Chapter 5
Premium principles and Risk measures

Actuaries have long been tasked with tackling difficult quantitative problems. Loss estimates, reserves requirements, and other quantities have been their traditional domain, but the actuary of the future has an opportunity to broaden his/her scope of knowledge to include other risks facing corporations around the globe. While this may seem like a daunting task at first, the reality is that the skills required to analyze business risks are not a significant stretch of the traditional actuary's background — Timothy Essaye, `imageoftheactuary.org`

5.1 Introduction

The activities of an insurer can be described as a system in which the acquired capital increases because of (earned) premiums and interest, and decreases because of claims and costs. See also the previous chapter. In this chapter we discuss some mathematical methods to determine the premium from the distribution of the claims. The actuarial aspect of a premium calculation is to calculate a minimum premium, sufficient to cover the claims and, moreover, to increase the expected surplus sufficiently for the portfolio to be considered stable.

Bühlmann (1985) described a top-down approach for the premium calculation. First we look at the premium required by the total portfolio. Then we consider the problem of spreading the total premium over the policies in a fair way. To determine the minimum annual premium, we use the ruin model as introduced in the previous chapter. The result is an exponential premium (see Chapter 1), where the risk aversion parameter α follows from the maximal ruin probability allowed and the available initial capital. Assuming that the suppliers of the initial capital are to be rewarded with a certain annual dividend, and that the resulting premium should be as low as possible, therefore as competitive as possible, we can derive the optimal initial capital. Furthermore we show how the total premium can be spread over the policies in a fair way, while the total premium keeps meeting our objectives.

For the policy premium, a lot of premium principles can be justified. Some of them can be derived from models like the zero utility model, where the expected utility before and after insurance is equal. Other premium principles can be derived as an approximation of the exponential premium principle. We will verify to which extent these premium principles satisfy some reasonable requirements. We will also consider some characterizations of premium principles. For example, it turns out that the only utility preserving premium principles for which the total premium for independent policies equals the sum of the individual premiums are the net premium and the exponential premium.

As an application, we analyze how insurance companies can optimally form a 'pool'. Assuming exponential utility, it turns out that the most competitive total premium is obtained when the companies each take a fixed part of the pooled risk (coinsurance), where the proportion is inversely proportional to their risk aversion, hence proportional to their risk tolerance. See also Gerber (1979).

Mathematically speaking, a risk measure for a random variable S is just a functional connecting a real number to S. A prime example is of course a premium principle, another is the ruin probability for some given initial capital. The risk measure most often used in practice is the *Value-at-Risk* (VaR) at a certain (confidence) level p with $0 < p < 1$, which is the amount that will maximally be lost with probability p, therefore the inverse of the cdf. It is also called the quantile risk measure. Note the difference between the variance, written as $\mathrm{Var}[X]$, and the value-at-risk, written in CamelCase as $\mathrm{VaR}[X;p]$. Fortunately, Vector Autoregressive (VAR) models are outside the scope of this book.

The VaR is not ideal as a risk measure. One disadvantage is that it only looks at the probability of the shortfall of claims over capital being positive. But the size of the shortfall certainly matters; someone will have to pay for the remainder. Risk measures accounting for the size of the shortfall $(X-d)_+$ when capital d is available include the Tail-Value-at-Risk, the Expected Shortfall and the Conditional Tail Expectation.

Another possible disadvantage of VaR is that it is not subadditive: the sum of the VaRs for X and Y may be larger than the one for $X+Y$. Because of this, the VaR is not a *coherent* risk measure. But insisting on coherence is only justified when looking at complete markets, which the insurance market is not, since it is not always possible to diversify a risk.

5.2 Premium calculation from top-down

As argued in Chapter 4, insuring a certain portfolio of risks leads to a surplus that increases because of collected premiums and decreases in the event of claims. The following recurrent relations hold in the ruin model between the surpluses at integer times:

$$U_t = U_{t-1} + c - S_t, \quad t = 1,2,\ldots \tag{5.1}$$

Ruin occurs if $U_\tau < 0$ for some real τ. We assume that the annual total claims S_t, $t = 1,2,\ldots$, are independent and identically compound Poisson random variables, say $S_t \sim S$. The following question then arises: how large should the initial capital $U_0 = u$ and the premium $c = \pi[S]$ be to remain solvent at all times with a prescribed probability? The probability of ruin at integer times only is less than $\psi(u)$, which in turn is bounded from above by e^{-Ru}. Here R denotes the adjustment coefficient, which is the root of the equation $\mathrm{e}^{Rc} = \mathrm{E}[\mathrm{e}^{RS}]$, see (4.11). If we set the upper bound equal to ε, then $R = |\log\varepsilon|/u$. Hence, we get a ruin probability bounded by ε by choosing the premium c as

$$c = \frac{1}{R}\log\left(\mathrm{E}[\mathrm{e}^{RS}]\right), \quad \text{where} \quad R = \frac{1}{u}|\log\varepsilon|. \tag{5.2}$$

This premium is the exponential premium (1.20) with parameter R. From Example 1.3.1, we know that the adjustment coefficient can be interpreted as a measure for the risk aversion: for the utility function $-\alpha\mathrm{e}^{-\alpha x}$ with risk aversion α, the utility preserving premium is $c = \frac{1}{\alpha}\log(\mathrm{E}[\mathrm{e}^{\alpha X}])$.

A characteristic of the exponential premium is that choosing this premium for each policy also yields the right total premium for S. The reader may verify that if the payments X_j on policy j, $j = 1, \ldots, n$, are independent, then

$$S = X_1 + \cdots + X_n \quad \Longrightarrow \quad \frac{1}{R}\log(\mathrm{E}[\mathrm{e}^{RS}]) = \sum_{j=1}^{n}\frac{1}{R}\log(\mathrm{E}[\mathrm{e}^{RX_j}]). \tag{5.3}$$

Another premium principle that is additive in this sense is the variance principle, where for a certain parameter $\alpha \geq 0$ the premium is determined by

$$\pi[S] = \mathrm{E}[S] + \alpha\mathrm{Var}[S]. \tag{5.4}$$

In fact, every premium that is a linear combination of cumulants is additive. Premium (5.4) can also be obtained as an approximation of the exponential premium by considering only two terms of the Taylor expansion of the cgf, assuming that the risk aversion R is small, since

$$\pi[S] = \frac{1}{R}\kappa_S(R) = \frac{1}{R}\left(\mathrm{E}[S]R + \mathrm{Var}[S]\frac{R^2}{2} + \cdots\right) \approx \mathrm{E}[S] + \frac{1}{2}R\mathrm{Var}[S]. \tag{5.5}$$

So to approximate (5.2) by (5.4), α should be taken equal to $\frac{1}{2}R$. In view of (5.2) and $\widetilde{\psi}(u) \leq \mathrm{e}^{-Ru}$, we can roughly state that:

- doubling the loading factor α in (5.4) decreases the upper bound for the ruin probability from ε to ε^2;
- halving the initial capital requires the loading factor to be doubled if one wants to keep the same maximal ruin probability.

We will introduce a new aspect in the discrete time ruin model (5.1): how large should u be, if the premium c is to contain a yearly dividend iu for the shareholders who have supplied the initial capital? A premium at the portfolio level that ensures ultimate survival with probability $1-\varepsilon$ and also incorporates this dividend is

$$\pi[S] = \mathrm{E}[S] + \frac{|\log\varepsilon|}{2u}\mathrm{Var}[S] + iu, \tag{5.6}$$

that is, the premium according to (5.2) and (5.5), plus the dividend iu. We choose u such that the premium is as competitive as possible, therefore as low as possible. By setting the derivative equal to zero, we see that a minimum is reached for $u = \sigma[S]\sqrt{|\log\varepsilon|/2i}$. Substituting this value into (5.6), we see that (see Exercise 5.2.1) the optimal premium is a *standard deviation premium*:

$$\pi[S] = \mathrm{E}[S] + \sigma[S]\sqrt{2i|\log\varepsilon|}. \tag{5.7}$$

In the optimum, the loading $\pi[S] - \mathrm{E}[S] - iu$ equals the dividend iu; notice that if i increases, then u decreases, but iu increases.

Finally, we have to determine which premium should be asked at the down level. We cannot just use a loading proportional to the standard deviation. The sum of these premiums for independent risks does not equal the premium for the sum, and consequently the top level would not be in balance: if we add a contract, the total premium no longer satisfies the specifications. On the other hand, as stated before, the variance principle is additive, just like the exponential and the net premium. Hence, (5.6) and (5.7) lead to Bühlmann's recommendation for the premium calculation:

1. Compute the optimal initial capital $u = \sigma[S]\sqrt{|\log\varepsilon|/2i}$ for S, i and ε;
2. Spread the total premium over the individual risks X_j by charging the following premium:

$$\pi[X_j] = \mathrm{E}[X_j] + R\mathrm{Var}[X_j], \quad \text{where } R = |\log\varepsilon|/u. \tag{5.8}$$

Note that in this case the loading factor $R = \alpha$ of the variance premium, incorporating both a loading to avoid ruin and an equal loading from which to pay dividends, is twice as large as it would be without dividend, see (5.4) and (5.5). The total dividend and the necessary contribution to the expected growth of the surplus that is required to avoid ruin are spread over the policies in a similar way.

Bühlmann gives an example of a portfolio consisting of two kinds (A and B) of exponential risks with mean values 5 and 1:

Type	Number of risks	Expected value	Variance	Exponential premium	Variance premium	Stand. dev. premium
A	5	5	25	$-\frac{1}{R}\log(1-5R)$	$5+\frac{R}{2}25$	
B	20	1	1	$-\frac{1}{R}\log(1-R)$	$1+\frac{R}{2}1$	
Total	25	45	145			$45+(2i\lvert\log\varepsilon\rvert 145)^{\frac{1}{2}}$

Choose $\varepsilon = 1\%$, hence $|\log\varepsilon| = 4.6052$. Then, for the model with dividend, we have the following table of variance premiums for different values of i.

	Portfolio premium	Optimal u	Optimal R	Premium for A	Premium for B
$i = 2\%$	50.17	129.20	0.0356	5.89	1.0356
$i = 5\%$	53.17	81.72	0.0564	6.41	1.0564
$i = 10\%$	56.56	57.78	0.0797	6.99	1.0797

The portfolio premium and the optimal u follow from (5.7), R from (5.2), and the premiums for A and B are calculated according to (5.8). We observe that:

- the higher the required return i on the supplied initial capital u, the lower the optimal value for u;

- the loading is far from proportional to the risk premium: the loading as a percentage for risks of type A is 5 times the one for risks of type B;
- the resulting exponential premiums are close to the variance premiums given: if $i = 2\%$, the premium with parameter $2R$ is 6.18 for risks of type A and 1.037 for risks of type B.

5.3 Various premium principles and their properties

In this section, we give a list of premium principles that can be applied at the policy level as well as at the portfolio level. We also list some mathematical properties that one might argue a premium principle should have. Premium principles depend exclusively on the marginal distribution function of the random variable. Consequently, we will use both notations $\pi[F]$ and $\pi[X]$ for the premium of X, if F is the cdf of X. We will assume that X is a *bounded* non-negative random variable. Most premium principles can also be applied to unbounded and possibly negative claims. When the result is an infinite premium, the risk at hand is uninsurable.

We have encountered the following five premium principles in Section 5.2:

a) **Net premium**: $\pi[X] = \mathrm{E}[X]$
Also known as the equivalence principle; this premium is sufficient for a risk neutral insurer only.
b) **Expected value principle**: $\pi[X] = (1+\alpha)\mathrm{E}[X]$
The loading equals $\alpha\mathrm{E}[X]$, where $\alpha > 0$ is a parameter.
c) **Variance principle**: $\pi[X] = \mathrm{E}[X] + \alpha\mathrm{Var}[X]$
The loading is proportional to $\mathrm{Var}[X]$, and again $\alpha > 0$.
d) **Standard deviation principle**: $\pi[X] = \mathrm{E}[X] + \alpha\sigma[X]$
Here also $\alpha > 0$ should hold, to avoid getting ruined with probability 1.
e) **Exponential principle**: $\pi[X] = \frac{1}{\alpha}\log(\mathrm{m}_X(\alpha))$
The parameter $\alpha > 0$ is called the risk aversion. We already showed in Chapter 1 that the exponential premium increases if α increases. For $\alpha \downarrow 0$, the net premium arises; for $\alpha \to \infty$, the resulting premium equals the maximal value of X, see Exercise 5.3.11.

In the following two premium principles, the 'parameter' is a function.

f) **Zero utility premium**: $\pi[X] \leftarrow u(0) = \mathrm{E}[u(\pi[X] - X)]$
This concept was introduced in Chapter 1. The function $u(x)$ represents the utility a decision maker attaches to his present capital plus x. So, $u(0)$ is the utility of the present capital and $u(\pi[X] - X)$ is the utility after insuring a risk X against premium $\pi[X]$. The premium solving the utility equilibrium equation is called the zero utility premium. Each linear transform of $u(\cdot)$ yields the same premium. The function $u(\cdot)$ is usually assumed to be non-decreasing and concave. Accordingly it has positive but decreasing marginal utility $u'(x)$. The special choice $u(x) = \frac{1}{\alpha}(1 - \mathrm{e}^{-\alpha x})$ leads to exponential utility; the net premium results for linear $u(\cdot)$.

g) **Mean value principle**: $\pi[X] = v^{-1}(\mathrm{E}[v(X)])$
The function $v(\cdot)$ is a convex and increasing valuation function. Again, the net premium and the exponential premium are special cases with $v(x) = x$ and $v(x) = \mathrm{e}^{\alpha x}$, $\alpha > 0$.

The following premium principles are chiefly of theoretical importance:

h) **Percentile principle**: $\pi[X] = \min\{p \,|\, F_X(p) \geq 1 - \varepsilon\}$
The probability of a loss on contract X is at most ε, $0 \leq \varepsilon \leq 1$.
i) **Maximal loss principle**: $\pi[X] = \min\{p \,|\, F_X(p) = 1\}$
This premium arises as a limiting case of other premiums: (e) for $\alpha \to \infty$ and (h) for $\varepsilon \downarrow 0$. A 'practical' example: a pregnant woman pays some premium for an insurance contract that guarantees that the baby will be a girl; if it is a boy, the entire premium is refunded.
j) **Esscher principle**: $\pi[X] = \mathrm{E}[X\mathrm{e}^{hX}]/\mathrm{E}[\mathrm{e}^{hX}]$
Here, h is a parameter with $h > 0$. This premium is actually the net premium for a risk $Y = X\mathrm{e}^{hX}/m_X(h)$. As one sees, Y results from X by enlarging the large values of X, while reducing the small values. The Esscher premium can also be viewed as the expectation with the so-called Esscher transform of $\mathrm{d}F_X(x)$, which has as a 'density':

$$\mathrm{d}G(x) = \frac{\mathrm{e}^{hx}\,\mathrm{d}F_X(x)}{\int \mathrm{e}^{hy}\,\mathrm{d}F_X(y)}. \tag{5.9}$$

This is the differential of a cdf with the same support as X, but for which the probabilities of small values are reduced in favor of the probabilities of large values. The net premium for Y gives a loaded premium for X.

5.3.1 Properties of premium principles

Below, we give five desirable properties for premium principles $\pi[X]$. Some other useful properties such as order preserving, which means that premiums for smaller risks should also be less, will be covered in Chapter 7.

1) **Non-negative loading**: $\pi[X] \geq \mathrm{E}[X]$
A premium without a positive loading will lead to ruin with certainty.
2) **No rip-off**: $\pi[X] \leq \min\{p \,|\, F_X(p) = 1\}$
The maximal loss premium (i) is a boundary case. If X is unbounded, this premium is infinite.
3) **Consistency**: $\pi[X + c] = \pi[X] + c$ for each c
If we raise the claim by some fixed amount c, then the premium should also be higher by the same amount. Synonyms for consistency are *cash invariance*, *translation invariance*, and more precisely, *translation equivariance*. Note that in general, a 'risk' need not be a non-negative random variable, though to avoid certain technical problems sometimes it is convenient to assume it is bounded from below.

4) **Additivity**: $\pi[X+Y] = \pi[X] + \pi[Y]$ for independent X, Y
Pooling independent risks does not affect the total premium needed.
5) **Iterativity**: $\pi[X] = \pi\big[\pi[X \,|\, Y]\big]$ for all X, Y
The premium for X can be calculated in two steps. First, apply $\pi[\cdot]$ to the conditional distribution of X, given $Y = y$. The resulting premium is a function $h(y)$, say, of y. Then, apply the same premium principle to the random variable $\pi[X \,|\, Y] := h(Y)$.

For the net premium, iterativity is the corresponding property for expected values (2.7). Otherwise, the iterativity criterion is rather artificial. As an example, assume that a certain driver causes a Poisson number X of accidents in one year, where the parameter λ is drawn from the distribution of the structure variable Λ. The number of accidents varies because of the Poisson deviation from the expectation λ, and because of the variation of the structure distribution. In case of iterativity, if we set premiums for both sources of variation one after another, we get the same premium as if we determined the premium for X directly.

Example 5.3.1 (Iterativity of the exponential principle)
The exponential premium principle is iterative. This can be shown as follows:

$$\begin{aligned}\pi[\pi[X \,|\, Y]] &= \frac{1}{\alpha}\log \mathrm{E}\big[\mathrm{e}^{\alpha\pi[X \,|\, Y]}\big] = \frac{1}{\alpha}\log \mathrm{E}\big[\exp(\alpha\frac{1}{\alpha}\log \mathrm{E}[\mathrm{e}^{\alpha X} \,|\, Y])\big] \\ &= \frac{1}{\alpha}\log \mathrm{E}\big[\mathrm{E}[\mathrm{e}^{\alpha X} \,|\, Y]\big] = \frac{1}{\alpha}\log \mathrm{E}\big[\mathrm{e}^{\alpha X}\big] = \pi[X].\end{aligned} \tag{5.10}$$

After taking the expectation in an exponential premium, the transformations that were done before are successively undone. ∇

Example 5.3.2 (Compound distributions)
Assume that $\pi[\cdot]$ is additive as well as iterative, and that S is a compound random variable with N terms distributed as X. The premium for S then equals

$$\pi[S] = \pi[\pi[S \,|\, N]] = \pi[N\pi[X]]. \tag{5.11}$$

Furthermore, if $\pi[\cdot]$ is also *proportional*, (or *homogeneous*), which means that $\pi[\alpha N] = \alpha\pi[N]$ for all $\alpha \geq 0$, then $\pi[S] = \pi[X]\pi[N]$. In general, proportionality does not hold, see for example Section 1.2. However, this property is used as a local working hypothesis for the calculation of the premium for similar contracts; without proportionality, the use of a tariff is meaningless. ∇

In Table 5.1, we summarize the properties of our various premium principles. A "+" means that the property holds in general, a "−" that it does not, while especially an "*e*" means that the property only holds in case of an exponential premium (including the net premium). We assume that S is bounded from below. The proofs of these properties are asked in the exercises, but for the proof of most of the characterizations that zero utility and mean value principles with a certain additional property must be exponential, we refer to the literature. See also the following section.

Summarizing, one may state that only the exponential premium, the maximal loss principle and the net premium principle satisfy all these properties. The last

Table 5.1 Various premium principles and their properties

Principle $\longrightarrow$ Property $\downarrow$	a) μ	b) $1+\lambda$	c) σ^2	d) σ	e) exp	f) $u(\cdot)$	g) $v(\cdot)$	h) %	i) max	j) Ess
1) $\pi[X] \geq \mathrm{E}[X]$	+	+	+	+	+	+	+	−	+	+
2) $\pi[X] \leq \max[X]$	+	−	−	−	+	+	+	+	+	+
3) $\pi[X+c] = \pi[X]+c$	+	−	+	+	+	+	e	+	+	+
4) $\pi[X+Y] = \pi[X]+\pi[Y]$	+	+	+	−	+	e	e	−	+	+
5) $\pi\big[\pi[X\,\vert\,Y]\big] = \pi[X]$	+	−	−	−	+	e	+	−	+	−

two are irrelevant in practice, so only the exponential premium principle survives this selection. See also Section 5.2. A drawback of the exponential premium has already been mentioned: it has the property that a decision maker's decisions do not depend on the capital he has acquired to date. On the other hand, this is also a strong point of this premium principle, since it is very convenient not to have to know one's current capital, which is generally either random or simply not precisely known at each point in time.

5.4 Characterizations of premium principles

In this section we investigate the properties marked with "e" in Table 5.1, and also some more characterizations of premium principles. Note that linear transforms of the functions $u(\cdot)$ and $v(\cdot)$ in (f) and (g) yield the same premiums. A technique to prove that only exponential utility functions $u(\cdot)$ have a certain property consists of applying this property to risks with a simple structure, and derive a differential equation for $u(\cdot)$ that holds only for exponential and linear functions. Since the linear utility functions are a limit of the exponential utility functions, we will not mention them explicitly in this section. For full proofs of the theorems in this section, we refer to Gerber (1979, 1985) as well as Goovaerts et al. (1984).

The entries "e" in Table 5.1 are studied in the following theorem.

Theorem 5.4.1 (Characterizing exponential principles)

1. A consistent mean value principle is exponential.
2. An additive mean value principle is exponential.
3. An additive zero utility principle is exponential.
4. An iterative zero utility principle is exponential.

Proof. Since for a mean value principle we have $\pi[X] = c$ if $\Pr[X = c] = 1$, consistency is just additivity with the second risk degenerate, so the second assertion follows from the first. The proof of the first, which will be given below, involves applying consistency to risks that are equal to x plus some Bernoulli(q) random variable, and computing the second derivative at $q = 0$ to show that a valuation function $v(\cdot)$ with the required property necessarily satisfies the differential equation $\frac{v''(x)}{v'(x)} = c$ for some constant c. The solutions are the linear and exponential valuation functions. The final assertion is proved in much the same way. The proof that an additive zero utility principle is exponential proceeds by deriving a similar equation, for which it turns out to be considerably more difficult to prove that the exponential utility functions are the unique solutions.

To prove that a consistent mean value principle is exponential, assume that $v(\cdot)$, which is a convex increasing function, yields a consistent mean value principle. Let $P(q)$ denote the premium, considered as a function of q, for a Bernoulli(q) risk S_q. Then, by definition,

$$v(P(q)) = qv(1) + (1-q)v(0). \tag{5.12}$$

The right-hand derivative of this equation in $q = 0$ yields

$$P'(0)v'(0) = v(1) - v(0), \tag{5.13}$$

so $P'(0) > 0$. The second derivative in $q = 0$ gives

$$P''(0)v'(0) + P'(0)^2 v''(0) = 0. \tag{5.14}$$

Because of the consistency, the premium for $S_q + x$ equals $P(q) + x$ for each constant x, and therefore

$$v(P(q)+x) = qv(1+x) + (1-q)v(x). \tag{5.15}$$

The second derivative at $q = 0$ of this equation yields

$$P''(0)v'(x) + P'(0)^2 v''(x) = 0, \tag{5.16}$$

and, since $P'(0) > 0$, we have for all x that

$$\frac{v''(x)}{v'(x)} = \frac{v''(0)}{v'(0)}. \tag{5.17}$$

Consequently, $v(\cdot)$ is linear if $v''(0) = 0$, and exponential if $v''(0) > 0$. ∇

Remark 5.4.2 (Continuous and mixable premiums)
Another interesting characterization is the following one. A premium principle $\pi[\cdot]$ is continuous if $F_n \to F$ in distribution implies $\pi[F_n] \to \pi[F]$. If furthermore $\pi[\cdot]$ admits mixing, which means that $\pi[tF + (1-t)G] = t\pi[F] + (1-t)\pi[G]$ for cdfs F and G, as well as $\pi[c] = (1+\lambda)c$ for all real c and some fixed λ, then it can be shown that $\pi[\cdot]$ must be the expected value principle $\pi[X] = (1+\lambda)\mathrm{E}[X]$. Note that the last condition can be replaced by additivity. To see this, observe that additivity implies that $\pi[n/m] = n/m\,\pi[1]$ for all natural n and m, and then use continuity. ∇

Finally, the Esscher premium principle can be justified as follows.

Theorem 5.4.3
Assume an insurer has an exponential utility function with risk aversion α. If he charges a premium of the form $\mathrm{E}[\varphi(X)X]$ where $\varphi(\cdot)$ is a continuous increasing

function with $\mathrm{E}[\varphi(X)] = 1$, his utility is maximized if $\varphi(x) \propto \mathrm{e}^{\alpha x}$, hence if he uses the Esscher premium principle with parameter α.

Proof. The proof of this statement is based on the technique of variational calculus and adapted from Goovaerts et al. (1984). Let $u(\cdot)$ be a convex increasing utility function, and introduce $Y = \varphi(X)$. Then, because $\varphi(\cdot)$ increases continuously, we have $X = \varphi^{-1}(Y)$. Write $f(y) = \varphi^{-1}(y)$. To derive a condition for $\mathrm{E}[u(-f(Y) + \mathrm{E}[f(Y)Y])]$ to be maximal for all choices of continuous increasing functions when $\mathrm{E}[Y] = 1$, consider a function $f(y) + \varepsilon g(y)$ for some arbitrary continuous function $g(\cdot)$. A little reflection will lead to the conclusion that the fact that $f(y)$ is optimal, and this new function is not, must mean that

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\mathrm{E}\big[u\big(-f(Y)+\mathrm{E}[f(Y)Y]+\varepsilon\{-g(Y)+\mathrm{E}[g(Y)Y]\}\big)\big]\Big|_{\varepsilon=0} = 0. \tag{5.18}$$

But this derivative is equal to

$$\begin{aligned}&\mathrm{E}\Big[u'\big(-f(Y)+\mathrm{E}[f(Y)Y]+\varepsilon\{-g(Y)+\mathrm{E}[g(Y)Y]\}\big)\\ &\quad\big\{-g(Y)+\mathrm{E}[g(Y)Y]\big\}\Big].\end{aligned} \tag{5.19}$$

For $\varepsilon = 0$, this derivative equals zero if

$$\begin{aligned}&\mathrm{E}\left[u'\big(-f(Y)+\mathrm{E}[f(Y)Y]\big)g(Y)\right] =\\ &\mathrm{E}\left[u'\big(-f(Y)+\mathrm{E}[f(Y)Y]\big)\right]\mathrm{E}[g(Y)Y].\end{aligned} \tag{5.20}$$

Writing $c = \mathrm{E}\left[u'(-f(Y)+\mathrm{E}[f(Y)Y])\right]$, this can be rewritten as

$$\mathrm{E}\left[\{u'\big(-f(Y)+\mathrm{E}[f(Y)Y]\big)-cY\}\{g(Y)\}\right] = 0. \tag{5.21}$$

Since the function $g(\cdot)$ is arbitrary, by a well-known theorem from variational calculus we find that necessarily

$$u'\big(-f(y)+\mathrm{E}[f(Y)Y]\big)-cy = 0. \tag{5.22}$$

Using $x = f(y)$ and $y = \varphi(x)$, we see that

$$\varphi(x) \propto u'\big(-x+\mathrm{E}[X\varphi(X)]\big). \tag{5.23}$$

Now, if $u(x)$ is exponential(α), so $u(x) = -\alpha\mathrm{e}^{-\alpha x}$, then

$$\varphi(x) \propto \mathrm{e}^{-\alpha(-x+\mathrm{E}[X\varphi(X)])} \propto \mathrm{e}^{\alpha x}. \tag{5.24}$$

Since $\mathrm{E}[\varphi(X)] = 1$, we obtain $\varphi(x) = \mathrm{e}^{\alpha x}/\mathrm{E}[\mathrm{e}^{\alpha X}]$ for the optimal standardized weight function. The resulting premium is an Esscher premium with parameter $h = \alpha$. ∇

Notice that the insurer uses a different weighting function for risks having different values of $\mathrm{E}[\varphi(X)]$; these functions differ only by a constant factor.

5.5 Premium reduction by coinsurance

Consider n cooperating insurers that individually have exponential utility functions with parameter α_i, $i = 1, 2, \ldots, n$. Together, they want to insure a risk S by defining random variables $S_1, \ldots, S_n$ with

$$S \equiv S_1 + \cdots + S_n, \tag{5.25}$$

with S_i denoting the risk insurer i faces. S might for example be a new risk they want to take on together, or it may be their combined insurance portfolios that they want to redistribute. The total premium they need is

$$P = \sum_{i=1}^{n} \frac{1}{\alpha_i} \log \mathrm{E}\left[\mathrm{e}^{\alpha_i S_i}\right]. \tag{5.26}$$

This total premium depends on the choice of the S_i. How should the insurers split up the risk S in order to make the pool as competitive as possible, hence to minimize the total premium P?

It turns out that the optimal choice $\widetilde{S}_i$ for the insurers is when each of them insures a fixed part of S, to be precise

$$\widetilde{S}_i = \frac{\alpha}{\alpha_i} S \quad \text{with } \frac{1}{\alpha} = \sum_{i=1}^{n} \frac{1}{\alpha_i}. \tag{5.27}$$

So, the optimal allocation is to let each insurer cover a fraction of the pooled risk that is proportional to the reciprocal of his risk aversion, hence to his risk tolerance. By (5.26) and (5.27), the corresponding total minimum premium is

$$\widetilde{P} = \sum_{i=1}^{n} \frac{1}{\alpha_i} \log \mathrm{E}\left[\mathrm{e}^{\alpha_i \widetilde{S}_i}\right] = \frac{1}{\alpha} \log \mathrm{E}\left[\mathrm{e}^{\alpha S}\right]. \tag{5.28}$$

This shows that the pool of cooperating insurers acts as one insurer with an exponential premium principle with as risk tolerance the total risk tolerance of the companies involved.

The proof that $\widetilde{P} \leq P$ for all other appropriate choices of $S_1 + \cdots + S_n \equiv S$ goes as follows. We have to prove that (5.28) is smaller than (5.26), so

$$\frac{1}{\alpha} \log \mathrm{E}\left[\exp\left(\alpha \sum_{i=1}^{n} S_i\right)\right] \leq \sum_{i=1}^{n} \frac{1}{\alpha_i} \log \mathrm{E}\left[\mathrm{e}^{\alpha_i S_i}\right], \tag{5.29}$$

which can be rewritten as

$$\mathrm{E}\left[\prod_{i=1}^{n} \mathrm{e}^{\alpha S_i}\right] \leq \prod_{i=1}^{n} \left(\mathrm{E}\left[\mathrm{e}^{\alpha_i S_i}\right]\right)^{\alpha/\alpha_i}. \tag{5.30}$$

This in turn is equivalent to

$$\mathrm{E}\left[\prod_{i=1}^{n} \frac{\mathrm{e}^{\alpha S_i}}{\mathrm{E}\left[\mathrm{e}^{\alpha_i S_i}\right]^{\alpha/\alpha_i}}\right] \leq 1, \tag{5.31}$$

or

$$\mathrm{E}\left[\exp \sum_i \frac{\alpha}{\alpha_i} T_i\right] \leq 1, \quad \text{with } T_i = \log \frac{\mathrm{e}^{\alpha_i S_i}}{\mathrm{E}\left[\mathrm{e}^{\alpha_i S_i}\right]}. \tag{5.32}$$

We can prove inequality (5.32) as follows. Note that $\mathrm{E}[\exp(T_i)] = 1$ and that by definition $\sum_i \alpha/\alpha_i = 1$. Since e^x is a convex function, we have for all real $t_1, \ldots, t_n$

$$\exp\left(\sum_i \frac{\alpha}{\alpha_i} t_i\right) \leq \sum_i \frac{\alpha}{\alpha_i} \exp(t_i), \tag{5.33}$$

and this implies that

$$\mathrm{E}\left[\exp\left(\sum_i \frac{\alpha}{\alpha_i} T_i\right)\right] \leq \sum_i \frac{\alpha}{\alpha_i} \mathrm{E}\left[\mathrm{e}^{T_i}\right] = \sum_i \frac{\alpha}{\alpha_i} = 1. \tag{5.34}$$

Hölder's inequality, which is well-known, arises by choosing $X_i = \exp(\alpha S_i)$ and $r_i = \alpha/\alpha_i$ in (5.30). See the exercises for the case $n = 2$, $r_1 = p$, $r_2 = q$.

5.6 Value-at-Risk and related risk measures

A risk measure for a random variable S is mathematically nothing but a functional connecting a real number to S. There are many possibilities. One is a premium, the price for taking over a risk. A ruin probability for some given initial capital measures the probability of becoming insolvent in the near or far future, therefore associates a real number with the random variable S representing the annual claims. It might be compound Poisson(λ, X), or $\mathrm{N}(\mu, \sigma^2)$. The risk measure most often used in practice is simply the *Value-at-Risk* at a certain (confidence) level q with $0 < q < 1$, which is the amount that will maximally be lost with probability q, therefore the argument x where $F_S(x)$ crosses level q. It is also called the *quantile risk measure*. Some examples of the practical use of the VaR are the following.

- To prevent insolvency, the available economic capital must cover unexpected losses to a high degree of confidence. Banks often choose their confidence level according to a standard of solvency implied by a credit rating of A or AA. These target ratings require that the institution have sufficient equity to buffer losses over a one-year period with probability 99.90% and 99.97%, respectively, based on historical one-year default rates. Typically, the reference point for the allocation of economic capital is a target rating of AA.

- For a portfolio of risks assumed to follow a compound Poisson risk process, the initial capital corresponding to some fixed probability ε of ultimate ruin is to be computed. This involves computing $\text{VaR}[L;1-\varepsilon]$, which is the VaR of the maximal aggregate loss L, see Section 4.7.
- Set the IBNR-reserve to be sufficient with a certain probability: the regulatory authorities prescribe, for example, the 75% quantile for the projected future claims as a capital to be held.
- Find out how much premium should be paid on a contract to have 60% certainty that there is no loss on that contract (*percentile premium principle*).

The definition of VaR is as follows:

Definition 5.6.1 (Value-at-Risk)
For a risk S, the *Value-at-Risk* (VaR) at (confidence) level p is defined as

$$\text{VaR}[S;p] \stackrel{\text{def}}{=} F_S^{-1}(p) \stackrel{\text{not}}{=} \inf\{s : F_S(s) \geq p\}. \tag{5.35}$$

So the VaR is just the inverse cdf of S computed at p. ∇

Remark 5.6.2 (VaR is cost-optimal)
Assume that an insurer has available an economic capital d to pay claims from. He must pay an annual compensation $i \cdot d$ to the shareholders. Also, there is the shortfall over d, which is valued as its expected value $\text{E}[(S-d)_+]$ (this happens to be the net premium a reinsurer might ask to take over the top part of the risk, but there need not be reinsurance in force at all). So the total costs amount to

$$i \cdot d + \text{E}[(S-d)_+]. \tag{5.36}$$

From the insurer's point of view, the optimal d turns out to be $\text{VaR}[S;1-i]$. This is easily seen by looking at the derivative of the cost (5.36), using the fact that $\frac{\text{d}}{\text{d}t}\text{E}[(S-t)_+] = F_S(t) - 1$. So the VaR is the cost minimizing capital to be held in this scenario. ∇

Since the typical values of i are around 10%, from this point of view the VaR-levels required for a company to get a triple A rating (99.975% and over, say) are suboptimal. On the other hand, maintaining such a rating might lead to a lower cost of capital i and increase production of the company. Also, it is a matter of prestige.

Note that the cdf F_S, by its definition as $F_S(s) = \Pr[S \leq s]$, is continuous from the right, but F_S^{-1} as in (5.35) is left-continuous, with $\text{VaR}[S;p] = \text{VaR}[S;p-0]$. It has jumps at levels p where F_S has a constant segment. In fact, any number s with $\text{VaR}[S;p] \leq s \leq \text{VaR}[S;p+0]$ may serve as the VaR. See Figure 5.1 for an illustration. We see that $F_S(s)$ has a horizontal segment at level 0.2; at this level, the VaR jumps from the lower endpoint of the segment to the upper endpoint. Where F_S has a jump, the VaR has a constant segment (between $p = 0.5$ and $p = 0.7$).

Remark 5.6.3 (Premiums are sometimes incoherent risk measures)
Relying on the VaR is often snubbed upon in financial circles these days, not as

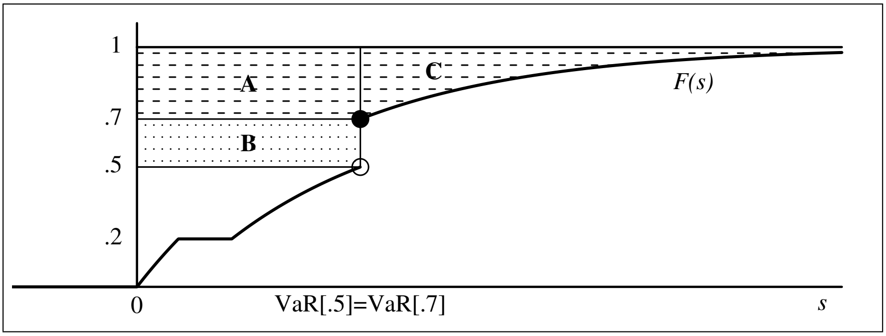

Fig. 5.1 Illustrating VaR and related risk measures for a cdf F_S

much because there is a danger in doing so, see Exercise 5.6.3, but because it is not a *coherent* risk measure in the sense of Artzner et al. (1997). To be called coherent, a risk measure must have respect for stochastic order, be positive homogeneous and translative, see Exercise 5.6.2, but it must also be *subadditive*. This last property means that the sum of the risk measures for a split-up portfolio is automatically an upper bound for the risk in the total portfolio. From Exercise 1.2.9 it is known that a risk-averse individual aiming to keep his utility at the same level or better is sometimes prepared to pay a premium for $2S$ that is strictly larger than twice the one for S. This is for example the case for exponential premiums, see Exercise 1.3.12, and the same superadditivity holds for Esscher premiums, see Exercise 5.6.11. Therefore, determining zero utility premiums may measure risks in a non-subadditive way, but to label this procedure incoherent is improper. Requiring subadditivity makes sense in complete markets where it is always possible to diversify a risk, but the insurance market simply is incomplete. ∇

Example 5.6.4 (VaR is not subadditive)
The following is a counterexample for the subadditivity of VaR: if S and $T \sim$ Pareto$(1,1)$ independent, then for *all* $p \in (0,1)$ we have

$$\mathrm{VaR}[S+T;p] > \mathrm{VaR}[S;p] + \mathrm{VaR}[T;p]. \tag{5.37}$$

To see that this is true, first verify that, since $F_S(x) = 1 - 1/x$, $x > 1$, we have $\mathrm{VaR}[S;p] = \frac{1}{1-p}$. Next, using convolution (see Exercise 5.6.10), check that

$$\Pr[S+T \le t] = 1 - \frac{2}{t} - 2\frac{\log(t-1)}{t^2}, \quad t > 2. \tag{5.38}$$

Now

$$\Pr\left[S+T \le 2\mathrm{VaR}[S;p]\right] = p - \frac{(1-p)^2}{2}\log\left(\frac{1+p}{1-p}\right) < p, \tag{5.39}$$

so for every p, we have $\mathrm{VaR}[S;p]+\mathrm{VaR}[T;p]<\mathrm{VaR}[S+T;p]$. This gives a counterexample for VaR being subadditive; in fact we have proved that for this pair (S,T), VaR is *superadditive*.

Note that from (1.33) with $d=0$ one might infer that $S+T$ and $2S$ should have cdfs that cross, so the same holds for the VaRs, being the inverse cdfs. Since we have $\mathrm{VaR}[2S;p]=\mathrm{VaR}[S;p]+\mathrm{VaR}[T;p]$ for all p, this contradicts the above. But this reasoning is invalid because $\mathrm{E}[S]=\infty$. To find counterexamples of subadditivity for VaR, it suffices to take $S\sim T$ with finite means. Then $S+T$ and $2S$ cannot have uniformly ordered quantile functions, so VaR must be subadditive at some levels, superadditive at others. ∇

The VaR itself does not tell the whole story about the risk, since if the claims S exceed the available funds d, someone still has to pay the remainder $(S-d)_+$. The Expected Shortfall measures 'how bad is bad':

Definition 5.6.5 (Expected shortfall)
For a risk S, the *Expected Shortfall* (ES) at level $p\in(0,1)$ is defined as

$$\mathrm{ES}[S;p]=\mathrm{E}[(S-\mathrm{VaR}[S;p])_+]. \tag{5.40}$$

Thus, ES can be interpreted as the net stop-loss premium in the hypothetical situation that the excess over $d=\mathrm{VaR}[S;p]$ is reinsured. ∇

In Figure 5.1, for levels $p\in[0.5,0.7]$, $\mathrm{ES}[S;p]$ is the stop-loss premium at retention $\mathrm{VaR}[S;0.7]$, therefore just the area C; see also (1.33).

As we noted, VaR is not subadditive. Averaging high level VaRs, however, does produce a subadditive risk measure (see Property 5.6.10).

Definition 5.6.6 (Tail-Value-at-Risk (TVaR))
For a risk S, the *Tail-Value-at-Risk* (TVaR) at level $p\in(0,1)$ is defined as:

$$\mathrm{TVaR}[S;p]=\frac{1}{1-p}\int_p^1 \mathrm{VaR}[S;t]\,\mathrm{d}t. \tag{5.41}$$

So the TVaR is just the arithmetic average of the VaRs of S from p on. ∇

Remark 5.6.7 (Other expressions for TVaR)
The Tail-Value-at-Risk can also be expressed as

$$\begin{aligned}\mathrm{TVaR}[S;p]&=\mathrm{VaR}[S;p]+\frac{1}{1-p}\int_p^1\left\{\mathrm{VaR}[S;t]-\mathrm{VaR}[S;p]\right\}\mathrm{d}t\\&=\mathrm{VaR}[S;p]+\frac{1}{1-p}\mathrm{ES}[S;p].\end{aligned} \tag{5.42}$$

It is easy to see that the integral in (5.42) equals $\mathrm{ES}[S;p]$, in similar fashion as in Figure 1.1.

Just as VaR, the TVaR arises naturally from the cost optimization problem considered in Remark 5.6.2. In fact, TVaR is the optimal value of the cost function, divided by i. To see why this holds, simply fill in $d=\mathrm{VaR}[S;p=1-i]$ in the cost function $i \cdot d + \mathrm{E}[(S-d)_+]$, see (5.36). As a consequence, we have the following characterization of TVaR:

$$\mathrm{TVaR}[S;p] = \inf_d \left\{ d + \frac{1}{1-p}\pi_S(d) \right\}. \tag{5.43}$$

Filling in other values than $d = \mathrm{VaR}[S;p]$ gives an upper bound for $\mathrm{TVaR}[S;p]$. ∇

In Figure 5.1, in view of (5.42) above, for level $p = \frac{1}{2}$, $(1-p)\mathrm{TVaR}[S;p]$ equals the total of the areas marked A, B and C. Therefore, the TVaR is just the average length of the dashed and the dotted horizontal lines in Figure 5.1.

TVaR resembles but is not always identical to the following risk measure.

Definition 5.6.8 (Conditional Tail Expectation)
For a risk S, the *Conditional Tail Expectation* (CTE) at level $p \in (0,1)$ is defined as

$$\mathrm{CTE}[S;p] = \mathrm{E}\left[S \,|\, S > \mathrm{VaR}[S;p]\right]. \tag{5.44}$$

So the CTE is the 'average loss in the worst $100(1-p)\%$ cases'. ∇

Writing $d = \mathrm{VaR}[S;p]$ we have

$$\mathrm{CTE}[S;p] = \mathrm{E}[S \,|\, S > d] = d + \mathrm{E}[(S-d)_+ \,|\, S > d] = d + \frac{\mathrm{E}[(S-d)_+]}{\Pr[S > d]}. \tag{5.45}$$

Therefore we have for all $p \in (0,1)$

$$\mathrm{CTE}[S;p] = \mathrm{VaR}[S;p] + \frac{1}{1 - F_S(\mathrm{VaR}[S;p])}\mathrm{ES}[S;p]. \tag{5.46}$$

So by (5.42), the CTE differs from the TVaR only when $p < F_S(\mathrm{VaR}[S;p])$, which means that F_S jumps over level p. In fact we have

$$\mathrm{CTE}[S;p] = \mathrm{TVaR}[S;F_S(F_S^{-1}(p))] \geq \mathrm{TVaR}[S;p] \geq \mathrm{VaR}[S;p]. \tag{5.47}$$

In Figure 5.1, levels p outside $(0.5, 0.7)$ lead to TVaR and CTE being identical. In case $p \in [0.5, 0.7]$, the CTE is just the total of the areas A and C, divided by $1 - F_S(\mathrm{VaR}[S;0.7])$. So the CTE at these levels p is the average length of the dashed horizontal lines in Figure 5.1 only. Diagrams of the VaR, ES, TVaR and CTE for risk S in Figure 5.1 are given in Figure 5.2. Observe that

- the VaR jumps where F_S is constant, and it is horizontal where F_S jumps;
- the ES, being a stop-loss premium at levels increasing with p, is decreasing in p, but not convex or even continuous;
- TVaR and CTE coincide except for p-values in the vertical part of F_S;
- TVaR is increasing and continuous.

Remark 5.6.9 (Terminology about ES, TVaR and CTE)
Some authors use the term Expected Shortfall to refer to CTE (or TVaR). As we define it, the ES is an *unconditional* mean of the shortfall, while TVaR and CTE

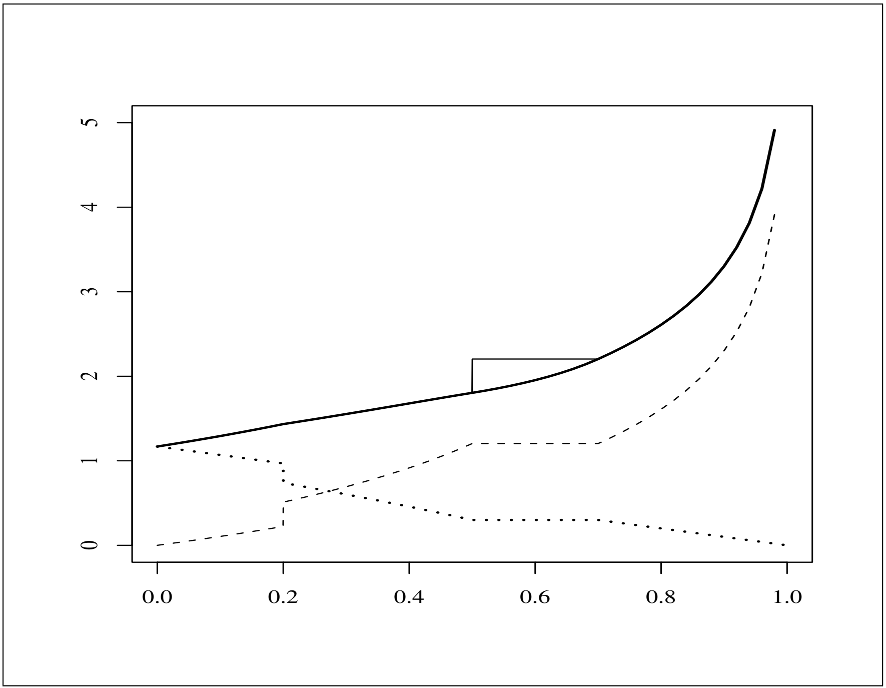

Fig. 5.2 TVaR (thick), CTE (thin), ES (dotted) and VaR (dashed) for the cdf F_S in Fig. 5.1, plotted against the level $p \in (0,1)$

involve the expected shortfall, given that it is positive. CTE and TVaR are also often used interchangeably. Defined as above, these quantities are the same only in case of continuous distributions. Another synonym for CTE that one may encounter in the literature is Conditional Value-at-Risk (CVaR). ∇

Unlike the VaR, see Exercise 5.6.9 and Example 5.6.4, the Tail-Value-at-Risk is subadditive.

Property 5.6.10 (TVaR is subadditive)
$\text{TVaR}[S+T;p] \leq \text{TVaR}[S;p] + \text{TVaR}[T;p]$ holds for all pairs of risks (S,T) and for all $p \in (0,1)$.

Proof. We mentioned that in the characterization of TVaR (5.43) above, an upper bound for $\text{TVaR}[S;p]$ is obtained by filling in any other value for d. Now let especially $d = d_1 + d_2$, with $d_1 = \text{VaR}[S;p]$ and $d_2 = \text{VaR}[T;p]$. Then we have

$$
\begin{aligned}
\mathrm{TVaR}[S+T;p] &\leq d + \frac{1}{1-p}\pi_{S+T}(d) \\
&= d_1 + d_2 + \frac{1}{1-p}\mathrm{E}[(S+T-d_1-d_2)_+] \\
&\leq d_1 + d_2 + \frac{1}{1-p}\mathrm{E}[(S-d_1)_+ + (T-d_2)_+] \\
&= \mathrm{TVaR}[S;p] + \mathrm{TVaR}[T;p].
\end{aligned} \tag{5.48}
$$

Here we used that $s_+ + t_+ = (s_+ + t_+)_+ \geq (s+t)_+$ for all real s,t. ∇

In Exercise 5.6.5, we will see an example where CTE is not subadditive. For continuous risks, CTE and TVaR coincide, but there is also a neat more direct proof that CTE is subadditive in this case.

Property 5.6.11 (CTE is subadditive for continuous risks)
If S and T are random variables with continuous marginal cdfs and joint cdf, CTE is subadditive.

Proof. If $s = \mathrm{VaR}[S;p]$, $t = \mathrm{VaR}[T;p]$ and $z = \mathrm{VaR}[S+T;p]$, by the assumed continuity we have $1-p = \Pr[S+T>z] = \Pr[S>s] = \Pr[Y>t]$. It is not hard to see that if for some s, the events A and $S>s$ are equally likely, for the conditional means we have $\mathrm{E}[S\,|\,A] \leq \mathrm{E}[S\,|\,S>s]$. Apply this with $A \equiv S+T>z$, then we get

$$
\mathrm{E}[S+T\,|\,A] = \mathrm{E}[S\,|\,A] + \mathrm{E}[T\,|\,A] \leq \mathrm{E}[S\,|\,S>s] + \mathrm{E}[T\,|\,T>t]. \tag{5.49}
$$

This is tantamount to

$$
\mathrm{CTE}[S+T;p] \leq \mathrm{CTE}[S;p] + \mathrm{CTE}[T;p], \tag{5.50}
$$

so subadditivity for the CTE's of continuous risks is proved. ∇

`R` provides nine ways to estimate a VaR at level p from a sample `S` of size n, differing subtly in the way the interpolation between the order statistics close to np of the sample is performed. See `?quantile` for more information on this. Leaving out the type parameter is equivalent to choosing the default method `type=7` in `R`. An example for a lognormal(0, 1) random sample:

```
set.seed(1); S <- exp(rnorm(1000,0,1))
levels <- c(0.5,0.95,0.975,0.99,0.995)
quantile(S, levels, type=7)
##      50%        95%      97.5%        99%      99.5%
##0.9652926  5.7200919  7.4343584 10.0554305 11.5512498
```

The reader is asked to computes estimates of the other risk measures in this section in Exercise 5.6.4.

5.7 Exercises

Section 5.2

1. Show that (5.7) is valid.
2. What are the results in the table in case of a dividend $i = 2\%$ and $\varepsilon = 5\%$? Calculate the variance premium as well as the exponential premium.

Section 5.3

1. Let $X \sim$ exponential(1). Determine the premiums (a)–(e) and (h)–(j).
2. [♠] Prove that $\pi[X;\alpha] = \log(\mathrm{E}[\mathrm{e}^{\alpha X}])/\alpha$ is an increasing function of α, by showing that the derivative with respect to α is positive (see also Example 1.3.1).
3. Assume that the total claims for a car portfolio has a compound Poisson distribution with gamma distributed claims per accident. Determine the expected value premium if the loading factor equals 10%.
4. Determine the exponential premium for a compound Poisson risk with gamma distributed individual claims.
5. Calculate the variance premium for the claims distribution as in Exercise 5.3.3.
6. Show that the Esscher premium equals $\kappa_X'(h)$, where κ_X is the cgf of X.
7. What is the Esscher transformed density with parameter h for the following densities: exponential(α), binomial(n,p) and Poisson(λ)?
8. Show that the Esscher premium for X increases with the parameter h.
9. Calculate the Esscher premium for a compound Poisson distribution.
10. Show that the Esscher premium for small values of α boils down to a variance premium principle.
11. Assume that X is a finite risk with maximal value b, hence $\Pr[X \leq b] = 1$ but $\Pr[X \geq b - \varepsilon] > 0$ for all $\varepsilon > 0$. Let π_α denote the exponential premium for X. Show that $\lim_{\alpha\to\infty} \pi_\alpha = b$.
12. Show that the exponential premium $\pi[X;\alpha]$ with risk aversion α is the difference quotient $\big(\kappa_X(\alpha) - \kappa_X(0)\big)/\alpha$. Prove that it also can be written as a uniform mixture of Esscher premiums $\int_0^\alpha \pi[X;h]\,\mathrm{d}h/\alpha$. From the fact that Esscher premiums increase with h, what can be concluded about the Esscher($h = \alpha$) premium compared with the exponential(α) premium?
13. In Table 5.1, prove the properties that are marked "+".
14. Construct counterexamples for the first 4 rows and the second column for the properties that are marked "−".
15. Investigate the additivity of a mixture of Esscher principles of the following type: $\pi[X] = p\pi[X;h_1] + (1-p)\pi[X;h_2]$ for some $p \in [0,1]$, where $\pi[X;h]$ is the Esscher premium for risk X with parameter h.
16. Formulate a condition for dependent risks X and Y that implies that $\pi[X+Y] \leq \pi[X] + \pi[Y]$ for the variance premium (*subadditivity*). Also show that this property holds for the standard deviation principle, no matter what the joint distribution of X and Y is.

Section 5.5

1. For a proof of Hölder's inequality in case of $n = 2$, let $p > 1$ and $q > 1$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$. Successively prove that
   - if $u > 0$ and $v > 0$, then $uv \le \frac{u^p}{p} + \frac{v^q}{q}$; (write $u = e^{s/p}$ and $v = e^{t/q}$);
   - if $E[U^p] = E[V^q] = 1$ and $\Pr[U > 0] = \Pr[V > 0] = 1$, then $E[UV] \le 1$;
   - $|E[XY]| \le E[|X|^p]^{1/p} E[|Y|^q]^{1/q}$.
2. Whose inequality arises for $p = q = 2$ in the previous exercise?

Section 5.6

1. Compute estimates for the TVaR and ES based on the same lognormal sample `S` as given in the `R`-example at the end of this section and at level 0.95.
2. Prove that the following properties hold for TVaR and VaR:
   - no rip-off — $(T)VaR[S;p] \le \max[S]$
   - no unjustified loading — $(T)VaR[S;p] = c$ if $S \equiv c$
   - non-decreasing in p — if $p < q$, then $(T)VaR[S;p] \le (T)VaR[S;q]$
   - non-negative safety loading — $TVaR[S;p] \ge E[S]$
   - translative — $(T)VaR[S + c;p] = (T)VaR[S;p] + c$
   - positively homogeneous — $(T)VaR[\alpha S] = \alpha \times (T)VaR[S;p]$ for all $\alpha > 0$
   - TVaR is continuous in p — but VaR and CTE are not
3. Let X_i, $i = 1, \dots, 100$ be iid risks with $\Pr[X_i = -100] = 0.01$, $\Pr[X_i = +2] = 0.99$. Compare the VaRs of $S_1 = \sum_{i=1}^{100} X_i$ (diversified) and $S_2 = 100X_1$ (non-diversified). The diversified risk is obviously safer, but the VaRs cannot be uniformly smaller for two random variables with the same finite mean. Plot the cdfs of both these risks. Investigate which of these risks has a smaller VaR at various levels p.
4. Write `R`-functions computing TVaRs for (translated) gamma and lognormal distributions. Make plots of the TVaRs at levels $1\%, 3\%, \dots, 99\%$.
5. Prove that CTE is not subadditive by looking at level $p = 0.9$ for the pair of risks (X,Y) with:

$$X \sim \text{uniform}(0,1); \qquad Y = \begin{cases} 0.95 - X & \text{if } X \le 0.95 \\ 1.95 - X & \text{if } X > 0.95 \end{cases}$$

   Note that $X \sim Y \sim$ uniform$(0,1)$ are continuous, but $X + Y$ is discrete.
6. Give expressions for ES, TVaR and CTE in case $S \sim$ uniform(a,b).
7. As the previous exercise, but now for $S \sim N(\mu, \sigma^2)$.
8. As the previous exercise, but now for $S \sim$ lognormal(μ, σ^2).
9. Show that VaR is not subadditive using the example $X, Y \sim$ Bernoulli(0.02) iid and $p = 0.975$. Consider also the case that (X,Y) is bivariate normal.
10. Derive (5.38).
11. If $h > 0$, prove that $\pi[2S;h] > 2\pi[S;h]$ for Esscher premiums.