ARC Centre of Excellence in Population Ageing Research

Working Paper 2018/7

Heterogeneity in mortality: a survey with an actuarial focus[1]

Ermanno Pitacco [2]

[1] Part of the research reported herein was performed in the framework of activities of the Mortality Working Group of the International Actuarial Association (http://www.actuaries.org/), and presented, as an invited lecture, at the UNSW-Macquarie University Workshop “Risk: modelling, optimization and inference with applications in Finance, Insurance and Superannuation”, University of New South Wales, Sydney, 7th-8th December 2017. A seminar on the research achievements was delivered at the Department of Mathematics, Stockholm University, on 18th April, 2018

[2] Dpt. DEAMS ‘B. de Finetti”, University of Trieste (Italy) and CEPAR Associate Investigator

Heterogeneity in mortality: a survey with an actuarial focus [1]

Ermanno Pitacco [2]

ermanno.pitacco@deams.units.it

Abstract

Heterogeneity of a population in respect of mortality 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. The impact of observable risk factors on individual mortality, in particular when they also constitute "rating factors" in the calculation of premiums and other actuarial values, is usually expressed approximately, according to some pragmatic approach. For example, additive or multiplicative adjustments to the average age-specific mortality are frequently adopted. Heterogeneity due to unobservable risk factors can conversely be quantified by adopting the concept of individual "frailty". However, individual frailty can be interpreted and consequently modeled in several ways, according to the causes which are considered as originating the frailty itself: congenital characteristics, environmental features, lifestyle aspects, etc. It follows that the individual frailty can, in particular, be assumed either constant or variable throughout the lifetime.

In this paper, we provide a survey of scientific contributions on mortality heterogeneity, focusing on modeling both observable and unobservable heterogeneity. We start with an overview of methodological

[1] Part of the research reported herein was performed in the framework of activities of the Mortality Working Group of the International Actuarial Association (http://www.actuaries.org/), and presented, as an invited lecture, at the UNSW-Macquarie University Workshop "Risk: modelling, optimization and inference with applications in Finance, Insurance and Superannuation", University of New South Wales, Sydney, 7th - 8th December, 2017. A seminar on the research achievements was delivered at the Department of Mathematics, Stockholm University, on 18th April, 2018.

[2] Dpt. DEAMS "B. de Finetti", University of Trieste (Italy)

contributions to heterogeneity and frailty modeling, coming from both the demographical and the actuarial context. We then shift to contributions analyzing the impact of frailty, in its various interpretations, on the results (cash flows, profits, etc.) of life insurance and life annuity portfolios and related risk profiles.

Keywords: Heterogeneity, Frailty, Risk factors, Force of mortality, Mortality laws, Parametric models, Special-rate annuities

1 Introduction and motivation

Heterogeneity of a population in respect of mortality 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. For example, age, gender and geographical area of residence are observable risk factors commonly accounted for when analyzing mortality in national populations. Possibly, marital status and working vs retired position can constitute further risk factors for national population analysis.

More risk factors can be observed and allowed for when referring to life insurance and life annuity portfolios, e.g. occupation, past and current health conditions, etc., while others are unobservable, e.g. individual attitude towards health, some congenital personal characteristics, etc. It is worth noting that, whatever the type of population or portfolio concerned, a residual heterogeneity inside each homogeneous "group" remains because of unobservable factors.

The impact of observable risk factors on individual mortality, in particular when they also constitute "rating factors" in the calculation of premiums and other actuarial values, is usually expressed approximately, according to some pragmatic approach. For example, additive or multiplicative adjustments to the average age-specific mortality are frequently adopted.

Heterogeneity due to unobservable risk factors can conversely be quantified by adopting the concept of individual "frailty". However, individual frailty can be interpreted and consequently modeled in several ways, according to the causes which are considered as originating the frailty itself: congenital characteristics, environmental features, lifestyle aspects, etc. It follows that the individual frailty can, in particular, be assumed either constant or variable throughout the lifetime.

A huge number of scientific and technical contributions in the field of heterogeneity in respect of mortality have been proposed, in particular:

- in demography, e.g. to explain mortality deceleration at high ages in terms of unobservable heterogeneity factors (a controversial issue!);
- in actuarial science, to represent the impact of observable heterogeneity factors on individual mortality, aiming at the construction of appropriate pricing and reserving models, and to assess the impact of unobservable heterogeneity factors on the risk profile of life insurance and life annuity portfolios to determine solvency requirements and hence capital allocation.

This paper aims at providing some guidelines, hopefully useful in exploring the complex network of contributions to the analysis of heterogeneity in respect of mortality. The remainder of the paper is organized as follows.

In Sect. 2 we suggest an approach to reading and interpreting the extensive demographical and actuarial literature on mortality heterogeneity. Sect. 3 focusses on some forerunners in the demographic and actuarial fields. Formal approaches to heterogeneity are proposed in Sect. 4, while Sect. 5 addresses seminal contributions in the field of observable and unobservable heterogeneity. More recent contributions are presented in Sect. 6. Allowing for heterogeneity in actuarial calculations is discussed in Sect. 7, with specific reference to pricing, product design, capital allocation and risk transfers; special attention is placed on the impact of unobservable heterogeneity. Sect. 8 concludes the paper with some final remarks.

2 Exploring heterogeneity: a risk-oriented approach

A key to the reading and interpretation of the extensive demographical and actuarial literature on mortality heterogeneity can be provided, in line with the ultimate focus of this paper, by the Risk Management (RM) logic, more precisely by the phases of the RM process. We focus on the following phases (see Fig. 1), and we specifically address heterogeneity issues in life insurance and life annuity business.

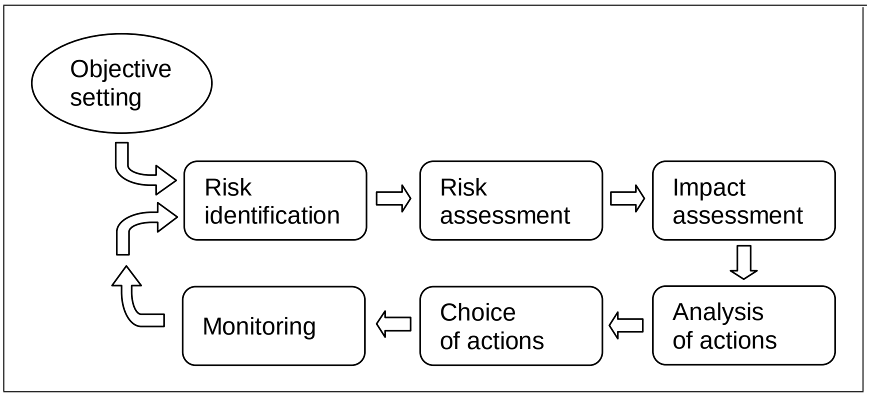

Figure 1: The RM process

In the preliminary *Objective setting* phase, targets of the organization, an insurance company in particular, must be specified. Among these, we find:

profit, value creation, risk mitigation, solvency, market share (for details, see e.g. Olivieri and Pitacco (2015)).

In the *Risk identification* phase, the risk causes, i.e. the causes of profits gained or losses suffered by an organization are singled out. The awareness of heterogeneity with respect to mortality is the first result, for an insurance company, of the risk identification phase, while the second result consists in recognizing that the heterogeneity may be due to both observable and unobservable risk factors.

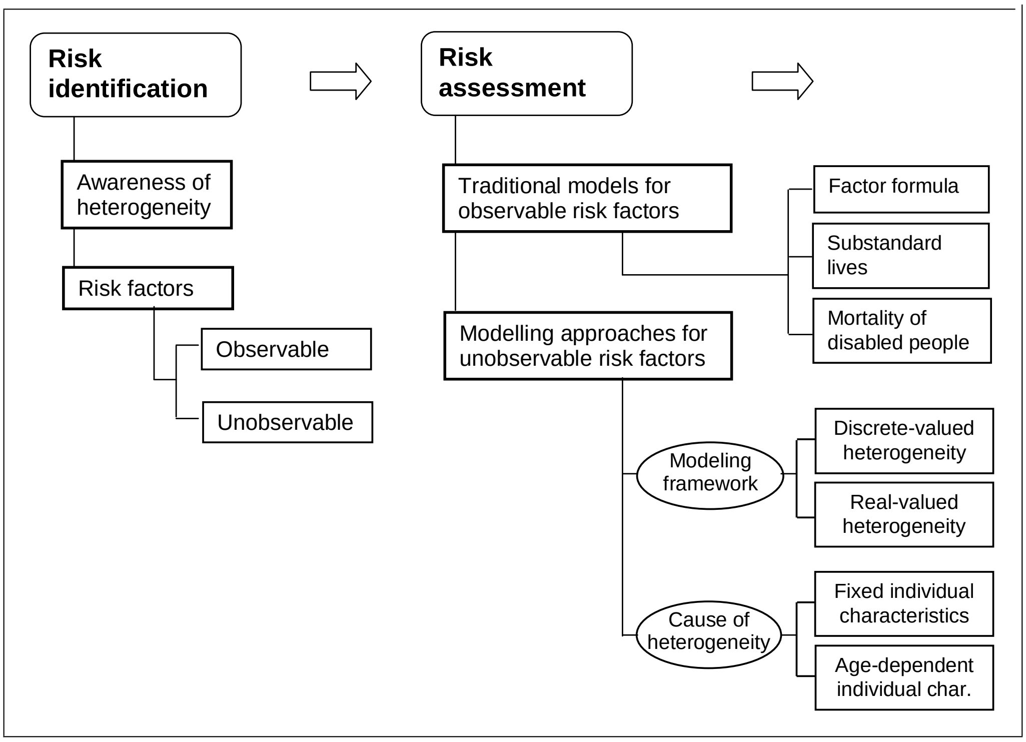

Figure 2: The biometric side: risk identification and risk assessment

Risk causes are expressed in quantitative terms via appropriate models (viz probability distributions, or, at least, typical values, e.g. expected values) in the *Risk assessment* phase. This phase can rely on well established approaches when observable risk factors are concerned. Indeed, a number of mortality models, suggested by medical statistics, have been proposed and implemented to express, in particular, the impact of health conditions. Heterogeneity models suitable in this regard can be classified as *individual models*, aiming at representing "differential mortality".

Conversely, in order to formally modeling the overall impact of unobservable risk factors, *collective models* are needed. Different approaches can be adopted: discrete-valued or real-valued measures of heterogeneity can be used, and different individual characteristics can be considered as causes of heterogeneity. Thus, the risk assessment phase involves several biometric issues.

Various aspects of the risk identification and the risk assessment phases are sketched in Fig. 2.

The *Impact assessment* phase aims at quantifying the effect of risk causes on results of interest (cash-flows, assets, profits, value creation, etc.) in terms of probability distributions of the results themselves, and relevant typical values (expected values, variances, etc.). The impact assessment phase constitutes the preliminary step to the analysis and the choice of RM actions. See Fig. 3.

The *Analysis of actions* consists in listing the available RM tools, then comparing the relevant costs and benefits. In the insurance and annuity business, actions can consist in:

- the product design and related pricing and reserving;
- risk hedging actions, i.e. capital allocation, reinsurance, Alternative Risk Transfers (ARTs, that is, transfers via mortality-linked securities, swaps, etc.).

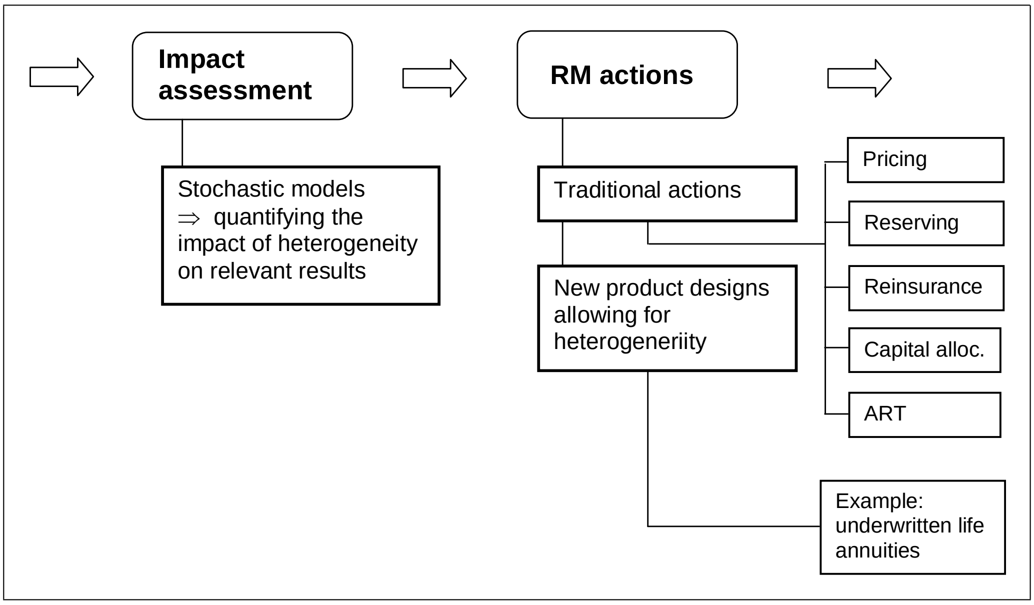

Figure 3: The actuarial side: impact assessment and risk management actions

The *Choice of actions* usually results in an appropriate mix of actions (e.g. combining risk transfer and capital allocation).

Impact assessment, analysis of actions and choice of actions imply the use of an appropriate actuarial toolkit.

The *Monitoring* phase should involve both the results achieved by the organization and the assumptions about the scenario (e.g. mortality trends, behavior of the capital markets, tax legislation, etc.) adopted when choosing RM actions.

3 Heterogeneity: the awareness

3.1 Among the antecedents: Francis Corbaux

The presence of heterogeneity in respect of mortality is intuitive and supported by statistical evidence. Well known examples are given by male mortality versus female mortality, by the impact of environmental features on the age-pattern of mortality, etc.

Heterogeneity has been recognized since the early contributions in demography and actuarial science. As Haberman (1996) notes, the earliest life tables separately developed for males and females and based on annuitants' mortality were constructed in 1740 by Nicholas Struyck.

In Corbaux (1833) we find: "*The object of consideration ... various classes susceptible of being discriminated amongst any extensive population, ...*". Corbaux proposed to split a population into five classes, relying on twelve risk factors (or proxies), concerning health, lifestyle, environment, etc.

We note that schemes in which rating classes are defined on the basis of several risk factors constitute a pillar in risk classification for life insurance business; see Sect. 4.3.1, and in particular, among the earliest contributions, Rogers and Hunter (1919). Recently, risk factors have also been adopted in defining rating criteria for underwritten (or special-rate) life annuities; see Sect. 7.4, and e.g. Rinke (2002) as regards rating criteria.

3.2 Wilfred Perks and the logistic models

As mentioned in Sect. 1, significant efforts have been devoted to explore possible relationships between heterogeneity and mortality deceleration at high ages, in particular mortality leveling-off. For a detailed survey, the reader should refer to Olshansky (1998) and Olshansky and Carnes (1997); a brief survey is provided by Pitacco (2016a).

Several models have been proposed, that are strictly related each other and share the purpose of representing a mortality leveling-off. In formal terms, the common feature of these models consists in a horizontal asymptote of the (instantaneous) force of mortality (or hazard function) μ_x. We note that, in all the following logistic-type models, the numerator of the fraction is given by a Makeham (or Gompertz) -type term (see Gompertz (1825), Makeham (1867)).

In 1932 W. Perks, aiming at the graduation of population mortality data, proposed two mortality laws; see Perks (1932). The first Perks law is as follows:

$$\mu_x = \frac{\alpha\,\mathrm{e}^{\beta x} + \gamma}{\delta\,\mathrm{e}^{\beta x} + 1} \tag{1}$$

while the second Perks law has the more general structure:

$$\mu_x = \frac{\alpha\,\mathrm{e}^{\beta x} + \gamma}{\delta\,\mathrm{e}^{\beta x} + \epsilon\,\mathrm{e}^{-\beta x} + 1} \tag{2}$$

The following law results from the approach to unobservable heterogeneity proposed by Beard (1959) (as we will see in Sect. 5.2.1):

$$\mu_x = \frac{\alpha\,\mathrm{e}^{\beta x}}{\delta\,\mathrm{e}^{\beta x} + 1} \tag{3}$$

We note that the Beard law (3) can simply be obtained from the first Perks law (1) by setting $\gamma = 0$.

Various models have been proposed more recently, still aiming at the graduation of population mortality data. We first recall the law proposed by Kannisto (1994):

$$\mu_x = \frac{\alpha\,\mathrm{e}^{\beta x}}{\alpha\,\mathrm{e}^{\beta x} + 1} \tag{4}$$

which can be obtained by setting $\gamma = 0$ and $\delta = \alpha$ in the first Perks law.

Thatcher (1999) proposed the following expression for the force of mortality:

$$\mu_x = \frac{\nu\,\alpha\,\mathrm{e}^{\beta x}}{\alpha\,\mathrm{e}^{\beta x} + 1} + \kappa \tag{5}$$

The simplified version of (5), used in particular for studying long-term trends and forecasting mortality at very old ages, has $\nu = 1$ and hence only three parameters, namely α, β and κ:

$$\mu_x = \frac{\alpha\,\mathrm{e}^{\beta x}}{\alpha\,\mathrm{e}^{\beta x} + 1} + \kappa \tag{6}$$

All the above models (together with other models) have been tested fitting mortality data of diverse populations. As regards mortality at old ages, see for example Thatcher et al. (1998) and references therein.

3.3 Some forerunners in the Fifties

As mentioned in Sect. 3.2, Robert Eric Beard (1959) provided a seminal contribution to the modeling of heterogeneity due to unobservable risk factors. According to Beard's approach, individual mortality is described by a Gompertz or a Makeham law, with parameters depending on a specific *longevity factor* (corresponding to what we currently call individual frailty), whose presence in the population is described by a gamma distribution. It follows that the mortality in the population follows a particular Perks law, hence implying deceleration of mortality at high ages. The Beard's contribution constitutes the starting point of the *fixed-frailty theory* (see Sect. 5.2).

The contribution by Louis Levinson (1959) probably constitutes the first attempt of modeling heterogeneity in a *dynamic setting* (see Sect. 5.3). The approach adopted by Levinson relies, in modern terms, on a "multistate" structure. Every population is heterogeneous in respect of mortality, and even if split into groups (e.g. the group of insureds accepted as "normal risks", the group of "substandard risks", etc.), each group is affected by some degree of heterogeneity. Levinson proposed the concept of *mortality strata*: each stratum consists of individuals with the same probability of death (regardless of age). Individuals move from one stratum to another one, in particular because of ageing, and in general because of health *deterioration*. The ultimate purpose of Levinson's work was the construction of life tables allowing for strata, and the relevant application to US mortality.

Claudio de Ferra (1954), generalizing a previous contribution by de Finetti and Taucer (1952), focused on a heterogeneous population split into a given number of homogeneous groups. The age-pattern of mortality in each group follows a Makeham law, with group-specific parameters. The population structure is described by the distribution of the parameters. The paper addresses the calculation of the (non-Makeham) law which expresses the age-pattern of mortality in the heterogeneous population, given the distribution of the Makeham parameters, and the construction of a Makeham approximation to the above law. A specific application, which places this contribution in the actuarial framework, refers to an insured population consisting of normal risks and substandard risks, with different extra-mortality levels.

4 Heterogeneity: formal approaches

Looking at demographical and actuarial literature, we can recognize two basic approaches to heterogeneity in mortality.

4.1 An "intuitive" approach

A heterogeneous population can be considered as a (finite) set of (more or less) homogeneous groups. The age-pattern of mortality in the population can then be represented as a (finite) mixture of the age-patterns of mortality in the various groups.

Formally, refer to a biometric function f; 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 m groups, the function f is then expressed as a mixture of the functions $f^{(i)}$, $i = 1, \dots, m$, pertaining to the various groups:

$$f = w_1 \, f^{(1)} + w_2 \, f^{(2)} + \cdots + w_m \, f^{(m)} \tag{7}$$

Several specific models can be placed in the framework described by Eq. (7). For example:

- the functions $f^{(i)}$, $i = 1, \dots, m$, can be suggested by various risk factors (e.g. individual health status, individual occupation, geographical area, etc.), and may be either known or unknown, depending on information available about the age-pattern of mortality inside each group;
- the weights w_i, $i = 1, \dots, m$, may be either known or unknown, depending on information available about the (relative) group sizes;
- 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.

Remark

> The construction of a uni-gender life table (requested by the European legislation for pricing insurance products) can be seen as a particular implementation of the scheme described by Eq. (7) with $m = 2$. We also note that the relative sizes w_1 and w_2 have to be assessed according to previous portfolio experience, but must be looked at as random quantities when referring to the future portfolio composition.

We note that Eq. (7) represents a discrete (finite) approach to heterogeneity in mortality.

4.2 Parametric representation: a (rather) general setting

A rather general setting, suggested by models developed in demography and in actuarial sciences, can be defined by adopting a parametric approach. In formal terms:

- choose a biometric function f to represent the "standard" (e.g. average) age-pattern of mortality in a given population (e.g. the general population in a country, the members of a pension fund, the insureds in a portfolio, etc.);
- express a "specific" age-pattern of mortality (in particular, mortality of people in poorer or better conditions than the average) as a transform Φ of f, involving various parameters.

Examples of function f have been given in Sect. 4.1. Referring to the force of mortality, μ, a specific age-pattern of mortality can be expressed as follows:

$$\mu_{x,t}^{[\mathrm{spec}]} = \Phi\left[\mu_{[x]+s+t}; \rho_{x,t}^{(1)}, \dots, \rho_{x,t}^{(r)}; z_{x,t}\right] \tag{8}$$

where:

x is a given age;

t is the past duration, that is, the time elapsed since a given event (occurred at age x); thus, $x + t$ is the current age;

$\mu_{[x]+s+t}$ denotes (according to the traditional actuarial notation) the standard select force of mortality at age $x + s + t$, where s is the "years-to-age" addition, also called "age-shift" parameter, summarizing the impact of some observable risk factors; we note that $\mu_{[x]+s+t}$ is a function of the two variables x and $s + t$ separately;

$\rho_{x,t}^{(j)}$ represents the impact of the observable risk factor j, $j = 1, \dots, r$;

$z_{x,t}$ denotes the overall impact of unobservable risk factors.

As regards the meaning of x, we note what follows:

- x can denote the age at policy issue; hence the model (8) is an *issue-select model*, expressing the "duration-since-initiation" dependence;
- x can represent the age at disability inception; the model (8) is then an *inception-select model*, which expresses the "duration-in-current-state" dependence;

- $x = 0$ can be assumed in a demographical analysis, addressing e.g. the age-pattern of mortality over the whole life span; in this case, $\mu_{[x]+s+t}$ will simply be replaced by μ_t.

4.3 Parametric representation: examples

Some examples of parametric representation follow, referring to either observable or unobservable risk factors.

4.3.1 Observable risk factors (disregarding unobservable risk factors)

In the context of life insurance technique, assessing the impact of observable risk factors fits into the framework of risk classification (see e.g. Haberman and Olivieri (2014)).

We consider the following (rather) general model:

$$\mu_{x,t}^{(h)} = A_{x,t}^{(h)} \, \mu_{x+s^{(h)}+t} + B_{x,t}^{(h)}; \quad A_{x,t}^{(h)}, B_{x,t}^{(h)}, s^{(h)} \geq 0 \tag{9}$$

where:

x is the age at policy issue, or the age at disability inception, while $x + t$ is the current age;

μ denotes the standard force of mortality;

h denotes a "group", i.e. a rating class;

$A_{x,t}^{(h)}, B_{x,t}^{(h)}, s^{(h)}$ summarize the impact of the observable risk factors $\rho_{x,t}^{(j)}$, $j = 1, \dots, r$, (see formula (8)) by assigning the individual risk to the group h;

the past-duration effect is accounted for via parameters $A_{x,t}^{(h)}$ and $B_{x,t}^{(h)}$.

The model described by Eq. (9) encompasses several simpler formulae adopted in the actuarial practice. In several implementations, the annual probabilities of death q are referred to, instead of the force of mortality μ. We focus on the following examples.

1. The traditional risk classification scheme in life insurance splits the insured population into *standard* risks and *sub-standard* risks, for which an extra-mortality is detected. Simple models for the *mortality of sub-standard lives* belong to the actuarial tradition, and are currently adopted for term insurance rating and, more recently, for underwritten

life annuities rating (see, for example, Ainslie (2000)). Disregarding, in particular, the past duration effect, we find the so-called linear model:

$$\mu_{x+t}^{(h)} = A^{(h)}\,\mu_{x+t} + B^{(h)}; \quad A^{(h)} \geq 1,\; B^{(h)} \geq 0 \tag{10}$$

where h denotes the substandard category. Simplified implementations are as follows:

$$\mu_{x+t}^{(h)} = A^{(h)}\,\mu_{x+t} \tag{11}$$

$$\mu_{x+t}^{(h)} = \mu_{x+t} + B^{(h)} \tag{12}$$

respectively called multiplicative model and additive model. If μ_{x+t} is (as usual) an increasing function of the attained age $x+t$, then Eq. (11) expresses an increasing extra-mortality, while a constant extra-mortality is expressed by Eq. (12).

Mortality of substandard lives can also be represented via the age-shift model, frequently used in the life insurance practice:

$$\mu_{x+t}^{(h)} = \mu_{x+s^{(h)}+t} \tag{13}$$

where the ageing parameter $s^{(h)} > 0$ summarizes the impact of observable risk factors via the substandard category h. If the standard mortality follows the Gompertz law, i.e. $\mu_{x+t} = \alpha\,\mathrm{e}^{\beta\,(x+t)}$, then the mortality pattern expressed by Eq. (13) coincides with that expressed by the multiplicative model (11) with $A^{(h)} = \mathrm{e}^{\beta\,s^{(h)}}$.

2. The use of the above models calls for criteria to assess the parameter values. In particular, referring to the multiplicative model (11), the impact of the relevant risk factors can be quantified via the *Factor formula* of the "Numerical rating system", proposed by Rogers and Hunter (1919) (see also Hunter (1917)), and originally adopted by the New York Life Insurance in 1919. A set of r risk (and rating) factors is referred to, and the individual specific mortality is then expressed as follows:

$$q_{x+t}^{[\text{spec}]} = q_{x+t}\Big(1 + \sum_{j=1}^{r} \rho^{(j)}\Big) \tag{14}$$

where q_{x+t} represents the standard mortality pattern, and, of course:

$$-1 < \sum_{j=1}^{r} \rho^{(j)} < \frac{1}{q_{x+t}} - 1$$

We note that formula (14) constitutes a particular implementation of (8), in terms of q instead of μ and, of course, disregarding the impact

z of unobservable risk factors.
Each parameter $\rho^{(j)}$ in (14) can take either a positive value (then expressing a "debit") or a negative value ("credit"). The total effect of the parameters leads to a higher or lower death probability for the individual with respect to the standard probability q_{x+t}. This way, formula (14) can also represent the mortality patterns of risks better than the standard ones, i.e. the so-called *preferred risks* (see, for example: Hughes (2012), Munich Re (1999), and Werth (1995)). Hence, the more detailed classification follows:

- preferred risks:
- standard risks;
- sub-standard risks.

For more details, see: Cummins et al. (1983); an alternative criterion for risk classification is discussed by Ingle (2013).

3. The mortality pattern of annuitants, who purchased a standard life annuity and are hence supposed in very good health conditions, can be expressed, via a multiplicative adjustment, as follows:

$$\mu_{x,t}^{[\mathrm{ann}]} = A_{x,t}\, \mu_{x+t} \tag{15}$$

where:

μ_{x+t} denotes the (projected) population mortality, or the mortality of pensioners who are members of an occupational pension plan;

x is the age at policy issue, while $x+t$ is the current age;

$A_{x,t}$ ($0 < A_{x,t} \leq 1$) is an increasing function of the policy past duration t, expressing a self-selection whose impact on the annuitant's mortality is assumed strong at policy issue and then decreasing.

4. *Mortality of disabled people* is a key input item in actuarial calculations for many life and health insurance products, in particular disability annuities providing income protection and Long-term Care (LTC) insurance products paying annuities to the elderly who need assistance. See, for example: Pitacco (2014).
According to statistical evidence, the extra-mortality of a disabled individual has a peak immediately after the disability inception, then

decreases. Following, for example, the approach proposed by Venter et al. (1991), we set:

$$q^{[\text{disab}]}_{[x]+t} = a + (q_{x+t})^b\, f(t) \tag{16}$$

where x is the age at disability inception, $q^{[\text{disab}]}_{[x]+t}$ is the inception-select probability of death, q_{x+t} is the standard probability of death, and $f(t)$ ($f(t) \geq 1$) is a definitely decreasing function of t. Specific statistical data concerning US disabled workers suggested the simpler model:

$$q^{[\text{disab}]}_{[x]+t} = q_{x+t}\, f(t) \tag{17}$$

which provides an example of simplified implementation of the model (9), in terms of the probabilities q.

5. A formula with additive extra-mortality, applied to the q's (instead of the μ), has been proposed by Rickayzen and Walsh (2002) for modeling the *mortality of LTC-disabled people*; see also Rickayzen (2007), and the sensitivity analysis in Pitacco (2016b).
   The mortality pattern is expressed as follows:

$$q^{(h)}_{x+t} = q_{x+t} + \Delta(x+t;\alpha,h) \tag{18}$$

   with:

$$\Delta(x+t;\alpha,h) = \frac{\alpha}{1+1.1^{50-(x+t)}}\,\frac{\max\{h-5,0\}}{5} \tag{19}$$

   where:

   - the parameter h expresses the LTC severity category summarizing the risk factors $\rho^{(j)}_{x,t}$, $j = 1,\dots,r$, according to the UK OPCS scale; in particular:

     $0 \leq h \leq 5$ denotes less severe LTC states, with no significant impact on mortality;

     $6 \leq h \leq 10$ denotes more severe LTC states, implying an extra-mortality;

   - the parameter α is chosen according to the type of the standard mortality q (e.g. population mortality versus insured lives mortality).

   Also Eq. (18) provides an example of simplified implementation of the general model (9).

4.3.2 Unobservable risk factors (disregarding observable risk factors)

Assume, for simplicity of notation, that μ (or q, etc.) refers to a group of individuals which is homogeneous in respect of the observable risk factors.

Then, disregarding the past-duration effect on the standard force of mortality μ, Eq. (8) reduces to:

$$\mu_{x,t}^{[\text{spec}]} = \Phi\big[\mu_{x+t}; z_{x,t}\big] \tag{20}$$

Eq. (20) encompasses a number of models. In particular, two basic assumptions can be adopted in relation to the impact of unobservable risk factors on individual mortality, the choice depending, to a large extent, on what individual characteristics are considered as causes of heterogeneity.

The assumption $z_{x,t} = z_x$ (where x denotes a given age, e.g. $x = 0$) expresses an impact independent of the attained age. For example, we can set:

$$\mu_{x+t}^{[\text{spec}]} = \mu_{x+t}\, z_x \tag{21}$$

where issue-select and inception-select effects are disregarded. Fixed-frailty models implement this assumption (see Sect. 5.2).

Conversely, if we assume $z_{x,t} = z_{x+t}$, then the impact of unobservable heterogeneity depends on the attained age $x + t$ and hence can vary over the individual lifetime. Variable-frailty models implement this assumption (see Sect. 5.3).

Further, the variable $z_{x,t}$ expressing the impact can either be real-valued or can take a finite set of values. The above possible settings are sketched in Fig. 2; several examples are provided in the following Sections.

5 Seminal contributions

A survey on seminal contributions to the modeling of observable and unobservable heterogeneity is provided in this Section.

5.1 Discrete approaches to heterogeneity and frailty

5.1.1 One-year mortality: a range of settings

A population of Bernoulli risks is addressed in an interesting paper by A. H. Pollard (1970). For each risk, a given event (e.g. the individual's death) may either occur or not in a one-year period. All individuals are assumed aged x; q_x generically denotes the probability of the event.

According to Pollard (1970): *"The population value of q_x may be considered as the weighted sum of the rates of mortality of groups of persons suffering from particular disabilities. If there is any variation in the proportion suffering, for example, from particular heart conditions or if there is any variation in the degree of such impairments then variations in the population value of q_x must be expected in addition to the random variations which occur in the observed rate of mortality when q_x is constant."*

Two aspects in particular emerge: (1) the population can be split into (homogeneous) groups; (2) (possible) "uncertain" features are allowed for, in particular concerning the (relative) size of the groups and the impact of risk factors on the event probability within some groups.

Various settings are considered by Pollard (1970):

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

2. The population is split into m groups of independent risks, each one with given size $n^{(i)}$ and given probability $q_x^{(i)}$. This setting expresses a situation of "known" 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 a random q_x, with given expected value and variance, affects all the individuals; the risks are assumed independent conditionally on any possible outcome of q_x. The variance of the total number of events is higher than in case 1.

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

5. The population is split into m groups of independent risks, each group with random size $n^{(i)}$ but given probability $q_x^{(i)}$. This setting expresses a situation of "unknown" heterogeneity. The variance of the total number of events, compared to case 1, is lowered thanks to the splitting into groups but increased because their random sizes.

6. The population is split into m groups, each one with random size $n^{(i)}$ and random probability $q_x^{(i)}$. Also this setting expresses a situation of "unknown" heterogeneity. Cases 1 to 5 can be recognized as particular implementations of this general setting.

We note that all the above cases can be traced back to the scheme defined by Eq. (7), referred, for example, to one-year mortality.

The above model can be generalized. In particular, distributional hypotheses (not considered by Pollard (1970)) can be assumed for the unknown q_x and the random sizes $n^{(i)}$. Further, frailty assumptions can be adopted for the individual probability of death. See Pitacco (2018).

5.1.2 Other discrete models

According to Redington (1969), a heterogeneous population can be split into a given number of homogeneous groups (i.e. subpopulations). The mortality in each group is described by a Gompertz law with group-specific parameters. The distribution of the two Gompertz parameters is assumed symmetric. Then, the average force of mortality in the population can be calculated and compared to the "central" force of mortality obtained by assigning to the parameters their modal values. Because of heterogeneity, the average force of mortality is expected to be lower than the central one, especially when old ages are addressed. A similar result is obtained in a continuous model, as we will see in Sect. 5.2.1.

The approach proposed by Keyfitz and Littman (1979) leads to a very simplified model, anyhow valuable because it marks some significant features of heterogeneity in mortality. A heterogeneous population can be split into a given number of homogeneous groups but the relative sizes of the groups are unknown because of unobservable heterogeneity. The paper in particular focusses on the impact of heterogeneity on the average expected lifetime, concluding that, if the presence of heterogeneity is disregarded, an underestimation of the average expected lifetime follows. This result constitutes an important issue in the management of life annuity portfolios and pension plans.

5.2 Unobservable heterogeneity: fixed-frailty models

5.2.1 The basic fixed-frailty model

The fixed-frailty approach was first proposed by Beard (1959), but formally defined by Vaupel et al. (1979).

Assume that:

- the heterogeneity due to unobservable risk factors is expressed by the individual *frailty*;
- the individual frailty (an unknown positive real number) remains constant over the whole life span.

Let denote by $\mu_y(z)$ the (conditional) force of mortality of a person current age y, with a generic frailty level z ($z > 0$). Further, let define the standard force of mortality μ_y, in a given cohort, as the force of mortality conditional on a given frailty level z^*. Assuming, as usual, $z^* = 1$ we have:

$$\mu_y = \mu_y(1) \tag{22}$$

Let denote by $g_y(z)$ the probability density function (pdf) of the frailty distribution in the cohort at age y. Then the average force of mortality in the cohort is given by:

$$\bar{\mu}_y = \int_0^{+\infty} \mu_y(z)\, g_y(z)\, \mathrm{d}z \tag{23}$$

Specific models and results rely on:

1. the model for the standard force of mortality μ_y;
2. the relation between $\mu_y(z)$ and μ_y;
3. the pdf of the frailty distribution at a given age x, e.g. $x = 0$: $g_0(z)$.

In particular, combining:

1. the Gompertz law for the individual standard force of mortality, that is: $\mu_y = \alpha\, \mathrm{e}^{\beta\, y}$,
2. the multiplicative model for the force of mortality:
$$\mu_y(z) = z\, \mu_y \tag{24}$$
3. the Gamma distribution of the frailty, e.g. at age $x = 0$, with parameters δ, θ,

we obtain:

$$\bar{\mu}_y = \frac{\alpha'\, \mathrm{e}^{\beta y}}{\delta'\, \mathrm{e}^{\beta y} + 1} \tag{25}$$

We note what follows.

- Eq. (24) expresses the so called *proportional frailty*, and constitutes a simple implementation of model (20), with $x + t = y$ and $z_{x,t} = \text{const}$.
- The above model, usually named the *Gompertz - Gamma model*, leads to the Beard law (25) (see also Eq. (3)), that is, a particular case of the first Perks law (see Eq. (1)), with parameters α', δ' depending on the parameters δ, θ of the frailty distribution.

- The Beard law belongs to the logistic class, and hence implies a deceleration in the cohort mortality (see Sect. 3.2).

Hence, deviations of the individual mortality from the cohort mortality is a straightforward consequence of the assumed heterogeneity inside the cohort.

As noted by Vaupel and Yashin (1985), *"... the deviation of individual death rates from population rates implies some surprising and intriguing results"*, among which: *"Death rates for individuals increase more rapidly than the observed death rate for cohort"*.

For a formal presentation of the above results, see, for example: Haberman and Olivieri (2014), Pitacco et al. (2009).

5.2.2 An intuitive interpretation of the logistic shape

Assume that a cohort only consists of low-frailty individuals; mortality observation then leads to estimate the force of mortality $\mu_y(z_1)$, and to determine the maximum attained age $\omega(z_1)$, where z_1 represents the (hypothetical) frailty level.

Assume, conversely, that a cohort only consists of medium-frailty (high-frailty) individuals; mortality observation then leads to estimate the force of mortality $\mu_y(z_2)$ ($\mu_y(z_3)$), and to determine the maximum attained age $\omega(z_2)$ ($\omega(z_3)$), where z_2 (z_3) represents the (hypothetical) frailty level inside the cohort. The results of the above hypothetical observations are sketched in Fig. 4.

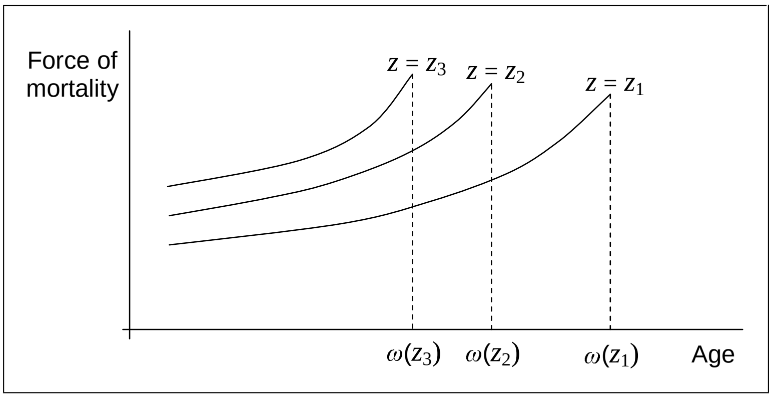

Figure 4: A set of forces of mortality depending on the parameter z

Real mortality observations address, of course, multi-frailty cohorts. Then, an "average" force of mortality can only be estimated. The average force of mortality, i.e. the cohort force of mortality, progressively moves towards

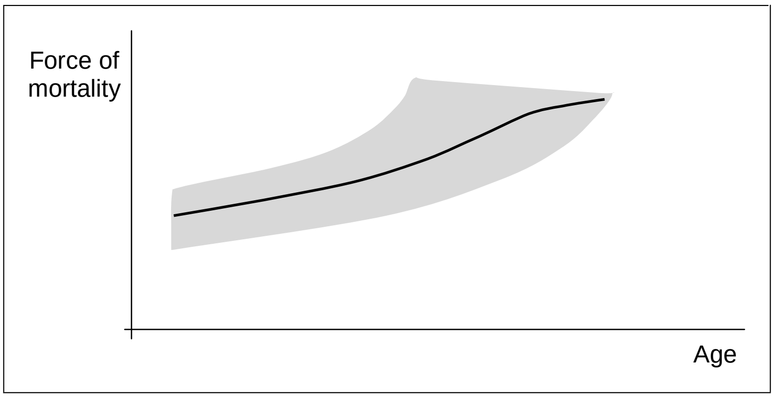

Figure 5: The average force of mortality in the cohort

low-frailty individual forces as the average frailty level decreases, and, as a consequence, the increase in the average force of mortality slows down. See Fig. 5, in which the idea of decelerating force of mortality is sketched; a numerical example is conversely provided by Fig. 6.

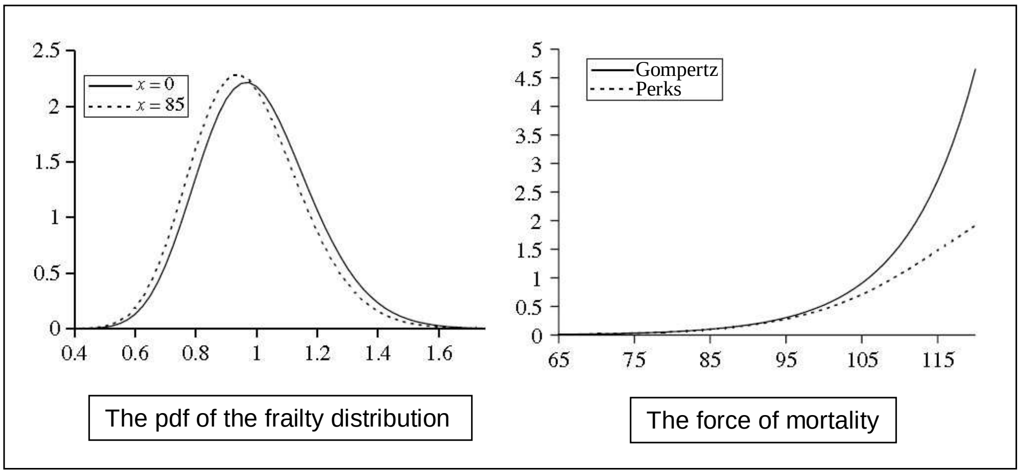

Figure 6: The Gompertz-Gamma model (Source: Pitacco et al. (2009))

5.2.3 Implementing and generalizing the basic model

A number of contributions followed the seminal proposals described in Sect. 5.2.1. See, in particular, Hougaard (1984, 1986), Manton et al. (1986), Steinsaltz and Wachter (2006), Yashin et al. (1985) and Yashin and Iachine (1997).

A compact review is provided by Haberman and Olivieri (2014). It is worth noting that several generalizations have been proposed; for example:

- the force of mortality expressed by the Makeham law (Beard (1959)), or the Weibull law (Manton et al. (1986));
- the frailty distribution given by the inverse Gaussian (Hougaard (1984, 1986), Manton et al. (1986), Butt and Haberman (2004, 2002)), the shifted Gamma distribution (Martinelle (1987)), the generalized Gamma distribution (which includes, as particular cases, the lognormal and the Weibull distributions; see e.g. Balakrishnan and Peng (2006)).

For a more general framework, the reader should refer to: Duchateau and Janssen (2008) and Wienke (2003).

Among the most recent implementations of the Beard law, see Dodd et al. (2018); see also the references therein.

5.3 Unobservable heterogeneity: variable individual frailty

As noted in Sect. 1, individual frailty can be interpreted in several ways, according to the causes which are considered as originating the frailty itself. It follows that the individual frailty can be assumed either constant or variable (in particular age-dependent) throughout the lifetime. The modeling framework proposed by Beard (1959) and Vaupel et al. (1979) (see Sect.5.2) relies on the assumption of constant individual frailty.

An influential contribution to the modeling of age-dependent frailty was provided by Hervé Le Bras (1976). The contribution can be placed in the framework of mortality modeling aiming to define the limit age. According to the proposed approach, a mortality law must be the result of assumptions on the structure of a process describing the evolution of individual mortality throughout the whole life. The following basic assumptions were considered by Le Bras (1976):

1. each individual has an "initial frailty" (denoted as *"faute"*);
2. the "transition" probabilities, i.e. the probabilities of increase in frailty (*"nouvelle faute"*) and the probability of death, are proportional to the current frailty level.

The resulting mortality law approximately follows the Gompertz age-pattern up to some age, then tending to a limit, and hence resulting in a logistic-like shape.

We note that:

- according to assumption 2, the individual frailty is modeled as a Markov process;
- an implementation of model (20) can be recognized, in terms of annual probabilities q (instead of the force of mortality μ).

6 Some recent contributions

A strengthened interest in heterogeneity modeling has recently emerged both in demography and in actuarial science as well. Heterogeneity due to unobservable risk factors has been invoked as a possible cause of mortality deceleration at high ages (although the relation between heterogeneity and mortality deceleration constitutes a controversial issue; see for example Pitacco (2016a), and references therein).

In the actuarial technique, the assessment of the portfolio risk profile in presence of heterogeneity calls for appropriate stochastic models (see Sects. 7.1 to 7.3). Further, an appropriate modeling of observable heterogeneity underpins the design and pricing of underwritten (or "special-rate") annuity products (see Sect. 7.4).

6.1 The Markov framework

The Markov modeling framework, originally proposed by Le Bras (1976), has been adopted by many Authors. Interesting applications in the actuarial field will be cited in Sect. 7. Here we only refer to two papers based on the ideal of *physiological age*.

In Lin and Liu (2007) a finite-state Markov process is adopted to model human mortality. The individual health status is represented by the physiological age, and modeled by the Markov process, each state of the process representing a possible outcome of the physiological age. The random time of death then follows a phase-type distribution. Hence, the frailty is measured by the physiological age, and is distributed according to the distribution of individuals among age classes.

The model proposed by Lin and Liu (2007) has been generalized by Liu and Lin (2012). Uncertainty in mortality has been introduced, and the generalization results in a subordinated Markov model.

6.2 Splitting a heterogeneous population into (homogeneous) subpopulations

Avraam et al. (2013) note that deviations from the Gompertz age-pattern of mortality can be observed in any population, and these deviations are particularly significant at very low ages and at highest ages. How to explain deviations from the Gompertz pattern? Heterogeneity with respect to mortality is considered as an important cause of deviations. The Authors' proposal then consists in splitting a generic population into homogeneous groups, i.e. subpopulations, and assessing the mortality of each group by adopting a Gompertz law with specific parameters. Referring to mortality data of the Swedish and US populations, deviations from the Gompertz pattern are explained by both heterogeneity and stochastic effects, the latter having an important impact particularly when the number of deaths is small, that is in early- and late-life age intervals.

We note that the modeling framework proposed by Avraam et al. (2013) can be traced back to the general structure defined by Eq. (7).

Generalizations of the previous model have been proposed by Avraam et al. (2014, 2016), by allowing for evolution of the parameters over time, so that a mortality trend can be represented.

6.3 From Gompertz to Makeham: a frailty-based interpretation

As is well known, Makeham (1867) generalized the Gompertz law by adding a constant term, that is, a term independent of the attained age. Makeham's generalization can be interpreted in terms of a "shock" model (see, for example, Doray (2008)). Assume that: (a) the random lifetime T of a person has a Gompertz distribution, (b) the random time Y to a fatal accident has an exponential distribution, and (c) T and Y are independent; then, the random variable $\min\{T, Y\}$ has a Makeham distribution.

An interesting alternative interpretation of Makeham's generalization has been provided by Lindholm (2017), which can be summarized as follows. Assume that all the individuals in a cohort follow a common baseline force of mortality $\mu_y = \mathrm{e}^{\beta\, y}$, that is, a particular case of the Gompertz law, combined with an individual proportional frailty Z, which follows a specific translated gamma distribution. Then, the force of mortality in the cohort is given by the Makeham law, $\mu_y = \gamma + \alpha\, \mathrm{e}^{\beta\, y}$, whose parameters depend on the parameters of the translated gamma distribution.

7 Heterogeneity in actuarial evaluations

The actuarial models for all life insurance products (and, more generally, for all products in the area of the insurances of the person, that is, including e.g. health insurance products) call for assumptions about the insureds' or annuitants' mortality. Hence, both observable and unobservable heterogeneity should be taken into account.

The assessment of observable heterogeneity belongs to the actuarial tradition, as far as pricing and reserving for insurance products providing death benefits (in particular, term assurances) are concerned. Simple formulae are commonly adopted, for example, to express extra-mortality of substandard lives or disabled insureds (see Sect. 4.3.1).

Of course, the longer the duration of the insurance contract, the higher the impact of mortality assumptions may be. It follows that particular attention should be placed on these assumptions when implementing actuarial models for life annuities, and specifically when assessing the risk profile of an annuity portfolio.

Allowing for observable risk factors when pricing life annuities is a recent issue in the life insurance context. The premiums for underwritten (or special-rate) life annuities are actually based on the applicant's health status, assessed via observable risk (and rating) factors. Conversely, analyzing the impact of unobservable heterogeneity constitutes a lively topic in the current research work.

In what follows, we focus on mortality heterogeneity in life annuities. Table 1 summarizes the main topics addressed in the next sections, which, from a Risk Management perspective, can be interpreted as "actions" in the management of a life annuity portfolio.

Table 1: RM Objectives & Actions

Objectives	**Actions**
Profit, value creation Market share	Product design, Pricing
Solvency	Capital allocation, Reinsurance, ART

7.1 Pricing and reserving

Does disregarding (unobservable) heterogeneity in a life annuity portfolio leads to wrong pricing and reserving? This problem has been attacked by Olivieri (2006). A portfolio of life annuities is referred to; all the annuitants are assumed initially aged $x = 65$; the portfolio is closed to new entrants, and death is the only cause of decrement. The same annual benefit b is paid to all the annuitants. Mortality in the portfolio is alternatively described by the Gompertz law and the Beard law (that is, according to the Gompertz-Gamma model, a fixed individual frailty is assumed; see Sect. 5.2.1). The Gompertz parameters are assessed so to represent a best-estimate pattern of mortality in a life annuity portfolio. According to the main achievements, disregarding heterogeneity in the portfolio (that is, assuming the best-estimate Gompertz mortality pattern) leads to:

- underestimation of the actuarial values and hence, in particular, of premiums and policy reserves;
- underestimation of the (relative) riskiness in the portfolio (expressed by the coefficient of variation of the net present value of benefits), and hence underestimation of the adequacy requirements, in terms of risk margin and/or solvency capital.

A portfolio of life annuities is also referred to by Su and Sherris (2012). Various mortality assumptions have been considered. In particular, the following heterogeneity models have alternatively been adopted:

- fixed individual frailty, Gamma distributed or inverse Gaussian distributed;
- variable frailty, according to the Markov ageing model proposed by Lin and Liu (2007).

The Authors note that life annuity rates are usually stated assuming a significant annuitants' self-selection, with a negative impact on the annuitization propensity. The above models are used to illustrate the financial impact of heterogeneity on the probability distribution of life annuity values. Both the models for heterogeneity have implications for annuity markets, and the results show clearly that heterogeneity needs to be taken into account if a viable life annuity market is to be developed.

7.2 Capital allocation

What is the appropriate capital allocation in face of a heterogeneous life annuity portfolio? Solvency requirements, and in particular capital allocation, should allow for all the risk components which affect the results (cash flows, profits, etc.) of a life annuity portfolio.

In Sherris and Zhou (2014) all the biometric risk components are taken into account:

1. the idiosyncratic longevity risk, that is, the risk of random fluctuations around the relevant expectations;

2. the aggregate longevity risk, that is the risk of systematic deviations from the relevant expectations;

3. the risk originated by unobservable mortality heterogeneity among the annuitants.

As is well known, the aggregate longevity risk cannot be diversified via pooling, while the idiosyncratic risk can be diversified thanks to appropriate pool sizes. However, the unobservable heterogeneity weakens the diversification of the idiosyncratic risk, and hence should not be disregarded when assessing the portfolio risk profile and the related solvency requirements. Sherris and Zhou (2014) represent mortality heterogeneity using both the fixed frailty and the variable frailty model, and show that, when a mortality model also includes systematic risk, a larger pool size results in a heavier tail risk, and then in a higher capital requirement. This effect is not captured by standard models of heterogeneity.

7.3 ART

Alternative Risk Transfers (swaps, mortality-linked securities, etc.) must be implemented in order to hedge biometric risks which are non-diversifiable via pooling. As specifically regards the longevity risk, longevity bonds transfer this risk from annuity providers to capital markets. In order to evaluate and price mortality-linked securities, appropriate mortality models are required, allowing for all the risk components.

Liu and Lin (2012) adopt a subordinated Markov model for modeling stochastic mortality. In particular, the aging process of an individual is assumed to follow a finite-state Markov process, while the stochasticity of mortality is governed by a subordinating gamma process. Finally, a valuation framework for pricing longevity bonds is proposed, based on the above model

7.4 Product design: the underwritten annuities

Can mortality heterogeneity in a population of potential annuitants suggest rating procedures in order to enlarge life annuity portfolios? A positive response relies on the possibility of implementing annuity products which, unlike the standard immediate annuities, can also attract people in non-optimal health conditions.

To this purpose, *underwritten annuities* (or *special-rate annuities*, or *substandard annuities*) have been defined, i.e. life annuities with lower premiums for individuals in non-optimal health conditions (see, for example, Pitacco (2017) and references therein). An underwriting procedure is required in order to assess the applicant's health status.

As noted by Meyricke and Sherris (2013), standard life annuities are priced assuming above-average longevity (as regards the related mortality assumption, see point 3 in Sect. 4.3.1), whereas underwritten annuity rates can reflect individual risk factors based on underwriting results. Of course, mortality risk still varies within each rating class due to unobservable individual risk factors, that is, because of individual frailty. Meyricke and Sherris (2013) quantify the impact of heterogeneity due to underwriting factors and frailty on the values of both standard and underwritten annuities, and propose a method to adjust annuity rates in order to allow for frailty.

Classification systems for underwritten annuities are addressed by Gatzert et al. (2012). The Authors propose a theoretical model to determine the optimal classification system, aiming at the maximization of the insurer's profits. In the optimization model, the following features are accounted for:

- the interactions between annuity price and demand;
- the classification costs;
- the underwriting risk, expressed by the cost of "insufficient" risk assessment, which can result in assigning applicants to the "wrong" rating classes.

Olivieri and Pitacco (2016) refer to a portfolio consisting of both standard annuities and underwritten annuities. Heterogeneity is expressed by the individual fixed frailty model. The underwriting process aims at providing, for each applicant, an estimate of his/her frailty level. Looking at the effects of portfolio enlargement, on the one hand, the larger portfolio size contributes to lower the variance in portfolio results (as regards the idiosyncratic risk, i.e. the risk of random fluctuations), on the other heterogeneity in the combined portfolio contributes to raise the variance in portfolio results.

Besides the heterogeneity among subportfolios, actually some degree of residual heterogeneity affects each subportfolio, because of residual unobservable risk factors: indeed, the underwriting process only provides a proxy for the health status assessment. What about the "balance" between improvement and worsening of the portfolio risk profile? Numerical examples show that an appropriate definition of the rating classes can provide a net improvement of the portfolio risk profile.

8 Concluding remarks

Heterogeneity in mortality is due to both observable and unobservable risk factors. The awareness of heterogeneity due to observable risk factors can be traced back to the mid of the Eighteenth century, as noted in Sect. 3.1. Conversely, the effect of unobservable risk factors has later been taken into account: early models for unobservable heterogeneity indeed date back to the mid of the Twentieth century (see Sects. 3.2 and 3.3).

Different modeling approaches are required when addressing observable or unobservable risk factors respectively (although a formal representation of the joint impact of all the factors is in principle conceivable, as seen in Sect. 4.2). It follows that, usually, practicable models only focus on either observable or unobservable factors.

Nevertheless, significant efforts have recently been devoted to analyze the combined effect of observable and unobservable risk factors on the mortality pattern. The term "inequality" has commonly been used, especially in the socio-economic framework, to encompass the set of causes which can be considered as determinants of mortality (income, education, access to health care, home and living environment, social connections, etc.).

Although the topic of inequality is outside the scope of this survey, some bibliographic references can help the reader in understanding the main aspects of this field of research. Interesting results of a statistical analysis at an international level are presented and discussed by Rodgers (1979). Among more recent contributions, the reader can refer, for example, to Daly and Wilson (2013) and Deaton and Paxson (2001) (focussing on dynamic aspects of income inequalities and mortality patterns), and Scharf and Shaw (2017) (addressing inequalities at high ages). All the cited papers also provide extensive lists of references.

Back to actuarial implications of heterogeneity, specifically due to unobservable risk factors, we stress that special attention should be devoted to the risk profile of heterogeneous portfolios, particularly when life annuities are concerned. Indeed, assessing and hedging biometric risks (and tail risk in particular) constitutes a complex setting in which all the risk components (heterogeneity included) should be accounted for.

References

Remark

> *Where links are provided, they were active as of the time this paper was completed but may have been updated since then.*

Ainslie, R. (2000). Annuity and insurance products for impaired lives. Working Paper. Presented to the Staple Inn Actuarial Society.

Avraam, D., Arnold (-Gaille), S., Jones, D., and Vasiev, B. (2014). Time-evolution of age-dependent mortality patterns in mathematical model of heterogeneous human population. *Experimental Gerontology*, 60:18–30.

Avraam, D., Arnold, S., Vasieva, O., and Vasiev, B. (2016). On the heterogeneity of human populations as reflected by mortality dynamics. *Aging*, 8(11):3045–3064.

Avraam, D., de Magalhaes, J. P., and Vasiev, B. (2013). A mathematical model of mortality dynamics across the lifespan combining heterogeneity and stochastic effects. *Experimental Gerontology*, 48:801–811.

Balakrishnan, N. and Peng, Y. (2006). Generalized gamma frailty model. *Statistics in Medicine*, 25(16):2797–2816.

Beard, R. E. (1959). Note on some mathematical mortality models. In Wolstenholme, C. E. W. and Connor, M. O., editors, *CIBA Foundation Colloquia on Ageing*, volume 5, pages 302–311, Boston.

Butt, Z. and Haberman, S. (2002). Application of frailty-based mortality models to insurance data. Actuarial Research Paper No. 142, Faculty of Actuarial Science & Insurance, City University, London. Available at: `http://openaccess.city.ac.uk/2279/1/142-ARC.pdf`.

Butt, Z. and Haberman, S. (2004). Application of frailty-based mortality models using generalized linear models. *ASTIN Bulletin*, 34(1):175–197.

Corbaux, F. (1833). On the Natural and Mathematical Laws concerning Population, Vitality, and Mortality. (Chaps. XXV, XXVI). London. Available at: `https://ia802703.us.archive.org/32/items/b21955207/b21955207.pdf`.

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

Daly, M. C. and Wilson, D. J. (2013). Inequality and mortality: New evidence from U.S. County Panel Data. Working Paper No. 2013-13, Federal Reserve Bank of San Francisco. Available at: https://www.frbsf.org/economic-research/files/wp2013-13.pdf.

de Ferra, C. (1954). Tavola di mortalità per una popolazione inomogenea. *Giornale dell'Istituto Italiano degli Attuari*, 17:62–76.

de Finetti, B. and Taucer, R. (1952). Sur l'emploi d'une base technique générale pour le risques tares. In *Conférence Internationale des Risques Aggravés*, pages 86 – 104, Stockholm.

Deaton, A. and Paxson, C. (2001). Mortality, income, and income inequality over time in Britain and the United States. NBER Working Paper No. 8534, National Bureau of Economic Research, Cambridge, MA. Available at: http://www.nber.org/papers/w8534.pdf.

Dodd, E., Forster, J. J., Bijak, J., and Smith, P. W. F. (2018). Smoothing mortality data: the English Life Tables, 2010-2012. *Journal of the Royal Statistical Society - Series A*, 181(Part 3):717–735.

Doray, L. G. (2008). Inference for logistic-type models for the force of mortality. Presented at the Living to 100 and Beyond Symposium. Orlando, FL. Available at: https://www.soa.org/library/monographs/retirement-systems/living-to-100-and-beyond/2008/january/subject-toc.aspx.

Duchateau, L. and Janssen, P. (2008). *The Frailty Model*. Springer.

Gatzert, N., Schmitt-Hoermann, G., and Schmeiser, H. (2012). Optimal risk classification with an application to substandard annuities. *North American Actuarial Journal*, 16(4):462–486.

Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of Royal Society (Series A)*, 115:513–583.

Haberman, S. (1996). Landmarks in the history of actuarial science (up to 1919). Actuarial Research Paper No. 84, Faculty of Actuarial Science & Insurance, City University, London. Available at: http://openaccess.city.ac.uk/2228/1/84-ARC.pdf.

Haberman, S. and Olivieri, A. (2014). Risk Classification / Life. In *Wiley StatsRef: Statistics Reference Online*. Wiley.

Hougaard, P. (1984). Life table methods for heterogeneous populations: distributions describing heterogeneity. *Biometrika*, 71(1):75–83.

Hougaard, P. (1986). Survival models for heterogeneous populations derived from stable distributions. *Biometrika*, 73(2):387–396.

Hughes, M. (2012). Preferred risk in life insurance. SCOR inFORM. Available at: `https://www.scor.com/en/files/preferred-risk-life-insurance`.

Hunter, A. (1917). Insurance on sub-standard lives. *The Annals of the American Academy of Political and Social Sciences*, 70:38–53.

Ingle, D. (2013). Implications for use of credits applied to preferred knock-out criteria. *On the Risk*, 29(4):42–52.

Kannisto, V. (1994). *Development of Oldest-Old Mortality, 1950-1990: Evidence from 28 Developed Countries*. Odense Monographs on Population Aging, 1. Odense University Press. Available at: `http://www.demogr.mpg.de/Papers/Books/Monograph1/OldestOld.htm`.

Keyfitz, N. and Littman, G. (1979). Mortality in a heterogeneous population. *Population Studies*, 33:333–342.

Le Bras, H. (1976). Lois de mortalité et age limite. *Population*, 31(3):655–692.

Levinson, L. H. (1959). A theory of mortality classes. *Transactions of the Society of Actuaries*, 11:46–87.

Lin, X. S. and Liu, X. (2007). Markov aging process and phase-type law of mortality. *North American Actuarial Journal*, 11:92–109.

Lindholm, M. (2017). A note on the connection between some classical mortality laws and proportional frailty. *Stastistics & Probability Letters*, 126:76–82.

Liu, X. and Lin, X. S. (2012). A subordinated Markov model for stochastic mortality. *European Actuarial Journal*, 2(1):105–127.

Makeham, W. M. (1867). On the law of mortality. *Journal of the Institute of Actuaries*, 8:301–310.

Manton, K. G., Stallard, E., and Vaupel, J. W. (1986). Alternative models for the heterogeneity of mortality risks among the aged. *Journal of the American Statistical Association*, 81:635–644.

Martinelle, S. (1987). *A generalized Perks formula for old-age mortality*. R & D report / Statistics Sweden. Statistiska Centralbyrån, Stockholm.

Meyricke, R. and Sherris, M. (2013). The determinants of mortality heterogeneity and implications for pricing annuities. *Insurance: Mathematics & Economics*, 53(2):379–387.

Munich Re (1999). The concept of preferred lives. Münchener Rückversicherungs - Gesellschaft.

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. (2015). *Introduction to Insurance Mathematics. Technical and Financial Features of Risk Transfers*. EAA Series. Springer, 2nd edition.

Olivieri, A. and Pitacco, E. (2016). Frailty and risk classification for life annuity portfolios. *Risks*, 4(4):39. Available at: `http://www.mdpi.com/2227-9091/4/4/39`.

Olshansky, S. J. (1998). On the biodemography of aging: A review essay. *Population and Development Review*, 24:381–393.

Olshansky, S. J. and Carnes, B. E. (1997). Ever since Gompertz. *Demography*, 34(1):1–15.

Perks, W. (1932). On some experiments in the graduation of mortality statistics. *Journal of the Institute of Actuaries*, 63:12–57.

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

Pitacco, E. (2016a). High age mortality and frailty. Some remarks and hints for actuarial modeling. CEPAR Working Paper 2016/19. Available at: `http://www.cepar.edu.au/working-papers/working-papers-2016.aspx`.

Pitacco, E. (2016b). Premiums for long-term care insurance packages: Sensitivity with respect to biometric assumptions. *Risks*, 4(1):1–22. Available at: http://www.mdpi.com/2227-9091/4/1/3.

Pitacco, E. (2017). Life annuities. Products, guarantees, basic actuarial models. CEPAR Working Paper 2017/6. Available at: http://cepar.edu.au/sites/default/files/Life_Annuities_Products _Guarantees_Basic_Actuarial_Models_Revised.pdf.

Pitacco, E. (2018). Can frailty models improve actuarial calculations in health insurance? Presented at the 31st International Congress of Actuaries, Berlin.

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

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

Redington, F. M. (1969). An exploration into patterns of mortality. *Journal of the Institute of Actuaries*, 95:243–298.

Rickayzen, B. D. (2007). An analysis of disability-linked annuities. Actuarial Research Paper No. 180, Faculty of Actuarial Science & Insurance, City University, London. Available at: http://openaccess.city.ac.uk/2312/1/180ARP.pdf.

Rickayzen, B. D. and Walsh, D. E. P. (2002). A multi-state model of disability for the United Kingdom: Implications for future need for Long-Term Care for the elderly. *British Actuarial Journal*, 8(36):341–393.

Rinke, C. R. (2002). The variability of life reflected in annuity products. Hannover Re's Perspectives - Current Topics of International Life Insurance. Issue No. 8.

Rodgers, G. B. (1979). Income and inequality as determinants of mortality: an international cross-section analysis. *Population Studies*, 33:343–351.

Rogers, O. and Hunter, A. (1919). Numerical rating. The numerical method of determining the value of risks for insurance. *Transactions of the Actuarial Society of America*, XX - Part 2(62):273–331.

Scharf, T. and Shaw, C. (2017). Inequalities in later life. Technical report, Centre for Ageing Better, London. Available at: https://www.ageing-better.org.uk/sites/default/files/2017-12/Inequalities 20scoping 20review 20full 20report.pdf.

Sherris, M. and Zhou, Q. (2014). Model risk, mortality heterogeneity, and implications for solvency and tail risk. In Mitchell, O. S., Maurer, R., and Hammond, P. B., editors, *Recreating Sustainable Retirement: Resilience, Solvency, and Tail Risk*. Oxford University Press.

Steinsaltz, D. R. and Wachter, K. W. (2006). Understanding mortality rate deceleration and heterogeneity. *Mathematical Population Studies*, 13(1):19–37.

Su, S. and Sherris, M. (2012). Heterogeneity of Australian population mortality and implications for a viable life annuity market. *Insurance: Mathematics & Economics*, 51(2):322–332.

Thatcher, A. R. (1999). The long-term pattern of adult mortality and the highest attained age. *Journal of the Royal Statistical Society - Series A*, 162:5–43.

Thatcher, A. R., Kannisto, V., and Vaupel, J. W. (1998). *The Force of Mortality at Ages 80 to 120*. Odense Monographs on Population Aging, 5. Odense University Press. Available at: http://www.demogr.mpg.de/Papers/Books/Monograph5/ForMort.htm.

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.

Vaupel, J. W. and Yashin, A. I. (1985). The deviant dynamics of death in heterogeneous populations. *Sociological Methodology*, 15:179–211.

Venter, G. G., Schill, B., and Barnett, J. (1991). Review of Report of Committee on Mortality for Disabled Lives. Casualty Actuarial Society. Arlington, Virginia. Available at: https://www.casact.org/pubs/forum/91wforum/91wf115.pdf.

Werth, M. (1995). Preferred lives. A more complete method of risk assessment. Working Paper. Presented to the Staple Inn Actuarial Society.

Wienke, A. (2003). Frailty models. MPIDR Working Paper WP 2003-032. Max Planck Institute for Demographic Research, Rostock, Germany. Available at: `http://www.demogr.mpg.de/papers/working/wp-2003-032.pdf`.

Yashin, A. I. and Iachine, I. A. (1997). How frailty models can be used for evaluating longevity limits: Taking advantage of an interdisciplinary approach. *Demography*, 34:31–48.

Yashin, A. I., Manton, K. G., and Vaupel, J. W. (1985). Mortality and ageing in a heterogeneous population: A stochastic process model with observed and unobserved variables. *Theoretical Population Biology*, 27(2):154–175.