_1, n_2, \ldots, n_r$; uncertainty in disability rate affects all individuals in each subportfolio; hence, the simulation of random variables W_j, beta-distributed with $\text{Beta}(\alpha_j, \beta_j, a_j, b_j)$, $j = 1, \ldots, r$, is needed.

In all the above cases, step ⓐ then follows.

5.1.4 Specific procedures: Case 5

In Case 5 the sizes of the subportfolios are random. So the first step in the calculation procedure consists in simulating the subportfolio sizes according to a multinomial distribution with parameters $n, f_1, f_2, \ldots, f_r$; see step © in Fig. 5.3.

In practice, the values of parameters $f_1, f_2, \ldots, f_r$ can be suggested, as already noted, by similar portfolio experiences.

An alternative procedure to determine the values of the parameters $f_1, f_2, \ldots, f_r$ is the following one.

- Split the range of outcomes, $c - a$, of the random variable W, with distribution $\text{Beta}(\alpha, \beta, a, c)$, into r subintervals, each with size $s = (c - a)/r$.
- Let $f(w)$ denote the probability density function of the beta distribution; calculate, for $j = 1, 2, \ldots, r$, the following probabilities:

$$f_j = \Pr\{(j-1)\, s < W \leq j\, s\} = \int_{(j-1)\, s}^{j\, s} f(w)\, \mathrm{d}w \tag{5.1}$$

- Calculate the conditional expected value related to each subinterval, and set:

$$w_j = \mathbb{E}(W | (j-1)\, s < W \leq j\, s) = \frac{1}{f_j} \int_{(j-1)\, s}^{j\, s} w\, f(w)\, \mathrm{d}w \tag{5.2}$$

- Finally, the subportfolio sizes are simulated according to the multinomial distribution with parameters $n, f_1, f_2, \ldots, f_r$.

We note that, this way, a continuous-frailty setting can be approximated by adopting a discrete setting, for which the frailty levels are given by the quantities w_j, $j = 1, 2, \ldots, r$ as defined by Eq. (5.2).

5.2 Defining the portfolio: input data

The numerical examples presented and discussed in Sects. 5.3 and 5.4 are based on the following input data.

5.2.1 General data

The following input data are used in all the Cases.

- Number of simulations 100 000 for each case.

- Portfolio initial size $n = 10\,000$.
- Insureds' initial age $x_0 = 40$.
- Policy term $m = 25$.
- Benefits $B_1 = B_2 = B_3 = 1\,000$ monetary units.
- Mortality of active people following the Heligman-Pollard law (see Eq. (4.1)), with the parameters suggested by Olivieri and Pitacco (2015) for mortality of assured people, given in Table 5.1 (see also Fig. 5.4).

TABLE 5.1: Parameters of the first Heligman-Pollard law

A	B	C	D	E	F	G	H
0.00054	0.01700	0.10100	0.00013	10.7200	18.67	$1.46400 * 10^{-5}$	1.11000

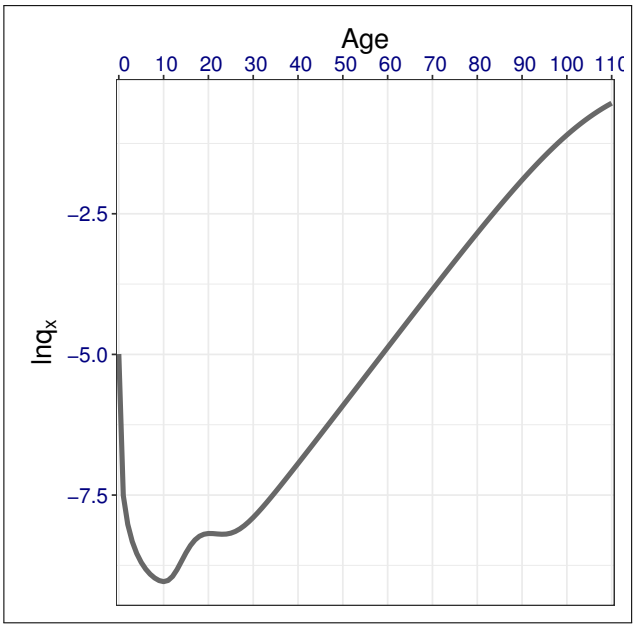

FIGURE 5.4: Mortality of active people following Heligman-Pollard law

- Mortality of disabled people according to the multiplicative model (see Eq. (4.2)), with $\mu = 0.30$.

5.2.2 Case-specific data

Case 1 only requires the value w of the probability of disablement, assumed independent of the attained age:

- $w = 0.02$

In *Case 2* the portfolio is split into r subportfolios, with different given probabilities of disablement:

- $r = 3$;
- subportfolio sizes and probabilities of disablement: see Table 5.2.

TABLE 5.2: Case 2 - Subportfolio sizes and probabilities of disablement

j	n_j	w_j
1	2 000	0.015
2	6 000	0.020
3	2 000	0.025

Case 3 represents the uncertainty situation in terms of probability of disablement W, following a beta distribution:

- $W \sim \text{Beta}(2.2, 3.3, 0, 0.05)$.

Case 4 represents the situation of random individual frailty, $W^{(i)}$, following a beta distribution:

- $W^{(i)} \sim \text{Beta}(2.2, 3.3, 0, 0.05)$, for $i = 1, 2, \ldots, n$.

Case 5 requires the simulation of subportfolio random sizes, according to the multinomial distribution with probabilities f_j, and related probability of disablements w_j; see Table 5.3.

In *Case 6* the portfolio is split into r subportfolios, with different random probability of disablement beta-distributed:

- $r = 3$;
- parameters of the beta distributions: see Table 5.4 and Fig.5.5

TABLE 5.3: Case 5 - Parameters of the multinomial distribution and probabilities of disablement

j	f_j	w_j
1	0.043139459	0.003373745
2	0.123561732	0.007722715
3	0.171613562	0.012573115
4	0.185302973	0.017495855
5	0.170889089	0.022437165
6	0.137297730	0.027378369
7	0.094414976	0.032304236
8	0.052093299	0.037187443
9	0.019244005	0.041943619
10	0.002443175	0.046184713

TABLE 5.4: Case 6 - Parameters of the beta distributions

j	α_j	β_j	a_j	c_j
1	2.2	5.13	0	0.05
2	2.2	3.30	0	0.05
3	2.2	2.20	0	0.05

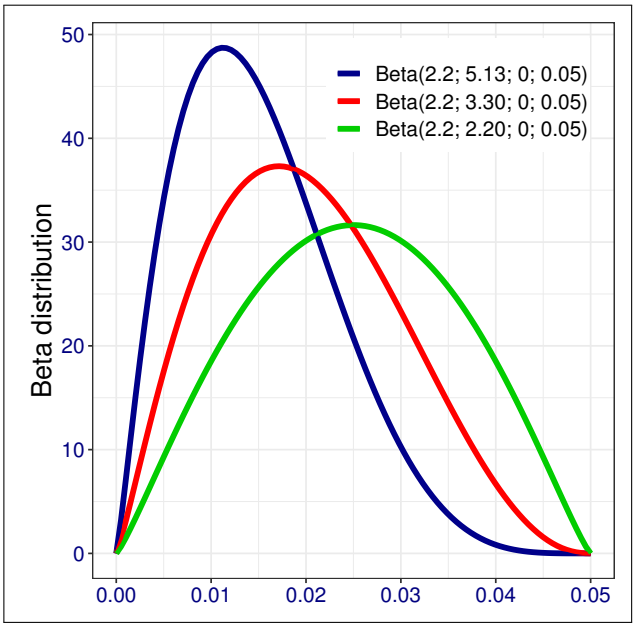

FIGURE 5.5: Beta Distribution

5.3 Simulated distributions

In this Section we compare the simulated distributions corresponding to various cases. In particular, for each couple of cases, we report the distributions at times $t = 5$ and $t = 25$ of $X(t)$, $X_1(t)$ and $X_2(t)$. We do not report the distributions of $X_3(t)$ because, due to the very small probability of the compound event death following the disablement, the related annual payouts are not significant. We note that, in all the comparisons, the shape of the distribution of $X(t)$ is very close to the shape of the distribution of $X_2(t)$.

5.3.1 Case 1 vs Case 2

Here we assess the effect of splitting the portfolio into homogeneous subportfolios thanks to observable risk factors. We see that, in our multi-year, three-state framework, splitting the portfolio is not effective for all the distributions considered, unlike the one-year, two-state framework (see 2.1 and, for more details, Valente (2017)). We note that, in both the cases, the setting is totally deterministic (in terms of probability of disablement, number of groups and related sizes): neither uncertainty nor individual frailty have been considered so far.

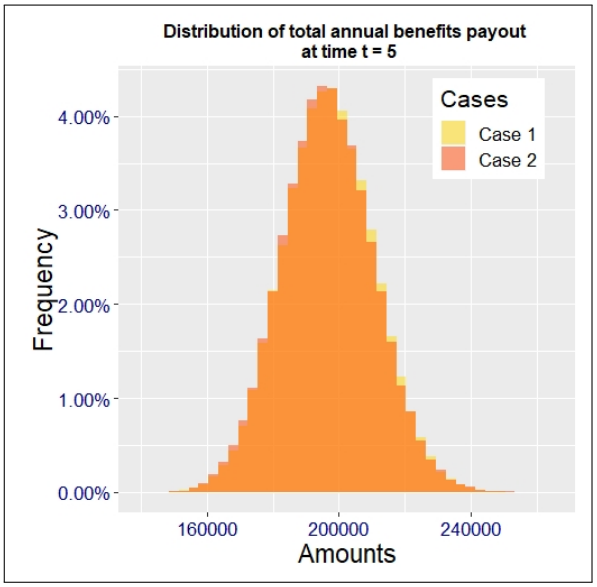

FIGURE 5.6: Case 1 vs Case 2, distribution of $X(5)$.

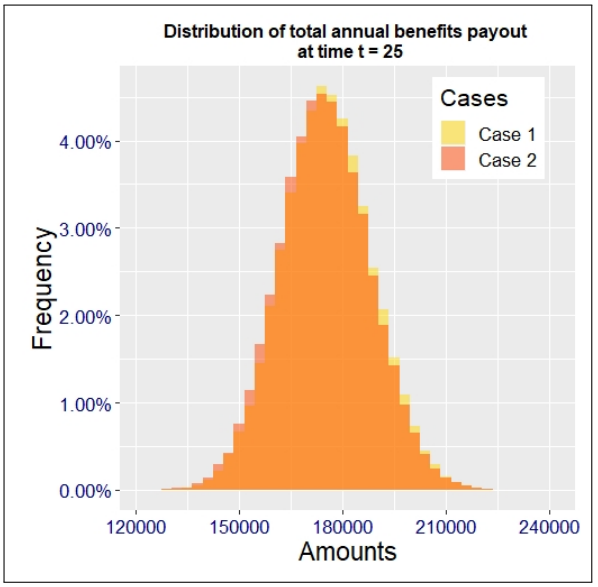

FIGURE 5.7: Case 1 vs Case 2, distribution of $X(25)$.

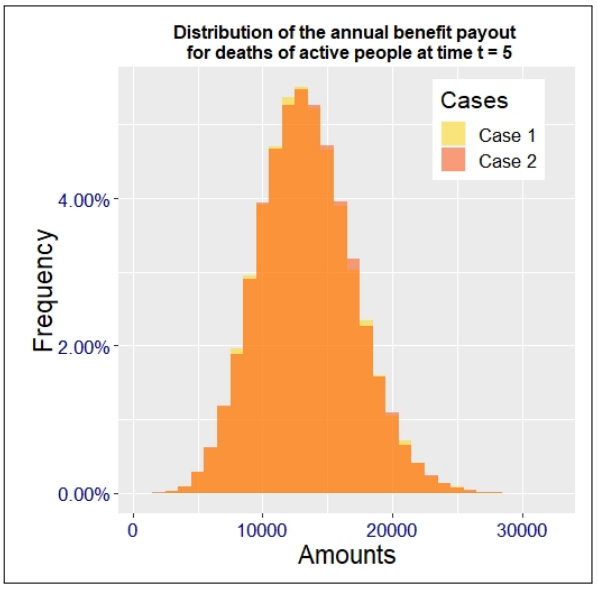

FIGURE 5.8: Case 1 vs Case 2, distribution of $X_1(5)$.

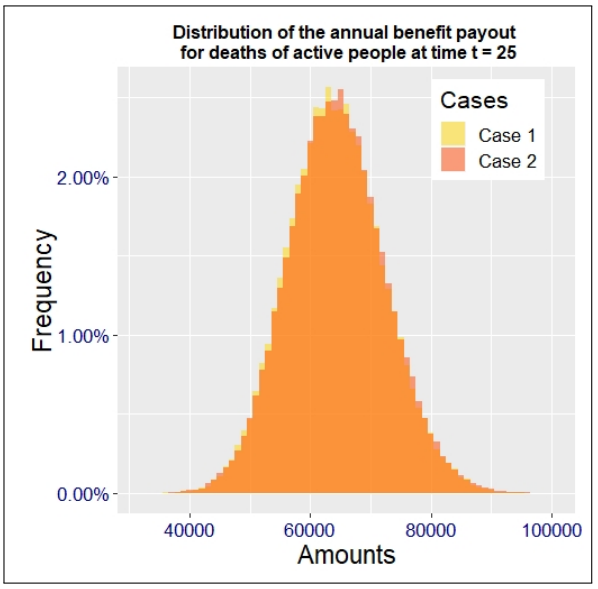

FIGURE 5.9: Case 1 vs Case 2, distribution of $X_1(25)$.

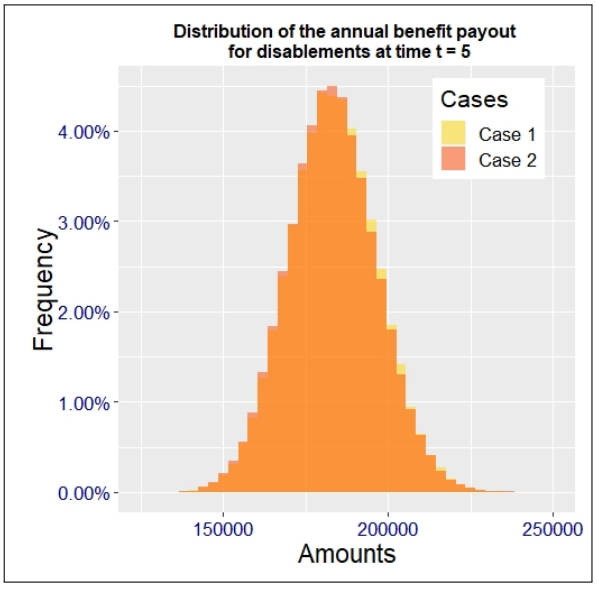

FIGURE 5.10: Case 1 vs Case 2, distribution of $X_2(5)$.

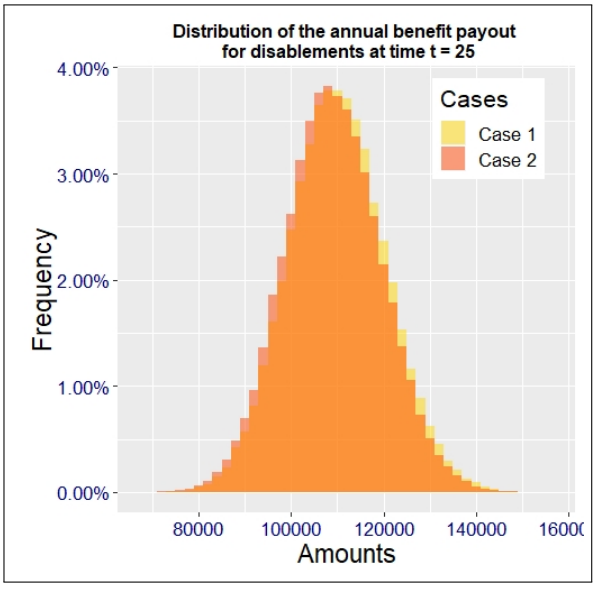

FIGURE 5.11: Case 1 vs Case 2, distribution of $X_2(25)$.

5.3.2 Case 1 vs Case 3

In both the cases the portfolio is not split into subportfolios. We analyze the impact of uncertainty in the assessment of the probability of disablement. The impact is evident in Figs. 5.12, 5.13, 5.16 and 5.17. Conversely, in the distributions shown by Figs. 5.14 and 5.15 the probability of disablement only intervenes in determining the number of insured still in portfolio, i.e. in the active state, and this effect increases throughout time (compare distribution of $X_1(5)$ to distribution of $X_1(25)$). We note that uncertainty in the probability of disablement affects, at the same time and to the same extent all the individuals in the portfolio. On the contrary, the frailty (see the next comparison) separately affects each individual determining a lower impact on the portfolio risk profile.

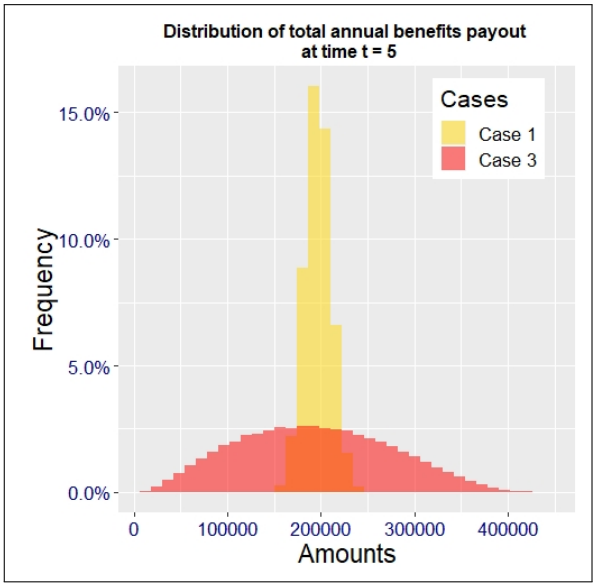

FIGURE 5.12: Case 1 vs Case 3, distribution of $X(5)$.

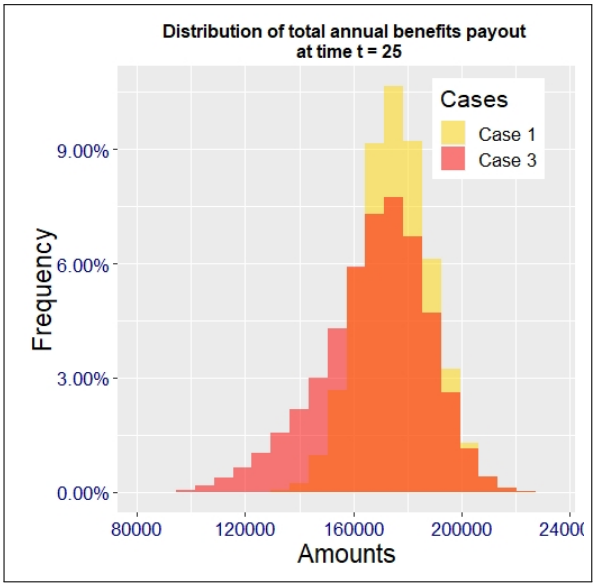

FIGURE 5.13: Case 1 vs Case 3, distribution of $X(25)$.

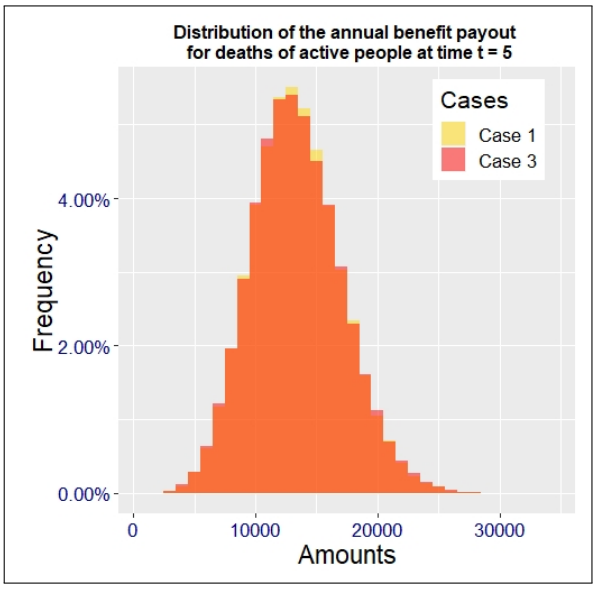

FIGURE 5.14: Case 1 vs Case 3, distribution of $X_1(5)$.

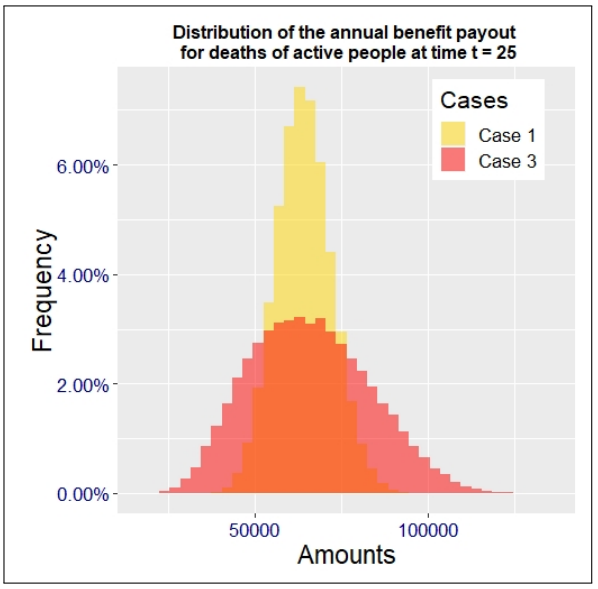

FIGURE 5.15: Case 1 vs Case 3, distribution of $X_1(25)$.

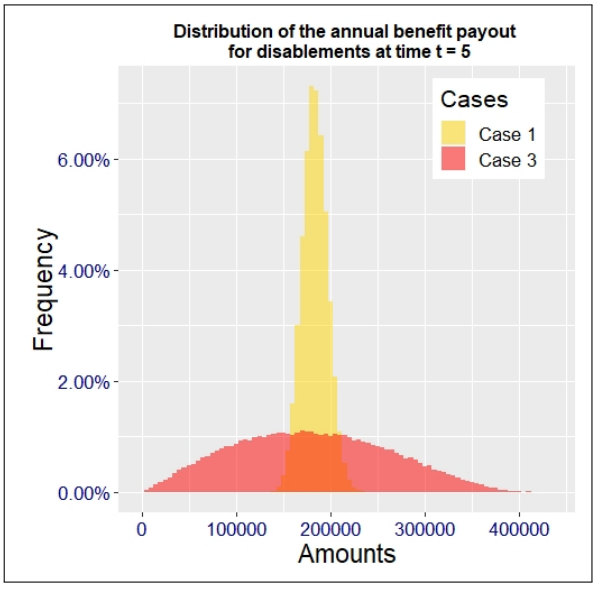

FIGURE 5.16: Case 1 vs Case 3, distribution of $X_2(5)$.

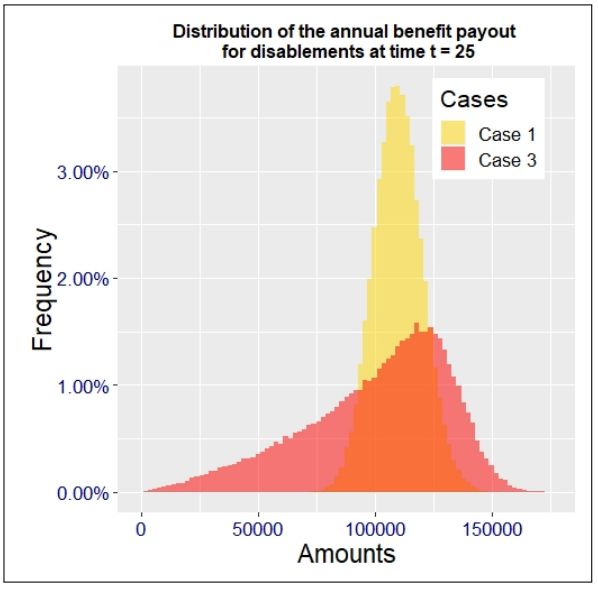

FIGURE 5.17: Case 1 vs Case 3, distribution of $X_2(25)$.

5.3.3 Case 1 vs Case 4

In Case 4 the portfolio is not split into subportfolios. Individual frailty is modeled as a continuous variable. Unlike the case of uncertainty, the presence of individual frailty does not heavily impact on the portfolio risk profile.

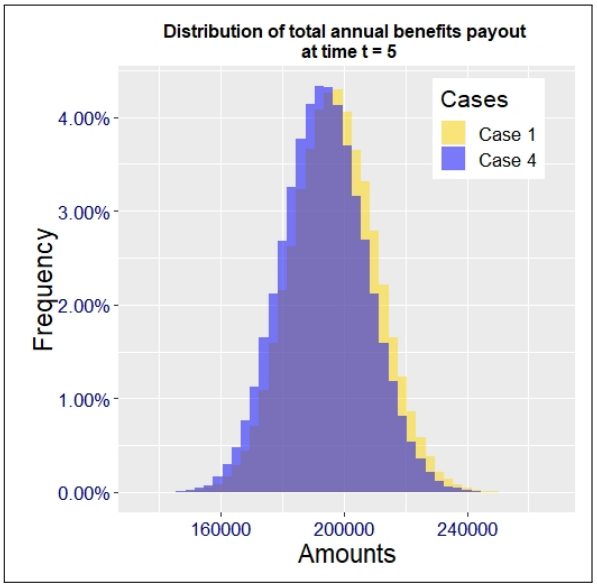

FIGURE 5.18: Case 1 vs Case 4, distribution of $X(5)$.

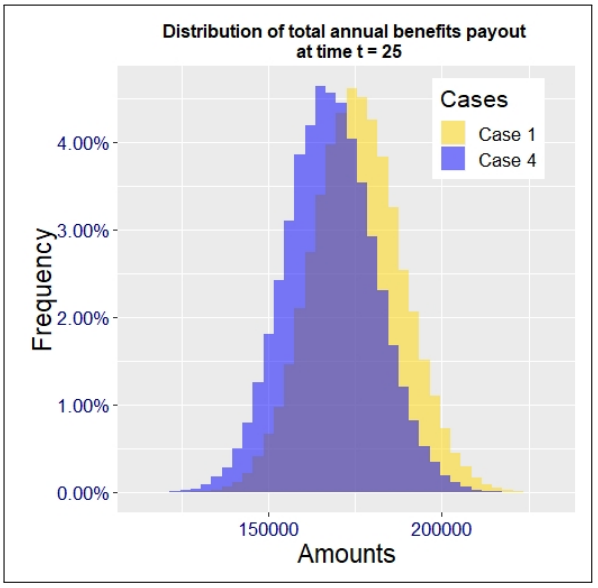

FIGURE 5.19: Case 1 vs Case 4, distribution of $X(25)$.

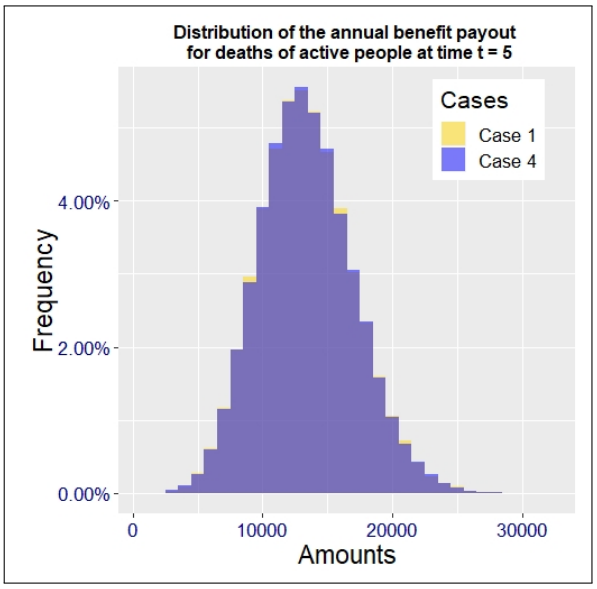

FIGURE 5.20: Case 1 vs Case 4, distribution of $X_1(5)$.

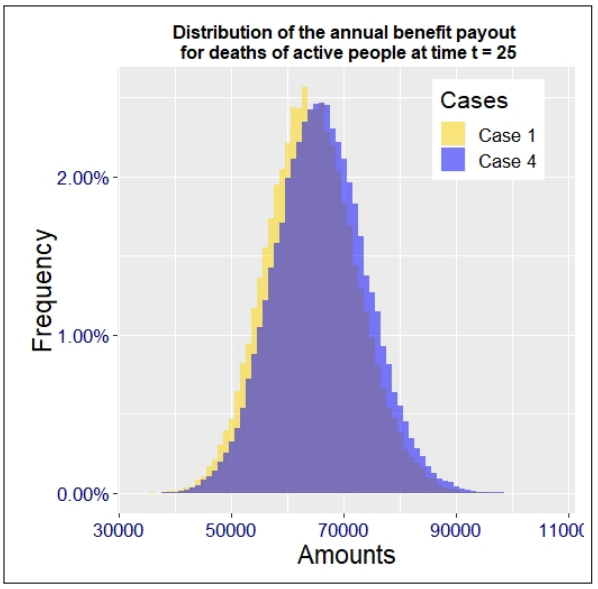

FIGURE 5.21: Case 1 vs Case 4, distribution of $X_1(25)$.

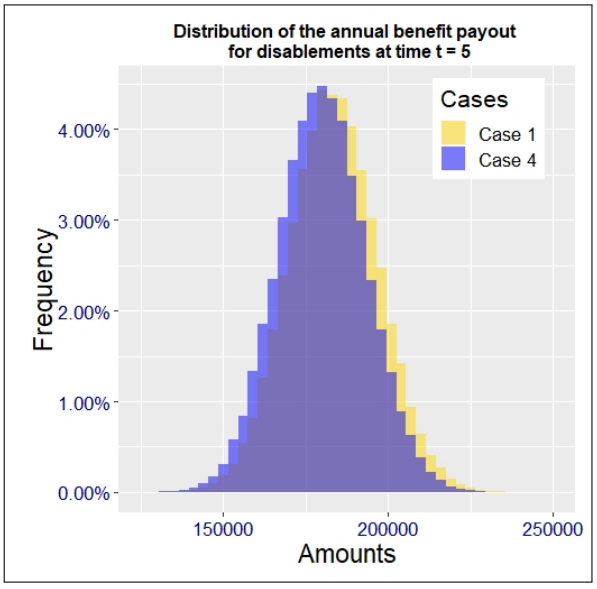

FIGURE 5.22: Case 1 vs Case 4, distribution of $X_2(5)$.

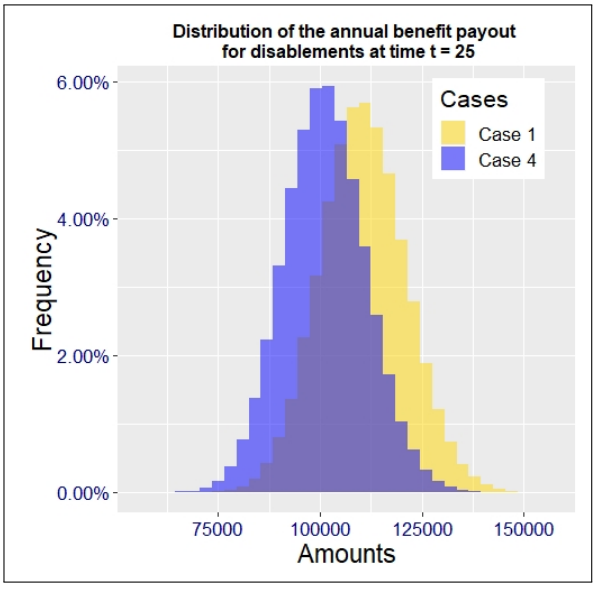

FIGURE 5.23: Case 1 vs Case 4, distribution of $X_2(25)$.

5.3.4 Case 4 vs Case 5

Frailty is modelled as a continuous or a discrete variable respectively. Discrete modelling is realized via grouping with random group sizes. We note that, in terms of simulated distributions, results almost coincide. Thanks to this fact, discrete frailty modelling can be considered as an interesting practical alternative.

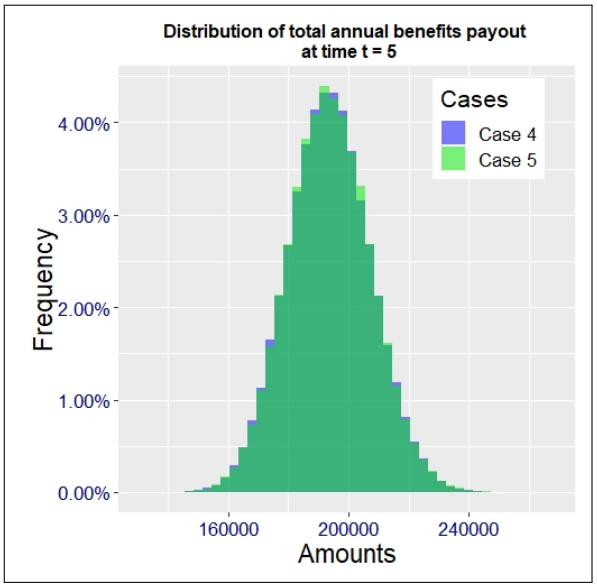

FIGURE 5.24: Case 4 vs Case 5, distribution of $X(5)$.

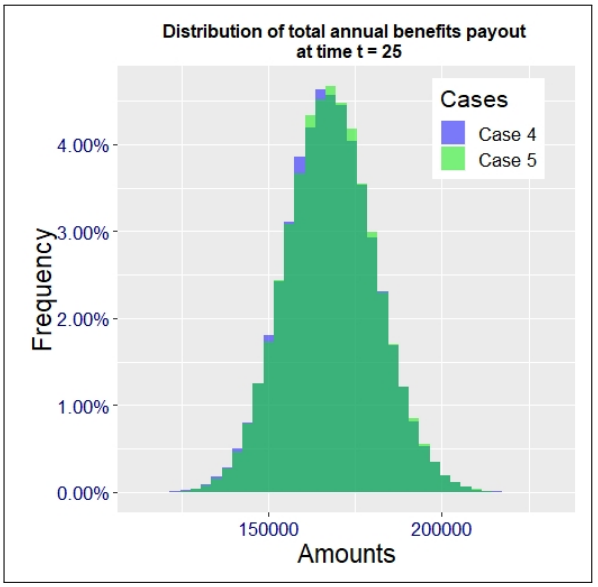

FIGURE 5.25: Case 4 vs Case 5, distribution of $X(25)$.

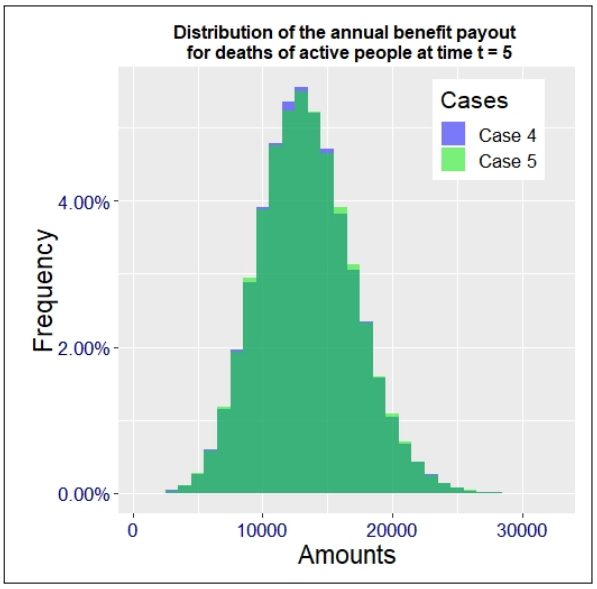

FIGURE 5.26: Case 4 vs Case 5, distribution of $X_1(5)$.

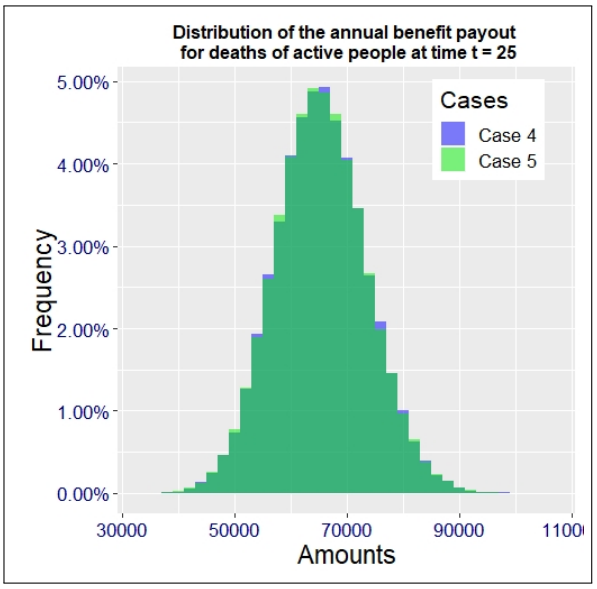

FIGURE 5.27: Case 4 vs Case 5, distribution of $X_1(25)$.

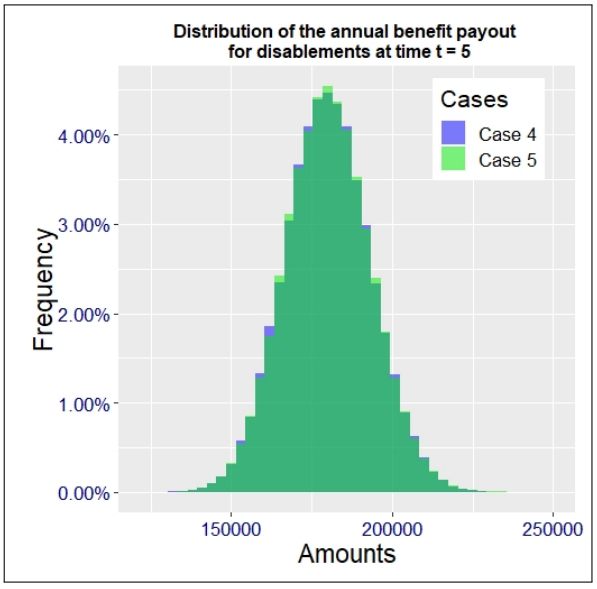

FIGURE 5.28: Case 4 vs Case 5, distribution of $X_2(5)$.

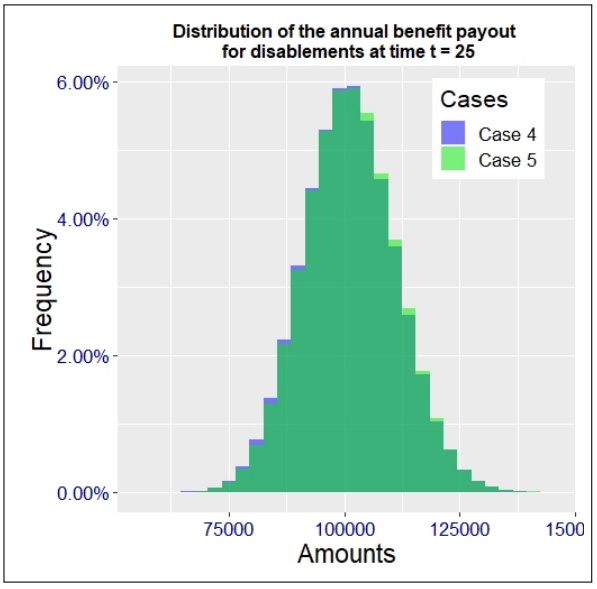

FIGURE 5.29: Case 4 vs Case 5, distribution of $X_2(25)$.

5.3.5 Case 2 vs Case 6

In both the cases, the portfolio is split into subportfolios. Case 2 is totally deterministic (see above), whereas in Case 6 uncertainty is allowed for and affects each subportfolio separately. Uncertainty heavily impacts on the portfolio risk profile. However the impact is much lower when only benefits in case of death of active people is involved (see Figs. 5.32 and 5.33); to this regard, see the remarks in the comparison between Case 1 and Case 3.

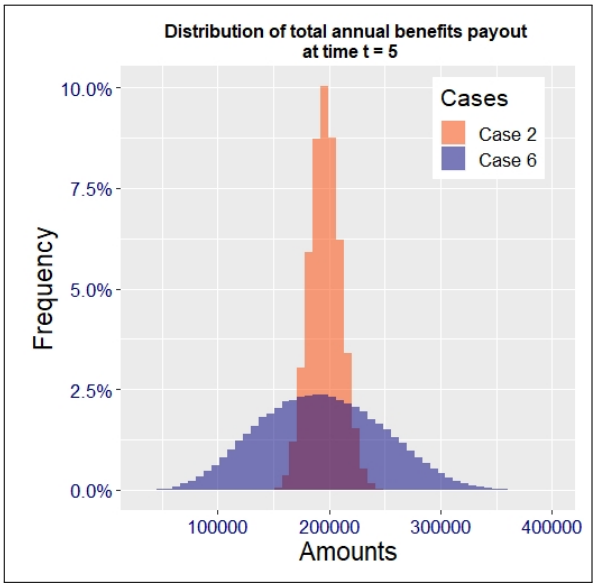

FIGURE 5.30: Case 2 vs Case 6, distribution of $X(5)$.

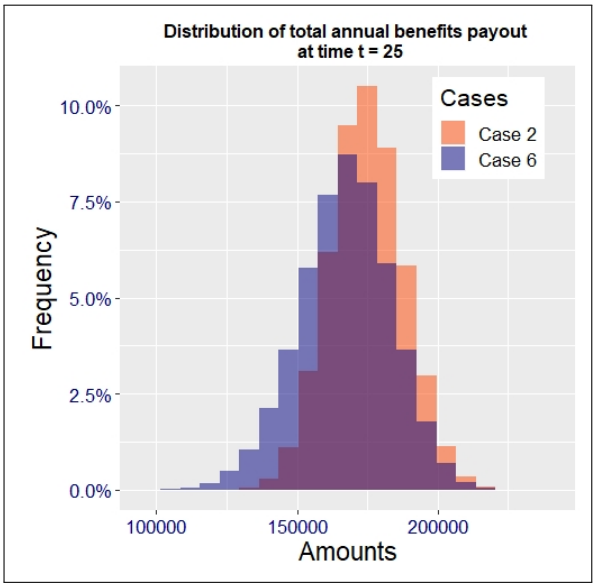

FIGURE 5.31: Case 2 vs Case 6, distribution of $X(25)$.

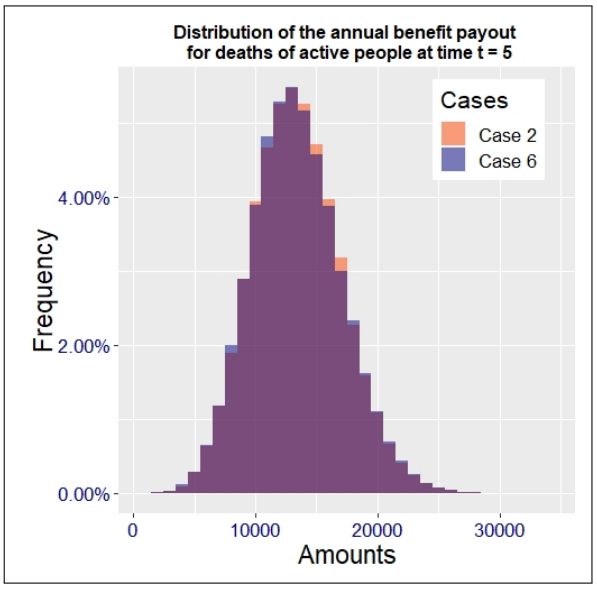

FIGURE 5.32: Case 2 vs Case 6, distribution of $X_1(5)$.

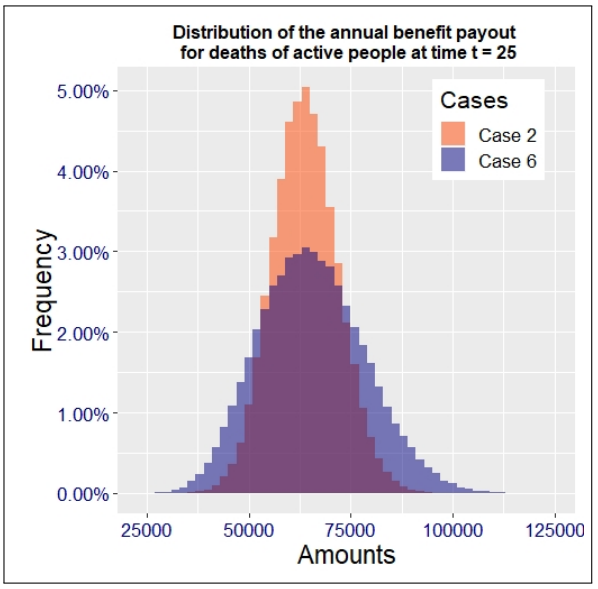

FIGURE 5.33: Case 2 vs Case 6, distribution of $X_1(25)$.

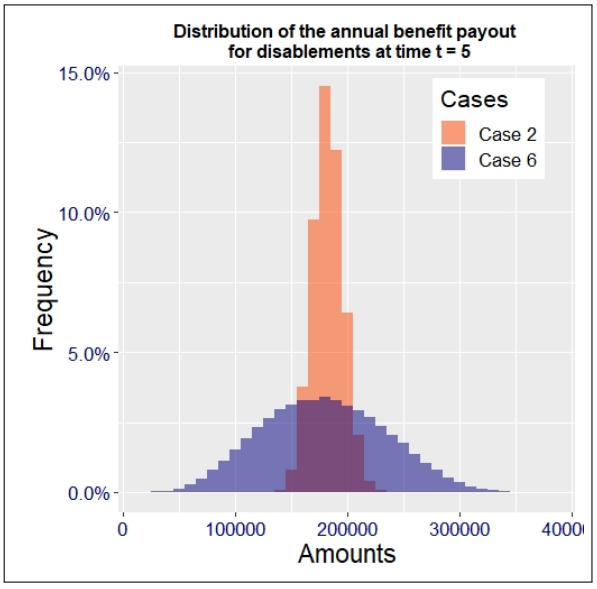

FIGURE 5.34: Case 2 vs Case 6, distribution of $X_2(5)$.

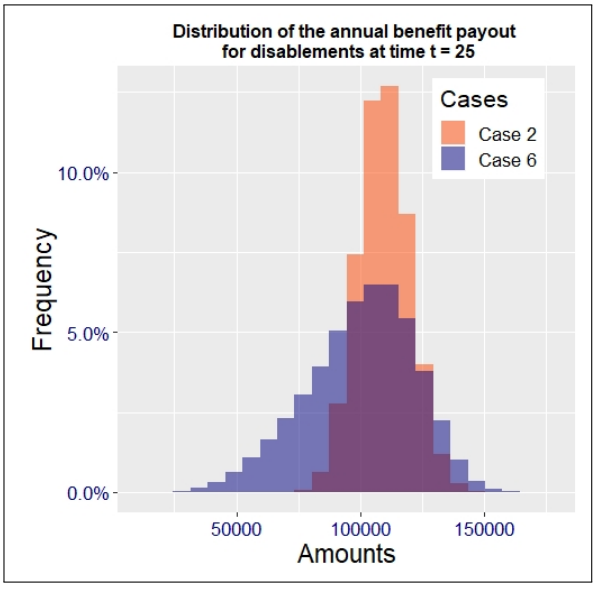

FIGURE 5.35: Case 2 vs Case 6, distribution of $X_2(25)$.

5.3.6 Case 3 vs Case 6

In present two cases we again call attention to the effect of splitting the portfolio into homogeneous subportfolios due to the observable risk factors. Unlike previous comparison (see Case 1 vs Case 2) in both the cases, the settings consider uncertainty

in the assessments of the probability of disablement. The impact is evident on Figs 5.36, 5.37 , 5.40 and 5.41. However the impact is not significant when refers to the benefits in case of death of active people is involved (see Figs. 5.38 and 5.39).

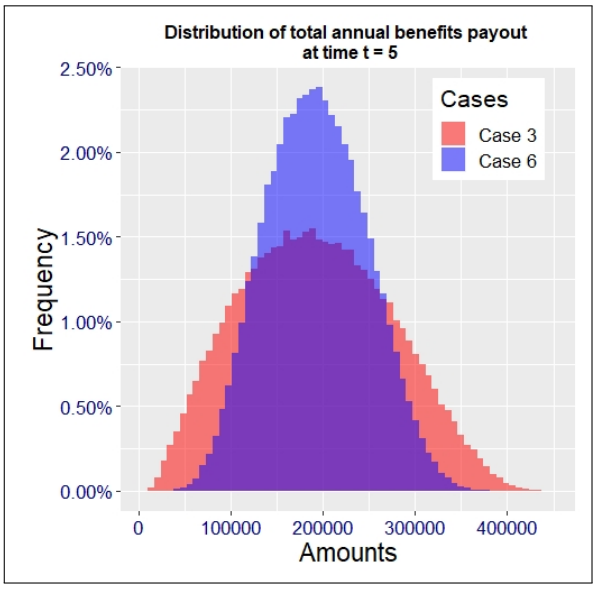

FIGURE 5.36: Case 3 vs Case 6, distribution of $X(5)$.

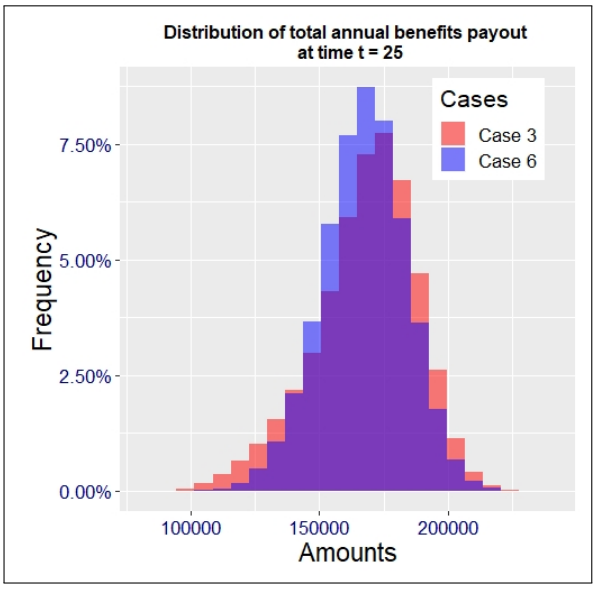

FIGURE 5.37: Case 3 vs Case 6, distribution of $X(25)$.

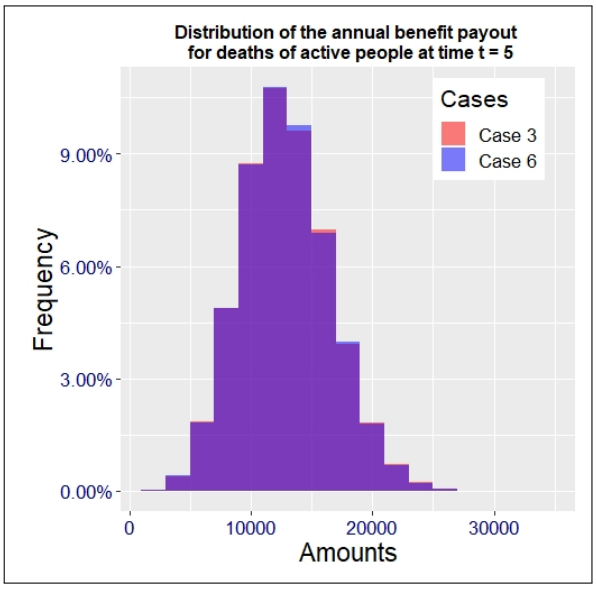

FIGURE 5.38: Case 3 vs Case 6, distribution of $X_1(5)$.

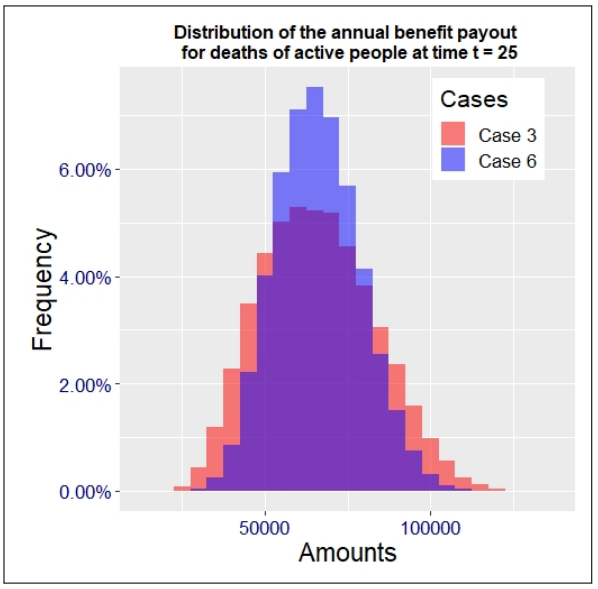

FIGURE 5.39: Case 3 vs Case 6, distribution of $X_1(25)$.

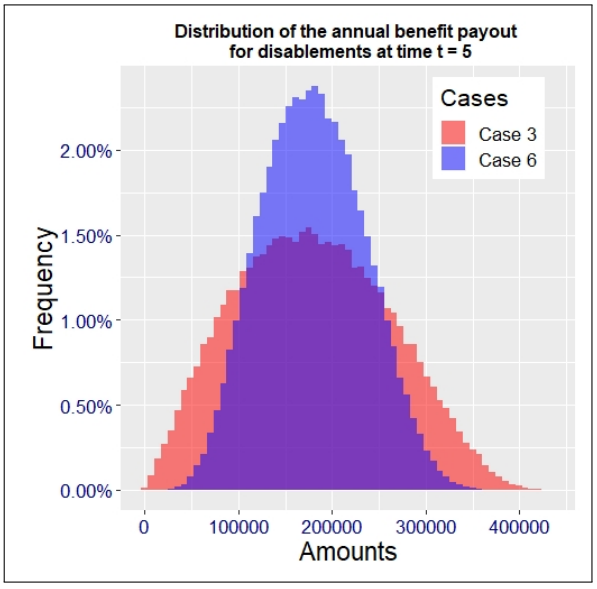

FIGURE 5.40: Case 3 vs Case 6, distribution of $X_2(5)$.

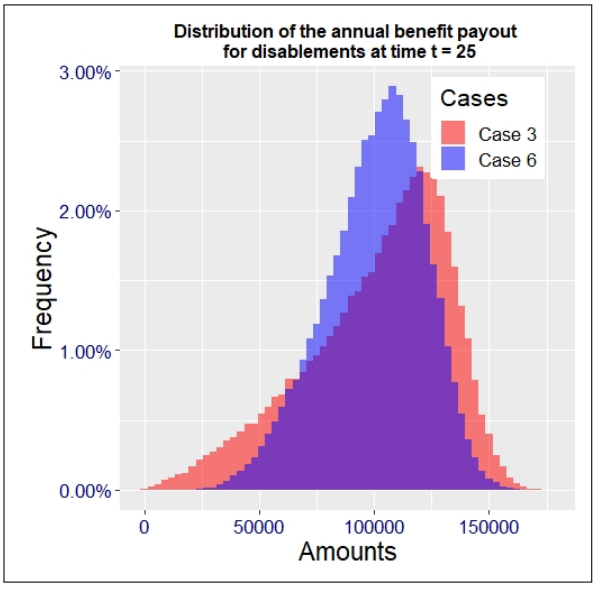

FIGURE 5.41: Case 3 vs Case 6, distribution of $X_2(25)$.

5.4 The percentile principle: some results

5.4.1 Assets

In Figs. 5.42 and 5.43 the total amounts of assets facing the liabilities are shown for various percentiles, at time $t = 1$ and $t = 5$ respectively. We note that, because of the uncertainty effect, the assets required are much higher in Case 3 than in Case 1. In Figs. 5.44, 5.45 and 5.46 the assets required for the percentile 95% are displayed. We note that requirements of course decrease with the decreasing portfolio size and thence the decreasing insurers liabilities. Finally, Fig. 5.47 shows the ratio between assets required (according to 95% percentile) and expected value of the total payout. We see, also from this perspective, the heavy impact of uncertainty in the probability of disablement.

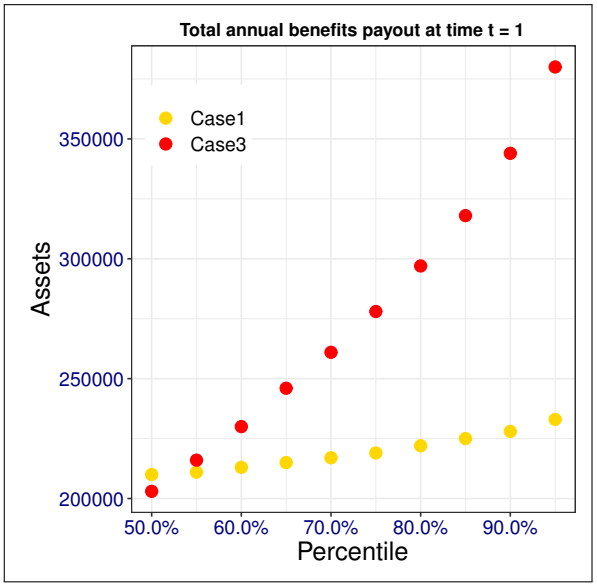

FIGURE 5.42: Percentiles for total payout $X(1)$, Case 1 vs Case 3

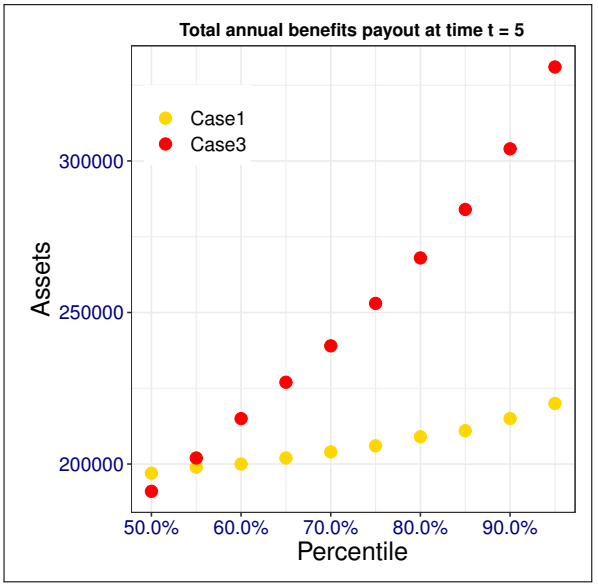

FIGURE 5.43: Percentiles for total payout $X(5)$, Case 1 vs Case 3

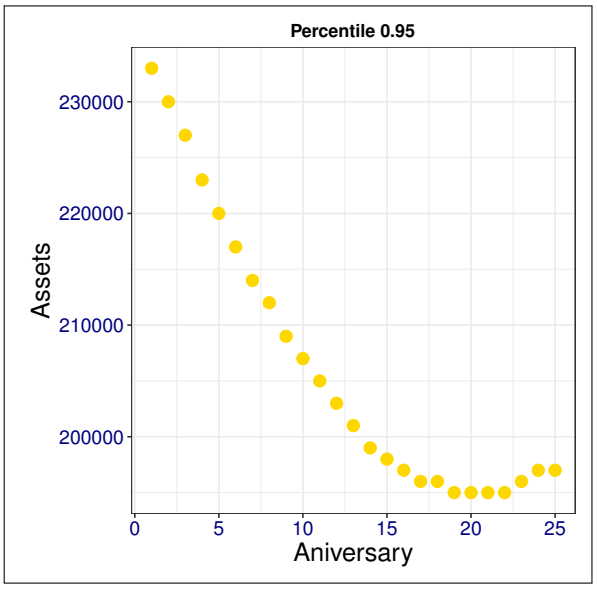

FIGURE 5.44: 95% percentiles for $X(t),\ t = 1, \ldots, 25$, Case 1.

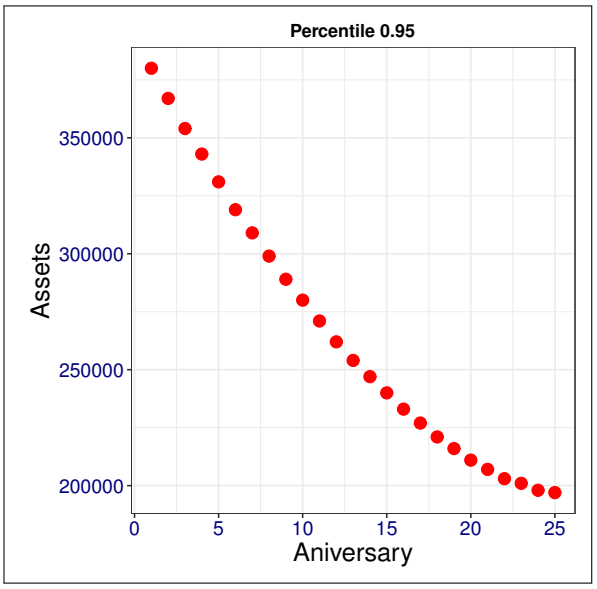

FIGURE 5.45: 95% percentiles for $X(t),\ t = 1, \ldots, 25$, Case 3.

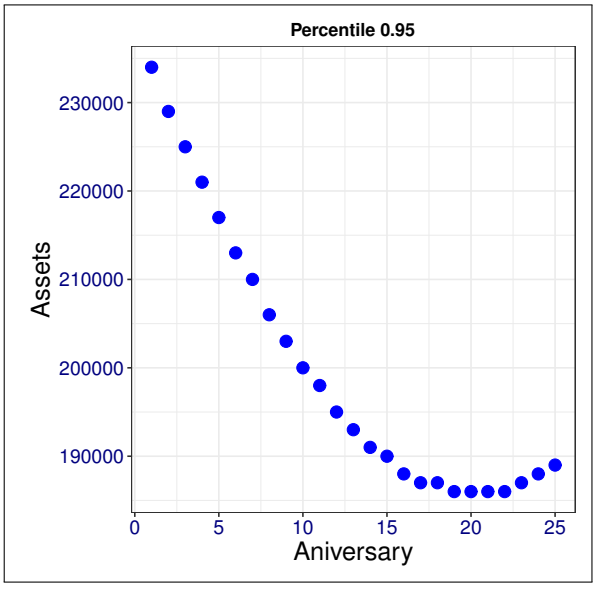

FIGURE 5.46: 95% percentiles for $X(t),\ t = 1, \ldots, 25$, Case 4.

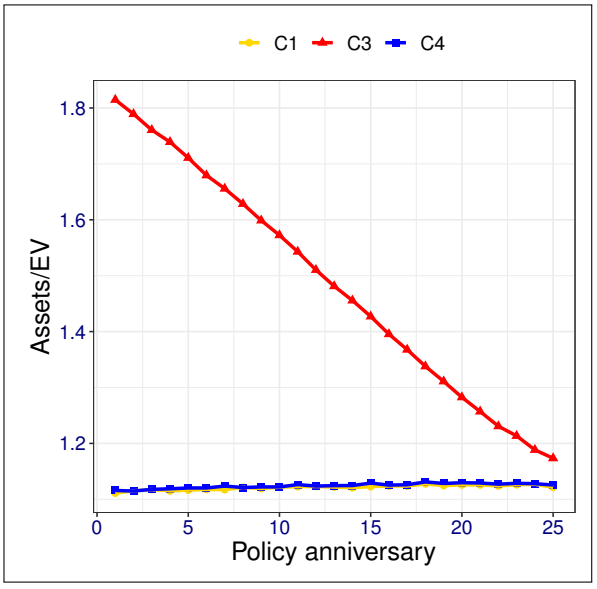

FIGURE 5.47: Assets backing the liabilities/ Expected value

Chapter 6

Concluding remarks and outlook

Probability distributions of the annual payouts in an insurance portfolio constitute the main topic dealt with in the present thesis. In particular, we have addressed a portfolio of term insurance policies providing lump sum benefits both in case of death and in case of permanent disability.

The analysis of probability distributions allows us to assess the riskiness inherent in an insurance portfolio, and hence to suggest appropriate actions in terms of premiums and capital allocation. In this regard, we have adopted the percentile principle.

As the baseline probabilistic structure, we have adopted a traditional three-state model in a Markov context. We have then proposed to implement the three-state model by allowing for:

- uncertainty in assigning transition probabilities;
- individual frailty, which causes heterogeneity inside the insurance portfolio because of unobservable risk factors.

Among our achievements, we stress the impact assessment of uncertainty and frailty on the risk profile of the portfolio, in particular in terms of assets to allocate in order to face the portfolio riskiness. Other significant results relate to the effect of splitting the portfolio into (more or less homogeneous) subportfolios.

The model we have proposed can be generalized in several ways, and can be implemented for other purposes. As regards possible generalizations:

1. the benefit structure can be extended, for example to include the payment of recurrent benefits such as disability annuities;
2. the multistate model can be extended from three to four states in order to represent diverse degrees of disability severity.

We note that generalization 1 can allow us to represent various insurance covers in the framework of Income Protection insurance.

Implementations may aim in particular at sensitivity analysis. For example:

1. to assess the impact of the (initial) portfolio size and hence the diversification effect produced by risk pooling (as regards the random fluctuation risk);

2. to assess the impact of a sudden jump in mortality and/or disability, thus according to the logic of stress testing.

We note that the above implementations do not require generalizations of the structure of our model.

Bibliography

Amsler, M. H. (1988) Sur la modélisation de risques vie par les chaînes de Markov. In *Transactions of the 23rd International Congress of Actuaries*, volume 3, pp. 1–17. Helsinki: .

Beard, R. E. (1959) Note on some mathematical mortality models. In *Ciba Foundation Symposium-The Lifespan of Animals (Colloquia on Ageing)*, volume 5, pp. 302–311.

Black, K. and Skipper, H. D. (1999) *Life & Health Insurance*. 13th edition. Pearson.

Bowers, N. L., Gerber, H. U., Hickman, J. C., Jones, D. A. and Nesbitt, C. J. (1997) *Actuarial Mathematics*. Schaumburg, Illinois: The Society of Actuaries.

Bülmann, H. (1970) *Mathematical Methods in Risk Theory*. Springer.

Cairns, J. G. A. (2000) A discussion of parameter and model uncertainty in insurance. *Insurance: Mathematics and Economics* **27**(3), 313–330.

Cummins, J. D., Smith, B. D., Vance, R. N. and VanDerhei, J. L. (1983) *Risk Classification in Life Insurance*. Kluwer - Nijhoff Publishing.

Denuit, M., Lucas, N. and Pitacco, E. (2019) Pricing and reserving in LTC insurance. In *Actuarial Aspects of Long Term Care*, eds E. Dupourqué, F. Planchet and N. Sator, Springer Actuarial, pp. 129–158. Springer.

Dickson, D. C. M., Hardy, M. R. and Waters, H. R. (2013) *Actuarial Mathematics for Life Contingent Risks*. Second edition. Cambridge University Press.

Gerber, H. U. (1995) *Life Insurance Mathematics*. Springer-Verlag.

Haberman, S. and Pitacco, E. (1999) *Actuarial Models for Disability Insurance*. Chapman & Hall / CRC.

Heligman, L. and Pollard, J. H. (1980) The age pattern of mortality. *Journal of the Institute of Actuaries* **107**, 49–80.

Hoem, J. M. (1969) Markov chain models in life insurance. *Blätter der deutschen Gesellschaft für Versicherungsmathematik* **9**, 91–107.

Hoem, J. M. (1972) Inhomogeneous semi-Markov processes, select actuarial tables, and duration-dependence in demography. In *Population Dynamics*, ed. T. N. E. Greville, pp. 251–296. London: Academic Press.

Hoem, J. M. (1988) The versatility of the Markov chain as a tool in the mathematics of life insurance. In *Transactions of the 23rd International Congress of Actuaries*, volume R, pp. 171–202. Helsinki: .

Norberg, R. (2002) Basic Life Insurance Mathematics. Laboratory of Actuarial Mathematics, Dpt. of Mathematical Sciences, University of Copenhagen. Available at: `http://www.math.ku.dk/~mogens/lifebook.pdf`.

Olivieri, A. (2006) Heterogeneity in survival models. Applications to pension and life annuities. *Belgian Actuarial Bulletin* **6**, 23–39. Available at: `http://www.actuaweb.be/frameset/frameset.html`.

Olivieri, A. and Pitacco, E. (2009) Stochastic mortality: the impact on target capital. *ASTIN Bulletin* **39**(2), 541–563.

Olivieri, A. and Pitacco, E. (2015) *Introduction to Insurance Mathematics. Technical and Financial Features of Risk Transfers*. Second edition. EAA Series. Springer.

Pitacco, E. (2014) *Health insurance: Basic actuarial models*. EAA Series. Springer.

Pitacco, E. (2019) Heterogeneity in mortality: a survey with an actuarial focus. *European Actuarial Journal* **9**(1), 3–30.

Pitacco, E., Denuit, M., Haberman, S. and Olivieri, A. (2009) *Modelling Longevity Dynamics for Pensions and Annuity Business*. Oxford University Press.

Pollard, A. (1970) Random mortality fluctuations and the binomial hypothesis. *Journal of the Institute of Actuaries* **96**(2), 251–264.

Valente, M. (2017) *Eterogeneitá per fattori di rischio non osservabili nella assicurazioni vita*. Master's thesis, Università di Trieste.

Vaupel, J. W., Manton, K. G. and Stallard, E. (1979) The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* **16**(3), 439–454.

Waters, H. R. (1984) An approach to the study of multiple state models. *Journal of the Institute of Actuaries* **111**, 363–374.