Health insurance

Abdul H. Rahman; Dick H. Harryvan

bookboon
The eBook company

DICK H. HARRYVAN AND
ABDUL H. RAHMAN

HEALTH INSURANCE

Health insurance
1st edition

ISBN 978-87-403-2147-0

CONTENTS

ABOUT THE AUTHORS

Drs. Dick Harryvan is Vice Chairman of the Supervisory Board of NN Group in the Netherlands an insurance and asset management group operating in 18 countries. He is a former Executive Board Member of ING Group and CEO of ING Direct globally. After obtaining a Masters degree in business economics from Erasmus University, an international career in insurance and retail banking followed. For Nationale Nederlanden Group, he spent nine years in various functions in the US and Canadian business units. The next four years at the international division included amongst others responsibility for setting up new life insurance greenfields in Hungary and Czech Republic after the Wall fell. He then moved to ING Bank to develop retail banking internationally for ING Group. In a three-year incubator period, various concepts were tested with responsibility for the ING Direct pilot in Canada. After the initial success in Canada this concept was rolled out to nine countries growing to 24 million customers in the space of 10 years. Since retirement from the ING Executive Board in 2010 several supervisory board positions for ING Group subsidiaries followed. Other positions include membership of the Board of the Dutch automobile association ANWB and partner and investment committee member at Finch Capital private equity fund investing in fintech startups.

I dedicate this book my wife Corrie without whom my career would not have been possible.

Rotterdam, The Netherlands, 2018.

Dr. Professor Abdul H. Rahman holds academic degrees that include MSc. (Mathematics); M.A. (Economics); and Ph.D. (Financial Economics). Over his career, he held positions as Full Professor of international economics and finance and Telfer Teaching Fellow at the Telfer School of Management, University of Ottawa, Canada; Associate Dean at what is now the John Molson School of Business, Concordia University in Montreal, Canada; Academic Director and Founding Director of the Retail Banking Academy, London, United Kingdom and Chairperson of Mathematics and Physics, Dawson College, Montreal, Canada.

He has published in several prestigious refereed academic journals including the Journal of Banking and Finance, Journal of Asset Management, Journal of International Financial Markets, Institutions and Money, Journal of Financial and Quantitative Analysis, Journal of Futures Markets and Review of Economics and Statistics.

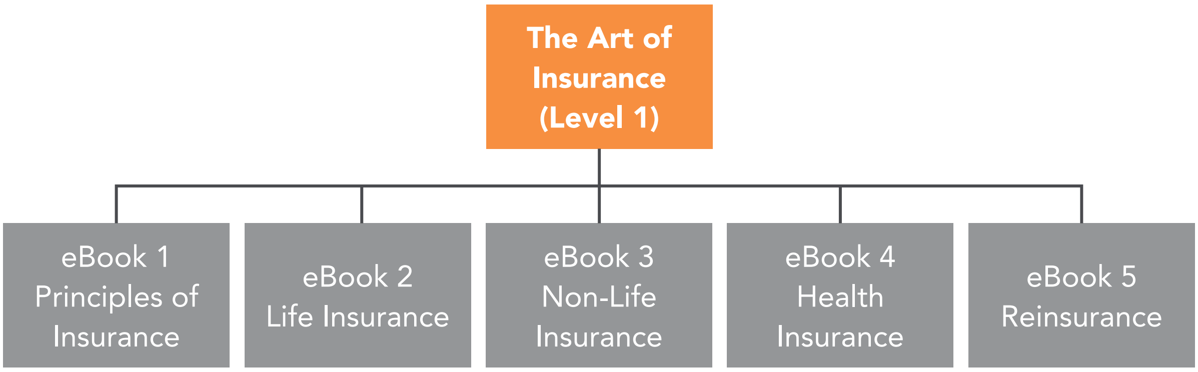
Dr. Rahman has consulted for global financial institutions in banking, asset management and insurance and currently facilitate high-level seminars for senior management professionals in retail and corporate banking, life and non-life insurance and institutional asset management.

I dedicate this eBook to my wife Ruth and daughters Sara and Lisa for their unfailing support of my professional pursuits.

Montreal, Canada, 2018.

PREFACE

The ART OF INSURANCE series is structured according to three levels – I, II and III each with increasing knowledge and understanding of insurance as a business. The architecture of The Art of Insurance (Level I) is schematically summarised as follows:



Each eBook in Level I is developed independently and is self-contained. Hence, the reader is not required to complete the five eBooks sequentially. Consequently, some information presented in the first eBook entitled the *Principles of Insurance* will be repeated so that there is a continuous flow in our presentation.

In addition, our approach focuses on the 'art' rather than the 'science' of insurance where we emphasise principles, concepts and intuition rather than mathematical proofs of complex theorems. While we recognise that the insurance business is founded on a high level of sophistication in probability and statistics, we avoid an overbearing level of jargon and a misplaced reliance on complex equations and formulae.

This eBook entitled the *Health Insurance* comprises six chapters.

Chapter 1 describes the fundamental model of health insurance called the three-party intermediation model and considers special cases where the third-party payer is a private insurer or a self-funded employer. Conventional health insurance products including sickness insurance, critical illness insurance as well as disability insurance are compared. We also describe the typical lifecycle of a health insurance claim wherein we define and explain common actuarial terminology such as waiting period, deferred period and link period.

Chapter 2 reviews deductibles and policy limits common in health insurance products (e.g., sickness insurance) that are similar to non-life techniques. We also discuss *stop-loss insurance* which is common in self-funded group health insurance.

The three main types of anti-selection in health insurance- external, internal and duration-related – are presented and linked to the combined ratio as well as a strategy of customer care. Finally, we show that *ex post* moral hazard and its risk of over-consumption is more important in health insurance compared to *ex ante* moral hazard which is common in non-life insurance (e.g., automobile insurance).

Health insurance is commonly combined with a generic life insurance contract. This facilitates modelling of health insurance that is similar to life techniques (i.e., Health SLT as named by Solvency II) within a multi-state framework. Chapter 3 considers this perspective and presents a procedure to calculate single premiums and annual level premiums based on the equivalence principle.

Chapter 4 considers a special case of the approach in chapter 3. This is called multi-decrement models and permits the use of multi-decrement life tables which consider the potential *causes* of death – for example, accidental death or death from cancer or death from other causes.

Chapter 5 considers sickness insurance as an example of health insurance *not* similar to life techniques (i.e., Health NSLT as named by Solvency II) within the collective risk model, common in actuarial science.

Chapter 6 is a list of references.

1 SETTING THE STAGE: THE THREE-PARTY INTERMEDIATION MODEL

1.1 INTRODUCTION

Consider the case of individuals who are faced with making savings decisions as part of their personal long-term financial planning. These decisions are facilitated by forecasts of future cash flows arising from employment and investment incomes over their expected remaining lifetimes. But in a world of uncertainty, there is a potential for unexpected interruptions in cash flows arising from random events that include an individual's premature death, partial or full disability which may be temporary or permanent, critical illness, outliving retirement resources and damage to his/her property.

The expected future evolution of an individual's *states of health* is an important factor in forecasting his/her long-term income potential. For example, a healthy person of age x years may become sick because of bacteria, viruses and similar pathogens that constantly invade the human body. The duration of sickness may be short-term resulting in unexpected medical expenses for the individual. Sickness may also be more chronic (i.e., long-term) leading to temporary or permanent disability or even death, thereby creating a potential for a substantial loss of employment income.

Disability is typically defined as a person having a physical or mental condition that limits their movements, senses, or activities. A more detailed discussion of stages of disability is provided later in this chapter. Clearly, the incidence of sickness can have unexpected long-term negative effects on the individual's expected future cash flow.

The second category of risk to an individual's long-term financial planning is the occurrence of injuries from accidents. Accidents are defined as unforeseen random events that may cause bodily harm with a potential loss of employment income and/or increase in medical expenses. For example, a healthy person of age x years may become sick, disabled or die from injuries sustained from an accident.

We present a formal definition of health insurance.

Olivieri and Pitacco (2010, page 440) states that "health insurance is the term used when the purpose is to compensate a person or his/her family for the economic consequences of an alteration of the health status originated by a sickness or an accident."

A similar definition provided by the Committee for European Insurance and Occupational Pension Supervisors (CEIOPS) places emphasis on sickness and accident as the causal factors of morbidity.

This definition highlights two points.

First, it shows that an individual's demand for health insurance is based primarily on an expected future demand for ***health care*** due to potential sickness and/or injuries from accidents. As a consequence, the *health care provider* is an important agent in a health insurance model.

While the probability of sickness is expected to increase at the individual gets older, accidents are random events and may occur at any age. However, injuries from accidents for individuals who are relatively older may be more severe and require more extensive medical treatments. This suggests that sickness and accidents may be correlated especially for older individuals. For example, an individual who suffered a coronary failure while driving a car may cause an accident. Conversely, a violent and sudden car accident may cause the driver to suffer coronary failure.

Second, the actual occurrence of sickness and/or accidents have similar but not necessarily, financial consequences for an individual. These include loss of employment income from disability or death and/or incurrence of medical and hospitalisation expenses.

As we discussed so far, the individual's demand for health insurance is based on an underlying need for health care provided by health care professionals. Hence, the market for health insurance is based on a three-party model comprising a third-party payer that intermediates between an individual or a select group of individuals (e.g., employees) and health care providers. This is called a *three-party intermediation model.*

1.2 THREE-PARTY INTERMEDIATION MODEL FOR HEALTH INSURANCE

Figure 1.1 illustrates a generic health insurance intermediation model. The demand for health care is created by individuals or a group (e.g., employees in a company) while health care providers (e.g., medical doctors, nurses, hospitals) create the supply of health care. The third-party payer receives insurance premiums from individuals or group of individuals and pays health care providers according to policy terms and conditions.

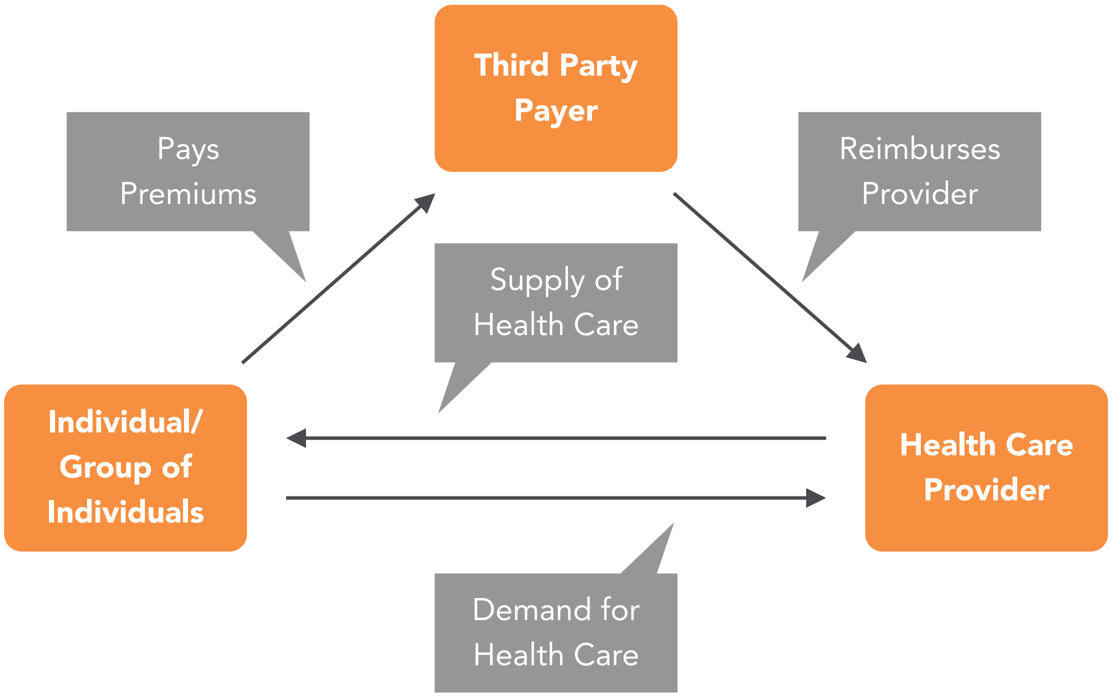


Figure 1.1: A Generic Intermediation Model of Health Insurance

Comment

The third-party payer may be a government entity in the case of publicly-funded health insurance. In this eBook, we do not consider this option and refer the interested reader to an excellent research paper by Randall P. Ellis, Tianxu Chen, and Calvin E. Luscombe, *Comparisons of Health Insurance Systems in Developed Countries*, Encyclopedia of Health Economics, Elsevier Press., Inc. 2014.

We now provide a more detailed discussion of the generic model in Figure 1.1 by focusing on the third-party payer in the case of a private insurance company and a self-funded employer.

We discuss each of these intermediation models in turn.

1.2.1 PRIVATE INSURANCE COMPANY AS THIRD-PARTY PAYER

We discuss *two main versions* of health insurance where the third-party payer is a private insurance company.

The ***first version*** is called *individual health insurance* and refers to the case where individuals purchase health insurance policies from private insurers. These insurance policies are customised reflecting the individual's expected future health care needs. The following diagram illustrates a typical intermediation model of individual health insurance.

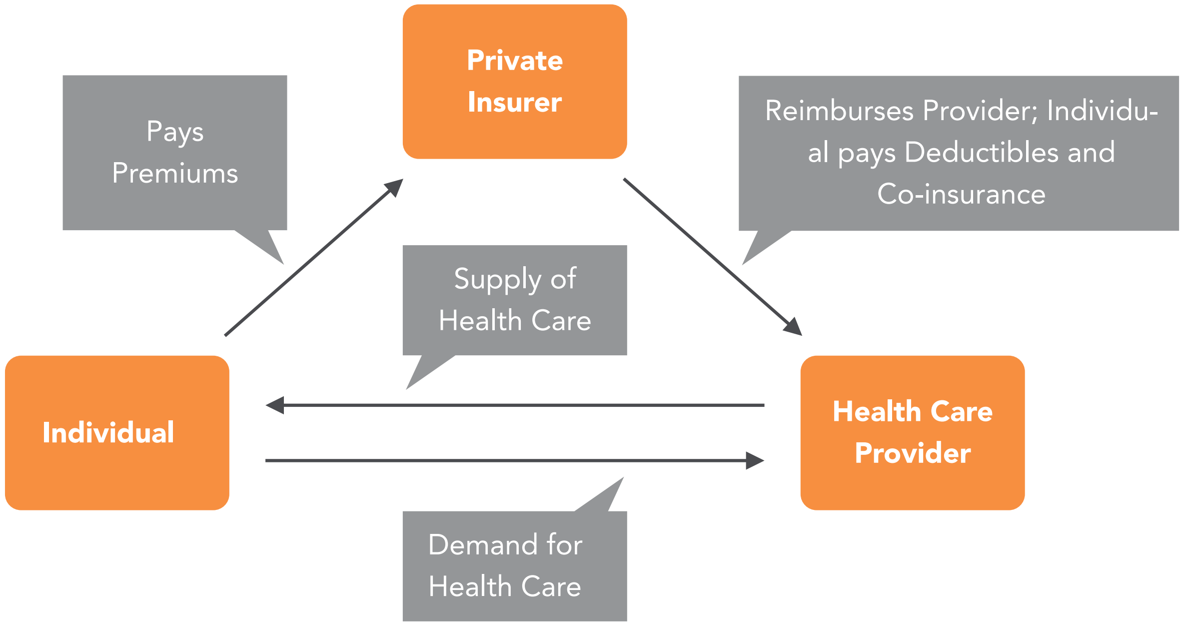


Figure 1.2: An Intermediation Model of Individual Health Insurance

Figure 1.2 shows that an individual creates a demand for health care that is provided by health care professionals. The individual pays premiums to an insurance company which reimburses the health care provider according to the policy conditions (e.g., deductibles and policy limits).

There are variations of this model. For example, in the so-called *reimbursement intermediation model,* the policyholder pays for services provided by the healthcare professional. Afterwards, the insurer reimburses the policyholder according to the terms and conditions of the policy.

The ***second version*** is called *group health insurance* which is designed mainly for policies that are employer-*sponsored.*

A *fully-funded group health insurance* is where the employer contracts an insurance company and pays a premium in direct relationship to the number of employees and eligible dependents (called, members) in the group and to the extent of policy coverage. The employer and employees jointly contribute to the required premium payments. The following diagram illustrates an intermediation model of a fully-funded group health insurance.

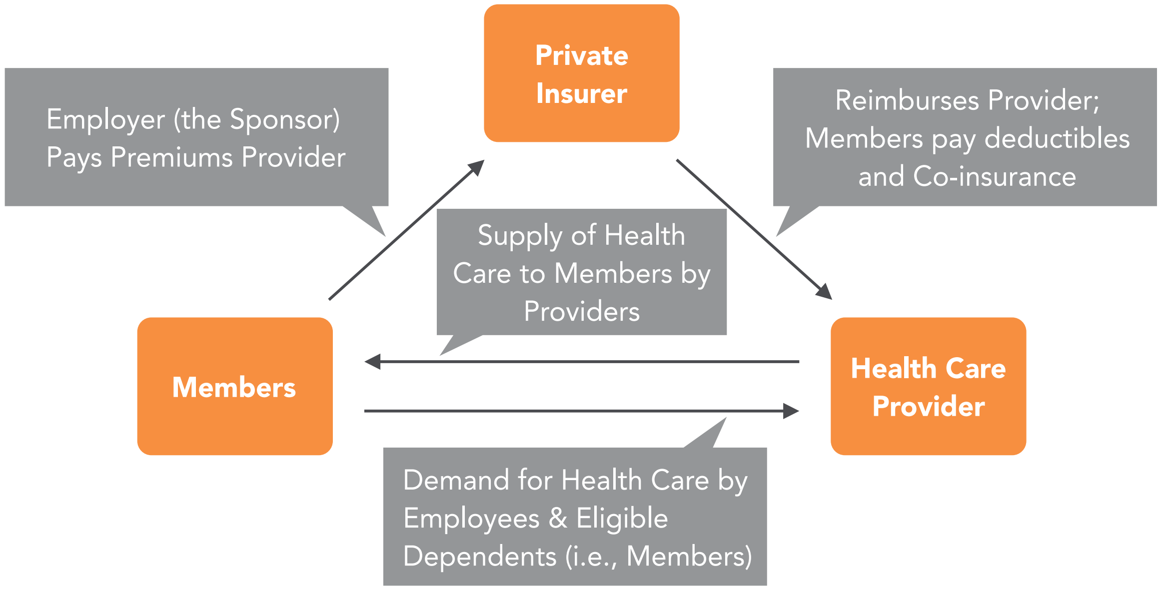


Figure 1.3: An Intermediation Model of a Fully-Funded Group Health Insurance

To this point, we considered the case where the third-party payer is a private insurance company. We considered two cases – individual health insurance and fully-funded group health insurance. We close this section by considering a version of group health insurance that is self-funded.

1.2.2 SELF-FUNDED EMPLOYER AS THIRD-PARTY PAYER

A self-funded health insurance (also called, fully-insured) is where an employer pays for health care benefits from funds of the company. This type of group health insurance is part of employee benefits.

As we discuss in chapter 2, self-funded employer is exposed to the potential of unexpected large claims from health providers on behalf of covered employees. The employer typically manages this financial risk by purchasing *stop loss insurance* that limits its out-of-pocket costs. In other words, the employer transfers its risk of large claims through stop loss insurance.

The following diagram illustrates an intermediation model of a self-funded group health insurance.

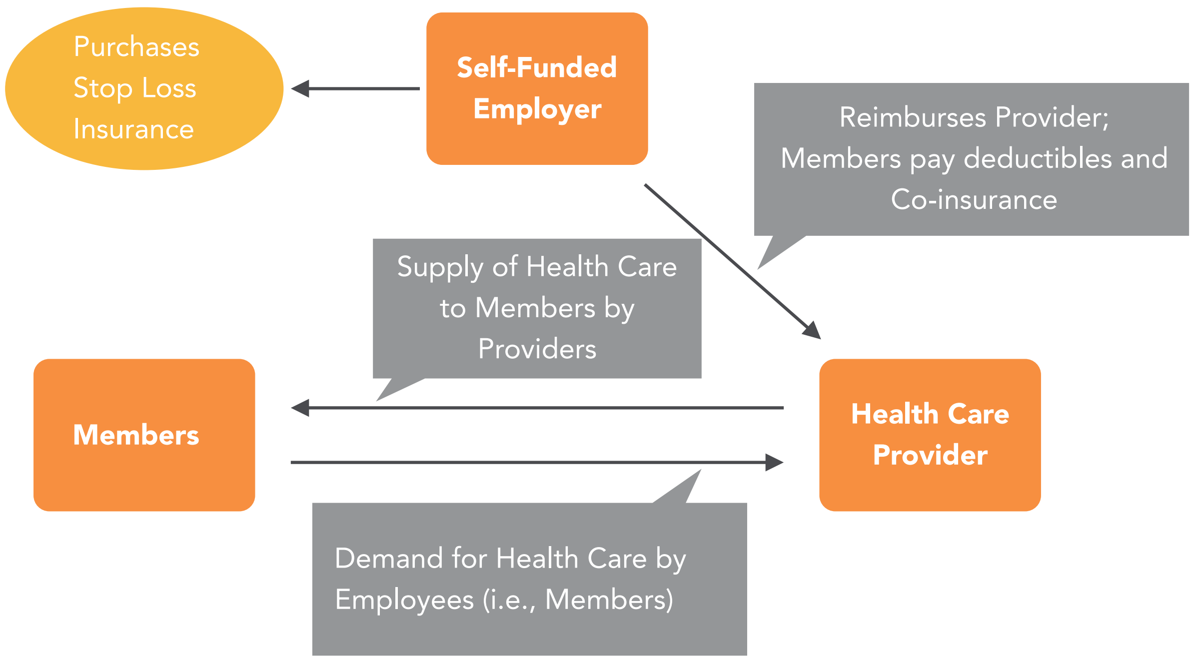


Figure 1.4: An Intermediation Model of a Self-Funded Group Health Insurance

The next section discusses the main attributes of conventional health insurance products.

1.3 CONVENTIONAL HEALTH INSURANCE PRODUCTS

Before we describe the fundamental characteristics of conventional health insurance products, we define the concept of *disability*. We consider the concept of disability from two perspectives – one that is based on physical and/or mental disablement and the second that considers the individual's ability to perform his/her usual occupation. Interestingly, these two perspectives enlighten the definition of different health insurance products.

First Perspective (physical and/or mental disablement)

From a physical and/or mental state of disablement, the Oxford English dictionary defines *disability* as a state in which person has a physical or mental condition that limits his/her movements, senses, or activities. This definition emphasises an individual's disablement as arising from physical injuries or mental impairment that constrains him/her from performing normal human functions. These constraints or limitations may range from broken bones, varying degrees of dismemberment (e.g., loss of a finger, leg amputation), total loss of sight to severe head injuries resulting in paralysis.

From an actuarial perspective, the time an individual spends in a state of disability (i.e., temporary or permanent) and the extent of the disablement (i.e., partial or total) are key dimensions that determine benefits payable by the insurer. In actuarial terminology, these dimensions of disability are labeled *duration* and *severity*.

Duration of a disability refers to whether the state of disablement is temporary or permanent. For example, an accident resulting in a broken leg is likely to render the individual temporarily disabled while an accident that causes the amputation of more than one limbs may be considered as permanent disability.

Typically, if an individual is in an initial state of temporary disability for a period of twelve consecutive months, then a *qualification period* comes into effect. During this period, a medical professional will determine whether the individual has transitioned from a temporary to a state of permanent disability.

Severity of a disability refers to whether an individual is partially or totally disabled. For example, an individual who lost a finger in an accident may be deemed by a medical professional to be partially disabled while the individual's loss of sight in both eyes may be deemed to be a state of total disability.

Comment

This perspective of personal disablement is the basis for *accident insurance* (also called *personal accident insurance*) described in section 1.3.1.

Second Perspective (functional disablement)

The definition of disability presented above considers the physical and/or mental state of the individual. The second perspective (i.e., functional disablement) considers the individual's inability to *partially* or *totally* to perform their usual occupation or similar occupation that matches his/her expertise and experience. The severity of the disability may also lead to a state where the individual is totally unable to perform any occupation whatsoever.

We summarise the two dimensions of disability (duration and severity) based on the second perspective in the box below:

	Temporary	Permanent
Partial	• Can perform usual or related occupation on a limited basis; • High probability (as assessed by a medical professional) of recovery and returning to full and continuous employment in his/her usual occupation.	• Can perform usual or related occupation on a limited basis; • Zero probability (as assessed by a medical professional) of recovery and returning to full and continuous employment in his/her usual occupation.
Total	• Unable to perform any work in his/her usual or related occupation, whatsoever for a limited time. • High probability (as assessed) by a medical professional) of recovery and returning to full and continuous employment in his/her usual occupation.	• Unable to perform any work in his/her usual or related occupation, whatsoever. • Zero probability (as assessed by a medical professional) of recovery and returning to full and continuous employment in his/her usual occupation.

The second perspective of disability is a basis for *disability insurance* (also called *income protection (IP) insurance*) described in section 1.3.2.

We now describe conventional health insurance products (CHIP) which are listed in Figure 1.5 below.

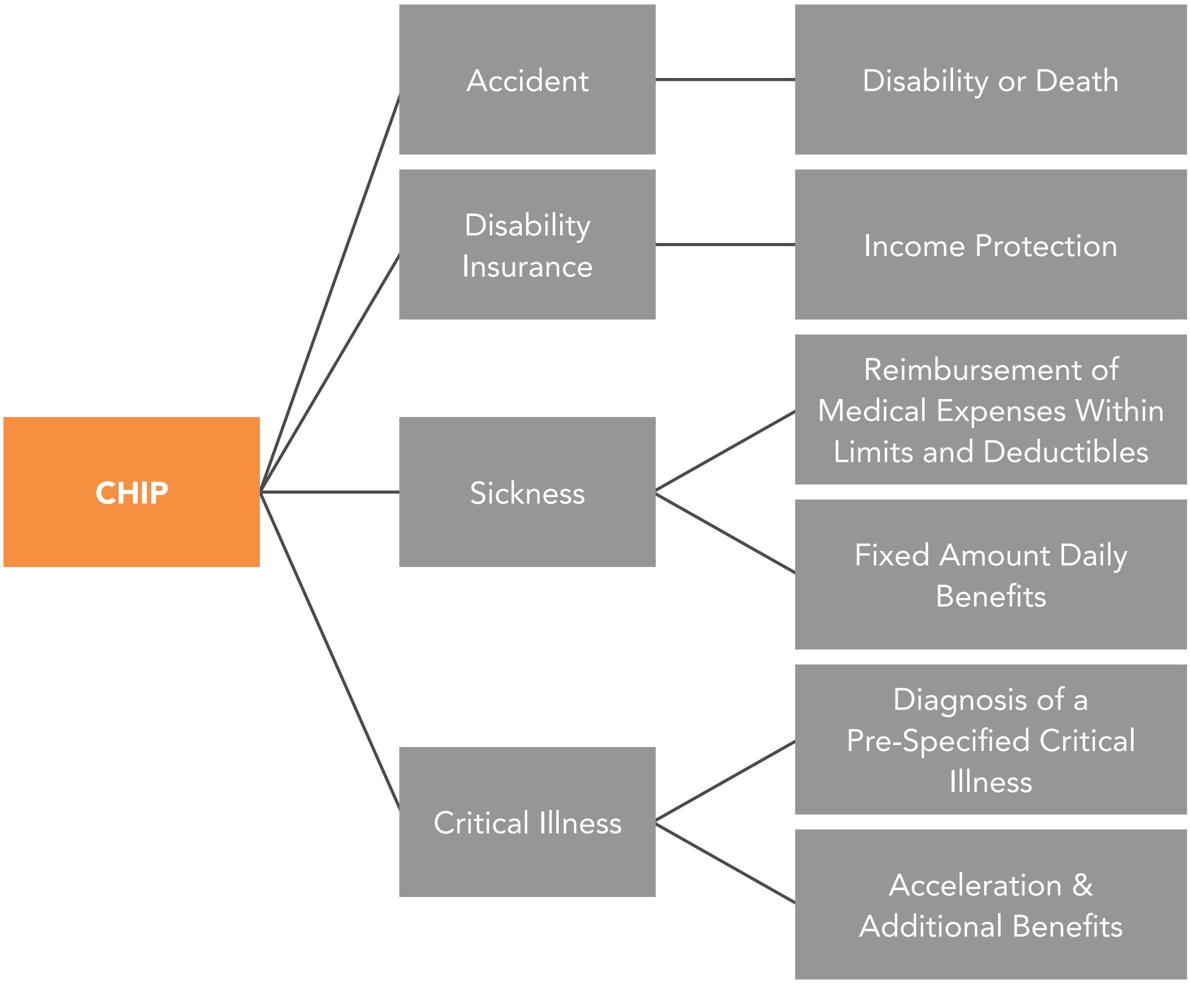


Figure 1.5: Conventional Health Insurance Products

1.3.1 ACCIDENT INSURANCE

In the context of accident insurance (also called *personal accident insurance*), the word *accident* is defined as a single unexpected or unforeseen event that is caused by ***violent, visible and external*** means resulting in bodily injuries or death. To illustrate this definition, consider the case of a policyholder who is driving his/her car to work and is involved in a sudden collision with another vehicle on the freeway resulting in severe bodily injuries to himself/herself. This is a single unexpected event that is caused by a violent collision which is visible and external to the policyholder. (Note that a heart attack may be violent and visible but it is not external (i.e., it is internal) to the person and does not qualify as an accident *per se.* In fact, it is an example of *sickness*.)

Typical covers for accident insurance include:

a) Death benefit which is equal to a sum insured (SI) determined at the policy effective date.
b) Benefit payment for a permanent disability that is equal to $\alpha \times$ SI where the benefit percentage (α) varies with the severity of the policyholder's disablement. For example, in case of amputation of both legs caused by a covered accident, the benefit percentage is typically 100% while it may be 30% for the loss of hearing in one ear.
c) Fixed-amount daily benefit for a maximum period of one year for temporary disability.
d) Reimbursement of medical expenses for short-term sickness and/or hospital stay resulting from a covered accident. Policy conditions typically include deductibles and/or policy limits.

1.3.2 DISABILITY INSURANCE

Disability insurance (also called *income protection (IP) insurance* or *workers' compensation in the USA*) is based on the definition of disability in terms of the second perspective described above where the focus in on the policyholder's ability to perform his/her usual or related occupation. In other words, an income-protection benefit is payable for a valid claim when the policyholder is unable to return to his/her usual or related occupation or cannot perform any work whatsoever.

The IP periodic (e.g., monthly) income payable by the insurer is less than the policyholder's pre-disability periodic employment income. In this context, a *replacement ratio* is defined as the IP monthly income payable expressed as a percentage of the pre-disability employment income. The value of IP is less than 100%. From the insurer's perspective, the replacement ratio less than 100% provides a financial incentive for him/her to return to work.

1.3.3 SICKNESS INSURANCE

Sickness insurance provides the policyholder two main categories of benefits. The first category refers to fixed-amount daily benefits for sickness periods and/or hospital stays, each subject to a maximum number of days. For example, temporary disability caused by sickness is typically covered for a maximum of one year. A lumpsum payment is typically payable on the determination by a medical professional that the policyholder has transitioned into a state of permanent disability.

The second category is based on a *reimbursement model* of health insurance where a benefit is payable for medical expenses incurred by the policyholder but subject to restrictions due to deductibles and/or policy limits.

1.3.4 CRITICAL ILLNESS INSURANCE

Critical illness insurance (also called *dread disease insurance*) provides a pre-specified lumpsum payment to a policyholder ***on the diagnosis*** of a critical illness that is included in a pre-determined list. This list typically includes severe illnesses such as heart failure, stroke, cancer, kidney failure, Alzheimer's disease etc. Upon payment of the sum insured, the policy is terminated.

While stand-alone critical illness insurance provide cover against critical illness, they are not commonly sold in this form. One reason is that if death occurs before the diagnosis of a critical illness, no benefit payable to the beneficiary. A more common insurance is a life insurance combined with a critical illness rider.

There are two typical such riders – *acceleration* and *additional covers* – which are respectively described and illustrated below.

a) Accelerated Critical Illness (ACI)

Consider the following diagram:

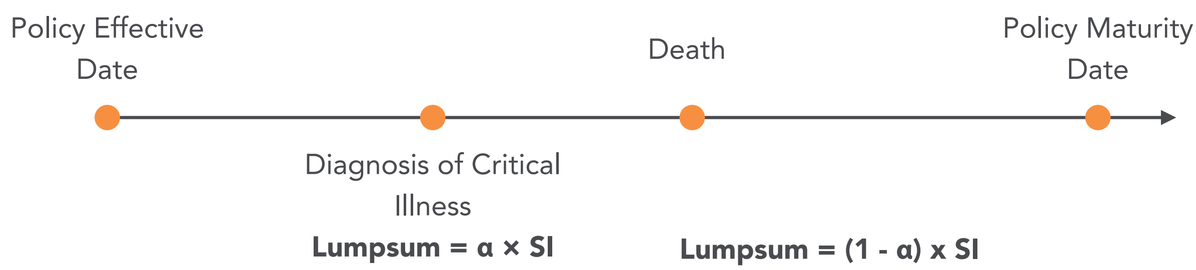


Figure 1.6 Accelerated Critical Illness

Let SI be the sum insured at the policy effective date. A percentage of SI is paid on the diagnosis of critical illness incurred by the policy holder. Let this percentage be represented by α, which is specified and known at the effective policy date. The remaining amount ($(1-\alpha) \times$ SI) is paid upon the death of the policyholder on the assumption that this event occurs on or before the policy maturity date. (Figure 1.6 illustrates this assumption).

For example, if the sum insured is SI = \$100,000 and α= 70%, then the policyholder is paid \$70,000 upon the diagnosis of a critical illness. The remaining \$30,000 on the death of the policyholder assuming that this event occurs on or before the policy maturity date.

Comment

The higher the value of α, the greater the acceleration of the death benefit towards the lump sum payment on the diagnosis of a covered critical illness. A value of α = 1 means that the acceleration is at a maximum and the cover for mortality risk is zero. Using the example above, a lumpsum of \$100,000 is paid on the diagnosis of a critical illness and the death benefit is reduced to zero.

If death occurs is earlier than the diagnosis of critical illness, then SI is paid to the beneficiary.

a) Additional Cover

We refer to Figure 1.7 below:

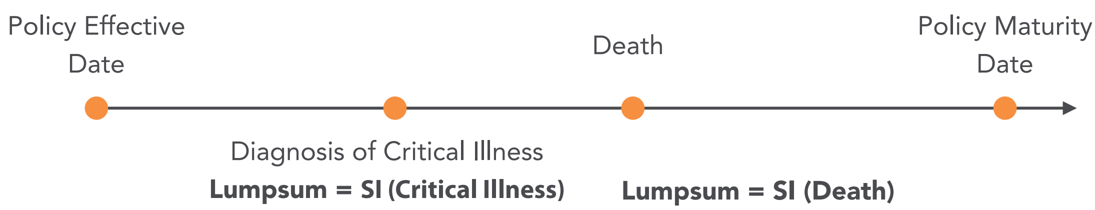


Figure 1.7 Additional Cover for Critical Illness

Different from the ACI described above, *additional cover* for critical illness comprises two independent covers – one on death of the policyholder and the other on the diagnosis of a pre-specified critical illness. For example, on the policy effective date, the sum insured for critical illness is \$100,000 and the sum insured on death of the policyholder is \$200,000, then a lumpsum of \$100,000 is paid on the diagnosis of the covered critical illness and a lumpsum of \$200,000 is paid upon death of the policyholder within the term of the policy.

Comment

To this point in this chapter, we have described the main features of conventional health insurance contracts. We do not consider *long-term care (LTC)* health insurance since this insurance has a complex structure and will be considered in Level II of the Art of Insurance series.

We conclude this chapter with a discussion and illustration of the typical life cycle of a health insurance claim.

1.4 LIFE CYCLE OF A HEALTH INSURANCE CLAIM

We describe the typical life cycle of a health insurance claim in two sequential steps. The first, presented Figure 1.8, considers the interval between the policy effective date and the date of claim inception. The second step considers the interval between the claim inception and the termination of the claim. This interval includes the date of benefit payment as well as a qualification period to assess the potential transition from a state of temporary disability to a state of permanent disability.

The following diagram illustrates the first step in our description of a typical life cycle of a health insurance claim.

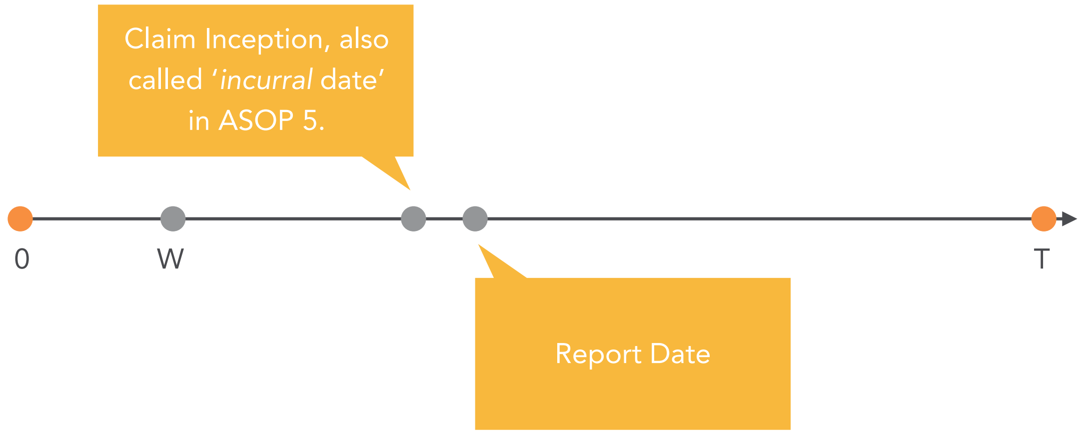


Figure 1.8: Life Cycle of Generic Health Insurance Policy (First Part)

Here is an explanation of Figure 1.8:

- The period [0, T] is the coverage period. The policy has an effective date at time 0 and terminates at time T. The termination date can be longer than one year (e.g., multi-year cover). The effective policy date is the date when health insurance coverage begins.
- The period [0, w] is between the policy effective date at time **0** and **w.** It is called a ***waiting period*** (also called a ***probationary period***). During this period, claims for covered insured events are ***not*** in effect. The actuarial rationale for a waiting time is based on the theory of asymmetric information which states that parties to a contract who have informational advantage can gain at the expense of the other party.

(The theory of asymmetric information and its related risk of anti-selection and moral hazard are discussed in chapter 2).

As an illustration, consider the case where a person applies for income protection (IP) insurance but hides the information that he/she has an internal back injury from a recent fall. The insurance company does not know this information but the individual does. The risk for the insurance company is that it provides an IP insurance policy to this individual who is charged a premium reflecting a relatively low risk of disability. The imposition of a waiting period is one way the insurer seeks to mitigate the risk of anti-selection.

- Another risk mitigation possibility is for the insurer to exclude pre-existing health conditions.
- The ***claim inception date*** or the ***incurral date*** is defined by the Actuarial Standard of Practice (ASOP 5) as the date a liability is created for the risk-bearing entity in accordance with the terms of the health insurance. It is the date a claim begins. This is the date, for example, when the policyholder visits the health care provider for medical services or when hospitalisation begins.
- A claim is typically reported to the insurer electronically by the policyholder or the health care provider on behalf of the policyholder via the insurer's call centre or website. This establishes the *report date.*

The second step in the life cycle of health insurance claim extends Figure 1.8 and considers the time interval after the report date. Here is a development of a life cycle of a health insurance claim in Figure 1.9 below with explanations following:

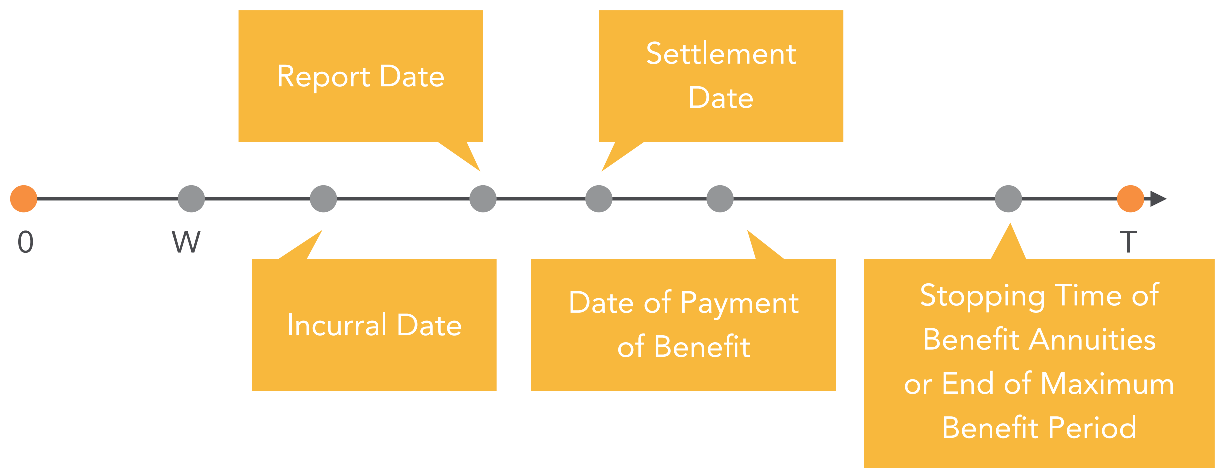


Figure 1.9: Life Cycle of Generic Health Insurance Policy

- The period between the report date and the settlement date is when the claim is being evaluated to determine its validity. These claims are described as 'in course of settlement' (ICOS).
- The period between the settlement date and the date of benefit payment is called a ***deferred*** period (also called an ***elimination*** period). The deferred period is called a period of self-insurance by the policyholder. The key point is that payment is not made during the period between the settlement date and the payment date.

  The policyholder chooses the duration of the deferred period (e.g., one week, one month, three months etc.) and all else equal, the longer the deferred period, the lower the premium for the policy.

Comment

Consider the case of a policyholder who makes a claim, receives payments and returns to work but subsequently suffers a recurrence of the *same disability* within a specified period, then the deferred period is waived. This specified period is called a ***link period***. A common link period is six months. Simply stated, if the ***same*** disability recurs within a period of six months, then payment of a valid claim is not subject to a deferred period. Payment is made upon settlement of a valid claim. Simply put, the deferred period is waived during the link period.

- Payment of benefits may be in a lumpsum or a stream of periodic payments (i.e., benefit annuities). In the latter case, benefit annuities are typically subject to a stopping time that corresponds to the policyholder's retirement age.

The following example illustrates the concepts described in this section.

Example 1 (adapted from April 2013 exam on Health and Care Specialist Technical, sponsored by the Institute of Faculty and Actuaries).

The following information is based on past claims from a health and care insurer for a policy that offers payments of £500 per month to a member who is absent from work due to sickness or accident and for a valid claim. There is a deferred period of four months and a link period of six months.

For claimant A, calculate the total benefit payment for each year.

Claimant A	Start of Absence (Claim Inception)	End of Absence (End of Payment Period)	Cause of Absence
Absence (#1)	1/4/15	1/4/16	Stress
Absence (#2)	1/6/16	1/7/16	Fall at Work
Absence (#3)	1/9/16	1/4/17	Stress

Answer

2015:

There is a deferred period of four months and so payments start four months after the claim inception date. Hence payments are made from 1/8/15 through 1/4/2016. For 2015 there are payments for five months amounting to 5 × £500 = £2,500.

2016:

There are three payments made in 2016 from absence (#1).
There is no payment for absence (#2) which lasted one month and shorter than the deferred period of four months.

Absence (#3) is for the same disability (stress) beginning on 1/9/16 which is five months after the end of the first absence on 1/4/16. Since this is shorter than the link period of six months, the deferred period is waived. There are four payments for 2016. In summary, there is a total of seven payments in 2016 for a total of £3,500.

2017:

Absence (#3) has three payments outstanding in 2017 totalling in £1,500.
This concludes chapter 1.

2 POLICY COVERAGE MODIFICATIONS, ANTI-SELECTION AND MORAL HAZARD

2.1 INTRODUCTION

This chapter considers typical policy coverage modifications in health insurance (i.e., deductibles and policy limits) and analyses their relationship in mitigating the risks of anti-selection and moral hazard. Specifically, we present the following main issues:

- Deductibles and policy limits for private health insurance and fully-funded group health insurance;
- Stop loss limits for self-insured group health insurance;
- External, internal and durational anti-selection due to Bluhm (2007);
- Ex-ante and ex-post moral hazard and the risk of over-consumption that is particularly endemic to health insurance.
- The mitigation of the risks of anti-selection an moral hazard by suitably-designed deductibles, policy limits and co-insurance factors.

We begin with a brief review of deductibles and policy limits which are fully discussed in the first eBook on the Principles of Insurance.

2.2 FIXED AMOUNT DEDUCTIBLE

The health insurer may be faced with high severity claims from policyholders who are permanently disabled from an automobile accident or frequent small claims from policyholders who experience short-term periods of sickness and thereby incur medical and/or hospitalisation expenses. Frequent small claims may negatively affect an insurer's profit. One reason is that average claim settlement costs include the insurer's fixed administrative costs and, for example, it may cost more than $100 to settle a claim of $100. The provision of a fixed amount deductible results in a relatively small policyholder's loss being retained by the policyholder.

Here is an example illustrating the concept of a fixed amount deductible.

Suppose that a sickness contract imposes an ordinary deductible of €500 and the policyholder incurs medical costs of €2,000. Since the policyholder's loss exceeds the deductible, the policyholder pays the deductible of €500 and the insurer's cost is equal to the €2,000 – €500 = €1,500. However, if the policyholder's loss is $300, the policyholder bears the full cost.

We consider implications of an ordinary deductible formally from the perspective of both the policyholder and the insurer as follows:

Policyholder's Perspective:

For an ordinary deductible (**d**) and a policyholder loss (**X**), the policyholder's cost per loss ($\mathbf{Z^L}$) is described as follows:

$$\mathbf{Z^L} = \begin{cases} \mathbf{X \text{ if } X \le d} \\ \mathbf{d \text{ if } X > d} \end{cases} \tag{2.1}$$

Comment

a) In equation (2.1), $\mathbf{Z^L}$ refers to the cost to the policyholder because of an ordinary deductible. This is the cost *retained* by the policyholder. It is noted that the maximum cost to the policyholder is the deductible in both cases.
b) Equation (2.1) can be rewritten as follows: $\mathbf{Z^L = \min(X, d)}$. Actuarial terminology for $\mathbf{Z^L}$ is as follows:

$$\mathbf{Z^L = X \wedge d} \tag{2.2}$$

The notation in equation (2.2) is common in the actuarial literature and we adopt this form for the remainder of this eBook. Here is an example illustrating equation (2.2).

Example 1

A policyholder is issued health insurance policy for one year with a deductible of $300. During the term, the policyholder incurred a covered loss of $800. A deductible of $300, means that the policyholder pays the *minimum* of $800 and $300. Hence, $\mathbf{Z^L = 800 \wedge 300} = 300$.

If the covered loss is $200, then $\mathbf{Z^L = 200 \wedge 300} = 200$. The policyholder pays the full amount of the cost since it is lower than the deductible.

We now consider the insurer's perspective:

Insurer's Perspective

The insurer's cost per loss (i.e., payment per loss) is typically represented by $\mathbf{Y^L}$ which is defined as follows:

$$\mathbf{Y^L} = \begin{cases} \textbf{not defined} & \textbf{if } \mathbf{X \le d} \\ \mathbf{X - d} & \textbf{if } \mathbf{X > d} \end{cases} \tag{2.3}$$

Comment

a) In actuarial terminology, Y^L is called ***cost per loss*** variable. Simply put, it is the cost to the insurer for policyholder loss. Equivalently, it is the insurer's payment per policyholder loss.

b) The insurer's cost per loss may be written as follows: $\mathbf{Y^L = X - X \wedge d}$ (2.4)

We explain (2.4) as follows:

Without a deductible, the insurer pays the policyholder loss **(X)**. When a deductible is provided, the policyholder pays a share of the loss described in equation (2.2). This share is $\mathbf{X \wedge d}$. Equation (2.4) states that the cost to the insurer **($\mathbf{Y^L}$)** is equal to the policyholder loss **X** minus the cost paid by the policyholder, $\mathbf{X \wedge d}$.

Here is a graphical illustration of an ordinary deductible that serves to highlight some additional insights.

Figure 2.1 Fixed Amount Deductible (d)

The orange-coloured line in Figure 2.1 represents the case of no deductible so that policyholder loss (X) is equal to the loss payment (Y^L) made by the insurer.

The horizontal part of the blue line segment represents the case where policyholders retain all losses because they are lower than the deductible. The upward-sloping blue line represents all claim amounts, X exceeding the deductible. There is a ***shift*** of the loss variable X to X-d.

Comment

If an ordinary deductible is imposed on a per-loss basis or an existing ordinary deductible is increased, there will be less payments made by the insurer since a reduced number of policyholder claims will be reported.

Example 2

Consider the following historical information for a portfolio of homogeneous health insurance policies issued on January 1, 2017 with maturity date on December 31, 2017. Each policy is provided with an ordinary deductible of $500. On April 1, 2017, the following information is reported:

a) 300 policyholders' medical losses each of an amount less than or equal to $500 with a total value of $60,000.
b) 100 policyholders' income losses each of an amount exceeding $500 with a total value of $180,000.

Case A: No Deductible

In the case of no deductible, the insurer pays the full amount of the insured losses (i.e., $Y^L = X$). The total loss for the insurer is = $60,000 +$180,000 = $240,000.

Case B: Deductible of $500 for each Policy

For an deductible of $500, the 300 policyholders' losses amounting to $60,000 will retained by policyholders.

For each policy with a policyholder loss above \$500, the policyholder pays the deductible. For 100 insured losses, the policyholder loss is equal to 100 × \$500 = \$50,000. Therefore, the total loss retained by policyholders is \$60,000 + \$50,000 = 110,000.

Here is a summary of the calculation of the loss that is eliminated by the deductible.

Total Amount of Claims ≤ Deductible	\$60,000
Total Amount of Claims > Deductible	Number of claims × Deductible = 100 × \$500 = \$50,000
Total Loss Eliminated by the Deductible (i.e., 100% out-of-pocket expenses for policyholders)	\$110,000

We now consider another type of deductible called a *franchise deductible.*

2.3 FRANCHISE DEDUCTIBLE

A franchise deductible (d) means that the policyholder pays the amount of losses that are less than or equal to d. For losses above the deductible, the insurer's cost is the *full amount* of the policyholder's loss.

Formally:

$$\mathbf{Y}^{\mathbf{L}} = \begin{cases} \mathbf{0} & \textbf{if } \mathbf{X} \leq \mathbf{d} \\ \mathbf{X} & \textbf{if } \mathbf{X} > \mathbf{d} \end{cases} \tag{2.5}$$

Comment

Losses less or equal to the deductible are not reported to the insurer. But different from the ordinary deductible, there is ***no shift*** of X to X-d. The insurer pays the full amount of losses that exceed the franchise deductible.

Example 3

For a policyholder loss of X = \$600 and a franchise deductible of \$150, the insurer's cost per loss is which one of the following?

a) $600
b) $150
c) $450
d) $250

The correct answer is a). The insurer pays the full amount of the loss of $600 since it exceeds the franchise deductible of $150.

We continue the review by considering the implications of modifying policyholders' losses by a fixed percentage deductible (also called proportional co-insurance).

2.4 PROPORTIONAL CO-INSURANCE (FACTOR = α%)

Here is an example illustrating the main characteristics of proportional co-insurance. Suppose that the policyholder makes a claim for a loss of €2,000 and the insurance contract has a co-insurance factor of **α** = 90%. In this case, the insurer pays 90% of the loss amount of €2,000 (i.e., €1,800) and the policyholder pays the remaining 10% (i.e.,€200).

Formally, for a co-insurance factor α and a policyholder loss (**X**), the amount (**Y**) paid by the insurer is calculated as follows:

$$\mathbf{Y} = \alpha \times \mathbf{X} \quad (2.6)$$

Equation (2.6) is illustrated in the following graph:

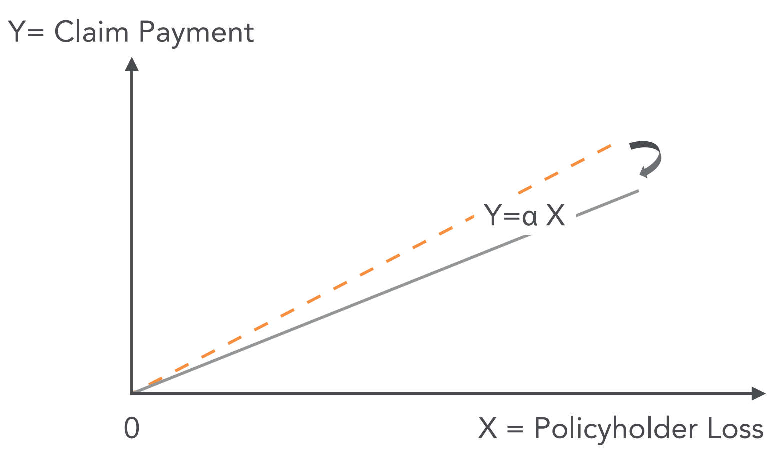


Figure 2.2: Coinsurance Factor is α%.

Comment

In a policy with a coinsurance factor $0 < \alpha < 1$ the insurer pays the portion $\alpha \times X$ of the loss. Proportional co-insurance does not eliminate the reporting of policyholders' small claims. *All claims are reported.*

We now consider policy limits where the insurer's cost per policyholder loss does not exceed a value equals to ***u.***

2.5 POLICY LIMITS

We provide an example to illustrate the main feature of a policy limit. Suppose that the policyholder who purchased a health insurance policy with a policy limit of €10,000 submits a claim for a loss that is equal to €20,000. The insurer pays the policyholder €10,000.

This is because the insurer pays the **minimum** of two values: the policyholder's claim and the policy limit. If the claim for payment were €8,000, then the insurer pays the full claim since is below the policy limit.

Formally, for a policy limit **u** and policyholder's loss **(X)**, the amount **Y** paid by the insurer is calculated as follows:

$$\mathbf{Y} = \begin{cases} \mathbf{X \text{ if } X < u} \\ \mathbf{u \text{ if } X \geq u} \end{cases} \tag{2.7}$$

Equivalently, $\mathbf{Y = X \wedge u}$ which is called a *limited loss variable* and **E(Y)** is called a *limited expected value.*

Here is a graph depicting the effect of a policy limit, u.

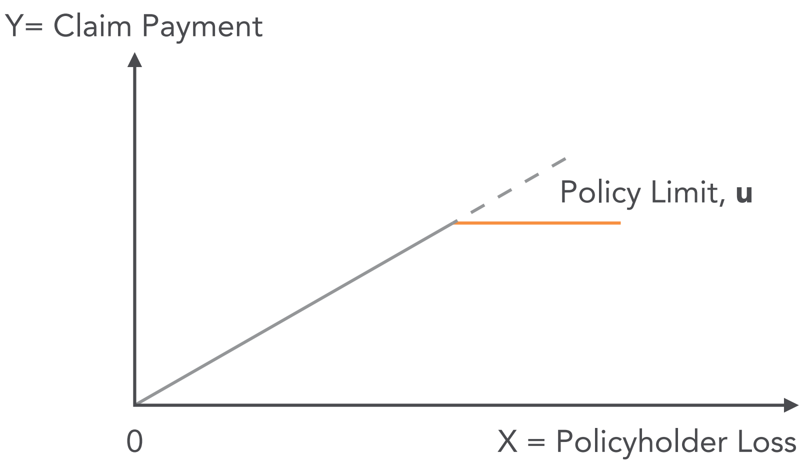


Figure 2.3: Policy Limit, u

Convention: The deductible is applied before the policy limit.

We now consider the case for a self-funded group health insurance that purchases stop loss insurance to mitigate the risk of unexpectedly large claims from members.

2.6 STOP LOSS INSURANCE (SELF-FUNDED GROUP HEALTH INSURANCE)

Consider the case of an employer-sponsored group health insurance that is self-funded (also called self-insured) because the employer assumes the financial risk of paying for health care benefits provided to its employees (i.e., members). As a result, the self-insured employer faces the risk of unpredictable and/or catastrophic employee claims. One way to mitigate this risk is to purchase *stop-loss insurance* that will reimburse the employer for claims with values above a specified monetary level. We formalise this procedure as follows:

Assume that L is the stop-loss limit. L is pre-determined and represents the maximum-out-of-pocket cost for the employer for each employee claim.

Also assume that there is a fixed amount deductible of **d** per member's claim. For example, if d = $500, then the member pays claim expenses up to and including $500. Other policy terms and conditions will take effect after the deductible is applied.

Finally, assume that **α** is the co-insurance factor so that the member pays a share of remaining claim expenses of **1-α**, after the deductible is implemented. It is convention to state the co-insurance factor as $\alpha/1\text{-}\alpha$. For example, the terminology 80/20 means that the employer/insurer (by convention) always pays the larger percentage. Accordingly, the member pays 20% of claim expenses after paying the fixed amount deductible.

Before, we present the formal model, here is an example that illustrates the priority of member's payment.

Example 4

For a self-funded group health insurance, a member incurs medical expenses of $1,000. There is a deductible of $200 and co-insurance factor of 80/20. What is the member's share of this expense?

Answer

The member first pays the deductible of $200 and then 20% of the remaining $800 which is equal to $160. The total member's share is $200 + $160 = $360.

We formalise this discussion below:

Let **d** be a fixed amount deductible, **α** = co-insurance factor or the employer's share, so that **1-α** is the member's share. Also, **L** = stop loss limit. Let **x** = monetary value of member's expenses that will become a valid claim according to policy coverage conditions. We analyse three cases as follows:

a) If the **x** is less than or equal to **d**, the employer pays zero. The member is fully responsible for all expenses up to and including the deductible.

b) If **x** is greater than **d**, the member pays the deductible and the employer pays a share of **α** of the remaining amount of **x-d**. That is, the employer pays **α** × **(x-d)**.

Since the maximum payment by the employer is L then α × (x-d) = L. Therefore, $x = \frac{L + \alpha d}{\alpha}$.

To summarise, the employer pays α × (x-d) for claims higher than the deductible and less than or equal to $\frac{L + \alpha d}{\alpha}$. For convenience we define $\mathbf{m} \equiv \frac{L + \alpha d}{\alpha} = d + \frac{L}{\alpha}$.

Note that the upper limit of this region (**m**) is higher than the employer's maximum payment (**L**).

c) If x is greater than **m**, the stop loss insurance provider reimburses the employer a value of **α** × **(x-m)**; As a result, the employer pays **L**

Here is a graphical illustration of these three possibilities for a member's claim expense.

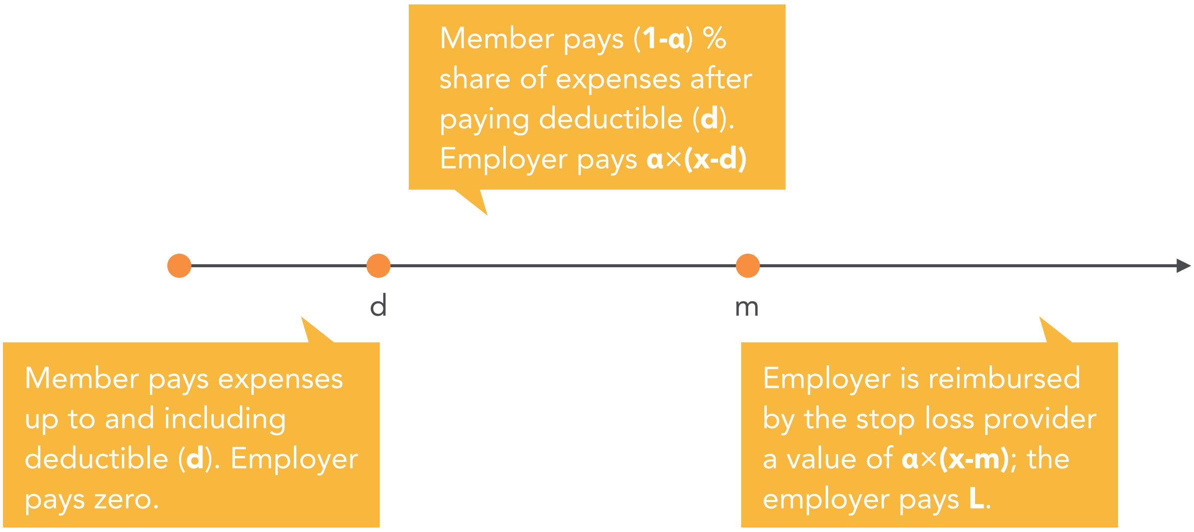


Figure 2.4: Stop Loss Insurance

From Figure 2.4, we state the employer's loss function (Y^L) for a member's expense (x) as follows:

$$Y^L = \begin{cases} 0 & \text{if } x \le d \\ \alpha \times (x - d) & \text{if } d < x \le m \\ L & \text{if } x > m \end{cases} \tag{2.8}$$

Comment

For member's expense (x) which exceeds the value of m, the stop loss provider reimburses the employer a value which is equal to **α×(x-m)**. This will ensure that that the employer pays a net amount of L.

Here are examples that illustrate equation (2.8).

Example 5

A self-insured group insurance has the following policy coverage conditions: fixed amount deductible = $200 and a co-insurance factor of 80%. The stop loss limit is L= $1,000.

For a member's expense amount of $100, what is the payment by the employer?

Answer

The expense of $100 is below the value of the deductible. The member pays the full amount of the expense and the employer pays zero.

Example 6

Referring to the information in Example 5, what is the payment by the employer for a member's expense of $500?

Answer

Since the value of the claim is higher than the deductible, we calculate the upper limit in this region which is m = \$200 + \$1,000/ 0.80 = \$1,450.

This means for a member's expense ranging between the deductible of \$200 and the upper level of m = \$1,450, the employer pays 0.80 × (\$500 – \$200) = \$240.

Example 7

Referring to the information in Example 5, what is the payment by the employer for a member's expense of \$1,800?

Answer

The claim of \$1,800 is above the value of m = \$1,450. In this case, the member pays the deductible of \$200 and then 20% of the remaining amount of \$1,600 for a value of \$320. Hence the member pays a total of \$520. The remaining amount is \$1,280. Hence, the stop loss provider pays \$280 to the employer so that the employer pays a net amount of \$1,000.

The same result is obtained by applying the formula **α×(x-m)** for the stop loss provider's payment. This gives a value of 0.80 × (1,800 – 1,450) = 0.80 × \$350 = \$280 as obtained above.

We now discuss the role of deductibles and policy limits in mitigating two important sources of risk arising from policyholder behaviour in the presence of asymmetric information. These are the risks of anti-selection and moral hazard.

2.7 ANTI-SELECTION RISK

The principle of anti-selection is linked to the concept of asymmetric information. In the context of insurance, asymmetric information is present when the insurance applicant has information that is not available to the insurer and importantly, this information is relevant to the outcome of their contractual relationship.

If a potential customer has information, prior to the effective date of an insurance contract, that is not revealed to the insurer, then anti-selection risk may exist. This risk is observed when the decision-maker inadvertently charges high-risk individuals premiums as if they were classified as lower risk. Evidently, anti-selection creates a potentially adverse financial impact for the insurer since these higher-risk policyholders are likely to make claims with greater frequency and severity as compared to their charged premiums.

In a typology proposed by Bluhm (2007), this is called ***external anti-selection***. This is a risk to the insurer which occurs ***before*** the policy effective date.

There is also the concept of ***internal anti-selection***. This risk occurs at the time when in-force policies for existing policyholders is considered for renewal – the reason for the adjective, *internal.*

To illustrate, consider the case where the insurer has decided to increase the premiums of policies subject to renewal, by 20%. Lower risk policyholders (also called *select risks*) who are likely to make fewer claims are faced with two choices.

They may not renew their policies and purchase policies from other insurers. In this case, the insurer is left with mostly high-risk policyholders (i.e., *non-select risks*) resulting in a loss of premium revenue.

Alternatively, at the renewal date, low risk policyholders may choose a policy with the same coverage but with higher deductibles so as to reduce the impact of the premium increase. This is because high-deductible policies have lower levels of premiums and so the 20% increase will be based on a lower base rate. The overall impact is to reduce the expected increase in premium revenue for the insurer.

If the insurer is forced to increase premiums further to combat this negative impact on the initial premium increase, then the probability that low risk policyholders will not renew their existing policies increases. The overall impact is a further reduction in premium income and the portfolio of policyholders is dominated principally by higher-risk policyholders. Economists label this phenomenon as the 'lemons problem'.

The third form proposed by Bluhm (2007) is called ***durational anti-*****selection** where the term 'durational' refers to the aging of the insured's coverage. This form of anti-selection occurs when policyholders make a decision to lapse the policy and thereby, end coverage. As stated by Bluhm (207, page 85), "higher cost insureds will tend to keep their coverage in-force more often than lower cost insureds, implying they find it of greater value compared to the cost."

Clearly, these three forms of anti-selection are interdependent in the following way:

If a high-risk individual is incorrectly selected and charge a low-risk premium (i.e., external anti-selection), then they are likely to renew their policies even if there is an increase in premiums (i.e., internal anti-selection) and finally, there is a high probability that they will not lapse their policies (i.e., durational anti-selection).

Important Consequence of Anti-Selection

The three forms of anti-selection have a combined effect of increasing loss ratios over the duration of the insured's coverage. This conclusion is explained as follows:

We showed in the second eBook on *Non-Life Insurance*, $\text{Loss Ratio} = \dfrac{\text{Claim Amount}}{\text{Earned Premium}}$.

Anti-selection creates a tendency for higher-risk policyholders to dominate the portfolio of insureds. This results in higher than expected claim costs compared to earned premiums and hence, higher loss ratios and a negative effect on underwriting profitability.

How can an insurer mitigate the risk of *external anti-selection* in its portfolio with a degree of confidence? This is a key issue since the incorrect selection of higher-risk policyholders who are charged lower-risk premiums will likely have a negative ***compound effect*** arising from internal and durational anti-selection.

Standard textbooks in health insurance agree with Folland, Goodman and Stano (2013) in *Economics of Health and Health Care*, 7th edition, page 201 who state that "Group insurance can often be a more useful mechanism to reduce adverse selection". This is because a portfolio of high-risk and low-risk policyholders result in a reduction of total portfolio (i.e., group) risk. The implication is that an insurer can, to some degree, mitigate the risk of external anti-selection by preferring to provide group health insurance over individual health insurance.

Assume that in the presence of asymmetric information, policyholders know their true level of risk (in terms of frequency and severity of claims) but the insurer does not. The central prediction of anti-selection theory is that high-risk policyholders tend to choose insurance policies with lower or no deductibles – that is, high insurance coverage. Similarly, low-risk policyholders tend to choose insurance policies with relatively high levels of deductibles.

This is a very important prediction since the actuary in the presence of a disadvantage in underwriting information has some difficulty in separating high risk applicants for insurance from low-risk applicants. This prediction suggests that the insurer should take note that high-risk insurance applicants have a greater incentive to opt for full insurance coverage.

Comment

There is evidence that superior customer service quality is related to a greater degree of customer loyalty to the company. Therefore, an insurer which focuses on achieving superior customer loyalty is likely to mitigate internal and durational anti-selection risk. The reason for this assertion is related to what is commonly called the *asymmetric information learning hypothesis.* This means that insurance professionals may acquire more information on policyholders and learn more about the riskiness of policyholders when they are more loyal to the company.

We conclude this section with a discussion of the concept of moral hazard in health insurance.

To illustrate the principle of moral hazard, consider the case of an individual who is issued an automobile insurance. After the policy is issued the individual takes actions (which where previously hidden from the insurer) that increase the probability of an accident and/or the loss amount from the accident. This is an example of moral hazard.

Moral hazard is classified as ex-ante moral hazard and ex-post moral hazard.

Ex-ante moral hazard refers to hidden actions taken by the policyholder that lead to an increase in risk-taking after the policy effective date. In the context of health insurance, it is hardly likely that a policyholder will take actions to become sicker than previously expected. Such actions would seem to be irrational.

Ex-post moral hazard refers to actions taken by the policyholder that lead to a greater consumption of health care than previously expected. It is called 'ex-post' which means that it takes place after the insured event occurred. Upon becoming sick or experiencing injuries from accidents, the policyholder may have a greater propensity to overly consume health service just because he/she is insured. This form of moral hazard is more common in health insurance.

The central prediction of moral hazard theory in health insurance is that policyholders engage less in over-consumption of health care when policy coverage is modified to include deductibles and policy limits. This means that relatively small claim amounts and incremental claim amounts above policy limits are retained by the policyholder. This tends to reduce the frequency and severity of claims made by the policyholder.

We conclude this section with an interesting behaviour issue that has implications for over-consumption in health insurance.

Splitting the Dinner Bill

The moral hazard risk of over-consumption in health insurance can be explained by an interesting article entitled, *The Inefficiency of Splitting the Bill* by Uri Gneezy *et al* in the Economic Journal (2004). This paper studies two ways by which a group of diners pays a restaurant bill. If each pays his/her own bill than each will make a menu choice that best matches his/her preference and budget. There is no free lunch!

But if they agree to split the bill equally, then a selfish diner can enjoy exceptional dinners at bargain prices. To see this point, assume that the group comprises five diners with one selfish diner. Then the selfish diner can enjoy fine menu choices and pay only 20% of the price. A large percentage of the cost is transferred to others. This tends to encourage over-consumption by the selfish diner either in terms of quantity or quality.

This concludes chapter 2.

The next two chapters present procedures for the calculation of single premiums and their equivalent annual level premiums for health insurance that is similar to life techniques commonly referred to as SLT. We utilise multistate models and their special case – multi-decrement models – as a foundation of our presentation. Chapter 5 discusses sickness insurance which is based on methods not similar to life techniques (i.e., NSLT). This health insurance best fits the common features of non-life insurance in terms of frequency and severity of covered risks.

3 HEALTH INSURANCE: SIMILAR TO LIFE TECHNIQUES (PART I)

3.1 INTRODUCTION

As we state in chapter 1, human beings transition between different states of health over their respective lifetimes. For example, some individuals who are currently in a state of good health may enter a state of sickness for a relatively short period of time and recover to a healthy state. Others in a healthy state may die from natural causes or from injuries sustained in an accident. Further, other individuals in a healthy state may incur temporary or permanent disability with varying degrees of severity to their health. For example, disability may be partial or total with implications for the person's ability to perform his/her own occupation. They may subsequently transition to death from a disabled state.

The random movements between states of health of an individual over his/her lifetime create a visualisation of a network of health states (called *directed graphs*) that is amenable to an approach called ***multi-state modeling***. Specifically**,** a multistate model for health insurance assumes that, at any point in time, an individual occupies one state of health. This may be a state of good health or states of varying degrees of sickness or disability. Random events cause the person to leave his/her current state and enter another state.

This discussion emphasises the importance of a concept called *decrement*.

The English Oxford dictionary defines *decrement* as a decrease or a reduction in a quantity. In the context of health insurance, a decrement is a decrease or reduction in an individual's current state of health. Simply put, a decrement creates a transition from the current state of health to another state of diminished health (e.g., sickness or disability) or to an exit state of death.

Decrement comprises two inherent characteristics. First, it indicates a *direction* of the individual's heath status – one of a deterioration or impairment. Second, it indicates a *level* of the deterioration (i.e., severity) which is a key factor in determining the benefit payment for a covered event. For example, total disability is associated with a higher benefit payment compared to partial disability.

It is interesting to compare *decrement* with *increment* which refers to an increase in a quantity. In health insurance, the transition from a state of sickness or temporary disability back to a healthy state is an example of an increment. An increment is also called *reactivation* in actuarial terminology.

We now present the fundamentals of the theory of directed graphs before proceeding to depict health insurance contracts as multi-state models. We keep the discussion at a non-technical level while emphasising its link to providing a visualisation of conventional health insurance contracts.

3.2 DIRECTED GRAPHS

A directed graph is a network which comprises a finite set of *states* and *arrows* that indicate the *direction* of transition between states. To clarify, consider the diagram in Figure 3.1 where each 'circle' is a state which is connected to another state by a *directed arrow* (also called an *edge*).

Figure 3.1: A Two-State Directed Graph

Here are some observations about Figure 3.1.

- The arrow with the tail at **u** and the head at **v** is called an **edge** and is represented by **(u, v).** In other words, the arrow originates at **u** and ends at **v**.

- In health insurance, it is typical for the nodes **u** and **v** to represent states of health for an individual of age *x* years. For example, u = healthy state and v = sick state for an individual who is *x* years old. The arrow = (u, v) = (healthy, sick) indicates a transition of an individual who is *x* years of age from a healthy state to a sick state.

  Similarly, if u = state of being alive and v= state of death, the (u, v) = (alive, death) indicates a transition of a life of age *x* years from being alive to death. This simple directed graph represents the basic insurance model which is discussed in the next section.

3.2.1 THE BASIC LIFE INSURANCE MODEL AS A DIRECTED GRAPH

In this model, a life of age *x* years is currently in state **A** (i.e., a state of being alive) and **D** is a state of death as shown below.

Figure 3.2: The Basic Life Insurance Model

The arrow (A, D) reflects the transition of an individual who is currently in a state of being alive to state of death. The state D is called ***absorbing*** since once a life enters this state, there is no possibility of leaving it.

Comment

Figure 3.2 represents the basic life insurance model as a two-state model with one level of decrement that indicates the transition to death.

The basic life insurance model does not consider the *causes* of death that may include natural causes due to age, critical illnesses including heart failure and cancer or accidental death. A life insurance actuary focuses on estimating the expected future lifetime of a life of x years.

An extension of this model is one that separates the cause of death – that is, accidental death that pays a higher benefit compared to the benefit payable for death by all other means. Here is a directed graph of this life insurance model with two levels of decrement – accidental death and death from all other causes:

In this model, a life of age x years is currently in state **A** (i.e., alive); $\mathbf{D}^o$ is an exit state due to death from other causes except by accident; $\mathbf{D}^a$ is an exit state due to death from accident. The directed graph for this insurance contract is shown below.

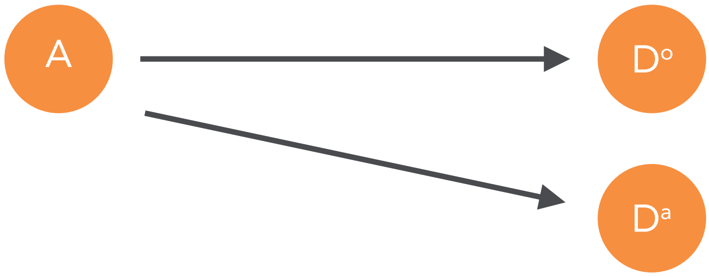


Figure 3.3: Accidental Death Insurance

The following comments for this accidental death insurance apply to Figure 3.3.

a) The life of x years is currently in a state of being alive at the effective policy date. From this state, there are two paths to enter exit states. Both of these exit states are ***absorbing***.

Comment

Figure 3.3 is an example of a **multiple-decrement model**. By this we mean that all directed arrows starting from the current state (being alive, A) lead to exit states. There are no intermediate states between the current state and the exit states. By intermediate states, we mean states which are not absorbing. States which are not absorbing are called ***transient*** states.

We consider multi-decrement models in Chapter 4.

b) The model in Figure 3.3 may be described as a three-state model with two levels of decrement.

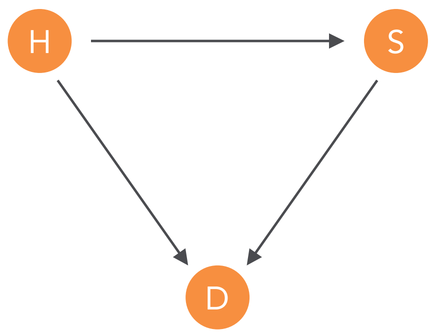
Comment

We present a key difference between multi-decrement models and multi-state models of health insurance. A multi-state model has at *least* one transient state – wherein individuals may enter and, importantly, are able to leave for another state A multi-decrement model has no transient state. All states are exit states (also called *stopping* or absorbing) in that the individual cannot leave these states.

We now provide directed graphs of health insurance to visualise multi-state models of health insurance.

3.2.2 THE POLLARD MODEL AS A DIRECTED GRAPH

Probably the simplest example of a multi-state model is the so-called Pollard (1980) model that permits an interaction between mortality and morbidity. Here is a depiction of this model.



Notes:
H = state of healthy; **S** = state of sickness; **D** = state of death.

Figure 3.4: The Pollard (1980) Model

This is three-state model wherein an individual of age x years is currently in the healthy state, H. There are two ways to exit the healthy state. These are represented by the edges (H, S) and (H, D). Both of these edges represent a deterioration in the health status of (x). In other words, each is a decrement. (Technically, there are two *causes of decrement.)* For example, the edge (H, S) is a transition from a healthy state to a state of sickness. The other edge (H, D) is a transition to death.

Comment

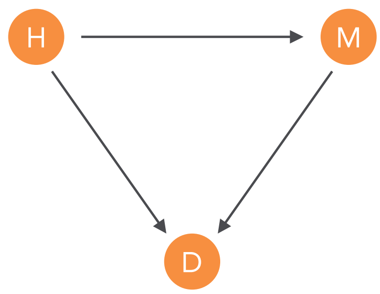
The Pollard model is a three-state model with two levels of decrement. *As a rule of thumb, the number of levels of decrement is the number of edges (i.e., number of directed arrows) that originates from the healthy state which is typically the starting state.* Note that sickness is a transient state since the life can leave this state to another state.

The edge (S, D) is a transition from sickness to death which is also a decrement since is a reduction in the current state, S. This is called a **second-order decrement**.

Here is an example illustrating the concept of decrement adapted from the following illness-death model presented by Artur Araujo *et al* in the Journal of Statistical software (2014).

Example 1

In this model of health insurance, an individual of age x years is currently in the healthy state, H.



Notes: H = healthy, M = diseased, D = death.

Which statement is correct?

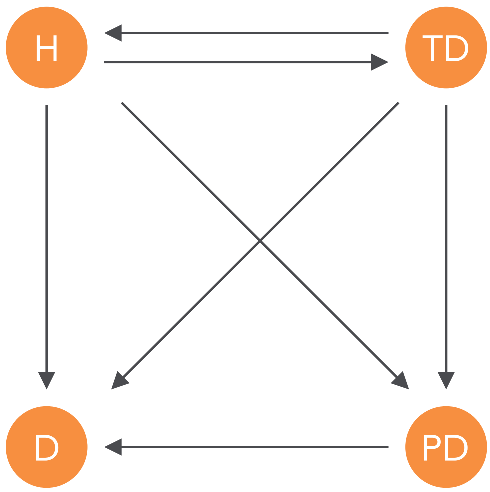
a) The edge (H, M) is an exit state.
b) The edge (H, D) is a transient state.
c) This is a 3-state model with 2 levels of decrement.
d) This a 3-state model with 3 levels of decrement.

Answer

This is a 3-state model since there are three nodes – healthy, diseased and dead. Since the individual is currently in state H, decrement refers to a deterioration in health status upon leaving the healthy state, H. Decrements are evidenced by (H, M) and (H, D). There are 2 levels of decrement. This is also obtained as the number of edges originating from the healthy state which is the starting state. The correct answer is c).

3.2.3 MODELING DISABILITY INSURANCE

Here is a diagram of a disability insurance model that begins with a healthy individual of age *x* years who can enter states of temporary disability, permanent disability death. Notes are provided in Figure 3.5 and explanations are provided immediately afterwards.



Notes:
H = state of a healthy individual of age *x* years;
TD = state of temporary disability;
PD = state of permanent disability; D = state of death.
Therefore, there are four states in this disability model.

Figure 3.5: Disability Insurance as a Directed Graph

We explain each 'edge' in the diagram as follows:

a) A healthy individual of age *x* years leaves state H and enters the absorbing state of death, D. *This is the edge (H, D) and is the basic life insurance model.*

b) A healthy individual of age*x* years leaves state H and enters the state of temporary disability, TD. The individual may recover from TD and re-enter state H. This is the reason for the edge (H, TD) and the recovery edge (TD, H) which is an example of *reactivation.*

This process of Heathy to Temporary Disability and recovery back to Healthy may occur more than once during the term of the policy. Hence, there is a similarity to non-life insurance where the policyholder may make more than once claim over the term of the policy. ***We consider the case for health insurance similar to non-life techniques in Chapter 5.***

c) Once in the state TD, the individual may not recover and his/her health status may worsen to permanent disability, PD or death, D. In our diagram, these are represented by the edges (TD, PD) and (TD, D).

d) If the individual enters the state PD then he/she can enter the state of death, D. This is represented by the edge (PD, D).

e) In b) above we considered the case where the individual leaves the state of healthy and enters temporary disability. But the individual may leave H and enter permanent disability, PD.
This is represented by the edge (H, PD). Finally, edge (PD, D) is also possible.

Comment

The examples of conventional health insurance described above as directed graphs show that the random movements of a life of x years can be quite complex to model. In addition, from an actuarial perspective, it is important to estimate the probability that a life of x years will leave a state and enter another state.

This brings us to the concept of *transition probability*.

3.3 TRANSITION PROBABILITIES

To explain the concept of transition probabilities, we refer to a two-state generic model for basic life insurance described in Figure 3.1.

Consider a life (x) currently in state **u**. This is the initial state from where the transition begins and time is equal to zero.

*We want to symbolise the probability that this individual is in state **v** after t years (i.e., at age x +t years) knowing that he/she is currently in state **u**.*

Definition

The symbol $_{t}\mathbf{p}_{x}^{uv}$ is the probability of a life of x years of age is in state **v** at age $x + t$ years**,** knowing the same life is currently in state **u**.

Comment

By convention, for t =1, $_{1}\mathbf{p}_{x}^{uv}$ is written as $\mathbf{p}_{x}^{uv}$and is interpreted the probability that a life of age *x* years is in state **v** after one year (or at age *x+1*), knowing that the life in currently in state **u**.

Important Assumption

There is a key assumption in the definition of transition probability. As stated above, the actuarial symbol $_{t}\mathbf{p}_{x}^{uv}$ is the probability that a life of *x* years of age is in state **v** at age $x + t$ years**,** knowing the same life is currently in state **u**. In simple words, transition probability applies only to the ***current state*** that the life occupies. The ***previous history*** of this life moving from one state of health to another state does not matter.

To provide some insight into this assumption, we consider a simple example. Suppose we want to calculate the probability that a life which is 60 years of age will have a stroke in one year from today. Then only the current health status of (60) matters. The previous health history does not matter for the calculation of the transition probability. Whether this life had a stroke in the past is not taken into account. This is one reason why the transition probability is sometimes called *memoryless.*

Comment

This is may be viewed as an unrealistic assumption. However, it is a common assumption in actuarial approaches to multi-state models and goes under the name of **Markov Chains**. The application of Markov chains in insurance to estimate premiums and loss reserves is common.

Here is an example illustrating this actuarial symbol for transition probabilities.

Example 2

Consider the basic life insurance model:

where **A** = healthy state and **D** = state of death.

Explain the symbol $\mathbf{p_{60}^{AD}}$.

Answer

The symbol $\mathbf{p_{60}^{AD}}$ is the probability that a life currently 60 years of age is dead at age 61 years knowing that this individual is currently alive. It is important to note that the symbol **AD** means that the current state is **A** and the state one year later is **D**. This is shown in the graph above with the arrow indicating that the direction of the transition is from **A** to **D**.

Comment

The symbol $\mathbf{p_{60}^{AA}}$ is the probability that a life currently 60 years of age is in the healthy state at age 61 years, knowing that this life is currently in the healthy state. Simply put, it is the probability that the life ***remains*** in the healthy state after one year. This is called the *occupancy probability*.

We now consider the Pollard model which we described in the previous section to illustrate the calculation of transition probabilities in discrete time.

3.4 CALCULATION OF TRANSITION PROBABILITIES (DISCRETE TIME MODEL)

The Pollard model has three states – H, S and D. It is convention to write the associated one-year transition probabilities in a matrix as follows:

$$\begin{pmatrix} & H & S & D \\ H & 0.75 & 0.15 & 0.10 \\ S & 0.20 & 0.66 & 0.14 \\ D & 0 & 0 & 1 \end{pmatrix}$$

The following information is obtained from this matrix of transition probabilities for a life of age 60 years.

- The rows in the matrix are the *current* states. The states in the columns are the states that the life of 60 years may enter in one year.

- The entry in the cell with value 0.75, is the probability that a life of age 60 years is in the healthy state (H) in one year, knowing that the life is currently in the healthy state (H). Similarly, the entry in the cell with value 0.15 is the probability that a life of age 60 years is in the sick state (S), knowing that the life is currently in the healthy state (H). The entry in the cell with value 0.10 is the probability that a life of age 60 years is in the state of death, knowing that the life is currently in the healthy state (H).

- The entry in the cell with value 0.20, is the probability that a life of age 60 years is in the healthy state (H) in one year, knowing that the life is currently in the sick state (S). Similarly, the entry in the cell with value 0.66 is the probability that a life of age 60 years is in the sick state (S) in one year, knowing that the life is currently in the sick state (S). The entry in the cell with value 0.14 is the probability that a life of age 60 years is in the state of death in one year, knowing that the life is currently in the sick state (S).

- It is noted that death is an absorbing state. Therefore, there are no further transitions from this state. This is the reason for the zero values in the third row. The entry with a value of 1 is the probability that a life of 60 years is in the death state in one year, knowing that the life is currently in the death state. This reflects the adage – *once dead, remains dead.*

- The sum of probabilities in each row is 1.

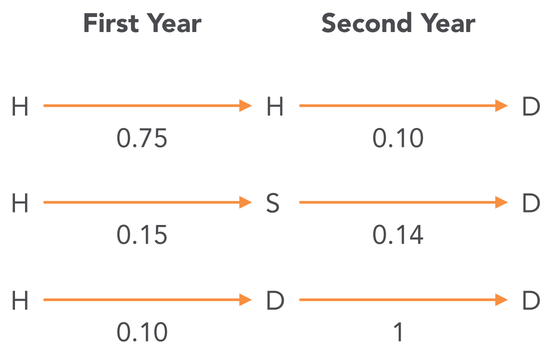
- The diagonal of the matrix with values (0.75, 0.66, 1) are occupancy probabilities. In each case, the values represent the probability that a life of age 60 years remains in the current state in one year.

Example 3

Refer to the matrix of transition probabilities presented above. Calculate the transition probability that a life of age 60 years is in the state of death (D) in two years, knowing that the life is currently in a healthy state (H). The actuarial symbol is ${}_2\mathbf{p}_{60}^{\mathbf{HD}}$.

Answer

There are three paths to death in two years for a life who is 60 years of age and who is currently in a health state. These are as follows where the corresponding transition probabilities are indicated.



The transition probability = $(0.75 \times 0.10) + (0.15 \times 0.14) + (0.10 \times 1) = 0.196$ =19.6%. This is the probability that a life of age 60 years will be in the state of death in two years, knowing that the life is in the health state currently.

Comment

These three paths to death may be expressed in terms of edges in a directed graph model. They are as follows:

a) The path H → H → D is equivalent to (H, H) + (H, D) with probabilities 0.75 and 0.10 respectively. Based on the *independence* of edges, the probability of this path to death is $0.75 \times 0.10 = 0.075$.
b) The path H → S →D is equivalent to (H, S) + (S, D) with probabilities 0.15 and 0.14 respectively. Based on the *independence* of edges, the probability of this path to death is $0.15 \times 0.14 = 0.021$.
c) The path H → D → D is equivalent to (H, D) + (D, D) with probabilities 0.10 and 1 respectively. Based on the *independence* of edges, the probability of this path to death is $0.10 \times 1 = 0.100$.

The probability that a life of age 60 years will be in the state of death in two years, knowing that the life is in the health state currently is the sum of the probability for each probability which is equal to $0.075 + 0.021 + 0.100 = 0.196$ which is 19.6%.

Example 4

The probability that a life of age 60 years will be in the state of death in two years, knowing that the life is in the sick state currently is which one of the following?

a) 0.2524
b) 0.0924
c) 0.0200
d) 0.1400

Answer

There are three paths from being currently sick (S) to be in the state of death in two years. These are:

a) The path S → S → D which is equivalent to (S, S) + (S, D) with probabilities 0.66 and 0.14 respectively. Based on the independence of edges, the probability of this path to death is 0.66 × 0.14 = 0.0924.
b) The path S → H →D which is equivalent to (S, H) + (H, D) with probabilities 0.20 and 0.10 respectively. Based on the independence of edges, the probability of this path to death is 0.20 × 0.10 = 0.020.
c) The path S → D → D is equivalent to (S, D) + (D, D) with probabilities 0.14 and 1 respectively. Based on the independence of edges, the probability of this path to death is 0.14 × 1= 0.140.

The probability that a life of age 60 years will be in the state of death in two years, knowing that the life is in the health state currently is the sum of the probability of each path which is equal to 0.0924 + 0.020 + 0.140 = 0.2524 = 25.24%. The correct answer is a).

We now describe the procedure to estimate premiums for health insurance contracts based on the equivalence principle.

3.5 PREMIUMS FOR HEALTH INSURANCE (BASED ON MULTI-STATE MODELS AND EQUIVALENCE PRINCIPLE)

We describe the procedure to estimate premiums based on multi-state models and equivalence principle by means of a generic example.

The second eBook on Life Insurance provides a detailed discussion on the application of the equivalence principle for estimating net premiums. In brief, this principle states that the net premium is calculated so that at the policy effective date, the expected value of future loss is zero. This implies that the expected present value of premiums paid to the insurer is equal to the expected present value of benefits paid by the insurer to the insured.

In the case of a single premium paid at the policy effective date, we have:

$$\mathbf{P_0 = EPV(B_T)}$$

Where P_0 = single premium payment by the insured; EPV (B_T) is the expected present value of the future benefit payable at time T, on the incurrence of the covered risk.

We provide a brief review of the equivalence principle, we present a comprehensive example showing the calculation procedure for single premium and annual level premium respectively.

Example 5

Consider a three-year term insurance with the following features (similar to the Pollard model).

a) Policyholders may be in one of three states at the beginning of each year – healthy (H), disabled (S), or dead (D). The starting state is healthy. The disabled life may recover to a healthy state or may enter the state of death.
b) The annual transition probabilities are provided in the following matrix:

$$\begin{pmatrix} & H & S & D \\ H & 0.8 & 0.1 & 0.1 \\ S & 0.1 & 0.7 & 0.2 \\ D & 0 & 0 & 1 \end{pmatrix}$$

c) A 100,000 benefit is payable at the end of the year of death.
d) The policyholder pays a premium only if he/she is alive at the beginning of each year.
e) The interest rate is 10% which results in a discount factor $v = \frac{1}{1+i} = 0.9091$.

Questions

a) What is the net single premium based on the equivalence principle?
b) What is the annual level net premium based on the equivalence principle?

Answer

We present all possible outcomes for each year as follows:

First Year:

The only path to death (D) from the starting state (H) is the edge (H,D) with probability = 0.1.

The probability of death in the first year is 0.10

Second Year:

There are two possible paths to death. These are:

i) Remaining in the healthy state in the first year and dying in the second year; this path is symbolised as (H, H) + (H, D) with probability (0.8) × (0.1) = 0.08.
ii) Entering the disability state in first year and dying in the second year; this path is symbolised as (H, S) + (S, D) = 0.1 × 0.2 = 0.02

The sum of the probabilities for year 2 is 0.08 + 0.02 = 0.10

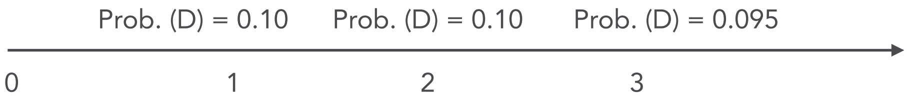
Third Year:

There are four possible paths to death. These are:

i) Remaining healthy for the first years and dying in the third year; this path is symbolised as (H, H) + (H, H) + H, D) with probability (0.8) × (0.8) × (0.1) = 0.064.
ii) Remaining healthy in the first year, becoming disabled in the second year and dying in the third year; this path is symbolised as (H, H) + (H, S) + (S, D) = (0.8) × (0.1) × (0.2) = 0.016.
iii) Becoming disabled in the first years and dying in the third year; this path is symbolised as (H, S) + (S, S) + (S, D) = (0.1) × (0.7) × (0.2) = 0.014.
iv) Becoming disabled in the first year, recovering to the healthy state in the second year and dying in the third year; this path is symbolised as (H, S)+ (S, H) + (H, D) = (0.1) × (0.1) × (0.1) = 0.001.

The sum of the probabilities for year 3 is 0.064 + 0.016 + 0.014 + 0.001 = 0.095

We can now graphically depict the three-year insurance with the probability of death at the end of each year as follows:



We now illustrate the calculation of the net single premium for this insurance as follows:

End of Year 1	Probability of Death = 0.10	Benefit Payment = $100,000	Expected Value = 0.10 × $100,000 = $10,000	Actuarial Discount Value = $10,000 × Discount Factor = $10,000 × 0.9091 = $9,090.91
End of Year 2	Probability of Death = 0.10	Benefit Payment = $100, 000	Expected Value = 0.10 × $100, 000 = $10,000	Actuarial Discount Value = $10,000 × Discount Factor = $10,000 × $(0.9091)^2$ = $8,264.46
End of Year 3	Probability of Death = 0.095	Benefit Payment = $100, 000	Expected Value = 0.095 × $100, 000 = $9,500	Actuarial Discount Value = $9,500 × Discount Factor = $9,500 × $(0.9091)^3$ = $7,513.15
Net Single Premium				$9,090.91+$8,264.46+$7,513.15= $24,868.52

To obtain the equivalent annual level premium, we note that the policyholder pays a premium only if he/she is alive at the beginning of each year.

Obviously, the probability of being alive at the beginning of the first year is 1.

The probability of being alive at the beginning of the second year is the probability of the path (H, H) +(H, S) = 0.8 +0.1 = 0.09.

To obtain, the probability of being alive at the beginning of the third year, there are four paths as follows:
(H, H) + (H, H); (H, H) + (H, S); (H, S) + (S, S) and (H, S) + (S, S). The sum of the corresponding probabilities is:

(0.8 × 0.8) + (0.8 × 0.1) + (0.1 × 0.1) + (0.1 ×0.7) = 0.80.

Hence the annual equivalent level premium (A), is calculated as follows:

$$\mathbf{A + (v \times 0.80 \times A) + (V^2 \times 0.80 \times A) = 24{,}868.52 \text{ (i.e., net single premium)}}$$

Noting that v = 0.9091, we obtain A × 2.38845 = 24,868.52. Hence A = $10,411.99

This concludes the answer to this comprehensive example.

Chapter 4 considers a special case of the multi-state model called the multi-decrement model for health insurance.

4 HEALTH INSURANCE: SIMILAR TO LIFE TECHNIQUES (PART II)

4.1 INTRODUCTION

This chapter discusses and applies a special case of multi-state models called ***multi-decrement models***. This is the case where ***all*** transitions are between a starting state and exit states. The policy expires in the exit state.

Comment

An example which is not an exit state is temporary disability since there is a potential for further transition to permanent disability or death. Simply put, it is possible to leave the state of temporary disability to enter another state of health. States from which further transitions are possible are called ***transient***.

Decrements that lead only to exit states include the following examples:

a) a life insurance policyholder is alive (i.e., starting state) and surrenders the policy (exit state); the policy expires.
b) a life insurance policyholder is alive (i.e., starting state) and dies (exit state); the policy expires on payment of sum insured.
c) a critical illness insurance policyholder is healthy (i.e., starting state) and a diagnosis reveals the presence of a covered critical illness (exit state); the policy expires on payment of sum insured for critical illness or death which ever comes first.
d) an insurance policy provides coverage for permanent disability and death whichever comes first; the policy expires upon payment of sum insured.

For the remainder of this eBook, we will adopt the common convention in the actuarial literature whereby states are identified by integers **0, 1, 2, …, k**. The initial state is labeled **0**.

As an example, we restate from the previous chapter, accidental death insurance as a directed graph using the notation based on successive positive integers. In this model, a life of age x years is currently in state **0** (i.e., alive); **1** is an exit state caused by death from other causes except by accident; **2** is an exit state due to death from accident. The directed graph for this insurance contract is shown below.

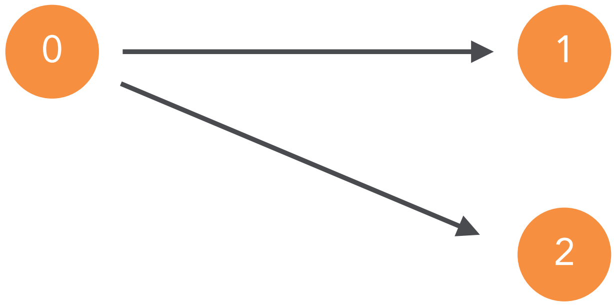


Figure 4.1: Accidental Death Insurance

The following comments on the accidental death insurance apply to Figure 4.1.

a) The life of x years of age is currently in a state of being alive at the effective policy date. From this state, there are two causes of decrement leading to exit states.

The following is an example of a **multiple-decrement model**. To review, this means that all directed arrows starting from the initial state lead to exit states. A life of x years of age may occupy the initial state of being alive for a long time, but any potential transition leads to an exit state.

A general multi-decrement model of k decrements is depicted in Figure 4.2 below.

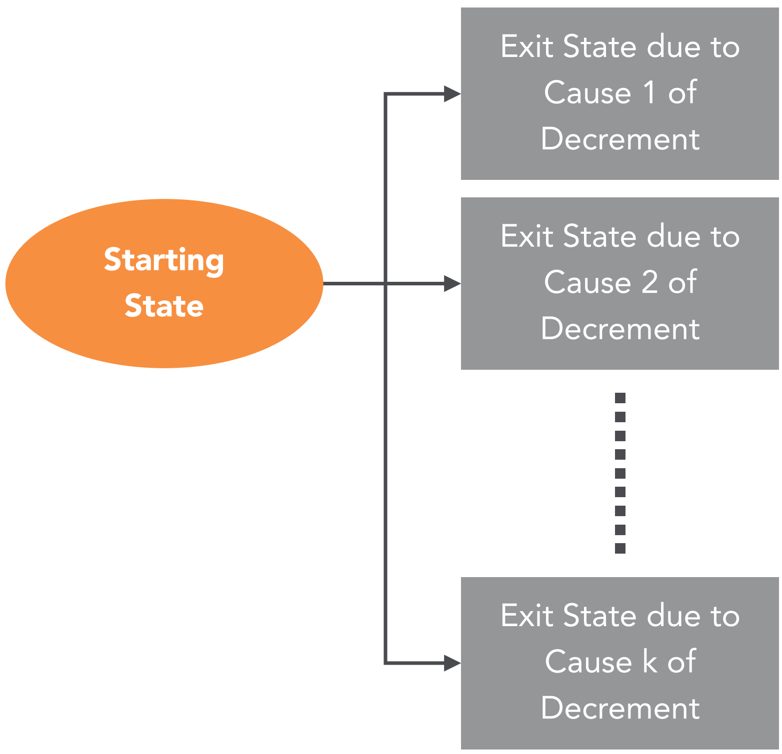


Figure 4.2: Multi-decrement Model (k Causes of Decrement)

This model has **k+1** states with **k** causes of decrement. There is a single starting state (e.g., state of being alive) and a life of x years of age has **k** possible modes of leaving the starting state and entering an exit state. It is called an exit state because in this state, the insurance contract expires. As stated above, reasons for this, include – a life of x years of age dies, a policyholder surrenders the contract, a diagnosis identifies a covered critical illness.

In keeping with our focus on the art of insurance, we consider only the discrete case of multi-decrement models for health insurance. The main reason is that the discrete case permits the utilisation of data from multi-decrement life tables.

We now present a typical multi-decrement life table, discuss its information content and show how to estimate transition probabilities. This will facilitate the estimation of premiums for health insurance based on the equivalence principle and multi-decrement models.

4.2 MULTI-DECREMENT LIFE TABLE

In the second eBook, we presented details on the information content of the Illustrative Life Table. The main feature of this conventional life table is that it provides information for an actuary to estimate mortality probabilities. The primary focus is on using the information provided in the life table to estimate the expected future lifetime of a life of x years of age.

Here is an example that is intended as review of the application of a conventional life table and as a backdrop for comparison with a multi-decrement life table.

Example 1

The following information is an extract from the Illustrative Life Table which is labeled as Table 4.1 below.

Age (x)	Number of living at Age x; l_x	Number of Deaths, d_x
0	100,000	632
1	99,368	
…		
20	96,178	99

Age (x)	Number of living at Age x; l_x	Number of Deaths, d_x
21	96,079	
…		
40	93,132	259
41	92,873	
…		
50	89,509	530
51	88,979	

Table 4.1 Excerpt from a Conventional Life Table

This table shows that there were 100,000 newborns (i.e., age 0) and by the end of the first year, there were 632 deaths. Hence the number of persons in the cohort living on there first birthday is 100,000 – 632 = 99,368. The probability that a newborn will die before his/her first birthday is 632/100,000 = 0.00632.

Similarly, the probability that a life of age 40 years will survive the next ten years is the ratio of the number of living at age 50 years divided by the number of living at age 40 years = 89,509 / 93,132 = 0.961098 = 96.1098%. Hence the probability that a life of age 40 years will die before the age of 50 years is (100 – 96.098) % = 3.8902%.

With this review of a conventional life table for life insurance, we present an excerpt from a typical multi-decrement life table.

Excerpt from a Multi-Decrement Life Table

For simplicity, we consider two causes of decrement. Cause 1 is accidental death and cause 2 is death except from accident.

In keeping with convention, we will label 'cause j of decrement' as 'decrement j'. In addition, 'deaths' are viewed as 'exits'. We also use the common actuarial notation (x) to represent a life of x years of age.

Age (x)	Number of living at Age, x	Number of Exits due to Decrement 1	Number of Exits due to Decrement 2
50	89,509	150	380
51	88,979	162	410
52	88,407	168	448
53	87,791		

Table 4.2: Excerpt from a Multi-Decrement Life Table

We now explain the data included in this multi-decrement life table.

- The number of individuals living at age 50 years is 89,509. During the year, there were 150 accidental deaths and 380 deaths not due to accidents. The number of deaths due to all causes is 150 + 380 = 530. Simply put, this is the total number of individuals who left the cohort within the year. Hence at age 51, the number of individuals living (after both causes of decrement) is equal to 89,509 – 530 = 88,979.

  Similarly, for the cohort of age 52 years, the total number of individuals who left within one year is 168+448 = 616.

- Using conventional actuarial notation, we restate the explanation made above as follows:

  $l_x^{(\tau)}$= number of individuals in a cohort who are *x* years of age after all decrements. In other words, it is the *surviving* cohort of age *x* years. The symbol Greek τ is pronounced 'tau' refers to decrements.

  From the table, $l_{50}^{(\tau)}$is the number of individuals who are 50 years of age after all causes of decrement. Table 4.2 shows that $l_{50}^{(\tau)}$= 89,509 and $l_{51}^{(\tau)}$= 88,979.

- The symbol ${}_t\mathbf{d}_x^{(j)}$ is the expected number exits (deaths in the case of mortality) in a cohort of age *x* years within the next *t* years due to decrement *j*. Simply put, it is the expected number of individuals who are *x* years of age who leave the group before the age of *x* + *t* years.

  For example, ${}_1\mathbf{d}_{50}^{(2)}$ is the number of individuals of *x* years old who leave the group within one year due to decrement 2. By convention, ${}_1\mathbf{d}_{50}^{(2)}$ is written as $\mathbf{d}_{50}^{(2)}$ and has a value of 380 in Table 4.2. This means that 380 individuals are expected to leave the group within the next year due to death except by accident.

  Similarly, ${}_3\mathbf{d}_{50}^{(2)}$ is the expected number of individuals who leave the cohort of age 50 years within the next three years due to death except by accident. This value is equal to $\mathbf{d}_{50}^{(2)}+\mathbf{d}_{51}^{(2)}+\mathbf{d}_{52}^{(2)}$ = 380+410+448 = 1,238.

  Finally, ${}_2\mathbf{d}_{51}^{(1)}$ is the expected number of individuals who leave the group of age 51 years within the next two years due to accidental death. This value is equal to $\mathbf{d}_{51}^{(1)}$ $+\mathbf{d}_{52}^{(1)}$ = 162+ 168 = 330.

Example 2

Fill the missing data in cells marked 'a' and 'b' in the multi-decrement table below:

Age (x)	Number in cohort (or group) at Age, x	Exits Due to Decrement 1	Exits due to Decrement 2
60	81,881	376	751
61	**(a)**	385	**(b)**
62	79,542		

Answer

The number in cohort at age 61 years = number in cohort at age 60 years – (exits due to both decrements) = 81,881 – (376+751) = 80,754.

In actuarial terminology, $l_{61}^{(\tau)} = l_{60}^{(\tau)} - d_{60}^{1} - d_{60}^{2}$ = 81,881- 376 -751= 80,754. This is the value for the cell marked (a).

To obtain the value for the cell marked (b), the number in the cohort at age 61 years – the number in the cohort at age 62 years = 80,754 – 79,542 = 1,212. This is the total number of individuals leaving the group due to both decrement 1 and decrement 2. Therefore, the number of individuals leaving due only to decrement 2 is 1,212 – 385 = 827.

Example 3

Referring to Table 4.2, which one of the following is the value of ${}_{2}\mathbf{d}_{51}^{(2)}$?

a) 330
b) 1,238
c) 858
d) 1,076

Answer

The actuarial symbol ${}_{2}\mathbf{d}_{51}^{(2)}$ represents the number of individuals of age 51 years old who are expected to leave the group within the next two years due to decrement 2. This is equal to $\mathbf{d}_{51}^{(2)} + \mathbf{d}_{52}^{(2)}$ = 410 + 448 = 858.

We discussed and defined the concept of transition probability in chapter 3. For multi-decrement models, the only possibility of transition is to an exit state and this specifies the concept of transition probability. In addition, the actuarial notation is different compared to multi-state models. We consider these issues next.

4.3 TRANSITION PROBABILITIES IN MULTI-DECREMENT MODELS

Before we consider the method for the calculation of transition probabilities using multi-decrement life tables, we introduce additional actuarial notation in Box 4.1 below. Examples illustrating the calculation of transition probabilities follow.

BOX 4.1 (Actuarial Notation and Transition Probabilities for Multi-Decrement Models)

- We already introduced and explained the actuarial symbol $l_x^{(\tau)}$ which is the surviving cohort of age x years after taking into account, exits due to all decrements.
- Similarly, the symbol ${}_t d_x^{(j)}$ is explained above. As a reminder, it represents the number of individuals of age x years who are expected to leave the current cohort within the next *t* years due to decrement *j*.
- The symbol ${}_t d_x^{(\tau)}$ represents the number of individual who are expected to leave the current cohort within the next *t* years due to **all** decrements. Therefore,

$${}_t d_x^{(\tau)} = l_x^{(\tau)} - l_{x+1}^{(\tau)}. \tag{4.1}$$

For example from Table 4.2, $d_{50}^{(\tau)} = l_{50}^{(\tau)} - l_{51}^{(\tau)}$ = 89,509 – 88,979 = 530 is the total number of individuals who are expected to leave within the next year due to both decrements.

- **Transition Probability**

The symbol ${}_t q_x^{(j)}$ is the probability that an individual of x years of age leaves the cohort within the next *t* years due to decrement *j*.

The formula for its calculation is as follows:

$${}_t q_x^{(j)} = \frac{{}_t d_x^{(j)}}{l_x^{(\tau)}}. \tag{4.2}$$

This equation states the probability that an individual of x years of age leaves the cohort within the next *t* years due to decrement *j* is equal to the number of individuals who leave within the next *t* years due to decrement *j* expressed as a ratio of the number of individuals in the (current) cohort.

For example, the symbol ${}_2 q_{60}^{(1)}$ represents the probability that an individual who is 60 years of age leaves the current cohort before the end of the next two years due to decrement 1.

- We now state the probability of leaving the current cohort ***due to all causes of decrement*** within the next t years. Recall that symbol τ is the actuarial notation for 'all causes of decrement'. Analogous to equation (4.2), we have:

$${}_t q_x^{(\tau)} = \frac{{}_t d_x^{(\tau)}}{l_x^{(\tau)}}. \tag{4.3}$$

This equation states that the probability that an individual of *x* years of age leaves the current cohort within the next *t* years due to all causes of decrement is equal to the number of individuals who leave due to all causes of decrement over the next *t* years expressed as a ratio of the number of individuals in the current cohort.

- If we consider the case for one year, (i.e., t = 1), equation (4.3) reduces to:

$$q_x^{(\tau)} = \frac{d_x^{(\tau)}}{l_x^{(\tau)}}. \qquad (4.4)$$

- We now consider the probability that the individual ***remains*** in the group for at least *t* years ***after taking into account all causes of decrement***. Simply, this is the probability that a life (x) is in the group at age x + t years.

Based on the interpretation of equation (4.3), we obtain:

$$ {}_tp_x^{(\tau)} = 1- 1- {}_tq_x^{(\tau)} \qquad (4.5)$$

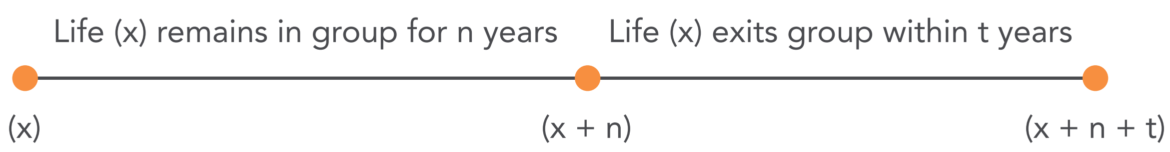
Equivalently, $ {}_tp_x^{(\tau)} = \frac{l_{x+t}^{(\tau)}}{l_x^{(\tau)}}.$ (4.6)

- The case where the individual ***remains*** in the group for at least *t* years ***after taking into account only decrement j*** is given by the equation:

$$ {}_tp_x^{(j)} = \frac{l_{x+t}^{(j)}}{l_x^{(\tau)}}. \qquad (4.7)$$

- **The Probability of Deferred Exit**

First recall the notation (x) which means a life of x years old. We consider the probability that a life of *x* years of age remains in the cohort (or group) for a period of time and exits in the subsequent period. This is an important case and used in estimating premiums for health insurance based on multi-decrement models. Here is a graphical depiction of deferred exit for decrement *j*. Life (x) remains in the group for *n* years and exits the group during the next *t* years.



The probability that a life (x) remains in the group at age x + n is given by the formula:

$_{t}p_x^{(j)} = \dfrac{l_{x+t}^{(j)}}{l_x^{(\tau)}}$ as stated in (4.7). The probability that a life (x) exits the group before the next *t* years is $_{t}q_{x+n}^{(j)} = \dfrac{_{t}d_{x+n}^{(j)}}{l_{x+n}^{(\tau)}}$ as stated in (4.2). By virtue of the independence of states, the probability of deferred exit is the product of the probability that (x) remains for *n* years and the probability that (x) exists in the next *t* years.

Conclusion: The probability that (x) remains in group for *n* years and exits within the next *t* years is given by the expression, $\dfrac{_{t}d_{x+n}^{(j)}}{l_x^{(\tau)}}$. (4.8)

What is the interpretation of (4.8)?

It states that the probability of deferred exit is the ratio of the number of exits in the subsequent period expressed as ratio of the size of the current group.

Here is an example based on Table 4.2 which we restate for reference.

Age (x)	Number of living at Age x *(After all Causes of Decrement)*	Number of Exits due to Decrement 1	Number of Exits due to Decrement 2
50	89,509	150	380
51	88,979	162	410
52	88,407	168	448
53	87,791		

Example

This example is based on decrement 1. Calculate the probability that a life (50) remains in the group until the age of 51 and exits during the next two years (i.e., ages, 51 and 52).

Based on (4.8), we have n=1 and t=2. Equation (4.8) states that we consider the number of exits at age 51 years and age 52 years which are 162 and 168 respectively. The current group size (i.e., at age 50 years) is 89,509. Hence the probability of life (50) remains in the group for one year and leaves the group within the next 2 years due to decrement 1 is (162 + 168) / 89,509 = 0.003687.

Here are examples illustrating the actuarial notation and equations for transition probabilities in Box 4.1.

Example 4

Referring to Table 4.2 above, the probability that an individual who is 50 years of age will leave the surviving cohort due to accidental death (decrement 1) within one year is which one of the following?

a) 0.001676
b) 0.004245
c) 0.005921
d) 0.006893

Answer

The answer to this question requires the calculation of $\mathbf{q_{50}^{(1)}}$ = 150/ 89,509 = 0.001676. The correct answer is a). Note that if the question asked for the same probability but for decrement 2, then the answer is 380/89,509 = 0.004245 which is b).

Example 5

Referring to Table 4.2, what is the probability that an individual of 50 years of age will leave the cohort after taking into account both causes of decrement within the next year?

Answer

Using equation (4.4) and noting that τ is the actuarial symbol for all causes of decrement, we calculate $\mathbf{q_{50}^{(\tau)}}$ = (150+380) / 89,509 = 0.005921.

Example 6

Referring to Table 4.2, what is the probability that an individual of 50 years of age will remain in the cohort regardless of both causes of decrement at age 51 years?

Answer

Reasonably, since the probability of leaving the group over the next year is 0.005951 (as shown in example 6), then the probability of remaining for at least one year is 1- 0.005951 = 0.994079. This result is formally stated in equation (3.5) in Box 4.1.

The same result is obtained from equation (4.6) in Box 4.1 which states that $p_{50}^{(\tau)} = \frac{l_{51}^{(\tau)}}{l_{50}^{(\tau)}}$ =88,979/89,509 = 0.994079.

Our examples to this point demonstrated the calculation of transition probabilities within the next year.

We now consider similar calculation procedures *for more than one year.*

Example 7

Refer to Table 4.2. What is the probability that an individual of age 51 years will leave the current cohort within the next two years due to decrement 2?

Answer

The key information required is the expected number of individuals who will leave the current cohort due to decrement 2 within the next two years. This was considered in example 5 above. There we calculated ${}_2d_{51}^{(2)}$ = 410 + 448 = 858. Since the size of the cohort of age 51 years old is equal to 88,979, the probability that an individual of age 51 years will leave within the next two years due to decrement 2 is 858/88,979 = 0.009643.

Formally, ${}_2q_{51}^{(2)} = \frac{{}_2d_{51}^{(2)}}{l_{51}^{(\tau)}}$ = 858/ 88979 = 0.009643. This is an application of equation (4.2).

Comment

From example 7, we infer that the probability of an individual of age 51 years will remain in the current cohort apart from decrement 2 for at least two years is 1 – 0.009643 = 0.990357.

We conclude this section with a more detailed example involving decrements in health insurance.

Example 8

Consider the following excerpt from a multi-decrement life table:

Age (x)	Number of living at Age x, $l_x^{(\tau)}$	Decrement 1 (Lapse)	Decrement 2 (Death from Other Causes)	Decrement 3 (Accidental Death)
60	81,881	376	415	251
61	80,839	411	456	289
62	79,683	459	501	314
63	78,409			

Questions

a) What is the probability of an individual of age 60 years will lapse within the next year?

Answer

We calculate $q_{60}^{(1)}$ = 376/81,881 = 0.00458= 0.458%.

b) What is the probability that an individual who is 60 years of year will leave the current cohort within the next year due to all causes of decrement?

Answer

We calculate $q_{60}^{(\tau)} = \dfrac{d_{60}^{(1)} + d_{60}^{(2)} + d_{60}^{(3)}}{l_{60}^{(\tau)}}$ = (376+415+251)/81,881=0.012726 = 1.2726%

c) What is the probability that an individual who is 60 years of age will remain in the current cohort for at least one year?

Answer

We calculate $p_{60}^{(\tau)}$ which is the complement of $q_{60}^{(\tau)}$ which is calculated in c) above to be 0.012726. Hence $p_{60}^{(\tau)}$ = 1 – 0.012726 = 0.987274 = 98.7274%. It is also obtained as $p_{60}^{(\tau)} = \dfrac{l_{61}^{(\tau)}}{l_{60}^{(\tau)}}$ = 80,839/81,881 = 0.987274.

d) What is the probability that an individual who is 60 years of age will leave the current cohort due to death from other causes within the next three years?

Answer

We calculate ${}_3q_{60}^{(2)} = \dfrac{{}_3d_{60}^{(2)}}{l_{60}^{(\tau)}} = (415+456+501)/81{,}881 = 0.016756 = 1.6756\%$

e) What is the probability that an individual who is 61 years will leave the cohort due to accidental death within the next two years?

Answer

We calculate ${}_2q_{61}^{(3)} = \dfrac{{}_2d_{61}^{(3)}}{l_{61}^{(\tau)}} = (289+314)/80{,}839 = 0.007459 = 0.7459\%$

We conclude this chapter with a discussion of a procedure to estimate single premiums for health insurance contracts which are described as multiple-decrement models.

4.4 ESTIMATING NET PREMIUMS BASED ON THE EQUIVALENCE PRINCIPLE

The main objective of this section is as follows:

Obtain the actuarial present value of this policy and obtain the single premium payment, (P_0).

Before we state a general formula for the net single premium payment, we illustrate by use of two examples – one for accidental death and the other for critical illness – which are amenable to multi-decrement modeling. Equation 4.8 in Box 4.1 for the probability of deferred exit is central to obtaining insurance premiums.

Example 1 (Term Life insurance with a Rider for Accidental Death – adapted from example 69.1 of Finan (see reference)

Here is a life table for a 3-year term life insurance subject to two levels of decrement – accidental death (decrement 1) and death from other causes (decrement 2). The policy pays $1 at the end of year of death from accident. The following information from a life table is provided.

Age, x	Number in Cohort (l_x^{τ})	Exits due to Decrement 1 ($d_x^{(1)}$)	Exits due to Decrement 2 ($d_x^{(2)}$)
30	40,750	2,145	2,550
31	36,055	2,276	2,671
32	31,108	2,501	3,004

Interest rate of 5% and the corresponding discount factor is $\mathbf{v} = \dfrac{1}{1+i}$ = 0.9524. This policy is issued to a life (30).

The procedure to obtain the single premium payment described as follows:

Step 1	The life (30) exits at within the first year. Payment is made at the end of the first year.	Payment = $\frac{d_{30}^{(1)}}{l_{30}^{(\tau)}}$ x \$1 = 2,145/40,750 = 0.052638 This calculation is based on equation (3.8)	Discounted Payment = **v** × $\frac{d_{30}^{(1)}}{l_{30}^{(\tau)}}$ = 0.9524 × 0.052638 = **0.050132**
Step 2	The life (30) remains in the group to the end of the first year and exits within the second year. Payment is made at the end of the second year	Payment = $\frac{d_{31}^{(1)}}{l_{30}^{(\tau)}}$ x \$1 = 2,276/40,750 =0.055853 This calculation is based on equation (3.8)	Discounted Payment = **v2** × $\frac{d_{31}^{(1)}}{l_{30}^{(\tau)}}$ = $(0.9524)^2$ × 0.055853 = **0.050662**
Step 3	The life (30) remains in the group to the end of the second year and exits within the third year. Payment is made at the end of the third year	Payment = $\frac{d_{32}^{(1)}}{l_{30}^{(\tau)}}$ x \$1 = 2,501/40,750 = 0.061374 This calculation is based on equation (3.8)	Discounted Payment = **v3** × $\frac{d_{32}^{(1)}}{l_{30}^{(\tau)}}$ = $(0.9524)^3$ × 0.061374 = **0.053020**
Single Premium			0.050132 + 0.050662 + 0.053020 = **0.153814**

The net single premium payment for this policy is 0.153814 for a benefit payment of \$1. For a sum insured of \$100,000, the single premium payment is 0.153814 × \$100,000 = \$15,381.40.

Comment

The previous discussion shows that a general formula for this 3-year insurance policy. It is as follows:

$$\mathbf{P0} = \mathbf{v} \times \frac{d_{30}^{(1)}}{l_{30}^{(\tau)}} + \mathbf{v2} \times \frac{d_{31}^{(1)}}{l_{30}^{(\tau)}} + \mathbf{v3} \times \frac{d_{32}^{(1)}}{l_{30}^{(\tau)}} \quad (4.9)$$

This formula is generalised for an *n*-year policy for a life (x) and decrement *j*. It is presented in Box 4.2 below.

Box 4.2 General Formula

Assumptions:

a) Maturity of policy at policy effective date is n years;
b) Policyholder is (x);
c) Exit is based on decrement *j*;
d) Discount factor is $v = \frac{1}{1+i}$ for interest rate *(i)*.

The single premium payment is: **P0** = **v** × $\frac{d_x^{(j)}}{l_x^{(\tau)}}$ + **v2** × $\frac{d_{x+1}^{(j)}}{l_x^{(\tau)}}$ + … + **v**^{n} × $\frac{d_{x+n-1}^{(j)}}{l_x^{(\tau)}}$ (4.10)

Question: What is the annual level premium that is equivalent to the single premium obtained above?

The answer to this question requires in part, the calculation of the probability that the policyholder will remain in the group at the beginning of each of the term of the policy. Equation (4.6) is used to conduct this calculation. It is noted that the policyholder is alive at the policy effective date and so the probability is equal to 1.

Let A = annual level premium and v = discount factor. Then similar to the calculation in chapter 3 for the comprehensive example, we have a general formula:

$$A \times \left(1 + v \frac{l_{x+1}^{(\tau)}}{l_x^{(\tau)}} + v^2 \frac{l_{x+2}^{(\tau)}}{l_x^{(\tau)}} + \ldots + v^{n-1} \frac{l_{x+n-1}^{(\tau)}}{l_x^{(\tau)}} \right) = P_0 \tag{4.11}$$

In our example, we obtain for *n = 3, x = 30 and v = 0.9524* that following:

A × (1 + { 0.9524 × (36,055/40,750)} + {0.9524^2 x (31,108/40750)} = \$15,381.40. This gives:

A × (2.5351) = \$15,381.40 and so A = \$ 15,381.40 / 2.5351 =\$6,067.37.

Example 2 (Multi -Decrement Model of Life Insurance – adapted from May 2013, MLC Examination, Problem No.2)

P&C Insurance Company is pricing a special fully discrete 3-year term insurance policy on (70). The policy will pay a benefit if and only if the insured dies as a result of an automobile accident. Benefit = $5,000 payable at the end of the year of accident.

- Decrement 1 = death due to cancer; Decrement 2 = death due to automobile accident; Decrement 3 = death due to all other causes.
- Interest rate = 6% so that $v = \frac{1}{1.06}$ = 0.9434

The multi-decrement life table is as follows:

Age, x	Number in Cohort (l_x^{τ})	Exits due to Decrement 1 ($d_x^{(1)}$)	Exits due to Decrement 2 ($d_x^{(2)}$)	Exits due to Decrement 3 ($d_x^{(3)}$)
70	1,000	80	10	40
71	870	94	15	60
72	701	108	18	82

Calculate the single premium based on the equivalence principle.

Calculate the annual level premium equivalent to the single premium.

Answer

Singe Premium

Based on equation (4.9), the single premium for a benefit of $1 is obtained as follows:

P_0 = (0.9434) × (10/1000) + (0.9434)^2 × (15/1000) + (0.9434)^3 × (18/1000) = 0.037897

For a benefit of $5,000 the single premium is 0.037897 × $5,000 = $189.48.

Annual Level Premium

Based on equation (4.11), we have:

A × (1+ {0.9434 × (870/1000)} + {0.9434^2 × (701/1000)) = \$189.48. Then A × 2.44465 = \$189.48. Then A = \$ 77.50

As we stated before and restate for emphasis, chapters 3 and 4 present a procedure for the calculation of single premiums and their equivalent annual level premiums for health insurance that is similar to life techniques commonly referred to as SLT. The next chapter discusses sickness insurance which is based on methods not similar to life techniques (i.e., NSLT).

5 HEALTH INSURANCE: NOT SIMILAR TO LIFE TECHNIQUES

5.1 INTRODUCTION

In this chapter, we focus on sickness insurance which is a considered by Solvency II as *NSLT Health* since it has a preponderance of features that makes it technically similar to non-life insurance. We describe two categories of benefit functions for sickness insurance and obtain single premiums based on the equivalence principle and within the collective risk model.

We advise that a review of the fundamentals of the Poisson and Exponential Distributions may be beneficial. We provide a detailed review in an appendix to chapter 2 of the third eBook entitled *Non-Life Insurance.* It may be advisable to review this appendix at this time.

5.2 SICKNESS INSURANCE BENEFIT FUNCTIONS

Sickness insurance typically provides the policyholder with two main categories of benefits. The first category comprises a fixed-amount daily benefit for short-term sickness periods and/or hospital stays, typically for a maximum duration of one year. The only source of uncertainty in forecasting the total amount of benefit payable for each covered event is the duration of sickness.

Specifically, the expected sickness benefit payable, $\mathbf{B_1}$ is given by the formula:

$$\mathbf{E(B_1) = f \times E(Z)} \tag{5.1}$$

where **f** = fixed-amount daily benefit and **Z** is the (stochastic) duration of sickness or length of stay in a hospital bed.

Example 1

An insurer has sold a sickness insurance which pays a fixed-amount daily benefit of $100 per day (adjusted for fractional days) for medical costs due to hospitalisation and related medical costs. The health insurance actuary predicts that the random variable **Z**, which is the duration of stay in a hospital bed is exponentially distributed with an expected value of 3.5 days.

Comment

The assumption of exponential distribution for the random variable 'length of time' is common in the health care insurance literature. For example, Fackrell (2009) in *Health Care Management Science (pages 11–26)* states, "The exponential distribution is ubiquitous in stochastic modelling, mainly because of its simplicity and ability to model random lengths of time reasonably well. For example, it has been used to model the length of stay in a hospital bed".

Answer for Example 1

Based on equation (5.1), expected benefit = \$100 × E(Z) = \$100 × 3.5 = \$350.

The second category is based on a *reimbursement model* of health insurance whereby a benefit is payable for medical expenses incurred by the policyholder subject to policy coverage modifications such as deductibles and/or policy limits. These are covered in chapter 2.

Example 2

A sickness insurance policy provides for a policy limit of \$300. The cost per doctor visit is exponentially distributed with mean (θ) = €200. Calculate the value of $\mathbf{E(X \wedge u)}$ which is the insurer's expected payment per policyholder loss. (Refer to the appendix below for the appropriate formula in the case of exponential distribution).

Then,

$$\mathbf{E(X \wedge u)} = \$200 \times \left(1 - e^{-\frac{300}{200}}\right) = \$200 \times 0.7769 = \$155.37.$$

The insurer's expected payment per doctor visit is \$155.37.

We present a model of sickness insurance in the next section.

5.3 MODELING SICKNESS INSURANCE

As a standalone insurance contract, sickness insurance may be modeled as follows:

Figure 5.1: Model of Sickness Insurance

This diagram is explained as follows:

a) **0** is the policy effective date and the maturity date is one year later;

b) A life (x) currently occupies the healthy state when the sickness insurance policy is issued. The life remains in this state for an uncertain duration and leaves for the state of sickness at a random time before the maturity date. Clearly, the life may remain in the healthy state for the entire coverage period of one year or may enter the state of sickness with random frequency. This random model of sickness insurance may be depicted as a directed graph as follows:

Figure 5.2: Sickness Insurance as a Directed Graph

In this model, H= healthy state and S= sick state. The edges are (H, S) and (S, H). Hence, the possibility exists that the life leaves H and enters S and subsequently recovers as indicated by the edge (S, H). This cycle may be repeated over the term of the insurance.

The diagrams above depict a generic form of a single sickness insurance policy. The main point is that the number of claims by the policyholder is random with a random claim amount. Hence the aggregate loss for this sickness insurance policy is a random sum of random claim amounts.

Comment

When viewed from a portfolio context so that claim frequency and claim amounts are not identified with any particular policy, the *collective risk model* is applied. We describe this model below.

5.4 COLLECTIVE RISK MODEL

A collective risk model considers claims arising from a portfolio of policies so that claim amounts are random and the number of such claims is also random. The collective risk model is described in Box 5.1 below:

Box 5.1 Collective Risk Model

Let N represent the random number of claims. N is called claim frequency. Each claim, $\mathbf{X_i}$ is random amount so $\{\mathbf{X_i}\}, \mathbf{i = 1, 2, ..., N}$ are independent and identically distributed (iid), and each $\mathbf{X_i}$ is independent of N. Let E(X) be the common expected value of each $\mathbf{X_i}$.

Each claim is not associated with a particular policy but with the portfolio.

Let $\mathbf{S = X_1 + X_2 + ... + X_N}$. Note that S is a random sum of random variables and represents the aggregate claims or loss to the insurer.

Note that policyholders' expenses are modified by possible deductibles and policy limits before claims are made.

By convention, if N = 0, then S = 0.

Then E(S) = E(N) E(X).

This result is key and its proof is found in S.K. Ross, *Introduction to Probability Models*, Third Edition, Academic Press, 1985 Chapter 3, pages 83–103.

In order to implement the collective risk model, we have to specify probability distributions for the random variables N and X.

Here is an example illustrating the collective risk model.

Example 3

Consider a one-year sickness insurance policy where the actuary predicts that claim frequency (N) is Poisson distributed with expected value 2 and that claim amount (X) is exponentially distributed with mean $300.

a) What is the expected aggregate loss for this policy?
b) What is the single annual premium based on the equivalence principle.

Answer

a) $E(S) = E(N) \times E\ (X) = 2 \times \$300 = \$600.$
b) $P_0 = E(S) = \$600.$

An advanced treatment of sickness insurance with deductibles and policy limits is included in Level II of the Art of Insurance series.

This concludes chapter 5 and this eBook.

6 REFERENCES

[1] Bluhm W.F., *Individual Health Insurance*. Winsted, CT: ACTEX Publications (2007).

Chapter 4 entitled *Managing Antiselection* is excellent for further reading.

[2] Ellis R.P., Chen T. and Luscombe C.E., *Comparisons of Health Insurance Systems in Developed Countries* in Encyclopedia of Health Economics, Elsevier Press., Inc. 2014.

This research covers the (public) health insurance systems of Canada, Germany, Japan, Singapore and the USA.

[3] Fackrell M., *Modelling healthcare systems with phase-type distributions*. Health Care Management Science 12: 11–26 (2009).

[4] Finan Marcel B. *A Reading of the Theory of Life Contingency Models: A Preparation for Exam MLC/3L,* Arkansas Tech University (2011).

This is an excellent book and the sections entitled *Multiple Decrement Models* starting on page 639 and *Multiple State Transition Models* starting on page 709 are especially informative.

[5] Folland S., Goodman A.C. and Stano M., *Economics of Health and Health Care*, 7th edition, (2013).

This is an excellent introductory textbook on the economics of health care required for an understanding of health care insurance.

[6] Gneezy U., Haruvy E. and Yafe H., *The Inefficiency of Splitting the Bill.* The Economic Journal, 114 (April), 265–280.

This is a technical paper but the result is important in the field of behavioural economics

[7] Haberman S. and Pitacco E., *Actuarial Models of Disability Models.* CRC Press, 1999.

There is extensive coverage of disability insurance beginning on page 80 as well as critical illness cover on page 174. This book is highly recommended for further reading.

[8] Olivieri A. and Pitacco E. *Introduction to Insurance Mathematics.* Second Edition, Springer, 2010.

This is an excellent textbook for life and non-life insurance. The definition of health insurance is on page 440.

[9] Pitacco E., *Health Insurance: Basic Actuarial Models.* Springer, 2014.

We strongly recommend this textbook for further reading. However, the final chapter comprises an extensive use of complex actuarial notation and may be avoided at the reader's discretion.

[10] Pollard A.H., *The interaction between morbidity and mortality. (with discussion*). J.I.A., 107, 233, 1980.

This is a classic paper that links states of health (morbidity) with mortality.