Review of statistical actuarial risk modelling
Hiroshi Shiraishi1*

Abstract: In this paper, we review some results for insurance risk theory. We first introduce a variety of the insurance risk models proposed thus far. Then, we show that the expected discounted penalty function (the so-called Gerber–Shiu function) can describe some risk indicators. Next, the dividend problem is discussed; more precisely, the (approximated) optimal dividend barrier is derived and other extended dividend strategies introduced. In addition, some modified models depending on reinsurance or tax are introduced. Finally, we discuss the statistical estimation of the ruin probability and the Gerber–Shiu function.

1. Introduction
In the actuarial literature, the classical approach is to represent an insurance risk as a stochastic process with a certain number of control variables. The term “insurance risk” here means the risk for non-life insurance such as motor vehicle insurance, home and contents insurance, and travel insurance rather than for life or health insurance, for example. Once an insurance policy is insured against the insurance risk, the insured party pays an appropriate premium to the insurer at the start of the period of insurance cover. Then, the insured party will make a claim under the insurance policy each time he or she has an insurance accident. From an insurer’s point of view, the surplus is generally defined by

\[
\text{initial surplus} + (\text{aggregated) premium income} - (\text{aggregated) claim payment}).
\]

ABOUT THE AUTHOR

Hiroshi Shiraishi received the BS degree in mathematics in 1998 and the MS and Dr degrees in mathematical science from Waseda University, Japan in 2004 and 2007, respectively. He joined the GE Edison Life Insurance Company, the Prudential Life Insurance Company of Japan and the Hannover-Re Reinsurance Company, in 1998, 2000 and 2005, respectively. His research interests are actuarial science, time series analysis, econometric theory and financial engineering. In particular, he investigates the statistical estimation of optimal dividend problems in the field of the actuarial science. He is currently an associate professor in the Department of Mathematics, Keio University, Japan. He is a fellow of the Institute of Actuary of Japan (FIAJ).

PUBLIC INTEREST STATEMENT

In this paper, we review some results for insurance risk theory. Once an insurance policy is insured against the insurance risk, the insured party pays an appropriate premium to the insurer at the start of the period of insurance cover. Then, the insured party will make a claim under the insurance policy. From an insurer’s point of view, the (pure) surplus is defined by “initial surplus + (aggregated) premium income – (aggregated) claim payment”. We introduce the insurance risk models proposed thus far. In classical insurance risk analysis, it is interested in the ruin probability which is the probability that the surplus is below zero. We introduce its extended quantity the so-called Gerber–Shiu function. In addition, we discuss some applications for the insurance risk theory such as the dividend, reinsurance, tax problems. Finally, we discuss the statistical estimation of the Gerber–Shiu function.
Here, there are two sources of uncertainty for the insurer as follows:

- How many claims will the insured party make?
- What will the amounts of those claims be?

This general framework is used to model “(aggregated) claim payments”. The Cramér-Lundberg model introduced by Lundberg (2009), Cramér (1930), and Cramér (1955) represents the (aggregated) claim payment part of the surplus as a compound Poisson process. This model is dealt with in most introductory books on risk theory such as Bühlmann (1970), Gerber (1979), Bowers, Gerber, Hickman, Jones, and Nesbitt (1997), and Dickson (2005).

Various extensions of the Cramér-Lundberg model have been proposed. For instance, Dufresne and Gerber (1991) added a diffusion process to the original model and Huzak and Perman (2004) included a Lévy process with mean 0 and no positive jumps. Andersen (1957) generalized the claim number process to a renewal process that included a Poisson process. Albrecher and Teugels (2006) considered the possible dependence between the interclaim time and claim size.

In classical insurance risk analysis, if the above surplus falls to zero or below, we say that ruin occurs, and the probability of ultimate ruin (i.e. the probability that ruin eventually occurs) is discussed. In particular, by denoting the ruin probability as a function of the initial surplus \( u \), the effect of \( u \) for the ruin probability is analysed. Gerber (1987) argued that the ruin probability is a crude stability criterion and discussed the seriousness of the situation when ruin occurs (the so-called “severity of ruin”). Gerber and Shiu (1998) introduced the expected discounted penalty function (the so-called “Gerber–Shiu function”) to treat both the deficit at ruin (which is an indicator of the severity of ruin) and the time of ruin simultaneously. The Gerber–Shiu function is described as a function of certain control variables such as the initial surplus, dividend barrier and reinsurance parameter.

De Finetti (1957) argued that it is unrealistic to minimize the ruin probability for an infinite length of time assuming that companies can grow their surplus without limit. He proposed an alternative formulation to avoid the infinite growth of the surplus, that is, a payment of a part of the surplus to the shareholder as a dividend (i.e. a refund). However, two fundamental questions have to be addressed:

- When should dividends be distributed?
- How much of the surplus should be distributed?

The answer to these questions is based on the dividend strategy. Avanzi (2009) reviewed a variety of dividend strategies such as the constant barrier strategy, bond strategy, linear barrier strategy and non-linear barrier strategy. Bühlmann (1970), Gerber and Shiu (2004), Gerber, Shiu, and Smith (2006), and Dickson and Drekic (2006) among others derived optimal dividend barriers, which maximize the expectation of the discounted dividends under the constant barrier strategy. In addition, Gerber et al. (2006) and Dickson and Drekic (2006) discussed the dividend problem under the optimization of the expected discounted dividends excluding the present value of the deficit at ruin. Further, because the explicit form of the optimal dividend barrier is obtained only in special cases, Gerber, Shiu, and Smith (2008) approximated some formulae for this case.

In this study, we introduce related topics such as “reinsurance” and “tax”. For the former, we adopt an insurer’s point of view to illustrate how utility theory can be applied to determine the optimal reinsurance parameters, which correspond to the optimal retention level under both the proportional and the excess-of-loss reinsurance. For the latter, we explain the “loss-carry forward system” discussed by Albrecher and Hipp (2007) and Albrecher, Badescu, and Landriault (2008).
The statistical estimation of the ruin probability and Gerber–Shiu function has been discussed by Grandell (1978/1979), Csörgö and Teugels (1990), Deheuvels and Steinebach (1990), Croux and Veraverbeke (1990), Csörgö and Steinebach (1991), Embrechts and Mikosch (1991), Bening and Korolev (2003), Mnatsakanov, Ruymgaart, and Ruymgaart (2008), and Shimizu (2012). Grandell (1978/1979), Csörgö and Teugels (1990), Deheuvels and Steinebach (1990), and Csörgö and Steinebach (1991) all discussed the statistical estimation of the adjustment coefficient, because the ruin probability is approximately expressed as a function of the adjustment coefficient based on the Cramér–Lundberg approximation. Croux and Veraverbeke (1990), Bening and Korolev (2003), Mnatsakanov et al. (2008), and Shimizu (2012) discussed the non-parametric estimation of the ruin probability and Gerber–Shiu function.

The remainder of this paper is organized as follows. In Section 2, we introduce a variety of the insurance risk models proposed thus far. Section 3 discusses whether the Gerber–Shiu function can describe some of the risk indicators. The dividend problem is reviewed in Section 4. We first show the expected discounted dividends under some distributions of the claim amount. Then, approximation methods for the optimal dividend barrier are shown. We next consider the effect of the deficit at ruin and, finally, extended dividend strategies are introduced. In Section 5, some modified models depending on reinsurance or tax are introduced. Finally, Section 6 discusses the statistical estimation of the ruin probability and Gerber–Shiu function.

2. Risk models
A variety of risk models have been introduced thus far. Here, the risk models show the modelling of the surplus at time $t$. The definition is as follows:

**Definition 2.1** The surplus at time $t$ is

$$U(t) = u + p(t) - S(t), \quad t \geq 0,$$

(2.1)

where $u \geq 0$ is the initial surplus of the company, $p(t)$ is the (predictable) premium income with $p(0) = 0$ and $p(t) > 0$ for $t > 0$, and $S(t)$ is the aggregate claim process with $S(t) \geq 0$ for $t > 0$.

In most cases, the premium income is assumed to be received continuously at a constant rate $c > 0$ (i.e. $p(t) = ct$), which implies that the possibility of gaining or losing customers is ignored. An insurance policy is a financial agreement between the insurance company and policyholder. The insurance company agrees to pay some benefits and policyholder agrees to pay premiums to the insurance company to secure these benefits. Suppose that the net premium (in which the insurance company’s expenses are excluded from the gross premiums) for a single payment per insurance policy is denoted $P$. If a constant number of new policies (denoted $L$) is continuously included and the same number of the existing policies is withdrawn from the portfolio of insurance policies, then the insurance company can continuously receive $LP(\equiv c)$ as the premium income and the aggregated premium income is $LP(\equiv ct)$. The insurance risk for the portfolio is maintained, because the volume of the insurance portfolio is the same over time. Premium calculation principles have many desirable properties such as non-negative loading, additivity, scale invariance, consistency and no ripoff (see Dickson, 2005). From these points of view, some premium principles have been introduced such as the expected value principle, variance principle, standard deviation principle, principle of zero utility, Esscher principle and Wan principle. In what follows, we suppose that the premium rate ($c$) is determined based on the expected value principle, that is, $c = (1 + \theta)E(S(1))$ with $\theta > 0$.

The initial surplus $u$ becomes a parameter of the function of the risk indicators in many cases. For example, in classical risk theory, the (ultimate) ruin probability is considered to be a function of the initial surplus, which implies that the investor can decide the initial surplus to satisfy a desirable condition.

The aggregated claim $(S(t))$ expresses the total amount of claims arising from the insurance portfolio. In classical risk theory, the individual claim amount is assumed to be independent and
identically distributed (i.i.d.), while the sequence of the claim amounts \(\{X_i\}\) and the number of claims \(N\) are assumed to be independent. However, the number of claims \(N \equiv N(t)\) that occurs in the fixed time interval \([0, t]\) is assumed to be a Poisson distribution with parameter \(\lambda t\) (i.e. \(\{N(t)\}\) is a Poisson process) to express that the interclaim time is random but independent.

Example 2.1 [The Cramér–Lundberg model] The Cramér–Lundberg model is the classical model of risk theory introduced by Lundberg (2009), Cramér (1930, 1955). Suppose that the surplus of an insurance company at time \(t\) is

\[
U(t) = u + ct - S(t), \quad S(t) = \sum_{i=1}^{N(t)} X_i, \quad t \geq 0, \tag{2.2}
\]

where the premium income \((p(t) = ct)\) is received continuously at a constant rate \(c > 0\); the aggregate claims process \(S(t) = \sum_{i=1}^{N(t)} X_i\) is a compound Poisson process; the claim frequency \(N(t)\) is a Poisson process with the intensity parameter \(\lambda > 0\); and \(\{X_i\}\) is a sequence of i.i.d. positive-valued random variables.

According to Dickson (2005), “This model is, of course, a simplification of reality. Some of the more important simplifications are that we assume claims are settled in full as soon as they occur, there is no allowance for interest on the insurer’s surplus, and there is no mention of expenses that an insurer would incur. Nevertheless, this is a useful model which can give us some insight into the characteristics of an insurance operation”.

Example 2.2 [De Finetti’s model] In this model, the surplus at time \(t\) is

\[
U(t) = u + t - \sum_{i=1}^{t} X_i, \quad t = 1, 2, 3, \ldots
\]

where \(X_1, X_2, \ldots\) are i.i.d. random variables with

\[
P(X_i = 0) = p, \quad P(X_i = 2) = q = 1 - p, \quad i = 1, 2, 3, \ldots
\]

and \(p > 1/2\) (i.e. \(p(t) = t\), \(S(t) = \sum_{i=1}^{t} X_i\)). This model, proposed by De Finetti (1957), was later reviewed by Seal (1969), Gerber and Shiu (2004), and Avanzi (2009).

Example 2.3 [Approximation of the Cramér–Lundberg model] Dickson and Waters (1991) discussed the compound Poisson process in discrete time of which the surplus at time \(t\) is

\[
U(t) = u + t - \sum_{i=1}^{N(t)} X_i, \quad t = 1, 2, \ldots
\]

De Vylder and Goovaerts (1988), Li and Garrido (2002), Claramunt, Mårmol, and Alegre (2003), and Dickson and Waters (2004) also dealt with this model.

The compound binomial process of which the surplus at time \(t\) is

\[
U(t) = u + t - \sum_{i=1}^{N(t)} X_i, \quad t = 1, 2, \ldots
\]

where \(\{N(t)\}\) is no longer a Poisson process, but a binomial process, that is, the increment \(N(t + h) - N(t)\) for \(h = 1, 2, \ldots\) is a binomial random variable with parameters \(h\) (number of trials) and \(p\) (probability of success). Gerber (1988), Shiu (1989), Dickson (1994), De Vylder and Marceau (1996), Cheng, Gerber, and Shiu (2000), Tan and Yang (2006), Bao (2007), and Landriault (2008b) also dealt with this model.
Example 2.4  [The Wiener process] The continuous counterpart of de Finetti’s model is the Wiener process. In this model, the surplus at time \( t \) is

\[
U(t) = u + \mu t + \sigma W(t), \quad t \geq 0,
\]

(2.3)

where \( \{W(t)\} \) is a standard Wiener process. As a consequence, \( U(t + h) - U(t) \) for \( h > 0 \) has mean \( h\mu \) and variance \( h\sigma^2 \). Gerber and Shiu (2003, 2004), and Leung, Kwok, and Leung (2008) dealt with this model.

Example 2.5  [With diffusion model (Wiener–Poisson risk process)] Dufresne and Gerber (1991) introduced the diffusion model, of which the surplus at time \( t \) is

\[
U(t) = u + ct - S(t) + \sigma W(t), \quad t \geq 0,
\]

(2.4)

where \( S(t) \) is the same as (Equation 2.2), \( \sigma > 0 \) is a constant volatility parameter and \( \{W(t)\} \) is a standard Wiener process. Li (2006), Li and Wu (2006), Jang (2007), Li, Feng, and Song (2007), Wan (2007), Frostig (2008), and Yuen, Lu, and Wu (2009) also dealt with this model.

Example 2.6  [The Sparre Andersen model] The Cramér–Lundberg model assumes that \( \{S(t)\} \) is a compound Poisson process, but in the Sparre Andersen model, \( \{S(t)\} \) is not necessarily a compound Poisson process. The idea is taken from Andersen (1957), who gave his name to the model. Li and Garrido (2004), Albrecher, Claramunt, and Mármol (2005), Albrecher, Hartinger, and Thonhauser (2007), Meng, Zhang, and Wu (2007), Albrecher and Hartinger (2006), and Yang and Zhang (2008) among others dealt with this model. Moreover, Zhang (2014) considered the Sparre Andersen model to be perturbed by a Brownian motion.

Example 2.7  [The dynamic control model] Højgaard (2002) considered a process with the dynamic control of the premium income, in which the classical Cramér–Lundberg model with premium rates \( c \equiv c_i \) is calculated by using the expected value principle (i.e. \( c_i = (1 + \theta_i)E(S(1)) \)). The company controls dynamically the relative safety loading \( \theta_i \) with the possibility of gaining or losing customers.

Example 2.8  [The interclaim-dependent claim size model] Albrecher and Teugels (2006), Boudreault, Cossette, Landriault, and Marceau (2006), Landriault (2008a), and Huang and Li (2011) considered an extension to the classical Cramér–Lundberg model in which a particular dependence structure among the interclaim time and subsequent claim size is introduced. When the claim number process \( \{N(t)\} \) is a (homogeneous) Poisson process with the intensity parameter \( \lambda > 0 \), the interclaim time \( \Delta_iT \) between the \( i \)th claim and \( (i - 1) \)th claim follows an independently exponential distribution with the parameter \( \lambda \). In this model, we suppose that the bivariate random vectors \( (X_i, \Delta_iT) \) for each \( j \in \mathbb{N} \) are mutually independent, whereas \( X_i \) and \( \Delta_iT \) are dependent.

Example 2.9  [The dual model] In the classical Cramér–Lundberg model, the surplus of a company at time \( t \) is

\[
U(t) = u + ct - S(t)
\]

where \( u \) is the initial surplus, \( c \) is the constant premium rate and \( S(t) \) are the aggregate claims at time \( t \). In the dual model, the surplus at time \( t \) is

\[
U(t) = u - ct + S(t)
\]

where \( u \) is also the initial surplus, \( c \) is the company’s expense rate and \( S(t) \) are the aggregate gains at time \( t \). Avanzi, Gerber, and Shiu (2007), Gerber and Smith (2008), and Ng (2009) discussed the
optimal barrier strategies in this model. Moreover, Avanzi and Gerber (2008) and Avanzi, Shen, and Wong (2011) introduced the dual model with diffusion.

Example 2.10 [The spectrally positive (negative) Lévy risk model] In this model, the surplus at time \( t \) is

\[
U(t) = u + ct - S(t) + Z(t), \quad t \geq 0,
\]

where \( u \) is the initial surplus, \( c \) is the constant premium rate and \( (S(t)) \) and \( (Z(t)) \) are independent with \( S(0) = Z(0) = 0, E\{Z(1)\} = 0 \) and \( E\{|Z(1)|\} < \infty \). We provide a more detailed description of the surplus by emphasizing the properties of \( (S(t)) \) and \( (Z(t)) \). The Laplace exponent of \( S \) is given by

\[
m_r(t) = \ln E\{e^{-rS}\} = \int_{0}^{\infty} (e^{-rx} - 1)\nu(dx), \quad r \geq 0,
\]

where \( \nu \) is a non-negative measure of \( \mathbb{R}_+ \) satisfying \( \int_{0}^{\infty} (1 + x)\nu(dx) < \infty \). In addition, let \( \check{Z} \) denote the dual process of \( Z \) (i.e. the process \( \{\check{Z}(t)\} \) has the same law as \( \{-Z(t)\} \)). Then, the Laplace exponent of \( \check{Z} \) is given by

\[
m_r(t) = \ln E\{e^{-r\check{Z}(t)}\} = \ln E\{e^{Z(t)}\} = \frac{\sigma^2}{2} + \int_{0}^{\infty} (e^{-rx} - 1 + rx)\nu_r(dx), \quad r \geq 0,
\]

where \( \sigma \geq 0 \) and \( \nu_r \) are non-negative measures of \( \mathbb{R}_+ \) satisfying \( \int_{0}^{\infty} (1 + x)\nu_r(dx) < \infty \) and \( \int_{0}^{\infty} x^2\nu_r(dx) < \infty \) to ensure that \( Z \) has finite expectations. The classical compound Poisson model can be recovered by assuming that \( \sigma = 0 \) and \( \nu_r \) are null measures and by setting \( \nu_1(dx) = \lambda f(x)dx \). Huzak and Perman (2004), Loeffen (2009), Biffis and Morales (2010), Loeffen and Renaud (2010), Yin and Wen (2013), and Feng and Shimizu (2013) among others dealt with this model.

3. Gerber–Shiu function

Suppose that the surplus of an insurance company \( U(t) \) follows the Cramér–Lundberg model:

\[
U(t) = u + ct - S(t), \quad S(t) = \sum_{i=1}^{N(t)} X_i, \quad t \geq 0,
\]

where the claim size \( X_i \) is an independent random variable with common distribution \( F(x) \) and probability density \( f(x) \), and \( \{N(t)\} \) is a Poisson process with the intensity parameter \( \lambda \).

Let \( T \) denote the time of ruin,

\[
T = \inf \{t \geq 0 | U(t) < 0\}
\]

(\( T = \infty \) if ruin does not occur). Classical ruin theory is interested in the probability of ultimate ruin as a function of the initial surplus \( u \geq 0 \),

\[
\psi(u) = P\{T < \infty | U(0) = u\}.
\]

(3.1)

We impose that following assumption for the premium rate \( c \).

Assumption 3.1 This assumption ensures that \( \{U(t)\} \) has a positive drift and \( \psi(u) < 1 \) (see Gerber & Shiu, 1998).

\[
c > \lambda \mu, \quad \mu = E(X_i) = \int_{0}^{\infty} xf(x)dx.
\]

Gerber (1987) discussed not only the probability of ruin, but also the severity of the situation when ruin occurs. Let \( U(T) \leq 0 \) denote the surplus at ruin. They introduced the following quantity:
ψ(u, y₀) = P\{T < ∞, |U(T)| ≤ y₀ |U(0) = u\}.

(3.2)

This quantity is an indicator of the severity of ruin, which corresponds to the probability that ruin will occur and states that the deficit at the time of ruin will be less than \(y_0\) for a given initial surplus \(u\). Dufresne and Gerber (1988) and Dickson (1992) also discussed the severity of ruin.

Furthermore, Gerber and Shiu (1997, 1999) derived the joint distribution of the time of ruin, the surplus immediately before ruin (denoted by \(U(T−)\)), and the deficit at ruin by using the Laplace transform, which leads to the introduction of the following generalized function by Gerber and Shiu (1998).

**Definition 3.1** (The expected discounted penalty function (Gerber–Shiu function))

\[
\phi(u) = E\{w(U(T−), |U(T)|)e^{−\delta T}I(T < ∞) |U(0) = u\} \tag{3.3}
\]

where \(w(x, y)\) is a non-negative function of \(x > 0, y > 0, \delta \geq 0\) is a force of interest, and \(I(A)\) is an indicator function \((I(A) = 1 if A is true and I(A) = 0 if A is false)\).

The Gerber–Shiu function is due at ruin, and it depends on the deficit at ruin as well as the surplus immediately prior to ruin. This function is thus useful for deriving results in connection with the joint and marginal distributions of \(T, U(T−)\) and \(|U(T)|\).

The following shows examples of the Gerber–Shiu function.

**Example 3.1** [Ruin probability] If \(w(x, y) = 1\) and \(\delta = 0\), then it follows that

\[
\phi(u) = E\{I(T < ∞) |U(0) = u\} = P\{T < ∞ |U(0) = u\} = ψ(u),
\]

which implies that the ruin probability defined by (Equation 3.1) is included in the class of the Gerber–Shiu function.

**Example 3.2** [Ruin probability with the deficit at the time of ruin] If \(w(x, y) = I(y \leq y_0)\) and \(\delta = 0\), then it follows that

\[
\phi(u; y_0) = E\{I(|U(T)| \leq y_0)I(T < ∞) |U(0) = u\} = P\{T < ∞, |U(T)| \leq y_0 |U(0) = u\} = ψ(u, y_0),
\]

which implies that the ruin probability defined by Equation (3.2) is included in the class of the Gerber–Shiu function.

**Example 3.3** [Conditional joint distribution function] If \(w(x, y) = I(x \leq x_0, y \leq y_0)\) and \(\delta = 0\), then it follows that

\[
\phi(u; x_0, y_0) = E\{I(U(T−) \leq x_0, |U(T)| \leq y_0)I(T < ∞) |U(0) = u\} = P\{T < ∞, U(T−) \leq x_0, |U(T)| \leq y_0 |U(0) = u\},
\]

which implies that the joint distribution of \((U(T−), |U(T)|)\) on the event \(T < ∞)\) is included in the class of the Gerber–Shiu function.

In addition, when the above joint distribution has the density (i.e. it satisfies a differentiable condition), the (conditional) joint density function of \((U(T−), |U(T)|)\) on the event \(T < ∞)\) is included in the class of the Gerber–Shiu function under \(w(x, y) = I(x = x_0, y = y_0)\) and \(\delta = 0\).

**Example 3.4** [Single premium for the ruin] If \(\delta > 0\) and \(w\) are defined in terms of the benefit amount of an insurance (or reinsurance) payable at the time of ruin, then \(\phi(u)\) is the single premium of this insurance.
Example 3.5 [Expected present value of a deferred continuous annuity] Define

\[ w(x, y) = \frac{1 - e^{-\rho y}}{\delta}, \quad \delta > 0, \]

where \( \rho \) is the positive solution to Lundberg’s fundamental equation satisfied by

\[ E\{e^{-\delta(T_b - T)} | T < \infty\} = \rho^{\delta(T)} \]

and \( T_b \) (called the time of recovery) is the first time the surplus reaches zero after ruin. Then, it follows that

\[
\phi(u) = E\left\{ \frac{1 - e^{-\rho(T_b - T)}}{\delta} e^{-\delta T} I(T < \infty) | U(0) = u \right\}
\]

\[
= \frac{1}{\delta} E\left\{ \left( 1 - \frac{1 - e^{-\delta T_b} T_b - T)}{T < \infty} \right) e^{-\delta T} I(T < \infty) | U(0) = u \right\}
\]

\[
= \frac{1}{\delta} E\left\{ E\{e^{-\delta T - e^{-\delta T_b} | T < \infty\} I(T < \infty) | U(0) = u \} \right\}
\]

\[
= E\left\{ \int_0^\infty e^{-\delta t} dt I(T < \infty) | U(0) = u \right\}.
\]

In this case, the Gerber–Shiu function corresponds to the expected present value of a continuous annuity at a rate of 1 per unit time, starting at the time of ruin and ending as soon as the surplus rises to zero (see Remark (v) of Section 6 in Gerber & Shiu, 1998).

Example 3.6 [Option pricing] For \( t \geq 0 \), let \( A(t) = e^{\lambda t} \) be the price of an asset such as a stock or a stock index at time \( t \). Gerber and Shiu (1999) considered the pay-off function

\[ w(x, y) = \max(K - e^{-\rho y}, 0), \]

where \( K \) is the exercise price of a perpetual American put option. Let \( T = T_b = \inf\{t | A(t) < e^y\} \) where \( b \) is a number satisfying \( e^b \leq \min(e^y, K) \). Then, it follows that

\[
\phi(u;b) = E\left\{ \max(K - e^{-\rho(T_b - T)}, 0) e^{-\delta T} I(T_b < \infty) | U(0) = u \right\}
\]

\[
= E\left\{ \max(K - A(T_b), 0) e^{-\delta T} I(T_b < \infty) | A(0) = e^y \right\},
\]

which implies that the value of the option-exercise strategy \( T_b \) is included in the class of the Gerber–Shiu function. Then, the optimal value of \( b \) is derived as the solution to

\[ \phi(b; b) = K - e^b. \]

4. Dividend problem

Suppose that in the absence of dividends, the surplus of an insurance company at time \( t \) is defined by (Equation 2.2) for \( t \geq 0 \) (in this section, we assume the Cramér–Lundberg model). The insurance company refunds the part that the surplus exceeds a constant barrier \( b \geq 0 \) as the dividend to the shareholder. In this model, let \( D_b(t) \) denote the aggregate dividend at time \( t \). Then, we can write

\[ dD_b(t) = \begin{cases} 0 & \text{if } U_b(t) < b \\ c dt & \text{if } U_b(t) = b \end{cases}, \quad (4.1) \]

where \( U_b(t) \) is the (modified) surplus with the dividend barrier \( b \) given by

\[ U_b(t) = U(t) - D_b(t). \quad (4.2) \]

In addition, the time of ruin with dividends is given by \( T_b = \inf\{t | U_b(t) < 0\} \).
4.1. Classical dividend problem

De Finetti (1957), Bühlmann (1970), Gerber (1979), and Gerber and Shiu (1998) dealt with the problem of the expectation of the discounted dividends, namely,

\[ V(u, b) = E \left[ \int_0^{T_b} e^{-\delta t} dD_b(t) \right] \]  \hspace{1cm} (4.3)

where \( \delta > 0 \) is a force of interest.

Suppose each claim amount \( X_i \) has the probability distribution function \( f(x) \). Then, the following representations are obtained (see Dickson, 2005):

\[ V(u, b) = e^{-\frac{b}{c}} V(b, b) + \frac{\lambda}{\delta} \int_u^b e^{-\frac{x}{c}} \int_0^x f(x) V(s - x, b) dx ds, \]  \hspace{1cm} (4.4)

\[ V(b, b) = \alpha + \frac{\lambda}{\delta + \lambda} \int_0^b f(x) V(b - x, b) dx. \]  \hspace{1cm} (4.5)

Equation (4.4) comes from the decomposition of the aggregate dividends as

(I) The case that the surplus attains \( b \) before the first claim occurs,

(II) The case that the first claim occurs before the surplus attains \( b \).

In case (I), since the “future aggregated dividends immediately before the first claim occur” corresponds to \( V(b, b) \), we can obtain the first term on the right-hand side of Equation (4.4). In case (II), since the “future aggregated dividends immediately after the first claim occur” corresponds to \( V(s - x, b) \), where \( s \in (u, b) \) is the amount of the “surplus immediately before the first claim occurs” and \( x \in (0, s) \) is the amount of the first claim, we can obtain the second term on the right-hand side of Equation (4.4).

Similarly, Equation (4.5) comes from the decomposition of the aggregate dividends as

(III) The part before the first claim occurs,

(IV) The part after the first claim occurs.

By using Equations (4.4) and (4.5), the following integro-differential equation is derived:

\[ \frac{\partial}{\partial u} V(u, b) = \frac{\lambda + \delta}{c} V(u, b) - \frac{\lambda}{c} \int_0^u f(x) V(u - x, b) dx. \]  \hspace{1cm} (4.6)

The optimal dividend barrier \( b_{\text{opt}} \equiv b_{\text{opt}}(u) \) is defined by the maximizer of \( V(u, b) \):

**Definition 4.1**

\[ b_{\text{opt}} := \arg \max_{b \geq u} V(u, b) \]  \hspace{1cm} (4.7)

for a constant value \( u \geq 0 \).

The explicit form of \( b_{\text{opt}} \) is obtained only if each claim amount \( X_i \) follows some special distributions such as exponential distribution and the diffusion process.

**Example 4.1** [Exponential distribution (Bühlmann, 1970)] Suppose that \( X \) follows an exponential distribution with \( E(X) = 1/\beta \) (i.e. \( f(x) = \beta e^{-\beta x} \)). Then, from Equation (4.6), the second-order differential equation
\[
\frac{\partial^2}{\partial u^2} V(u, b) + \left( \beta - \frac{\lambda + \delta}{c} \right) \frac{\partial}{\partial u} V(u, b) - \frac{\beta \delta}{c} V(u, b) = 0 \tag{4.8}
\]

is obtained. By solving this differential equation with the boundary condition \( \frac{\partial}{\partial u} V(u, b) \big|_{u=b} = 1 \) from Equations (4.4) and (4.5), we have

\[
V(u, b) = \frac{h(u)}{h'(b)}, \quad h(x) = (\beta + \rho) \exp(\rho x) - (\beta - R) \exp(-Rx), \tag{4.9}
\]

where \( \rho > 0 \) and \(-R < 0\) are the solutions to the following quadratic equation with respect to \( x \)

\[
x^2 + \left( \beta - \frac{\lambda + \delta}{c} \right)x - \frac{\beta \delta}{c} = 0.
\]

Finally, from Equation (4.9), we obtain by solving \( h''(b_{\text{opt}}) = 0 \)

\[
b_{\text{opt}} = \frac{1}{\rho + R} \ln \frac{R^2 (\beta - R)}{\rho^2 (\beta + \rho)}. \tag{4.10}
\]

**Example 4.2** [Diffusion process (Gerber & Shiu, 2004)] Suppose that the aggregated claim process \( S(t) \) is not a compound Poisson process but a diffusion process defined by \( S(t) = \sigma W(t) \), where \( \{W(t)\} \) is a standard Wiener process. Then, the second-order differential equation

\[
\frac{\sigma^2}{2} \frac{\partial^2}{\partial u^2} V(u, b) + c \frac{\partial}{\partial u} V(u, b) - \delta V(u, b) = 0
\]

is obtained. By solving this differential equation with the boundary conditions \( V(0, b) = 0 \) and \( \frac{\partial}{\partial u} V(u, b) \big|_{u=b} = 1 \), we have

\[
V(u, b) = \frac{h(u)}{h'(b)}, \quad h(x) = \exp(\rho x) - \exp(-Rx), \tag{4.11}
\]

where \( \rho > 0 \) and \(-R < 0\) are the solutions to the following quadratic equation with respect to \( x \)

\[
\frac{\sigma^2}{2} x^2 + c x - \delta = 0.
\]

Finally, from Equation (4.11), we obtain by solving \( h''(b_{\text{opt}}) = 0 \)

\[
b_{\text{opt}} = \frac{2}{\rho + R} \ln \frac{R}{\rho}. \tag{4.12}
\]

The factorization formula

\[
V(u, b) = \frac{h(u)}{h'(b)} \tag{4.13}
\]

is well known [(e.g. (7.5) of Gerber & Shiu, 1998)] and satisfied for general distributions with the p.d.f. \( f(x) \), where \( h(x) \) is a positive solution to the following integro-differential equation:

\[
ch'(x) - (\lambda + \delta)h(x) + \lambda \int_0^u h(x - y)f(y)dy = 0, \quad x > 0. \tag{4.14}
\]

Let \( f^*(s) \) and \( h^*(s) \) denote the Laplace transform of \( f(x) \) and \( h(x) \), namely,

\[
f^*(s) = \int_0^\infty e^{-sx} f(x)dx, \quad h^*(s) = \int_0^\infty e^{-sx} h(x)dx.
\]

Then, the Laplace transform of Equation (4.14) is (see e.g. Section 7.8 of Dickson, 2005)

\[
c \{-h(0) + sh^*(s)\} - (\lambda + \delta)h^*(s) + \lambda f^*(s)h^*(s) = 0, \tag{4.15}
\]
which yields \( h^*(s) = ch(0)/L(s) \) with \( L(s) = cs - (\lambda + \delta) + \lambda f^*(s) \). Because \( \lim_{s \to \infty} f^*(s) = 0 \) and \( \lim_{s \to \infty} h^*(s) = 0 \), it follows from Equation (4.15),

\[
\lim_{s \to \infty} sh^*(s) = h(0). \tag{4.16}
\]

**Example 4.3**  [Mixture of exponential distribution (Gerber et al., 2006)] Suppose that \( X_i \) follows a mixture of the exponential distribution with the p.d.f.

\[
f(x) = \sum_{i=1}^{h} A_i \beta_i e^{-\beta_i x},
\]

where \( 0 < \beta_1 < \cdots < \beta_h , A_i > 0 \) with \( A_1 + \cdots + A_h = 1 \). Ross (2014) and others called such a density hyperexponential. In this case, from Equations (4.15) and (4.16), we can write

\[
h^*(s) = h(0) \prod_{i=1}^{h} \frac{(s + \beta_i)}{(s - \rho_i)}.
\]

where \( \rho_1, \rho_2, \ldots, \rho_h \) are the solutions to \( L(s) = 0 \), which are called Lundberg’s fundamental equation (Gerber & Shiu, 1998), that is,

\[
s - \frac{\lambda + \delta}{c} + \frac{\lambda}{c} \sum_{i=1}^{h} \frac{A_i \beta_i}{s + \beta_i} = 0,
\]

with \( -\rho_h < \cdots < -\rho_1 < 0 < \rho_0 \). By the method of partial fractions, we have

\[
h^*(s) = h(0) \sum_{k=0}^{h} C_k \frac{1}{s - \rho_k},
\]

and it follows that

\[
h(x) = h(0) \sum_{k=0}^{h} C_k \exp(\rho_k x), \quad x > 0,
\]

where \( C_k = \prod_{i=1}^{h} (\rho_k + \beta_i) / \prod_{i=1}^{h} (\rho_k - \rho_i) \). The optimal dividend barrier \( b_{\text{opt}} \) is obtained as the solution to

\[
h''(b) = h(0) \sum_{k=0}^{h} C_k \rho_k^2 \exp(\rho_k b) = 0.
\]

However, we cannot obtain the explicit form of \( b_{\text{opt}} \) when \( h \geq 2 \).

**Example 4.4**  [Gamma distribution (Gerber et al., 2008)] Suppose that \( X_i \) follows a gamma distribution with the p.d.f.

\[
f(x) = \frac{\beta^m}{(m-1)!} x^{m-1} e^{-\beta x},
\]

where \( \beta > 0 , m \in \mathbb{N} \). In this case, we can write

\[
h^*(s) = \frac{ch(0)}{L(s)} = \frac{(s + \beta)^m}{\prod_{k=0}^{m} (s - \rho_k)}.
\]

where \( \rho_0, \rho_1, \ldots, \rho_m \) are the solutions to \( L(s) = cs - (\lambda + \delta) + \lambda \left( \frac{1}{s + \beta} \right)^m = 0 \). Then, in the same way as Example (4.3), we have
where \( C_k = (\rho_k + \beta)^m / \prod_{l \neq k} (\rho_k - \rho_l) \). The optimal dividend barrier \( (b_{\text{opt}}) \) is obtained as the solution to

\[
h''(b) = h(0) \sum_{k=0}^{m} C_k \rho_k^2 \exp(\rho_k b) = 0.
\]

However, we cannot obtain the explicit form of \( b_{\text{opt}} \) when \( m \geq 2 \).

From the factorization formula (4.11), the optimal dividend barrier \( (b_{\text{opt}}) \) is defined as the solution to \( h(\beta) = 0 \); however, the explicit form cannot be obtained except in special cases. From this point of view, some asymptotic formulae for \( b_{\text{opt}} \) have been considered by Højgaard (2002), Dickson and Drekic (2006), and Gerber et al. (2008).

Example 4.5  [De Vylder approximation (A) (Gerber et al., 2008)] This approximation method aims to approximate the original model by using a model with exponential claim amounts and the formula for the optimal dividend barrier. The De Vylder approximation has been used by Højgaard (2002), Dickson and Drekic (2006), and Gerber et al. (2008). Let

\[
U(t) = u - ct + S(t), \quad S(t) = \sum_{i=1}^{N(t)} X_i,
\]

denote the original model where \( X_i \) has the p.d.f. \( f(x) \) with the first three moments \( \mu_j = E(X_j) \) for \( j = 1, 2, 3 \), and \( N(t) \) has the intensity parameter \( \lambda \).

Suppose that the parameters \( c \) and \( \lambda \) are replaced by \( \bar{c} \) and \( \bar{\lambda} \), respectively, and \( f(x) \) is replaced by \( \beta e^{-\rho x} \), satisfying the first three moments per unit time of the process \( \{U(t)\} \). Let

\[
\bar{U}(t) = u - \bar{c} t + \bar{S}(t), \quad \bar{S}(t) = \sum_{i=1}^{\bar{N}(t)} \bar{X}_i,
\]

denote the approximated model where \( \bar{X}_i \) has the p.d.f. \( \beta e^{-\rho x} \) with the first three moments

\[
\bar{\mu}_j = E(\bar{X}_j) = \frac{\mu_j}{\beta} (j = 1, 2, 3),
\]

and \( \bar{N}(t) \) has the intensity parameter \( \bar{\lambda} \). Then, the conditions for \( E(\bar{U}(t)) = E(U(t)) \) \( (j = 1, 2, 3) \) are

\[
c - \lambda \mu_1 = \bar{c} - \frac{\bar{\lambda}}{\beta}, \quad \lambda \mu_2 = \frac{2 \bar{\lambda}}{\beta^2}, \quad \lambda \mu_3 = \frac{6 \bar{\lambda}}{\beta^3}.
\]

Therefore, \( \beta, \lambda, \bar{\lambda} \) and \( \bar{c} \) are given by

\[
\beta = \frac{3 \mu_2}{\mu_3}, \quad \bar{\lambda} = \frac{9}{2} \lambda \frac{\mu_3^2}{\mu_2^3}, \quad \bar{c} = c - \lambda \mu_1 + \frac{3 \lambda}{2} \frac{\mu_3^2}{\mu_2^3}.
\]

By using this result and Equation (4.10), we get the approximation

\[
b_{\text{opt}} \approx \frac{1}{\beta + \bar{R}} \ln \frac{R^2 (\beta - \bar{R})}{\beta^2 (\beta + \bar{R})},
\]

(4.17)

where \( \beta > 0 \) and \( -\bar{R} < 0 \) are the solutions to the quadratic equation

\[
x^2 + \left( \beta - \frac{\lambda + \delta}{\bar{c}} \right) x - \frac{\beta \delta}{\bar{c}} = 0.
\]
Example 4.6  [De Vylder approximation (B) (Gerber et al., 2008)] Another approximation method is not to modify c, but to choose $\tilde{\lambda}$ and $\tilde{\beta}$, satisfying the first two moments per unit time of the process $(U(t))$. In this case, the conditions are

$$\lambda_1 = \frac{\tilde{\lambda}}{\beta}, \quad \lambda_2 = \frac{2\tilde{\lambda}}{\beta^2},$$

Therefore, $\tilde{\beta}$ and $\tilde{\lambda}$ are given by

$$\tilde{\beta} = \frac{2\mu_1}{\mu_2}, \quad \tilde{\lambda} = \frac{2\lambda_1^2}{\mu_2}.$$ 

By using this result and Equation (4.10), we get the approximation

$$b_{\text{opt}} \approx \frac{1}{\tilde{\beta} + R} \ln \frac{\tilde{R}^2(\tilde{\beta} + \tilde{R})}{\tilde{\beta}^2(\tilde{\beta} + \tilde{R})},$$   \hfill (4.18)

where $\tilde{\beta} > 0$ and $-\tilde{R} < 0$ are the solutions to the quadratic equation

$$x^2 + \left(\tilde{\beta} - \frac{\tilde{\lambda} + \delta}{c}\right)x - \frac{\tilde{\beta}\delta}{c} = 0.$$ 

Example 4.7  [Diffusion approximation (Gerber et al., 2008)] This approximation method aims to approximate the original model by using a model with a diffusion process and the formula for the optimal dividend barrier. The idea is to replace the original model $U(t) = u + ct - S(t)$ with $U(t) = u + ct - \sigma W(t)$, where the drift parameter is $c = c - \lambda \mu$, and the volatility parameter is $\sigma^2 = \lambda \mu$. Then, from Equation (4.12), we get the approximation

$$b_{\text{opt}} \approx \frac{2}{\tilde{\beta} + R} \ln \frac{\tilde{R}}{\tilde{\beta}},$$   \hfill (4.19)

where $\tilde{\beta} > 0$ and $-R < 0$ are the solutions to the quadratic equation

$$\frac{\tilde{\beta}^2}{2} x^2 + \tilde{\sigma} x - \tilde{\sigma} = 0.$$ 

Example 4.8  [Diffusion approximation (Order 1, Order 2) (Gerber et al., 2008)] This approximation method is considered to develop a refined approximation compared with Example (4.7). Suppose that the original model $(U(t) = u + ct - S(t))$ is a compound Poisson model with the premium rate $c$, intensity parameter $\lambda > 0$, and p.d.f. $f(x)$ of the claim amount $X$. In this method, rescaling the p.d.f. of the claim amount is considered. For a fixed $m > 0$, we reconsider another compound Poisson model

$$U_m(t) = u + c_m t - S_m(t), \quad S_m(t) = \sum_{i=1}^{N_m(t)} X_{i,m},$$

where $(N_m(t))$ is a Poisson process with the intensity parameter $\lambda_m > 0$ and $X_{i,m}$ with the p.d.f. $f_m(x)$. Here, we assume that the p.d.f. $f_m(x)$ is defined by

$$f_m(x) = m f\left(\frac{x}{m}\right), \quad x > 0$$

and $c_m, \lambda_m$ are determined by the corresponding first two moments of the process $(U_m(t))$ and $(U(t))$. Then, it follows that $E(\lambda'_m) = mE(\lambda)' = m\lambda \mu$. Note that if $m \to 0$, $(U_m(t))$ converges to a diffusion process and if $m = 1$, $(U_m(t))$ corresponds to the original process.

Next, we consider a diffusion approximation for the model $(U_m(t))$. We replace the model $U_m(t) = u + c_m t - S_m(t)$ with $U_0(t) = u + c_0 t - \sigma_0 W(t)$, where the drift parameter and volatility parameter are given by
\( c_0 = c_m - \lambda_m(m\mu_1), \quad \sigma_0^2 = \lambda_m(m^2\mu_2). \)  \(\text{(4.20)}\)

Then, from Example (4.7), we have the diffusion approximation for \( b_{\text{opt}} \) as

\[
b_{\text{opt}}^{(b)} = \frac{2}{R_0 + \rho_0} \ln \frac{R_0}{\rho_0},
\]

where \( \rho_0 > 0 \) and \(-R_0 < 0\) are the solutions to the quadratic equation

\[
\mathcal{L}_0(x) = \frac{\sigma_0^2}{2} x^2 + c_0 x - \delta = 0. \quad \text{(4.21)}
\]

On the contrary, we can approximately express the optimal dividend barrier \( b_{\text{opt}} \) as \([\text{see (4.2) of Gerber et al. (2008)}]\)

\[
b_{\text{opt}} \approx \frac{1}{R + \rho} \ln \frac{-\mathcal{L}'(\rho)R^2}{-\mathcal{L}'(-R)\rho^2},
\]

where \( \rho, -R \) are the solutions to Lundberg’s fundamental equation \( (\mathcal{L}(s) = 0)\). Hence, we deal with the true optimal dividend barrier for the model \( (U_m(t)) \) as

\[
b_{\text{opt}}^{(m)} = \frac{1}{R_m + \rho_m} \ln \frac{-\mathcal{L}'(\rho_m)R_m^2}{-\mathcal{L}'(-R_m)\rho_m^2},
\]

where \( \rho_m, -R_m \) are the solutions to

\[
\mathcal{L}_m(s) = c_m s - (\lambda_m + \delta) + \lambda_m f_m^*(s) = 0. \quad \text{(4.22)}
\]

The Taylor expansion for the Laplace transform \( f_m^*(s) \) is obtained as follows:

\[
f_m^*(s) = \int_0^\infty e^{-sx} f_m(x) dx = \int_0^\infty e^{-msy} f(y) dy = f^*(ms) = \sum_{k=0}^\infty \frac{(-ms)^k}{k!} \mu_k.
\]

\(\text{(4.23)}\)

From Equations (4.21–4.23), the Lundberg function is written as

\[
\mathcal{L}_m(s) = \frac{\sigma_0^2}{2} (s - \rho_0)(s + R_0) - \frac{1}{6} \frac{\mu_1}{\mu_2} \sigma_0^2 s^3 + o(m),
\]

\(\text{(4.24)}\)

which leads to

\[
\rho_m = \rho_0 + \frac{1}{3} \frac{\mu_1}{\mu_2} \frac{\rho_0^3}{\rho_0 + R_0} + o(m), \quad -R_m = -R_0 + \frac{1}{3} \frac{\mu_1}{\mu_2} \frac{(-R_0)^3}{\rho_0 + R_0} + o(m),
\]

because of \( \mathcal{L}_m(R_m) = \mathcal{L}_m(-R_m) = 0 \). Furthermore, by differentiating Equation (4.24), we have

\[
\mathcal{L}_m'(\rho_m) = \frac{\sigma_0^2}{2} \left\{ \rho_0 + R_0 + \frac{2}{3} \frac{\mu_1}{\mu_2} \frac{\rho_0^3}{\rho_0 + R_0} - \frac{\mu_1}{\mu_2} R_0^2 + o(m) \right\},
\]

\[
\mathcal{L}_m'(-R_m) = \frac{\sigma_0^2}{2} \left\{ -\rho_0 - R_0 + \frac{2}{3} \frac{\mu_1}{\mu_2} \frac{(-R_0)^3}{\rho_0 + R_0} - \frac{\mu_1}{\mu_2} (-R_0)^2 + o(m) \right\},
\]

which implies that we obtain \( b_{\text{opt}} \approx b_{\text{opt}}^{(m)} \approx b_{\text{opt}}^{(b)} - m\Delta_1 \) with

\[
\Delta_1 = \frac{1}{3} \frac{\mu_1}{\mu_2} \left( \frac{\rho_0^3}{(\rho_0 + R_0)^2} b_{\text{opt}}^{(m)} + \frac{1}{3} \frac{\mu_1}{\mu_2} \frac{3\rho_0^2 - 2\rho_0 (-R_0) + 3(-R_0)^2}{(\rho_0 + R_0)^2} \right). \quad \text{(4.25)}
\]

In a similar way, we have \( b_{\text{opt}} \approx b_{\text{opt}}^{(m)} \approx b_{\text{opt}}^{(b)} - m\Delta_1 + m^2\Delta_2 \) with
\[
\Delta_2 = \frac{1}{6} \mu_2 \left( \rho_0 - R_0 \right) (2\rho_0 - 3\rho_0 (-R_0) + 2 (-R_0)^2) (\rho_0 + R_0)^4
\]
\[
- \frac{1}{18} \mu_2^2 (\rho_0 - R_0) (\rho_0 - 4\rho_0 (-R_0) + (-R_0)^2) (5\rho_0^2 - 4\rho_0 (-R_0) + 5 (-R_0)^2) (\rho_0 + R_0)^6
\]
\[
+ \left\{ \frac{1}{12} \mu_2 + (-R_0)^2 \left[ 1 - \frac{1}{2} \left( \rho_0^2 - 3\rho_0 (-R_0) - 2\rho_0^2 (-R_0)^2 - 3\rho_0 (-R_0)^3 + (-R_0)^4 \right) \right] \right\} b_{\text{opt}}^{(0)}.
\] (4.26)\\

4.2. The deficit at ruin
Dickson and Waters (2004), Gerber et al. (2006), and Dickson (2005) considered that the expected present value of the deficit at ruin \( E[e^{-R_t} | U_b(T_b)] \) and the initial surplus \( u \) should be excluded from the expected present value of net income to shareholders. In that case, the objective function becomes

\[
L(u, b) = V(u, b) - Y(u, b) - u,
\]
(4.27)

where \( Y(u, b) = E[e^{-R_t} | U_b(T_b)] \) with the deficit at ruin \( |U_b(T_b)| \).

Similar to Equation (4.6), the expected present value of the deficit at ruin \( Y(u, b) \) has the following integro-differential equation (see (8.34) of Dickson, 2005):

\[
\frac{\partial}{\partial u} Y(u, b) = \frac{1}{c} \int_0^u f(x) Y(u - x, b) dx - \frac{\lambda}{c} \int_u^\infty (x - u) f(x) dx.
\]
(4.28)

The optimal dividend barrier \( b_{\text{opt}}^{\text{opt}} = b_{\text{opt}}^{(0)}(u) \) is defined by the maximizer of \( V(u, b) \).

Definition 4.2

\[
b_{\text{opt}}^{\text{opt}} : = \arg \max_{b \geq u} L(u, b)
\] (4.29)

for a constant value \( u \geq 0 \).

In this case, the explicit form of \( b_{\text{opt}}^{\text{opt}} \) is not obtained even if each claim amount \( X_i \) follows an exponential distribution.

Example 4.9 [Exponential distribution (Dickson, 2005, Gerber et al., 2006)] Suppose that \( X_i \) follows an exponential distribution with \( E(X_i) = 1/\beta \) (i.e. \( f(x) = \beta e^{-\beta x} \)). Then, from Equation (4.28), we obtain the second-order differential equation

\[
\frac{d^2}{du^2} Y(u, b) + \left( \beta - \frac{\lambda + \delta}{c} \right) \frac{d}{du} Y(u, b) - \frac{\beta \delta}{c} Y(u, b) = 0,
\]
(4.30)

which yields

\[
Y(u, b) = D_0(b) e^{\alpha u} + D_1(b) e^{-\rho u},
\]
(4.31)

where \( \rho > 0 \) and \( -\beta < 0 \) are the same as in Example (4.1). From this result and the boundary condition \( \frac{d}{du} Y(u, b) \bigg|_{u=b} = 0 \), we have

\[
D_0(b) = R e^{-\rho b} D(b), \quad D_1(b) = \rho e^{\rho b} D(b), \quad D(b) = \frac{(\beta + \rho)(\beta - R)}{\beta^2 \left\{ \rho (\beta + \rho) e^{\rho b} + R (\beta - R) e^{-\rho b} \right\}}.
\]

In addition, since we can write from Example (4.1)

\[
V(u, b) = C_0(b) e^{\alpha u} + C_1(b) e^{-\rho u},
\]
where
\[ C_0(b) = (\beta + \rho)C(b), \quad C_1(b) = -(\beta - R)C(b), \quad C(b) = \frac{1}{\rho(\beta + \rho)e^{\beta t} + R(\beta - R)e^{-\beta t}}, \]

it follows that
\[ L(u, b) = \{ C_0(b) - D_0(b) \} e^{\beta u} + \{ C_1(b) - D_1(b) \} e^{-\beta u}. \] (4.32)

The optimal dividend barrier \( b^i_{\text{opt}} \) is obtained by solving
\[ \frac{\partial L(u, b)}{\partial b} = \{ C_0(b) - D_0(b) \} e^{\beta u} + \{ C_1(b) - D_1(b) \} e^{-\beta u} = 0 \]
or
\[ \rho^2 \{ C_0(b) - D_0(b) \} e^{\beta b} + R^2 \{ C_1(b) - D_1(b) \} e^{-\beta b} = 0, \]

which comes from another boundary condition \( \frac{\partial^2 L(u, b)}{\partial u^2} \bigg|_{u=b} = 0 \). However, they must be solved numerically.

**Example 4.10**  [Mixture of exponential distributions (Gerber et al., 2006)] Suppose that \( X_i \) follows a mixture of exponential distributions with the p.d.f.

\[ f(x) = \sum_{h=1}^{n} A_h e^{-\beta_h x}, \]

where \( 0 < \beta_1 < \cdots < \beta_n, A_1 + \cdots + A_n = 1 \). By applying the operator \( \prod_{i=1}^{n} \left( \frac{1}{\beta} + \beta_i \right) \) to the integro-differential equation (Equation 4.6) and

\[ \frac{\partial}{\partial u} Y(u, b) = \frac{\lambda + \delta}{C} Y(u, b) + \frac{\delta}{C} \int_{0}^{u} f(x)f(u-x,b)dx, \] (4.33)

we have

\[ V(u, b) = \sum_{k=0}^{n} C_k(b)e^{\rho_i u}, \quad Y(u, b) = \sum_{k=0}^{n} D_k(b)e^{\rho_i u}, \]

where \( \rho_0, \rho_1, \ldots, \rho_h \) are the same as in Example (4.3), and \( C_k(b), D_k(b), k = 0, \ldots, h \) are determined by

\[ \sum_{k=0}^{n} \frac{C_k(b)}{\lambda + \rho_k} = 0, \quad \sum_{k=0}^{n} \frac{D_k(b)}{\lambda + \rho_k} = \frac{1}{\beta_i}, \quad i = 1, \ldots, h \]

and

\[ \sum_{k=0}^{n} C_k(b)\rho_i e^{\rho_i b} = 1, \quad \sum_{k=0}^{n} D_k(b)\rho_i e^{\rho_i b} = 0. \]

Hence, we have

\[ L(u, b) = \sum_{k=0}^{n} \{ C_k(b) - D_k(b) \} e^{\rho_i u}. \] (4.34)

The optimal dividend barrier \( b^i_{\text{opt}} \) is obtained by solving

\[ \frac{\partial L(u, b)}{\partial b} = \sum_{k=0}^{n} \{ C_k(b) - D_k(b) \} e^{\rho_i u} = 0 \]
or
\[ \sum_{k=0}^{h} \rho^k \left( C_k(b) - D_k(b) \right) e^{\lambda b} = 0. \]

However, they must be solved numerically.

### 4.3. Other dividend strategies

Thus far, we have considered that when the surplus attains a constant level \( b \), the premium income is paid to shareholders as dividends. By contrast, Avanzi (2009) introduced other dividend strategies as follows.

**Example 4.11** [Band/ (Multiple) Threshold/ Constant Barrier Strategies] For a “band strategy”\(^2\), dividends are paid according to the region in which the surplus \( U_A(t) \) is located. Given a region \( A = \{(0, b_0), (a_1, b_1), \ldots, (a_h, b_h)\} \) for \( 0 < b_0 \leq a_1 \leq b_1 \leq \cdots \leq a_h \leq b_h, \ h \in \mathbb{N} \), while \( U_A(t) \) is located in the region \( A \), no dividends are paid. On the contrary, while \( U_A(t) \) is located in the region \( \mathbb{R}_+ \setminus A \), dividends are paid. In a band strategy on the region \( A \), the aggregate dividend \( D_A(t) \) is defined by

\[
D_A(t) = \begin{cases} 
0 & \text{if } U_A(t) \in A \\
\int_c^b dt & \text{if } U_A(t) \in \mathbb{R}_+ \setminus A
\end{cases}
\]

with \( 0 \leq c \leq c \). Note that if \( c = c, U_A(t) \leq b \). A “threshold strategy” is a special case of the band strategy. When we set \( h = 1, b_0 = b \) under the band strategy, the definition of the aggregate dividend \( D_A(t) \equiv D_A(t) \) is

\[
D_A(t) = \begin{cases} 
0 & \text{if } U_A(t) < b \\
\int_c^b dt & \text{if } U_A(t) \geq b
\end{cases}
\]

with \( 0 \leq c \leq c \). Frostig (2005), Gerber et al. (2006), Lin and Pavlova (2006), Fang and Wu (2007), Wan (2007), Frostig (2008), Cheung, Dickson, and Drekic (2008), Yuen et al. (2009), and Avanzi (2009) discussed the threshold strategy.

A “constant barrier strategy” is a special case of the threshold strategy, in which the definition of \( D_A(t) \) corresponds to Equation (4.1) when we set \( c = c \) under the threshold strategy. A “multiple threshold strategy” or “multilayer strategy” is a combination of the threshold strategy. When we set \( a_i = b_i \) and \( 0 \leq \xi_i \leq c \) for \( i = 1, \ldots, h \) under the band strategy, the definition of the aggregate dividend \( D_A(t) \) is

\[
D_A(t) = \begin{cases} 
0 & \text{if } U_A(t) \in A \\
\int_c^b dt & \text{if } U_A(t) \in [b_{i-1}, b_i], \ i = 1, \ldots, h
\end{cases}
\]

where \( b_h = \infty \). Albrecher and Hartinger (2007), Badescu and Landriault (2008), and Lin and Sendova (2008) discussed the multilayer strategy.

**Example 4.12** [Linear barrier strategies] Gerber (1974) argued that a (constant) barrier strategy does not sufficiently take into account the safety of the company, since the probability of ruin is 1 (i.e. \( P(T_b = \infty) = 1 \)). In a linear barrier strategy with the parameters \( (a, b) \), the aggregate dividend \( D_{a,b}(t) \) is defined by

\[
D_{a,b}(t) = \begin{cases} 
0 & \text{if } U_{a,b}(t) < b + at \\
(c - a)dt & \text{if } U_{a,b}(t) = b + at
\end{cases}
\]

(4.38)
Gerber (1981), Siegl and Tichy (1999) and Albrecher, Hartinger, and Tichy (2005) also discussed the linear strategy.

Example 4.13 [Non-linear barrier strategies] These are a general family of strategies allowing for a (reflecting) barrier $b(t)$ as a function of time. In this strategy, the aggregate dividend $D_k(t)$ is defined by

$$dD_k(t) = \begin{cases} 0 & \text{if } U_k(t) < b(t) \\ (c - b(t)')dt & \text{if } U_k(t) = b(t) \end{cases}$$

Example 5.1 [Proportional Reinsurance] Under a proportional reinsurance arrangement, the insurer pays a fixed proportion $\alpha$ of each claim that occurs during the period of the reinsurance arrangement. The remaining, $1 - \alpha$, of each claim is paid by the reinsurer, which implies that $X_i^{(\alpha)} = \alpha X_i$. On the contrary, the insurance company (i.e. insurer) effects reinsurance by paying a reinsurance premium continuously at a constant rate, then the process becomes a net of the reinsurance surplus process $\{U_k(t)\}_{t \geq 0}$ given by

$$U^{(\alpha)}(t) = u + c^{(\alpha)}t - \sum_{i=1}^{N(t)} X_i^{(\alpha)}$$

where $u$ is the initial surplus, $c^{(\alpha)}$ is the insurer’s net premium income per unit time, $\{N(t)\}$ is a Poisson process with the intensity parameter $\lambda > 0$ and $X_i^{(\alpha)}$ is the amount the insurer pays on the $i$-th claim. Note that the insurer’s net premium rate $c^{(\alpha)}$ and payable amount $X_i^{(\alpha)}$ depend on the reinsurance scheme and reinsurance parameter $\alpha \geq 0$.

Example 5.2 [Excess-of-loss Reinsurance] Under an excess-of-loss reinsurance arrangement, a claim is shared between the insurer and reinsurer only if it exceeds a fixed amount called the retention level $a \geq 0$. If the claim amount is smaller than $a$, the insurer has to pay the entire claim. The remaining, $\max(X_i - a, 0)$ of the post-retention claim is paid by the reinsurer. Hence, the aggregate claim amount $X_i^{(\alpha)} = \min(a, X_i)$ and the remaining $X_i - \min(a, X_i) = \max(X_i - a, 0)$ is paid by the reinsurer.

Let $f(x)$ denote the p.d.f. of the claim amount $X_i$. Then, the reinsurance premium for the reinsurance portion $\max(X_i - a, 0)$ is written as
As a consequence, the insurer's net premium rate \( c^{(a)} \) under the expected value principle is given by
\[
c^{(a)} = c - c^{(a)} = \{ (1 + \theta) \mu - (1 + \theta) \mu^{(a)} \} \lambda.
\]
Note that if \( 0 < \theta \eta < \theta \), the insurance company can earn a profit without any risk by \( a = 0 \).

The optimal reinsurance parameter \( (a_{\text{opt}} \equiv a_{\text{opt}}(u)) \) is defined by the maximizer of an objective function \( Q(a) \):

**Definition 5.1**

\[
a_{\text{opt}}^* = \arg \max_{a \geq 0} Q(a),
\]

where \( Q(a) \) is satisfied with \( Q(a) < Q(a_{\text{opt}}) < \infty \) for any \( a \geq 0 \).

**Example 5.3** [Adjustment Coefficient] Let \( M^{(a)}_X(r) \) denote the moment-generating function of the insurer's payable amount \( X^{(a)}_i \), that is, \( M^{(a)}_X(r) = E(e^{rX}) \) for \( r \in \mathbb{R} \). Then, the reinsurance adjustment coefficient \( R(a) \) is defined by the positive solution to
\[
c^{(a)} = \lambda \int_{0}^{\infty} (x - a) f(x) dx = (1 + \theta) \lambda \mu^{(a)}.
\]

Waters (1983) and Centeno (1986) among others discussed the optimality of the reinsurance parameter under \( Q(a) \equiv R(a) \).

**Example 5.4** [Ruin Probability] Define
\[
Q(a) = -P\{ T^{(a)} < \infty | F^{(a)}(0) = u \},
\]

which implies that the optimal reinsurance parameter \( a_{\text{opt}} \) is the minimizer of the ruin probability given \( u \geq 0 \). Dickson and Waters (1996) discussed the optimal proportion retained \( a_{\text{opt}} \) based on the ruin probability. According to Dickson (2005), the \( a_{\text{opt}} \) should be the maximizer of the adjustment coefficient rather than the minimizer of the ruin probability, at least for large values. In addition, the excess-of-loss reinsurance maximizes the adjustment coefficient (and thus minimizes the probability of ruin) (see Bowers et al., 1997).

Recently, many researchers have considered the optimal reinsurance and dividend control problem of an insurance company. As for proportional reinsurance, Højgaard and Taksar (1998), Taksar (2000), David (2005), and Cao and Wan (2009) among others considered optimal investment and reinsurance strategies to maximize (or minimize) the utility functions (included the ruin probability) for a diffusion risk model or extended models. As for the excess-of-loss reinsurance, Asmussen, Højgaard, and Taksar (2000), Beveridge, Dickson, and Wu (2007), Gu, Guo, Li, and Zeng (2012), Zhao, Rong, and Zhao (2013), and Liu and Hu (2014) among others discussed the optimization (i.e. the maximization or minimization of utility functions) of the excess-of-loss reinsurance and investment problems for an insurer with diffusion or jump-diffusion risk models.
5.2. Tax
Albrecher and Hipp (2007) discussed the effect of tax payments under a loss-carry forward system in the Cramér–Lundberg model without dividends. The loss-carry forward system is motivated by its realistic feature of carried forward losses (i.e., occurred losses can be deducted from later income and hence they reduce the taxable profit of the business). Under this system, the classical Cramér–Lundberg model is modified as follows.

Example 5.5  [The Cramér–Lundberg Model under the loss-carry forward system] Suppose that the surplus of an insurance company at time t is defined by (Equation 2.2) for t ≥ 0 (in this section, we assume the Cramér–Lundberg model):

\[ U(t) = u + ct - S(t), \quad S(t) = \sum_{n=1}^{N(t)} X_n. \]

We assume that tax is paid at a fixed rate γ ∈ (0, 1) of the insurer’s premium income whenever it is in a profitable situation. Suppose that the initial surplus is u. We now consider the risk process \( \{U(t); t ≥ 0\} \), which is the surplus process with tax. Before the first claim occurs, the (aggregated) premium income at time t becomes \( c(1 - γ)t \) and the (aggregated) tax payment at time t becomes \( cγt \).

Let \( Δ_t, T, \gamma \) be the time of the first claim. Then, the gains level (the surplus immediately before the first claim occurs) is set to \( M_1 := u + c(1 - γ)Δ_1 T \). The surpluses immediately before and after the first claim occurs are written as

\[ U_j(1_-) = u + c(1 - γ)Δ_1 I, \quad U_j(1_+) = U_j(1_-) - X_1, \]

respectively, if 0 < \( U(T_i) \) and \( cΔ_i > X_i \), the insurer does not have to pay any tax during \( [T_i, T_i + X_i/c] \), because the surplus is less than \( M_i \) and we set the time that the surplus equals \( M_i \) as \( σ_i := t_1 + X_i/c \). Then, we can see that \( U_j(σ_i) = M_j \) and from \( t = σ_i \) to \( t = T_i + Δ_i T = T_j \), the tax is paid. By denoting the term at which the tax is paid as \( W_{ij} \), we can write \( W_{ij} = T_j - T_i - σ_i \). Then, the new gains level is set to \( M_j = M_i + c(1 - γ)W_{ij} \). Let \( N(t) \) be the underlying claims arrival process. Albrecher and Hipp (2007) formally defined a sequence of gains levels \( \{M_n\} \), waiting times (term at which tax is paid) \( \{W_n\} \) and starting times of periods with profit \( \{σ_n\} \) as

\[ M_n = M_{n-1} + c(1 - γ)W_n, \]

\[ W_n = \inf\{t > 0; N(σ_{n-1} + t) > N(σ_{n-1})\}, \]

\[ σ_n = \inf\{t > σ_{n-1} + W_n; U_j(t) = M_n\}, \]

with \( σ_0 = 0 \) and \( M_0 = u \) for \( n = 1, 2, \ldots \) where the surplus process \( U_j(t) \) at time t is

\[ U_j(t) = M_{\Lambda(t)} + c(1 - γ)\min\{W_{\Lambda(t)}, t - σ_{\Lambda(t)}\} - \sum_{i=N(t)_{\Lambda(t)}}^{N(t)} X_i, \]

where \( \Lambda(t) = \sup\{n ∈ \mathbb{N}; σ_n ≤ t\} \).

Albrecher et al. (2008) introduced the following modified dual risk model, which is reasonable to model the surplus process for pharmaceutical and petroleum companies.

Example 5.6  [The Dual Risk Model under the Loss-carry Forward System] Let \( \{Y_j\} \) denote a sequence of i.i.d. random variables, where \( Y_j \) corresponds to the revenue associated with the \( j \)-th innovation, and let \( \{Δ_j, T\} \) denote a sequence of the inter-innovation times (i.i.d. random variables). Then, the sequence of innovation times \( \{T_j\} \) is defined by \( T_0 = 0 \) and \( T_j = \sum_{i=1}^{j} Δ_i T \). Suppose that the innovation number process \( \{N(t), t ≥ 0\} \) is a renewal process with the inter-innovation times \( Δ_1 T, Δ_2 T, \ldots \)

The following dual risk model is thus reasonable to model the surplus process (without tax) for pharmaceutical and petroleum companies:
\[ R(t) = u - ct + S(t), \quad S(t) = \sum_{i=1}^{N(t)} Y_i, \]

where \( u \) is the company’s initial surplus, \( c \) is now the constant expense rate and \( (S(t), t \geq 0) \) is the aggregate revenue process. Let \( \gamma \in (0, 1) \) be the tax rate at which the insurance company has to pay the profit times \( \gamma \) as the tax. In this system, the profit is defined as the excess of each new record high of the surplus over the previous one. The value of the \( n \)th record high \( (J_n) \) is defined as follows:

\[ J_0 = u, \quad J_n = J_{n-1} + (1 - \gamma) \sum_{j=k_{n-1}+1}^{k_n} (Y_j - c\Delta_j T_j), \quad n = 1, 2, \ldots \]

where \( k_n = \inf_{k \in \mathbb{N}} \left\{ k > k_{n-1} \mid \sum_{j=k_{n-1}+1}^{k} (Y_j - c\Delta_j T_j) > 0 \right\} \) with \( k_0 = 0 \) is the number of innovations up to the time of the \( n \)-th record high. The resulting surplus with tax is given by

\[ R_j(t) = J_{n(t)} - \sum_{j=k_{n(t)}+1}^{N(t)} (c\Delta_j T_j - Y_j) - c(t - T_{n(t)}), \]

\[ = J_{n(t)} - c(t - T_{n(t)}) + \sum_{j=k_{n(t)}+1}^{N(t)} Y_j, \]

where \( \Lambda(t) = \sup \{ n \in \mathbb{N} \mid T_{k_n} \leq t \} \) and \( \sum_{j=k_{n(t)}+1}^{N(t)} Y_j = 0 \) if \( k_{n(t)} = N(t) \).

Note that \( Y_j - c\Delta_j T_j \) is the “profit and loss” from the \((i - 1)\)-th innovation to \( i \)-th innovation, and thus \( k_n - k_{n-1} \) corresponds to the minimum term that the “aggregated profit and loss” is positive. Under the loss-carry forward system, the loss is carried forward, which implies that while the aggregated profit and loss is negative, the company does not have to pay the tax. That is the reason why \( k_n \) is calculated. In addition, \( \{J_n\} \) corresponds to the surplus at the time that the \( n \)-th tax payment is made.

Table 1 shows an example of the dual risk model under the loss-carry forward system.

| \( j \) | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | ... |
|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| \( \Delta_j \) | 1.0 | 2.0 | 1.0 | 1.0 | 3.0 | 1.0 | 1.0 | 2.0 | 1.0 | 1.0 | ... |
| \( Y_j \) | 1.0 | 1.0 | 2.0 | 2.0 | 10.0 | 1.0 | 2.0 | 1.0 | 10.0 | 11.0 | 11.0 |
| \( T_j \) | 1.0 | 3.0 | 4.0 | 5.0 | 8.0 | 9.0 | 10.0 | 12.0 | 13.0 | 14.0 | 14.0 |
| \( c\Delta_j T_j \) | 2.0 | 6.0 | 8.0 | 10.0 | 16.0 | 18.0 | 20.0 | 24.0 | 26.0 | 28.0 | 28.0 |
| \( S_i(T_i) \) | 1.0 | 11.0 | 12.0 | 14.0 | 24.0 | 25.0 | 27.0 | 28.0 | 38.0 | 49.0 | 49.0 |
| \( R_i(T_i) \) | 9.0 | 15.0 | 14.0 | 14.0 | 18.0 | 17.0 | 17.0 | 14.0 | 22.0 | 31.0 | 31.0 |
| \( Y_j - c\Delta_j T_j \) | -1.0 | 6.0 | -1.0 | 0 | 4.0 | -1.0 | 0 | -3.0 | 8.0 | 9.0 | ... |
| \( \sum_{j=1}^{k_i} (Y_j - c\Delta_j T_j) \) | -1.0 | 5.0 | - | -1.0 | 1.0 | -1.0 | -0.0 | 3.0 | - | ... |
| \( \sum_{j=k_i+1}^{N(t)} (Y_j - c\Delta_j T_j) \) | - | - | - | -1.0 | -1.0 | -0.0 | 4.0 | - | ... |
| \( \sum_{j=k_i+1}^{N(t)} Y_j \) | - | - | - | - | - | 9.0 | - | ... |
| \( \Lambda(T) \) | 0 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 3 | 4 | ... |
| \( J_{n(T)} \) | 10.0 | 14.0 | 14.0 | 14.0 | 16.4 | 16.4 | 16.4 | 16.4 | 19.6 | 26.8 | ... |
| \( c(T_j - T_{n(T)}) \) | 2.0 | 0 | 2.0 | 4.0 | 0 | 2.0 | 4.0 | 8.0 | 0 | 0 | ... |
| \( \sum_{j=k_{n(T)}+1}^{N(t)} Y_j \) | 1.0 | 0 | 1.0 | 3.0 | 0 | 1.0 | 3.0 | 4.0 | 0 | 0 | ... |
| \( R_j(T_j) \) | 9.0 | 14.0 | 13.0 | 13.0 | 16.4 | 15.4 | 15.4 | 12.4 | 19.6 | 26.8 | ... |
6. Statistical estimation

In this section, we discuss the statistical estimation of the ruin probability and Gerber–Shiu function. The statistical inference for the ultimate ruin probability has been studied by many authors. Grandell (1978/1979), Csörgő and Teugels (1990), Deheuvels and Steinebach (1990), Csörgő and Steinebach (1991), and Embrechts and Mikosch (1991) discussed the statistical estimation of the adjustment coefficient, because the ruin probability is approximately expressed as a function of the adjustment coefficient by using the Cramér–Lundberg approximation. On the contrary, Frees (1986), Croux and Veraverbeke (1990), Bening and Korolev (2003), and Mnatsakanov et al. (2008) discussed the non-parametric estimation of the ruin probability. Shimizu (2012) proposed a non-parametric estimator of the Gerber–Shiu function when the surplus process \( \{U_t\} \) is a compound Poisson claim process plus a diffusion perturbation (called the Wiener–Poisson risk model, defined by Equation (2.4)).

6.1. Statistical estimation of the adjustment coefficient

Suppose that the surplus of an insurance company at time \( t \) follows the Cramér–Lundberg model defined by Equation (2.2), where we assume that the claim size \( X_i \) has the probability density \( f(x) \) for \( x \geq 0 \). Let \( \hat{\mu} \) denote the ruin probability defined by Equation (3.1). We introduce the adjustment coefficient \( \dot{R} \), which is the positive solution to

\[
-c\dot{R} + \lambda \dot{M}_X = 0
\]

(6.1)

with respect to \( r \), where \( \dot{M}_X(r) = E(e^{rX}) = \int_0^\infty e^{rx}f(x)dx \). Hence, the adjustment coefficient satisfies

\[
\dot{M}_X(R) - \frac{cR}{\lambda} = 1.
\]

Here, the left-hand side of Equation (6.1) comes from the Laplace exponent of \( U(1) - u \), that is,

\[
\ell(r) = \ln E(e^{-r(U(1) - u)}) = -rc + \ln E\{M_X(r)^{N(1)}\} = -rc + \ln e^{i\dot{M}_X(r) - 1}
\]

which corresponds to Lundberg’s fundamental equation with \( \delta = 0 \). Then, the Cramér–Lundberg approximation

\[
\psi(u) \approx Ce^{-Ru}, \quad C = \frac{c/\dot{\lambda} - E(X)}{E(Xe^{RX}) - c/\dot{\lambda}}
\]

is derived (see Minkova, 2010). A natural problem is then to consider the estimation of the adjustment coefficient \( \dot{R} \). Grandell (1978/1979) proposed the following estimator for \( \dot{R} \).

**Definition 6.1** Suppose that \( X_1, \ldots, X_N(T) \) and \( T \) are observed, where \( X_i \) is the \( i \)-th claim amount, \( T \) is the ruin time and \( N(T) \) is the claim frequency up to the ruin. Then, an estimator \( \hat{R} \) of \( R \) is defined as the solution to

\[
\hat{M}_X(r) - \frac{cR}{\lambda} = 1,
\]

with respect to \( r \), where \( \hat{M}_X(r) = \frac{1}{N(T)} \sum_{k=1}^{N(T)} e^{RX_k} \) and \( \hat{\lambda} = \frac{N(T)}{T} \).

Furthermore, Grandell (1978/1979) showed its asymptotic normality as follows:

**Theorem 6.1** If \( E(X_i) - c/\dot{\lambda} < 0 \) and \( \lim_{r \to \infty} \{M_X(r) - cr/\dot{\lambda}\} = \infty \) are satisfied, then

\[
\sqrt{T}(\hat{R} - \dot{R}) \to N\left(0, \frac{\psi'(2\dot{R})}{\lambda(\psi'(\dot{R})^2)}\right)
\]
where \( y(r) = M_y(r) - cr/\lambda. \)

Csörgő and Teugels (1990) derived an estimator of \( R \) based on the sample of size \( n \). Deheuvels and Steinebach (1990) and Csörgő and Steinebach (1991) derived an estimator of \( R \) by using intermediate order statistics. Embrechts and Mikosch (1991) discussed a bootstrap procedure to estimate the adjustment coefficient \( R \). Furthermore, Shimizu (2009) considered the estimation of the adjustment coefficient \( R \) when the underlying process consists of a Brownian motion plus negative jumps where each jump size of the process does not necessarily correspond to each claim size.

### 6.2. Non-parametric estimation of the ruin probability

Suppose that the surplus \( U(t) \) is defined by (Equation 2.2), where the claim size \( X_i \) has the mean \( E(X_i) = \mu \), the distribution function \( F(x) \), and the p.d.f. \( f(x) \). Let \( \psi(u) \) denote the ruin probability defined by (Equation 3.1). Define \( \tilde{\psi}(u) = 1 - \psi(u) \) as the non-ruin probability. Then, it follows the following integrals (Dickson, 2005):

\[
\frac{d}{du} \tilde{\psi}(u) = \frac{\lambda}{c} \tilde{\psi}(u) - \frac{\lambda}{c} \int_0^u \tilde{\psi}(u-x)f(x)dx, \quad u \geq 0.
\]  

(6.2)

Integrating Equation (6.2) from 0 to \( t \) leads to (see e.g. Minkova, 2010.)

\[
\tilde{\psi}(t) = \tilde{\psi}(0) + \frac{\lambda}{c} \int_0^t \tilde{\psi}(t-x)(1-F(x))dx, \quad t \geq 0.
\]  

(6.3)

Let \( \tilde{\psi}^*(s) \), \( \tilde{\psi}^{**}(s) \) denote the Laplace transforms

\[
\tilde{\psi}^*(s) = \int_0^\infty e^{-sz} \tilde{\psi}(z)dz, \quad \tilde{\psi}^{**}(s) = \int_0^\infty e^{-sz} d\tilde{\psi}(z).
\]

Taking the Laplace transform in both sides of Equation (6.3) gives

\[
\tilde{\psi}^*(s) = \frac{\psi(0)}{s\left(1 - \frac{s}{c} f^*(s)\right)} = \frac{\psi(0)}{s} \sum_{r=0}^{\infty} \left(\frac{s}{c} f^*(s)\right)^r, \quad f^*(s) = \int_0^\infty f(x) e^{-sx} dx
\]

where \( f^*(s) \) is the Laplace transform of \( f(x) \). Since it follows that \( \tilde{\psi}^{**}(s) = s\tilde{\psi}^*(s) \), we have

\[
\tilde{\psi}^{**}(s) = \psi(0) \sum_{r=0}^{\infty} \left(\frac{s}{c} f^*(s)\right)^r \left(\frac{f^*(s)}{\mu}\right)^r.
\]  

(6.4)

The inverse Laplace transform of Equation (6.4) leads to the following formula (the so-called “Pollaczek–Khinchin formula”):

\[
\psi(u) = \left(1 - \frac{\lambda \mu}{c}\right) \sum_{r=1}^{\infty} \left(\frac{\lambda \mu}{c}\right)^r \{1 - G^{(r)}(u)\},
\]  

(6.5)

where \( G(u) = \frac{1}{\mu} \int_0^u (1 - F(y))dy \) and \( G^{(r)} \) denote the \( r \)-fold convolution of the distribution function \( G \) itself. The following formula is equivalent to Equation (6.5):

\[
\psi(u) = \frac{\lambda \mu}{c} - \left(1 - \frac{\lambda \mu}{c}\right) \sum_{r=1}^{\infty} \left(\frac{\lambda \mu}{c}\right)^r G^{(r)}(u)
\]

Assuming that \( \mu \) and \( \lambda \) are known, Croux and Veraverbeke (1990) proposed its estimator as

\[
\psi_n(u) = \frac{\lambda \mu}{c} - \left(1 - \frac{\lambda \mu}{c}\right) \sum_{r=1}^{m_0} \left(\frac{\lambda \mu}{c}\right)^r U_m(u),
\]  

(6.6)
where \( \{m_n\} \) is a sequence of natural numbers that depends on the sample size \( n \) satisfied with \( m_n \leq n, \lim_{n \to \infty} m_n = +\infty \), and \( U_{nr}(u) \) is a U-statistic given as follows.

**Definition 6.2** Suppose that \( X_1, \ldots, X_n \) are observed, where \( X_i \) is the \( i \)th claim amount. Then, the estimator \( U_{nr}(u) \) of \( G(r)(u) \) for a fixed \( r, u \) is defined by

\[
U_{nr}(u) = \left( \begin{array}{c} n \\ r \end{array} \right)^{-1} \sum_{1 \leq i_1 < \cdots < i_r \leq n} h_r(x_{i_1}, \ldots, x_{i_r}; u),
\]

where \( \sum_{1 \leq i_1 < \cdots < i_r \leq n} \) is the sum over all \( \left( \begin{array}{c} n \\ r \end{array} \right) \) distinct combinations of \( \{i_1, i_2, \ldots, i_r\} \) satisfied with \( 1 \leq i_1 < \cdots < i_r \leq n \), and \( h_r(x_{i_1}, \ldots, x_{i_r}; u) \) is a symmetric kernel

\[
h_r(x_{i_1}, \ldots, x_{i_r}; u) = \frac{1}{\mu} \int_0^\infty \cdots \int_0^\infty I(y_1 + \cdots + y_r \leq u) \prod_{j=1}^r I(x_j > y_j) dy_1 \cdots dy_r.
\]

Note that for \( r = 1, 2, \ldots \), we can interpret \( G(r)(u) = P(Y_1 + \cdots + Y_r \leq x) \), where \( Y_1, \ldots, Y_r \) are independent with a common density function \((1/\mu)(1 - F(y))\) for \( y \geq 0 \). Hence, we can write

\[
G(r)(u) = \frac{1}{\mu} \int_0^\infty \cdots \int_0^\infty I(y_1 + \cdots + y_r \leq u) \prod_{j=1}^r (1 - F(y_j)) dy_1 \cdots dy_r.
\]

Therefore, we can confirm that \( U_{nr}(u) \) is an unbiased estimator of \( G(r)(u) \), since

\[
E(U_{nr}(u)) = E(h_r(X_1, \ldots, X_r; u))
\]

\[
= \frac{1}{\mu} \int_0^\infty \cdots \int_0^\infty I(y_1 + \cdots + y_r \leq u) \prod_{j=1}^r E(I(x_j > y_j)) dy_1 \cdots dy_r
\]

\[
= \frac{1}{\mu} \int_0^\infty \cdots \int_0^\infty I(y_1 + \cdots + y_r \leq u) \prod_{j=1}^r (1 - F(y_j)) dy_1 \cdots dy_r = G(r)(u).
\]

Frees (1986) showed the consistency for the linear combination of U-statistics, i.e.

\[
\sum_{r=1}^m \left( \frac{\lambda \mu}{c} \right)^r U_{nr}(u) \overset{a.s.}{\rightarrow} \sum_{r=1}^m \left( \frac{\lambda \mu}{c} \right)^r G(r)(u).
\]

Furthermore, Croux and Veraverbeke (1990) showed the asymptotic normality for \( \psi_n(u) \) as follows.

**THEOREM 6.2** If \( \log n = o(m_n) \), \( n \to \infty \), then for each \( u > 0 \)

\[
\sqrt{n}(\psi_n(u) - \psi(u)) \overset{d}{\rightarrow} N\left( 0, \left( 1 - \frac{\lambda \mu}{c} \right)^2 \sum_{r,s=1}^m \left( \frac{\lambda \mu}{c} \right)^{r+s} \xi_n(1; u) \right).
\]

where

\[
\xi_n(1; u) = E\{h_r(X_1; u)h_s(X_r; u)\} - G(r)(u)G(s)(u),
\]

\[
h_r(x; u) = \frac{1}{\mu} \int_0^\infty G(r-1)(u-y)I(x < y)dy.
\]

Bening and Korolev (2003) mentioned that the above results can rarely be applied in practice since

- The Croux–Veraverbeke estimator Equation (6.6) is constructed from a sample with a non-random size,
• The property of the asymptotic normality of the proposed estimator cannot be directly used to construct the (asymptotic) confidence intervals since the limit distribution of the estimator (more precisely, its variance) depends on the unknown distribution of claims.

For these problems, they considered the following model. Let \( \Lambda(t), \ t \geq 0 \) be a random process independent of \( N(t) \) and possessing non-decreasing almost surely finite right-continuous sample paths starting from the origin. A Cox process, also called a doubly stochastic Poisson process, guided by the process \( \Lambda(t) \), is defined as

\[
\tilde{N}(t) = N(\Lambda(t)), \quad t \geq 0.
\]

Consider the (new) process

\[
U(t) = u + c\Lambda(t) - \tilde{S}(t), \quad \tilde{S}(t) = \sum_{i=1}^{R_t} X_i, \quad t \geq 0.
\]

Under this model, we suppose that the claims \( X_1, \ldots, X_{\tilde{N}_t} \) are obtained at time \( t \). Then, the following estimator for \( \psi(u) \) is introduced:

\[
\psi_{\tilde{N}_t}(u) = \frac{\lambda \mu}{c} \left( 1 - \frac{\lambda \mu}{c} \right) \sum_{r=1}^{n} \left( \frac{\lambda \mu}{c} \right)^r \tilde{S}^r_{\tilde{N}_t}(u),
\]

where \( m(n) \) is an integer with \( m(n) \leq n \), \( \lim_{n \to \infty} m(n) = +\infty \), and \( \tilde{S}^r_{\tilde{N}_t}(u) \) is the same as \( U_n(u) \) when \( \tilde{N}(t) = n \). Moreover, a consistent estimator of asymptotic variance is introduced as follows:

\[
\sigma^2_{\tilde{N}_t}(u) = \left( 1 - \frac{\lambda \mu}{c} \right)^2 \sum_{r,s=1}^{k} \frac{n}{\lambda \mu} \left( \frac{\lambda \mu}{c} \right)^{r+s} \tilde{S}^r_s_{\tilde{N}_t}(u),
\]

where \( k(n) \) is an integer with \( k(n) \leq n \), \( \lim_{n \to \infty} k(n) = +\infty \), and

\[
\tilde{S}^r_s_{\tilde{N}_t}(u) = \frac{1}{\lambda \mu} \int_{0}^{\infty} U_{n,r}(u-y)I(x > y)dy.
\]

Then, Bening and Korolev (2003) showed the asymptotic normality for \( \psi_{\tilde{N}_t}(u) \) as follows.

**Theorem 6.3** Suppose \( \log n = o(m(n)) \) and

\[
\frac{1}{\sqrt{n}} \sum_{i=1}^{m(n)} \left( \frac{u_i}{c} \right)^{r_i/2} \max \left\{ 1, \sum_{i=1}^{m(n)} \left( \frac{u_i}{c} \right)^{r_i} \right\} \to 0
\]

as \( n \to \infty \) are satisfied. Assume \( \Lambda(t) \to \infty \) as \( t \to \infty \). Then,

\[
\sqrt{N(t)} \frac{\psi_{\tilde{N}_t}(u) - \psi(u)}{\sigma_{\tilde{N}_t}(u)} \sim N(0, 1).
\]

By contrast, Mnatsakanov et al. (2008) discussed the estimation of the ruin probability by using the inverse of the Laplace transform. In the same way to Equation (4.15), the Laplace transform of Equation (6.2) is

\[
-\psi(0) + s\psi'(s) = \frac{\lambda}{c} \psi'(s) - \frac{\lambda}{c} \psi(s) f'(s),
\]
Assuming that $\lambda$ is known, Mnatsakanov et al. (2008) proposed its estimator as follows.

**Definition 6.3** Suppose that $X_1, \ldots, X_n$ are observed, where $X_i$ is the $i$th claim amount. Then, the estimator $(\hat{\psi}^*(s))$ of $\psi^*(s)$ for a fixed $s$ is defined by

\[
\hat{\psi}^*(s) = \frac{1 - \hat{\rho}}{D(s)}, \quad \hat{D}(s) = s - \frac{\lambda}{\zeta} \{1 - f^*(s)\}, \quad s \geq 0.
\]

(6.7)

Although $\hat{\psi}^*$ is close to $\psi^*$, there is no guarantee that the inverse of the Laplace transform $L^{-1}\hat{\psi}^*$ will be close to $L^{-1}\psi^*$ since $L^{-1}\psi^*$ is not continuous. Hence, Mnatsakanov et al. (2008) considered slightly modifying the functions $\psi^*$ and $\hat{\psi}^*$, that is, for an arbitrary fixed $\theta > 0$,

\[
\psi^*_\theta(s) = \psi^*(s + \theta), \quad \hat{\psi}^*_\theta(s) = \hat{\psi}^*(s + \theta), \quad s \geq 0
\]

are defined. Then, $\psi^*_\theta(s), \hat{\psi}^*_\theta(s) \in L^2(0, \infty)$ a.s. are satisfied because the non-continuity for $L^{-1}\hat{\psi}^*$ and $L^{-1}\psi^*$ is due to the behaviour of $D(s)$ and $\hat{D}(s)$ near 0. For $\psi^*_\theta(s), \hat{\psi}^*_\theta(s)$, the inverse of the Laplace transform is introduced as follows:

\[
(L^{-1}\hat{\psi}^*_\theta)(u) = \frac{1}{\pi^2} \int_{0}^{\infty} \int_{0}^{\infty} R_\theta(s) \frac{1}{\sqrt{s}} e^{-\sqrt{s}u} \hat{\psi}_\theta^*(t) dt ds,
\]

\[
(L^{-1} \psi^*_\theta)(u) = \frac{1}{\pi^2} \int_{0}^{\infty} \int_{0}^{\infty} R_\theta(s) \frac{1}{\sqrt{s}} e^{-\sqrt{s}u} \psi^*_\theta(t) dt ds
\]

for a fixed $\alpha > 0$, where

\[
R_\theta(s) = \frac{1}{2} \int_{[\log(s)]} \frac{\rho(x) \cos(x \log s) dx}{\cosh(x)}, \quad \rho(x) = \frac{\pi}{\cosh \pi x}, \quad s \in (0, \infty), \quad x \in \mathbb{R}.
\]

As a result, the following estimators are introduced.

**Definition 6.4**

\[
\hat{\psi}^*_\theta(u) = e^{\theta u} \hat{\psi}^*_\theta(u) = (L^{-1}\hat{\psi}^*_\theta)(u), \quad \hat{\psi}^*_\theta(u) = e^{\theta u} \psi^*_\theta(u) = (L^{-1} \psi^*_\theta)(u).
\]

Then, the following consistency in the ISE (integrated squared error) sense is shown.

**Theorem 6.4** Suppose that $0 < d\psi(u)/du \leq M$ is satisfied for a fixed $M$ and $\forall u \geq 0$, and Assumption 3.1 is made. Then, if $f$ is unknown,

\[
\|\psi^*_\theta - \psi\|^2 = O\left(\frac{1}{\log n}\right), \quad \text{as } n \to \infty,
\]

where $\alpha = a(n) = \sqrt{\log n/n}$ and if $f$ is known

\[
\|\hat{\psi}^*_\theta - \psi\|^2 = O\left(\frac{1}{\log(1/\alpha)}\right), \quad \text{as } \alpha \downarrow 0.
\]
Here, \( \| \cdot \|_g \) is the norm in \( L^2(0,B) \) for an arbitrary fixed \( 0 < B < \infty \), which implies that the comparison between \( \psi_n^g(\psi_n^g) \) and \( \psi_n \) is carried out in \( L^2(0,B) \) by restricting the functions to \([0,B]\).

### 6.3. Statistical estimation of the Gerber–Shiu function

Shimizu (2012) proposed an estimator of the Gerber–Shiu function for the Wiener–Poisson risk model (corresponding to the model in Example (2.5)) in a similar way to Mnatsakanov et al. (2008), and showed the consistency in the ISE sense.

The Laplace transform of the Gerber–Shiu function for the Wiener–Poisson risk process (Equation 2.4) was obtained by Tsai and Willmot (2002), Morales (2007), and Biffis and Morales (2010). The following result was given by Morales (2007).

**Theorem 6.5** Suppose that the penalty function \( w \) is uniformly bounded, or it satisfies \( |w(x,y) - w_0| \leq k(x+y)^\alpha \) for a universal constant \( k > 0 \) and \( \alpha > 1 \). Then,

\[
\phi^*(s) = \frac{K^*(\rho) - K^*(s) + w_0(s - \rho)D}{(c + (s + \rho)D)(s - \rho) - \lambda \left( f^*(\rho) - f^*(s) \right)}, \quad s \geq 0,
\]

(6.8)

where \( f \) is a p.d.f. of \( X_n, f^* \) and \( K^* \) are the Laplace transforms of \( f \) and \( K, D = \sigma^2/2 \), the function \( K \) is given by

\[
K(x) = \lambda \int_x^\infty w(x,y-x)f(y)dy,
\]

\( \rho \) is the unique non-negative solution to the following Lundberg’s fundamental equation with respect to \( s \):

\[
\mathcal{L}(s) = cs - \lambda(1 - f^*(s)) + Ds^2 - \delta = 0.
\]

In particular, \( \rho = 0 \) as \( \delta = 0 \).

Shimizu (2012) proposed its estimator as follows:

**Definition 6.5** Suppose that \( (U(t))_{t \equiv t^0_i = i\Delta_n, i = 0,1,2, \ldots, n} \) for some \( \Delta_n > 0 \) and the claim data \( X_1, \ldots, X_{N_T} \), where \( T_n = n\Delta_n \) are observed. Then, the estimator \( \hat{\phi}^*(s) \) of \( \phi^*(s) \) for a fixed \( s \) is defined by

\[
\hat{\phi}^*(s) = \frac{\hat{K}^*(\hat{\rho}) - \hat{K}^*(s) + w_0(s - \hat{\rho})\hat{D}}{\left( c + (s + \hat{\rho})\hat{D} \right)(s - \hat{\rho}) - \lambda \left( \hat{f}^*(\hat{\rho}) - \hat{f}^*(s) \right)},
\]

where

\[
\hat{f}^*(s) = \frac{1}{N_T} \sum_{i=1}^{N_T} e^{-sx}, \quad \hat{K}^*(s) = \frac{1}{T_n} \sum_{i=1}^{T_n} \int_0^X e^{-sy}w(y,x)dy, \quad \hat{\lambda} = \frac{N_T}{T_n}
\]

and

\[
\hat{D} = \frac{\hat{\sigma}^2}{2}, \quad \hat{\sigma}^2 = \frac{1}{T_n} \sum_{i=1}^{T_n} (U(t^0_i) - U(t^0_{i-1}))^2 I(r_j \notin (t^0_{i-1}, t^0_i), j = 1, \ldots, N_T).
\]

Here, \( r_j \) is the \( j \)-th claim occurring time and \( \hat{\rho} \) is defined as the following M estimator:

\[
\hat{\rho} = \arg \min_{s \in I} \mathcal{L}(s), \quad \mathcal{L}(s) = cs - \hat{\lambda}(1 - \hat{f}^*(s)) + \hat{D}s^2 - \delta,
\]

where \( I \subset (0, \infty) \) is a compact set in which \( \rho \) exists, and we set \( \hat{\rho} = 0 \) for \( \delta = 0 \).
In the same way to Mnatsakanov et al. (2008), Shimizu (2012) considered slightly modifying the functions \( \hat{\phi}^n \) for an arbitrary fixed \( \theta > 0 \), \( \hat{\phi}^n(\theta) = \hat{\phi}^n(s + \theta), \quad s \geq 0. \)

For \( \Phi_m \), the inverse of the Laplace transform is introduced as follows:

\[
(L_m^{-1} \Phi_m)(u) = \frac{1}{\pi^2} \int_0^\infty \int_0^\infty \Phi_m(s) \frac{1}{\sqrt{s}} e^{-uts} \delta(t) dt ds,
\]

where \( m > 0 \) is a constant and

\[
\Phi_m(s) = \int_0^{s_n} (\cosh \pi x) \cos(x \log s) dx, \quad \sigma_m = \frac{1}{\pi \cosh \pi m} > 0, \quad s \in (0, \infty).
\]

As a result, the following estimator is introduced.

**Definition 6.6**

\[
\hat{\Phi}_m^n(u) = e^{\frac{u}{m}} \hat{\Phi}_m^n(u) = (L_m^{-1} \Phi_m^n)(u)
\]

**Theorem 6.6** Suppose that

(i) \( E(|X|^p) < \infty \) for \( p > 0 \),

(ii) There exists a compact \( I \subset (0, \infty) \) such that \( \theta \) belongs to the interior of \( I \),

(iii) \( c > \lambda E(X) \) (i.e. Assumption 3.1),

(iv) \( T_n \Delta_n \to 0 \) as \( n \to \infty \),

(v) The penalty function \( w \) is uniformly bounded, or it satisfies \( |w(x,y) - w_0| \leq k(x+y)^\alpha \) for a universal constant \( k > 0 \) and \( \alpha > 1 \),

(vi) There exists a constant \( k' > 0, \alpha' > 0 \) such that \( \int_0^\infty w(x,dy) \leq k'(1+|x|^{\alpha'}) \) for any \( x \in \mathbb{R}^+ \),

(vii) The Gerber–Shiu function \( \phi \) and first derivative \( \phi' \) are of polynomial growth.

Then, for numbers \( m \equiv m(n) \) such that \( m(n) = \frac{(T_n/\log T_n)^{1/2}}{\log \log T_n} \) as \( n \to \infty \) and a number \( \theta > 0 \) such that \( (c - \lambda E(X)) \theta > \delta \), the following hold true:

(I) For any constant \( B > 0 \)

\[
\| \hat{\Phi}_m - \Phi \|_B = O_p\left( (\log T_n)^{-1/2} \right), \quad \text{as } n \to \infty,
\]

(II) If we take \( \theta \in (0, 1/2) \), it follows for a sequence \( T_n = \log \log T_n \) that,

\[
\| \hat{\Phi}_m - \Phi \|_{B_n} = O_p\left( (\log T_n)^{1/2} \right), \quad \text{as } n \to \infty.
\]

**Funding** This work was supported by JSPS KAKENHI Grant Number 24730193.

**Author details** Hiroshi Shiraishi;
E-mail: shiraishi@math.keio.ac.jp

1 Department of Mathematics, Keio University, Yokohama, Kanagawa, Japan.

**Citation information** Cite this article as: Review of statistical actuarial risk modeling, Hiroshi Shiraishi, Cogent Mathematics (2016), 3: 1123945.
Albrecher, H., Borst, S., Boxma, O., & Resing, J. (2009). The tax identity in risk theory: A simple proof and an extension. *Insurance: Mathematics and Economics, 44*, 304–306.

Albrecher, H., Claramunt, M. M., & Marmol, M. (2007). On the distribution of dividend payments in a sparre-andersen model with generalized erlang (n) interclaim times. *Insurance: Mathematics and Economics, 37*, 324–334.

Albrecher, H., & Hartinger, J. (2006). On the non-optimality of horizontal barrier strategies in the sparre-andersen model. *HERMES International Journal of Computer Mathematics and Its Applications, 7*, 1–14.

Albrecher, H., & Hartinger, J. (2007). A risk model with multilayer dividend strategy. *North American Actuarial Journal, 11*, 43–64.

Albrecher, H., Hartinger, J., & Thonhauser, S. (2007). On exact ruin probabilities in Sparre Andersen models. *Insurance: Mathematics and Economics, 41*, 249–257.

Albrecher, H., Renaud, J.-F., & Zhou, X. (2008). A Lévy insurance risk process with tax. *Journal of Applied Probability, 45*, 363–375.

Albrecher, H., & Teugels, J. L. (2006). Exponential behavior in the presence of dependence in risk theory. *Journal of Applied Probability, 43*, 257–273.

Alegre, A., Claramunt, M., & Marmol, M. (1999). Dividend policy and ruin probability. In Proceedings of the Third International Congress on *Insurance, Mathematics & Economics*. London.

Andersen, E. S. (1957). On the collective theory of risk in case of contagion between claims. *Bulletin of the Institute of Mathematics and Its Applications, 12*, 275–279.

Asmussen, S., Højgaard, B., & Taksar, M. (2000). Optimal risk control and dividend distribution policies. Example of excess-of-loss reinsurance for an insurance corporation. *Finance and Stochastics, 4*, 299–324.

Avanzi, B. (2009). Strategies for dividend distribution: A review. *North American Actuarial Journal, 13*, 217–251.

Avanzi, B., & Gerber, H. U. (2008). Optimal dividends in the dual model with diffusion. *Asin Bulletin, 38*, 653–667.

Avanzi, B., Gerber, H. U., & Shiu, E. S. (2007). Optimal dividends in the dual model. *Insurance: Mathematics and Economics, 41*, 111–123.

Avanzi, B., Shen, J., & Wong, B. (2011). Optimal dividends and capital injections in the dual model with diffusion. *Asin Bulletin, 41*, 611–644.

Bădescu, A., & Landriault, D. (2008). Recursive calculation of the dividend moments in a multi-threshold risk model. *North American Actuarial Journal, 12*, 74–88.

Bao, Z.-H. (2007). A note on the compound binomial model with randomized dividend strategy. *Applied Mathematics and Computation, 194*, 276–286.

Bening, V. E., & Koralov, V. Y. (2003). Nonparametric estimation of the ruin probability for generalized risk processes. *Theory of Probability & its Applications, 47*, 1–16.

Beveridge, C. J., Dickson, D. C., & Wu, X. (2007). Optimal dividends under reinsurance. *Actuarial Studies, Department of Economics, University of Melbourne*.

Biffis, E., & Morales, M. (2010). On a generalization of the Gerber–Shiu function to path-dependent penalties. *Insurance: Mathematics and Economics, 46*, 92–97.

Boudreault, M., Cossette, H., Landriault, D., & Marceau, E. (2006). On a risk model with dependence between interclaim arrivals and claim sizes. *Scandinavian Actuarial Journal, 2006*, 265–285.

Bowen, N., Eriksson, H., Broman, J., Jones, D., & Nesbitt, C. (1997). Actuarial mathematics. Itasca, MN: Society of Actuaries.

Bühlmann, H. (1970). Mathematical methods in risk theory (Die Grundlehren der mathematischen Wissenschaften, Band 172). New York, NY: Springer-Verlag.

Cao, Y., & Wan, N. (2009). Optimal proportional reinsurance and investment based on Hamilton–Jacobi–Bellman equation. *Insurance: Mathematics and Economics, 45*, 157–162.

Centeno, L. (1986). Measuring the effects of reinsurance by the adjustment coefficient. *Insurance: Mathematics and Economics, 5*, 169–182.

Cheng, S., Gerber, H. U., & Shiu, E. S. (2000). Discounted probabilities and ruin theory in the compound binomial model. *Insurance: Mathematics and Economics, 26*, 239–250.

Cheung, E. C., Dickson, D. C., & Drekic, S. (2008). Moments of discounted dividends for a threshold strategy in the compound Poisson risk model. *North American Actuarial Journal, 12*, 299–318.

Claramunt, M. M., Marmol, M., & Alegre, A. (2003). A note on the expected present value of dividends with a constant barrier in the discrete time model. *Bulletin of the Swiss Association of Actuaries, 2*, 149–159.

Cramer, H. (1930). On the mathematical theory of risk. *Skandia Jubilee Volume, Stockholm*.

Cramer, H. (1957). *Collective risk theory: A survey of the theory from the point of view of the theory of stochastic processes*. Stockholm: Nordiska bokhandeln.

Croux, K., & Veraverbeke, N. (1990). Nonparametric estimators for the probability of ruin. *Insurance: Mathematics and Economics, 9*, 127–130.

Csörgő, M., & Steinebach, J. (1991). On the estimation of the adjustment coefficient in risk theory via intermediate order statistics. *Insurance: Mathematics and Economics, 10*, 37–50.

Csörgő, S., & Teugels, J. L. (1990). Empirical Laplace transform and approximation of compound distributions. *Journal of Applied Probability, 27*, 88–101.

David Promislow, S., & Young, V. R. (2005). Minimizing the probability of ruin when claims follow Brownian motion with drift. *North American Actuarial Journal, 9*, 110–128.

De Finetti, B. (1957). Su un’ impostazione alternativa della teoria collettiva del rischio [On a setting of the alternative theory of collective risk]. *Transactions of the XVth International Congress of Actuaries, 2*, 433–443.

De Vylder, F., & Goovaerts, M. J. (1980). Recursive calculation of finite-time ruin probabilities. *Insurance: Mathematics and Economics, 7(1)*, 1–7.

De Vylder, F., & Marceau, E. (1996). Classical numerical ruin probabilities. *Scandinavian Actuarial Journal, 1996*, 109–123.

Deheuvels, P., & Steinebach, J. (1990). On some alternative estimates of the adjustment coefficient in risk theory. *Scandinavian Actuarial Journal, 135–159. doi:10.1080/03461238.1990.10413878*

Dickson, D. (1994). Some comments on the compound binomial model. *Asin Bulletin, 24*, 33–45.

Dickson, D., & Waters, H. R. (1991). Recursive calculation of survival probabilities. *Asin Bulletin, 21*, 199–221.

Dickson, D. C. (1992). On the distribution of the surplus prior to ruin. *Insurance: Mathematics and Economics, 11*, 191–207.

Dickson, D. C., & Drekic, S. (2006). Optimal dividends under a ruin probability constraint. *Annals of Actuarial Science, 1*, 291–306.

Dickson, D. C., & Waters, H. R. (1996). Reinsurance and ruin. *Insurance: Mathematics and Economics, 19*, 61–80.

Dickson, D. C. M. (2005). *Insurance risk and ruin. International series on actuarial science*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511624155

Deheuvels, P., & Steinebach, J. (1990). On some alternative estimates of the adjustment coefficient in risk theory. *Scandinavian Actuarial Journal, 135–159. doi:10.1080/03461238.1990.10413878*.
Dickson, D. C. M., & Waters, H. R. (2004). Some optimal dividends problems. Astin Bulletin, 34, 49–74. doi:10.2143/AST.34.1.504954

Dufresne, F., & Gerber, H. U. (1998). The surpluses immediately before and at ruin, and the amount of the claim causing ruin. Insurance: Mathematics and Economics, 7, 193–199.

Dufresne, F., & Gerber, H. U. (1991). Risk theory for the compound Poisson process that is perturbed by diffusion. Insurance: Mathematics and Economics, 10, 51–59.

Embrechts, P., & Mikosch, T. (1991). A bootstrap procedure for estimating the adjustment coefficient. Insurance: Mathematics and Economics, 10, 181–190.

Feng, Y., & Wu, R. (2007). Optimal dividend strategy in the compound poisson model with constant interest. Stochastic Models, 23, 149–166.

Feng, R., & Shimizu, T. (2011). On a generalization from ruin to default in a Lévy insurance risk model. Methodology and Computing in Applied Probability, 15, 773–802.

Frees, E. W. Nonparametric renewal function estimation. Annals of Statistics, 14, 1366–1378. doi:10.1214/aos/1176350163

Frostig, E. (2005). On the expected time to ruin and the expected dividends when dividends are paid while the surplus is above a constant barrier. Journal of Applied Probability, 42, 595–607.

Frostig, E. (2008). On model with dividends payments perturbed by a Brownian motion—An algorithmic approach. Astin Bulletin, 38, 183–206.

Gerber, H. (1979). The dilemma between dividends and safety and a generalization of the Lundberg–Cramér formulas. Scandinavian Actuarial Journal, 1974, 46–57.

Gerber, H. U. (1979). An introduction to mathematical risk theory, volume 8 of S.S. Heubner Foundation Monograph Series. Philadelphia: University of Pennsylvania, Wharton School, S.S. Heubner Foundation for Insurance Education. Distributed by Richard D. Irwin Inc, Homewood, Ill, With a foreword by James C. Hickman.

Gerber, H. U. (1981). On the probability of ruin in the presence of a linear dividend barrier. Scandinavian Actuarial Journal, 1981, 105–115.

Gerber, H. U. (1988). Mathematical fun with the compound binomial process. Astin Bulletin, 18, 161–168.

Gerber, H. U., Goovaerts, M. J., & Kaas, R. (1997). On the probability and severity of ruin. Astin Bulletin, 27, 151–163.

Gerber, H. U., & Shiu, E. S. (1998). On the time value of ruin. North American Actuarial Journal, 2, 48–72.

Gerber, H. U., & Shiu, E. S. (1999). From ruin theory to pricing reset guarantees and perpetual put options. Insurance: Mathematics and Economics, 24, 3–14.

Gerber, H. U., & Shiu, E. S. (2003). Geometric Brownian motion models for assets and liabilities: From pension funding to optimal dividends. North American Actuarial Journal, 7, 37–51.

Gerber, H. U., & Shiu, E. S. W. (1998). On the time value of ruin (With discussion and a reply by the authors). North American Actuarial Journal, 2, 48–78. doi:10.1080/1092077199.10595671

Gerber, H. U., & Shiu, E. S. W. (2004). Optimal dividends: Analysis with Brownian motion. North American Actuarial Journal, 8(1), 1–20. doi:10.1080/109207704.2005.10596125

Gerber, H. U., Shiu, E. S. W., & Smith, N. (2008). Maximizing dividends without bankruptcy. Astin Bulletin, 36, 5–23. doi:10.2143/AST.36.1.2014143

Gerber, H. U., Shiu, E. S. W., & Smith, N. (2008). Methods for estimating the optimal dividend barrier and the probability of ruin. Insurance: Mathematics and Economics, 42, 243–254. doi:10.1016/j.insmatheco.2007.02.002

Gerber, H. U., & Smith, N. (2008). Optimal dividends with incomplete information in the dual model. Insurance: Mathematics and Economics, 43, 227–233.

Grandell, J. (1978/1979). Empirical bounds for ruin probabilities. Stochastic Processes and their Applications, 8, 243–255. doi:10.1016/0304-4149(79)90001-2

Gu, A., Guo, X., Li, Z., & Zeng, Y. (2012). Optimal control of excess-of-loss reinsurance and investment for insurers under a CEX model. Insurance: Mathematics and Economics, 51, 674–684.

Højgaard, B. (2002). Optimal dynamic premium control in non-life insurance. Maximizing maximizing pay-out. Scandinavian Actuarial Journal, 2002, 225–245.

Højgaard, B., & Taksar, M. (1998). Optimal proportional reinsurance policies for diffusion models. Scandinavian Actuarial Journal, 1998, 166–180.

Huang, Y. J., & Li, A. Q. (2011). On a risk model with dependence between interclaim arrivals and claim sizes with multiple thresholds under stochastic interest. Advanced Materials Research, Vol. 179, pp. 1086–1090.

Højgaard, B. (2002). Optimal dynamic premium control in non-life insurance. Maximizing maximizing pay-out. Scandinavian Actuarial Journal, 2002, 225–245.

Højgaard, B., & Taksar, M. (1998). Optimal proportional reinsurance policies for diffusion models. Scandinavian Actuarial Journal, 1998, 166–180.

Li, L.-L., Feng, J., & Song, L. (2007). The dividend payments for general claim size distributions under interest rate. International Journal of Mathematical Sciences, 1, 135–138.

Li, S. (2006). The distribution of the dividend payments in the compound Poisson risk model perturbed by diffusion. Scandinavian Actuarial Journal, 2006, 73–85.

Li, S., & Garrido, J. (2002). On the time value of ruin in the discrete time risk model. Working Paper. Bussiness Economics. Retrieved from Name website: http://hdl.handle.net/10016/67

Li, S., & Garrido, J. (2004). On a class of renewal risk models with a constant dividend barrier. Insurance: Mathematics and Economics, 35, 691–701.

Li, S., & Wu, B. (2006). The diffusion perturbed compound Poisson risk model with a dividend barrier. Preprint. Retrieved from: http://hdl.handle.net/11343/4313

Lin, X. S., & Pavlova, K. P. (2006). The compound Poisson risk model with a threshold dividend strategy. Insurance: Mathematics and Economics, 38, 57–80.

Lin, X. S., & Sendra, K. P. (2008). The compound Poisson risk model with multiple thresholds. Insurance: Mathematics and Economics, 42, 617–627.

Liu, W., & Hu, Y. (2010). Optimal financing and dividend control of the insurance company with excess-of-loss reinsurance policy. Statistics & Probability Letters, 80, 121–130.

Loeffen, R. (2009). An optimal dividends problem with a terminal value for spectrally negative Lévy processes with a completely monotone jump density. Journal of Applied Probability, 46, 85–98.

Loeffen, R. L., & Renou, J.-F. (2010). De Finetti’s optimal dividends problem with an affine penalty function at ruin. Insurance: Mathematics and Economics, 46, 98–108.

Lundberg, F. (1909). Über die theorie der rückversicherung. Transactions VI

Meng, H., Zhang, C., & Wu, R. (2007). The expectation of insru-mathco.2007.02.002

Meng, H., Zhang, C., & Wu, R. (2007). The expectation of insru-mathco.2007.02.002

Ngo, T. T., & Stadje, W. M. (2008). The compound Poisson process that is perturbed by diffusion. Annals of Applied Probability, 18, 2273–2316.
Minkova, L. D. (2010). Insurance risk theory. TEMPUS Project SEE doctoral studies in mathematical sciences: Lecture notes.

Mnatsakanov, R., Ruymgaart, L. L., & Ruymgaart, F. H. (2008). Nonparametric estimation of ruin probabilities given a random sample of claims. Mathematical Methods of Statistics, 17, 35–43. doi:10.3103/S1066563008010031

Morales, M. (2007). On the expected discounted penalty function for a perturbed risk process driven by a subordinator. Insurance: Mathematics and Economics, 40, 293–301.

Ng, A. C. (2009). On a dual model with a dividend threshold. Insurance: Mathematics and Economics, 44, 315–324.

Ross, S. M. (2014). Introduction to probability models. Amsterdam: Academic Press.

Seal, H. L. (1969). Stochastic theory of a risk business. New York, NY: Wiley.

Shimizu, Y. (2009). A new aspect of a risk process and its statistical inference. Insurance: Mathematics and Economics, 44, 70–77.

Shimizu, Y. (2012). Non-parametric estimation of the Gerber-Shiu function for the Wiener–Poisson risk model. Scandinavian Actuarial Journal, 2012, 56–69.

Shiu, E. S. (1989). The probability of eventual ruin in the compound binomial model. Astin Bulletin, 19, 179–190.

Siegl, T., & Tichy, R. F. (1999). A process with stochastic claim frequency and a linear dividend barrier. Insurance: Mathematics and Economics, 24, 51–65.

Taksar, M. I. (2000). Optimal risk and dividend distribution control models for an insurance company. Mathematical Methods of Operations Research, 51(1), 1–42.

Tan, J., & Yang, X. (2006). The compound binomial model with randomized decisions on paying dividends. Insurance: Mathematics and Economics, 39(1), 1–18.

Tsai, C. C.-L., & Willmot, G. E. (2002). A generalized defective renewal equation for the surplus process perturbed by diffusion. Insurance: Mathematics and Economics, 30, 51–66.

Wan, N. (2007). Dividend payments with a threshold strategy in the compound Poisson risk model perturbed by diffusion. Insurance: Mathematics and Economics, 40, 509–523.

Waters, H. R. (1983). Some mathematical aspects of reinsurance. Insurance: Mathematics and Economics, 2, 17–26.

Yang, H., & Zhang, Z. (2008). Gerber–Shiu discounted penalty function in a sparre andersen model with multi-layer dividend strategy. Insurance: Mathematics and Economics, 42, 984–991.

Yin, C., & Wen, Y. (2013). Optimal dividend problem with a terminal value for spectrally positive Lévy processes. Insurance: Mathematics and Economics, 53, 769–773.

Yuen, K. C., Lu, Y., & Wu, R. (2009). The compound Poisson process perturbed by a diffusion with a threshold dividend strategy. Applied Stochastic Models in Business and Industry, 25, 73–93.

Zhang, Z. (2014). On a perturbed Sparre Andersen risk model with threshold dividend strategy and dependence. Journal of Computational and Applied Mathematics, 255, 248–269.

Zhao, H., Rong, X., & Zhao, Y. (2013). Optimal excess-of-loss reinsurance and investment problem for an insurer with jump-diffusion risk process under the Heston model. Insurance: Mathematics and Economics, 53, 504–514.