(d) Use (c) to prove that, if $e(x) \geq e(0)$, then

$$S(x) \leq \frac{\mathrm{E}(X)}{x + \mathrm{E}(X)}$$

and thus

$$S[k\mathrm{E}(X)] \leq \frac{1}{k+1},$$

which for $k = 1$ implies that the mean is at least as large as the (smallest) median.

(e) Prove that (b) may be rewritten as

$$S_e(x) = \frac{e(x)}{x + e(x)} \frac{\int_x^\infty y f(y) dy}{\mathrm{E}(X)}$$

and thus that

$$S_e(x) \leq \frac{e(x)}{x + e(x)}.$$

3.5 MEASURES OF RISK

3.5.1 Introduction

Probability-based models provide a description of risk exposure. The level of exposure to risk is often described by one number, or at least a small set of numbers. These numbers are functions of the model and are often called "key risk indicators." Such key risk indicators inform actuaries and other risk managers about the degree to which the company is subject to particular aspects of risk. In particular, Value-at-Risk (VaR) is a quantile of the distribution of aggregate losses. Risk managers often look at "the chance of an adverse outcome." This can be expressed through the VaR at a particular probability level. VaR can also be used in the determination of the amount of capital required to withstand such adverse outcomes. Investors, regulators, and rating agencies are particularly interested in the company's ability to withstand such events.

VaR suffers from some undesirable properties. A more informative and more useful measure of risk is Tail-Value-at-Risk (TVaR). It has arisen independently in a variety of areas and has been given different names including Conditional-Value-at-Risk (CVaR), Conditional Tail Expectation (CTE), and Expected Shortfall (ES).

While these measures have been developed in a risk management context, they are useful in assessing any random variable.

3.5.2 Risk measures and coherence

A risk measure is a mapping from the random variable representing the loss associated with the risks to the real line (the set of all real numbers). A risk measure gives a single number that is intended to quantify the risk exposure. For example, the standard deviation, or a multiple of the standard deviation of a distribution, is a measure of risk because it provides a measure of uncertainty. It is clearly appropriate when using the normal distribution. In the field of finance, the size of loss for which there is a small (e.g., 0.05%) probability of exceedence is a simple risk measure.

Risk measures are denoted by the function $\rho(X)$. It is convenient to think of $\rho(X)$ as the amount of assets required to protect against adverse outcomes of the risk X. Studies of risk measures and their properties have included the behavior of risk measures when several losses are combined and treated as a single loss. Combining risks is important in the study of capital needs of an insurance company when considering the overall risk exposure of the insurance company. The insurance company may have several divisions or departments specializing in different products; for example, individual life, homeowners, automobile, group life, annuities, and health. When risk measures are applied to the individual departments, the results should be consistent in some way with the results that are obtained when the risk measure is applied to the entire company. In what follows, it is useful to think of the random variables X and Y as the loss random variables for two divisions and $X+Y$ as the loss random variable for the entity created by combining the two divisions.

The study of risk measures and their properties has been carried out by numerous authors such as Wang [180] and [181]. Specific desirable properties of risk measures were proposed as axioms in connection with risk pricing by Wang, Young, and Panjer [182] and more generally in risk measurement by Artzner et al. [7]. The Artzner paper introduced the concept of *coherence* and is considered to be the groundbreaking paper in risk measurement.

We consider the set of random variables such that if X and Y are two members of the set, then both cX and $X+Y$ are also in the set. This is not very restrictive, but it does eliminate risks that are measured as percentages as with Model 1 of Chapter 2.

Definition 3.11 *A* ***coherent risk measure*** *is a risk measure $\rho(X)$ that has the following four properties for any two loss random variables X and Y:*

1. *Subadditivity:* $\rho(X+Y) \leq \rho(X) + \rho(Y)$.
2. *Monotonicity: If $X \leq Y$ for all possible outcomes, then $\rho(X) \leq \rho(Y)$.*
3. *Positive homogeneity: For any positive constant c, $\rho(cX) = c\rho(X)$.*
4. *Translation invariance: For any positive constant c, $\rho(X+c) = \rho(X) + c$.*

Subadditivity means that the risk measure (and, hence, the capital required to support it) for two risks combined will not be greater than for the risks treated separately. Subadditivity reflects the fact that there should be some diversification benefit from combining risks. In general, this is necessary at the corporate level. Otherwise, companies would find it to be an advantage to disaggregate into smaller companies. There has been some debate about the appropriateness of the subadditivity requirement. In particular, the merger of several small companies into a larger one exposes each of the small companies to the reputational risk of the others. We will continue to require subadditivity as it reflects the benefit of diversification.

Monotonicity means that if one risk always has greater losses than another risk under all circumstances,[3] the risk measure (and, hence, the capital required to support it) should always be greater. This requirement should be self-evident from an economic viewpoint.

[3] Technically, this means that for the joint distribution of (X, Y), $\Pr(X > Y) = 0$.

Positive homogeneity means that the risk measure (and, hence, the capital required to support it) is independent of the currency in which the risk is measured. Equivalently, it means that, for example, doubling the exposure to a particular risk requires double the capital. This is sensible because doubling the position provides no diversification.

Translation invariance means that there is no additional risk (and, hence, capital required to support it) for an additional risk for which there is no additional uncertainty. In particular, by making X identically zero, the value of the assets required for a certain outcome are exactly the value of that outcome. Also, when a company meets the capital requirement by setting up additional risk-free capital, the act of injecting the additional capital does not, in itself, trigger a further injection (or reduction) of capital.

Risk measures satisfying these four criteria are deemed to be coherent. There are many such risk measures.

EXAMPLE 3.12

(*Standard deviation principle*) The standard deviation is a measure of uncertainty of a distribution. Consider a loss distribution with mean μ and standard deviation σ. The quantity $\mu + k\sigma$, where k is the same fixed constant for all distributions, is a risk measure (often called the **standard deviation principle**). The coefficient k is usually chosen to ensure that losses will exceed the risk measure for some distribution, such as the normal distribution, with some specified small probability. □

In Exercise 3.31 you are asked to prove that the standard deviation principle is not coherent.

If X follows the normal distribution, a value of $k = 1.645$ results in an exceedence probability of $\Pr(X > \mu + k\sigma) = 5\%$ while, if $k = 2.576$, then $\Pr(X > \mu + k\sigma) = 0.5\%$. However, if the distribution is not normal, the same multiples of the standard deviation will lead to different exceedence probabilities. One can also begin with the exceedence probability, obtaining the quantile $\mu + k\sigma$ and the equivalent value of k. This is the key idea behind Value-at-Risk.

3.5.3 Value-at-Risk

Value-at-Risk (VaR) has become the standard risk measure used to evaluate exposure to risk. In general terms, the VaR is the amount of capital required to ensure, with a high degree of certainty, that the enterprise doesn't become technically insolvent. The degree of certainty chosen is arbitrary. In practice, it can be a high number such as 99.95% for the entire enterprise, or it can be much lower, such as 95%, for a single unit or risk class within the enterprise. This lower percentage may reflect the inter-unit or inter-risk type diversification that exists.

Suppose $F_X(x)$ represents the distribution function of outcomes over a fixed period of time, such as one year, of a portfolio of risks (such as a set of insurance risks or an entire insurance company). An adverse outcome is referred to as a "loss." In the notation used throughout of this book, positive values of the random variable X are adverse outcomes, that is, losses. The VaR of the random variable X is the $100p$th percentile of the distribution of X, denoted by $\mathrm{VaR}_p(X) = \pi_p$. This shows why VaR is often called a *quantile risk measure*. When the insurance company has this amount of capital available, it can absorb $100p\%$ of possible outcomes. When $p = 99.95\%$ for a one-year time period, the interpretation is that there is only a very small chance (0.05%) that the insurance company will be bankrupted by an adverse outcome over the next year.

Definition 3.12 *Let X denote a loss random variable. The* ***Value-at-Risk*** *of X at the $100p\%$ level, denoted $\mathrm{VaR}_p(X)$ or π_p, is the $100p$ percentile (or quantile) of the distribution of X.*

For continuous distributions, we can simply write $\mathrm{VaR}_p(X)$ for random variable X as the value of π_p satisfying

$$\Pr(X > \pi_p) = 1 - p.$$

It is well known that VaR does not satisfy one of the four criteria for coherence, the subadditivity requirement. The failure of VaR to be subadditive can be shown by a simple but extreme example inspired by a more complicated one from Wirch [188].

EXAMPLE 3.13

(*Incoherence of VaR*) Let Z denote a loss random variable of the continuous type with the following cdf values:

$$\begin{aligned} F_Z(1) &= 0.91, \\ F_Z(90) &= 0.95, \\ F_Z(100) &= 0.96. \end{aligned}$$

The 95% quantile, the $\mathrm{VaR}_{95\%}(Z)$ is 90 because there is a 5% chance of exceeding 90.

Suppose that we now split the risk Z into two separate (but dependent) risks X and Y such that the two separate risks in total are equivalent to risk Z, that is, $X + Y = Z$. One way to define them is:

$$X = \begin{cases} Z, & Z \le 100 \\ 0, & Z > 100 \end{cases}$$

and

$$Y = \begin{cases} 0, & Z \le 100 \\ Z, & Z > 100. \end{cases}$$

The cdf for risk X satisfies

$$\begin{aligned} F_X(1) &= 0.95, \\ F_X(90) &= 0.99, \\ F_X(100) &= 1, \end{aligned}$$

indicating $\mathrm{VaR}_{95\%}(X) = 1$.

Similarly, the cdf for risk Y satisfies $F_Y(0) = 0.96$, indicating that there is a 96% chance of no loss. Therefore the 95% quantile cannot exceed 0, and so

$\mathrm{VaR}_{95\%}(Y) \leq 0$. Consequently, the sum of the 95% quantiles for X and Y is less than the $\mathrm{VaR}_{95\%}(Z)$, which violates subadditivity. □

Although this example may appear to be somewhat artificial, the existence of such possibilities creates opportunities for strange or unproductive manipulation. Therefore we turn to a risk measure that is coherent.

3.5.4 Tail-Value-at-Risk

As a risk measure, VaR is used extensively in financial risk management of trading risk over a fixed (usually relatively short) time period. In these situations, the normal distribution is often used for describing gains or losses. If distributions of gains or losses are restricted to the normal distribution, VaR satisfies all coherency requirements. However, the normal distribution is generally not used for describing insurance losses, which are typically skewed. Consequently, the use of VaR is problematic because of the lack of subadditivity.

Definition 3.13 *Let X denote a loss random variable. The* ***Tail-Value-at-Risk*** *of X at the $100p\%$ security level, denoted $\mathrm{TVaR}_p(X)$, is the expected loss given that the loss exceeds the $100p$ percentile (or quantile) of the distribution of X.*

For the sake of notational convenience, we restrict consideration to continuous distributions to avoid ambiguity about the definition of TVaR. In general, we can extend the results to discrete distributions or distributions of mixed type by appropriately modifying definitions. For most practical purposes, it is sufficient to think in terms of continuous distributions.

We can write $\mathrm{TVaR}_p(X)$ as

$$\mathrm{TVaR}_p(X) = \mathrm{E}\left(X \mid X > \pi_p\right) = \frac{\int_{\pi_p}^{\infty} x f(x)dx}{1 - F(\pi_p)}.$$

Furthermore, if this quantity is finite, we can use integration by parts and substitution to rewrite it as

$$\mathrm{TVaR}_p(X) = \frac{\int_p^1 \mathrm{VaR}_u(X)du}{1-p}.$$

Thus, TVaR can be seen to average all VaR values above the security level p. This means that TVaR tells us much more about the tail of the distribution than does VaR alone.

Finally, TVaR can also be written as

$$\begin{aligned}
\mathrm{TVaR}_p(X) &= \mathrm{E}\left(X \mid X > \pi_p\right) \\
&= \pi_p + \frac{\int_{x_p}^{\infty} \left(x - \pi_p\right) f(x)dx}{1-p} \\
&= \mathrm{VaR}_p(X) + e(\pi_p), \qquad (3.21)
\end{aligned}$$

where $e(\pi_p)$ is the mean excess loss function evaluated at the $100p$th percentile. Thus TVaR is larger than the corresponding VaR by the average excess of all losses that exceed VaR. Furthermore, because $\pi_p = \mathrm{Var}_p(X)$, (3.21) expresses $\mathrm{TVaR}_p(X)$ as a function of $\mathrm{Var}_p(X)$, and in Exercise 3.37 the fact that $\mathrm{TVar}_p(X)$ is an non-decreasing function of $\mathrm{Var}_p(X)$ is established.

TVaR has been developed independently in the insurance field and is called *Conditional Tail Expectation* (CTE) by Wirch [188] and widely known by that term in North America. It has also been called *Tail Conditional Expectation* (TCE). In Europe, it has also been called *Expected Shortfall* (ES). (See Tasche [168] and Acerbi and Tasche [5].)

Overbeck [132] also discusses VaR and TVaR as risk measures. He argues that VaR is an "all or nothing" risk measure, in that if an extreme event in excess of the VaR threshold occurs, there is no capital to cushion losses. He also argues that the VaR quantile in TVaR provides a definition of "bad times," which are those where losses exceed the VaR threshold, thereby not using up all available capital when TVaR is used to determine capital. Then TVaR provides the average excess loss in "bad times," that is, when the VaR "bad times" threshold has been exceeded.

EXAMPLE 3.14

(*Normal distribution*) Consider a normal distribution with mean μ, standard deviation σ, and pdf

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right].$$

Let $\phi(x)$ and $\Phi(x)$ denote the pdf and the cdf of the standard normal distribution ($\mu = 0$, $\sigma = 1$). Then

$$\mathrm{VaR}_p(X) = \mu + \sigma\Phi^{-1}(p),$$

and, with a bit of calculus, it can be shown that

$$\mathrm{TVaR}_p(X) = \mu + \sigma\frac{\phi\left[\Phi^{-1}(p)\right]}{1-p}.$$

Note that, in both cases, the risk measure can be translated to the standard deviation principle with an appropriate choice of k. □

EXAMPLE 3.15

(*Exponential distribution*) Consider an exponential distribution with mean θ and pdf

$$f(x) = \frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right),\ x > 0.$$

Then

$$\mathrm{VaR}_p(X) = -\theta\ln(1-p)$$

and

$$\mathrm{TVaR}_p(X) = \mathrm{VaR}_p(X) + \theta.$$

The excess of TVaR over VaR is a constant θ for all values of p because of the memoryless property of the exponential distribution. □

■ EXAMPLE 3.16

(*Pareto distribution*) Consider a Pareto distribution with scale parameter θ, shape parameter $\alpha > 1$, and cdf

$$F(x) = 1 - \left(\frac{\theta}{x+\theta}\right)^{\alpha}, \quad x > 0.$$

Then

$$\mathrm{VaR}_p(X) = \theta\left[(1-p)^{-1/\alpha} - 1\right]$$

and

$$\mathrm{TVaR}_p(X) = \mathrm{VaR}_p(X) + \frac{\mathrm{VaR}_p(X) + \theta}{\alpha - 1}.$$

The excess of TVaR over VaR is a linear increasing function in the VaR. This means that a larger VaR results in a larger mean excess loss over the VaR indicating a dangerous distribution. □

It is well known that TVaR is a coherent measure. This has been shown by Artzner et al. [7]. Therefore, when using it, we never run into the problem of subadditivity of the VaR. TVaR is one of many possible coherent risk measures. However, it is particularly well-suited to insurance applications where you may want to reflect the shape of the tail beyond the VaR threshold in some way. TVaR represents that shape through a single number: the mean excess loss or expected shortfall.

■ EXAMPLE 3.17

(*Tail comparisons*) Consider three loss distributions for an insurance company, Losses for the next year are estimated to be \$100 million with a standard deviation \$223.607 million. You are interested in finding high quantiles of the distribution of losses. Using the normal, Pareto, and Weibull distributions, obtain the VaR at the 99%, 99.9%, and 99.99% security levels.

From the mean and standard deviation, using the moment formulas in Appendix A, the distributions and their parameters (in \$millions) are Normal(100, 223.607), Pareto(120, 2.2) and Weibull(50, 0.5). From the formulas for the cumulative distribution functions, the quantiles $\pi_{0.90}$, $\pi_{0.99}$, and $\pi_{0.999}$ are obtained. They are listed, in millions, in Table 3.1. □

From this example, it should be noted that the results can vary widely depending on the choice of distribution. The normal distribution has a lighter tail than the others. Therefore the probabilities at extreme outcomes are relatively small leading to smaller quantiles. The Pareto distribution and the Weibull distribution with $\tau < 1$ have heavy tails and thus relatively larger extreme quantiles. This