Chapter 7

TWO TYPES OF RISK

Jerzy A. Filar
Center for the Industrial and Applied Mathematics,
School of Mathematics & Statistics,
University of South Australia,
Mawson Lakes, SA 5095,Australia
jerzy.filar@unisa.edu.au

Boda Kang
Center for the Industrial and Applied Mathematics,
School of Mathematics & Statistics,
University of South Australia,
Mawson Lakes, SA 5095,Australia
boda.kang@unisa.edu.au

Abstract The risk encountered in many environmental problems appears to exhibit special "two-sided" characteristics. For instance, in a given area and in a given period, farmers do not want to see too much or too little rainfall. They hope for rainfall that is in some given interval. We formulate and solve this problem with the help of a "two-sided loss function" that depends on the above range. Even in financial portfolio optimization a loss and a gain are "two sides of a coin", so it is desirable to deal with them in a manner that reflects an investor's relative concern. Consequently, in this paper, we define Type I risk: "the loss is too big" and Type II risk: "the gain is too small". Ideally, we would want to minimize the two risks simultaneously. However, this may be impossible and hence we try to balance these two kinds of risk. Namely, we tolerate certain amount of one risk when minimizing the other. The latter problem is formulated as a suitable optimization problem and illustrated with a numerical example.

Keywords: Two-sided risk, rainfall, temperature, value-at-risk, conditional value-at-risk, Type I risk, value-of-gain, conditional value-of-gain, Type II risk, assurance, scenarios, portfolio optimization.

Introduction

The risk encountered in many environmental problems appears to exhibit special "two-sided" characteristics. The "fundamental security" in an environmental problem may be a variable such as a rainfall or a temperature. For instance, in a given area and in a given period, farmers do not want to see too much or too little rainfall. They hope for rainfall that is in some given interval. Similarly, we often hope that the temperature is neither too high nor too low. We formulate and solve this problem with the help of a "two-sided loss function" that depends on the above range.

In financial mathematics, there is an extensive literature discussing the risk of a financial portfolio using the value-at-risk concept, see [3–5, 14–18]. However, these authors consider only the "one-sided" risk using the return of the portfolio. We argue that - even in financial context - a loss and a gain are "two sides of a coin", so it is desirable to differentiate between the loss and the gain of a portfolio and deal with them in a manner that reflects an investor's relative concern about loss and gain. This is because different people have different attitudes toward a loss and a gain. Thus, it might be useful to provide models that trade-off the aversion to these two types of risk. Trying to minimize these kinds of risks is somewhat different from minimizing conditional value-at-risk using the usual loss or gain function (as is done in, for instance, [14, 15]).

Based on the above discussion, we define Type I risk: "the loss is too big" and Type II risk: "the gain is too small". Ideally, we would want to minimize the two risks simultaneously. However, this may be impossible and hence we try to balance these two kinds of risk. Namely, we tolerate certain amount of one risk when minimizing the other. In the financial context, investors can then suitably choose parameters according to their own attitude towards the loss and gain risk.

The paper is organized as follows: we provide a new loss function for the two-sided problem such as rainfall or temperature. Using the new loss function together with conditional value-at-risk, we show how to formulate such a risk in Section 1. In Section 2 we introduce two types of risk associated with the new loss and gain function. We suggest a way to balance a loss and a gain in a more general case. We also provide a criterion for users to choose parameters in these problems. Finally, in Section 3, together with the portfolio problem, we put forward the concept and some properties of conditional value-at-risk and conditional value-of-gain with the new loss and gain functions respectively. All proofs are provided in Section 4.

1. Two-sided risk

Risks in environmental problems are different from financial market risk in some aspects. For example, in the rainfall problem, too much rain or too little rain are both undesirable. Too much rain will lead to a flood, whereas too little rain will lead to a drought. A similar problem arises with temperature. We do not want the temperature in a location to be too high or too low. In this section we introduce one natural formulation of this "two-sided risk" problem.

Let a random variable (r.v.,for short) X denote the rainfall in a location during some specific period. Let us suppose, for instance, that farmers in this location hope that the rainfall in this season is in the interval $[\nu_1, \nu_2]$. That is, exceeding ν_2 or being lower than ν_1 are both risky in some sense. We shall call $[\nu_1, \nu_2]$ the *riskless interval.*

Let $X_1 := \max\{\nu_1 - X, 0\}$ define the *lower risk* random variable. Obviously, as X falls below ν_1, X_1 increases above 0. Since, insufficient rain is undesirable, so are large values of X_1. Similarly, we define $X_2 := \max\{X - \nu_2, 0\}$ as the *upper risk* random variable. Again, as X raises above ν_2, the r.v. X_2 increases above 0. Therefore, the smaller are the values of both X_1 and X_2, the better is the result for our risk sensitive farmers. Figure 7.1 illustrates this situation.

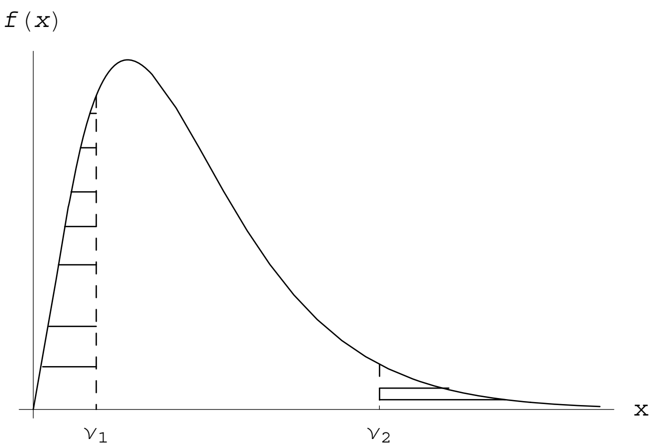

Figure 7.1. Two-sided risk.

It follows immediately that

$$P\{X \notin [\nu_1, \nu_2]\} = P\{X < \nu_1\} + P\{X > \nu_2\} = P\{X_1 > 0\} + P\{X_2 > 0\}.$$

While the above probability of an undesirable event is an entity that we would like to be "small", in practical situations, it could well be that the event $\{X_1 > 0\}$ is relatively more, or less, undesirable than the event $\{X_2 > 0\}$ (for instance, some crops may withstand drought better than excessive moisture). To capture this, unequal, concern about the lower and upper risks we now introduce a single *two-sided risk function* (also called loss function) parameterized by $\gamma \in [0,1]$:

$$h(X,\gamma,\nu_1,\nu_2) = \gamma X_1 + (1-\gamma)X_2 = \gamma \max\{\nu_1 - X, 0\} + (1-\gamma)\max\{X - \nu_2, 0\}.$$

Here, γ captures the relative importance of lower risk versus the upper risk.

Note that, if we assume the distribution of X is $F(x)$, namely $F(x) = P(X \le x)$, then the distribution function of $h(X,\gamma,\nu_1,\nu_2)$ is:

$$H(x,\gamma,\nu_1,\nu_2) = P(h(X,\gamma,\nu_1,\nu_2) \le x) = F\left(\nu_2 + \frac{x}{1-\gamma}\right) - F\left(\nu_1 - \frac{x}{\gamma}\right).$$

Given γ, ν_1 and ν_2, it is now easy to see $h(X,\gamma,\nu_1,\nu_2)$ is a convex function in X. It is also clear that the function $h(X,\gamma,\nu_1,\nu_2)$ is nonnegative everywhere in its domain. Since here we use nonnegative numbers to describe the risk, we can now use the loss function $h(X,\gamma,\nu_1,\nu_2)$ in place of the "portfolio $f(x,y)$" in [15]. Similarly to [15] we now define *value-at-risk (VaR)* and *conditional value-at-risk (CVaR)* based on this two-sided risk function as follows.

For a given distribution or given data sample of X and the confidence level α, we can obtain the VaR ($\zeta_\alpha(X,\gamma,\nu_1,\nu_2)$) and CVaR ($\phi_\alpha(X,\gamma,\nu_1,\nu_2)$) of $h(X,\gamma,\nu_1,\nu_2)$ as follows:

$$\phi_\alpha(X,\gamma,\nu_1,\nu_2) = \min_\zeta \; F_\alpha(\zeta, X,\gamma,\nu_1,\nu_2),$$

where $F_\alpha(\zeta, X,\gamma,\nu_1,\nu_2) = \zeta + \frac{1}{1-\alpha}E[h(X,\gamma,\nu_1,\nu_2) - \zeta]^+$, and

$$\zeta_\alpha(X,\gamma,\nu_1,\nu_2) \in \mathrm{argmin}_\zeta F_\alpha(\zeta, X,\gamma,\nu_1,\nu_2).$$

In fact, value-at-risk is the maximal loss the farmer will face with the confidence level $\alpha \in [0,1]$ and conditional value-at-risk is the mean loss in the $(1-\alpha)$ worst case of the two-sided risk function $h(X,\gamma,\nu_1,\nu_2)$.

Figure 7.2 below portrays the essence of these concepts in the case where the distribution function of $h(X,\gamma,\nu_1,\nu_2)$ is continuous. Note that the shaded area is on the right (rather than left) tail of the distribution because large values of our two-sided risk function are undesirable.

Based on these definitions, some limit properties of $\zeta_\alpha(X,\gamma,\nu_1,\nu_2)$ and $\phi_\alpha(X,\gamma,\nu_1,\nu_2)$ follow immediately:

$$\lim_{\alpha\to 0} \zeta_\alpha(X,\gamma,\nu_1,\nu_2) = 0, \quad \lim_{\alpha\to 0} \phi_\alpha(X,\gamma,\nu_1,\nu_2) = E[h(X,\gamma,\nu_1,\nu_2)];$$

$$\lim_{\alpha\to 1} \zeta_\alpha(X,\gamma,\nu_1,\nu_2) = \lim_{\alpha\to 1} \phi_\alpha(X,\gamma,\nu_1,\nu_2) = \sup_X h(X,\gamma,\nu_1,\nu_2).$$

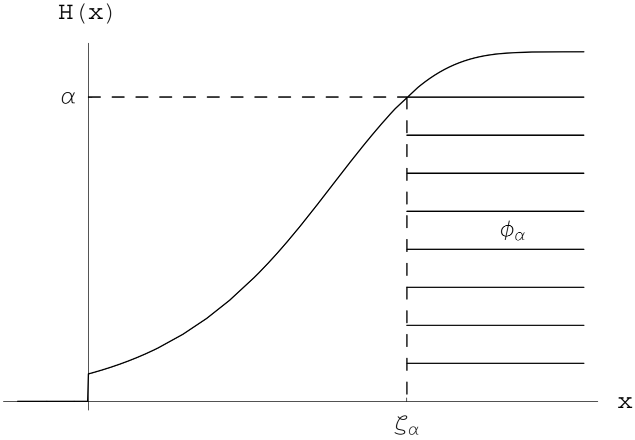

Figure 7.2. $\alpha-$VaR (ζ_α) and $\alpha-$CVaR (ϕ_α).

Two-sided risk as an optimization problem

Assume we obtain a sample of observations of X denoted by $x_1, x_2, \cdots, x_N$. After specifying γ, ν_1 and ν_2, using the method provided in [15] we can state the following mathematical programming problem.

$$\min_{\zeta} \ \zeta + \frac{1}{N(1-\alpha)} \sum_{k=1}^{N} u_k$$

subject to

$$m_k \geq 0,\ m_k - \gamma(\nu_1 - x_k) \geq 0, k = 1, \cdots, N;$$

$$n_k \geq 0,\ n_k - (1-\gamma)(x_k - \nu_2) \geq 0, k = 1, \cdots, N;$$

$$u_k \geq 0,\ u_k - m_k - n_k + \zeta \geq 0, k = 1, \cdots, N,$$

where $\zeta; u_1, u_2, \cdots, u_N; m_1, m_2, \cdots, m_N; n_1, n_2, \cdots, n_N$ are the decision variables of this optimization problem.

The optimal objective value of this mathematical program constitutes an estimate - based on the sample - of the conditional value-at-risk of the two-sided risk function $h(X, \gamma, \nu_1, \nu_2)$. Furthermore, the ζ^*_α entry of an optimal solution is an estimate of value-at-risk corresponding to this CVaR.

To explain how the above mathematical program arises, we note that after obtaining the sample $x_1, x_2, \ldots, x_N$ from the r.v. X, the sample mean $\frac{1}{N}\sum_{k=1}^{N}[h(x_k, \gamma, \nu_1, \nu_2) - \zeta]^+$ approximates the nonnegative deviation of the loss from ζ, that is, $E[h(X, \gamma, \nu_1, \nu_2) - \zeta]^+$. Hence, we can use following function to approximate the function $F_\alpha(\zeta, X, \gamma, \nu_1, \nu_2)$ defined above:

$$\tilde{F_\alpha}(\zeta, X, \gamma, \nu_1, \nu_2) = \zeta + \frac{1}{N(1-\alpha)} \sum_{k=1}^{N} [h(x_k, \gamma, \nu_1, \nu_2) - \zeta]^+ =$$

$$\zeta + \frac{1}{N(1-\alpha)} \sum_{k=1}^{N} [\gamma \max\{\nu_1 - x_k, 0\} + (1-\gamma) \max\{x_k - \nu_2, 0\} - \zeta]^+.$$

Thus, instead of minimizing $F_\alpha(\zeta, X, \gamma, \nu_1, \nu_2)$, we try to minimize $\tilde{F_\alpha}(\zeta, X, \gamma, \nu_1, \nu_2)$:

$$\min_{\zeta} \ \zeta + \frac{1}{N(1-\alpha)} \sum_{k=1}^{N} [\gamma \max\{\nu_1 - x_k, 0\} + (1-\gamma) \max\{x_k - \nu_2, 0\} - \zeta]^+.$$

In terms of auxiliary real variables u_k, m_k and n_k, for $k = 1, \cdots, N$, after setting $u_k = [\gamma \max\{\nu_1 - x_k, 0\} + (1-\gamma) \max\{x_k - \nu_2, 0\} - \zeta]^+$, $m_k = \gamma \max\{\nu_1 - x_k, 0\}$ and $n_k = (1-\gamma) \max\{x_k - \nu_2, 0\}$, the preceding is equivalent to minimizing the linear expression

$$\min_{\zeta} \ \zeta + \frac{1}{N(1-\alpha)} \sum_{k=1}^{N} u_k$$

subject to the linear constrains as follows:

$$m_k \geq 0, \ m_k - \gamma(\nu_1 - x_k) \geq 0, k = 1, \cdots, N;$$

$$n_k \geq 0, \ n_k - (1-\gamma)(x_k - \nu_2) \geq 0, k = 1, \cdots, N;$$

$$u_k \geq 0, \ u_k - m_k - n_k + \zeta \geq 0, k = 1, \cdots, N.$$

Note that the above linear constraints can be obtained from properties of the function $[x]^+ = \max\{x, 0\}$.

Numerical examples

In the first example we generated 1000 observations with log normal distribution $N(1, 2)$, namely, $\log X \sim N(1, 2)$, and set $\nu_1 = 15, \nu_2 = 50, \gamma = 0.5$. We obtained $\zeta_{0.95} = 62.93$ and $\zeta_{0.99} = 225.50$. If we choose $\nu_1 = 0.3, \nu_2 = 250$, we will know that $\zeta_{0.95} = 0.28$.

For instance, this means that with probability 0.95, the two-sided risk associated with our hypothetical rainfall r.v. X satisfies

$$P\{h(X, .5, 15, 50) \leq 62.93)\} = P\{.5X_1 + .5X_2 \leq 62.93)\} \geq 0.95.$$

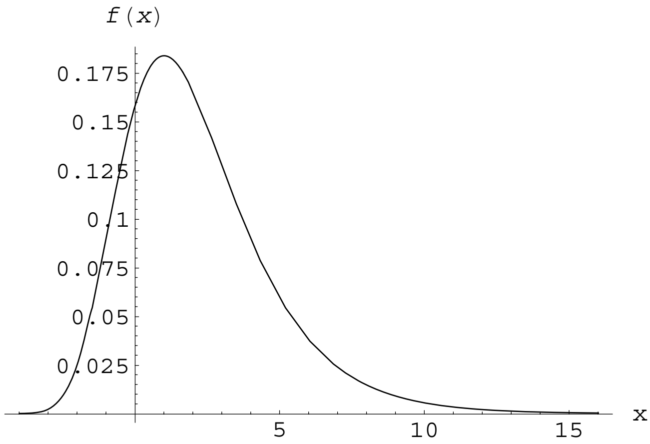

Figure 7.3. Density function of Extreme(1,2).

In the next example we will see how the parameter γ influences this problem. Here we use the asymmetric extreme distribution: $X \sim$ Extreme(1,2), see Figure 7.3.

Again, we obtained $N = 1,000$ observations for which $\max x_i = 16.33$, $\min x_i = -3.25$, range $= 19.59$. We chose $\alpha = 0.9$, then the following results were obtained from the optimization problem above:

$$\zeta_\alpha(X, 0.7, 10, 15) = 7.48,\ \zeta_\alpha(X, 0.3, 10, 15) = 3.21;$$

$$\zeta_\alpha(X, 0.7, 1, 15) = 1.18,\ \zeta_\alpha(X, 0.3, 1, 15) = 0.52.$$

Take $\zeta_\alpha(X, 0.7, 10, 15) = 7.48$ as an example to explain the meaning of value-at-risk here. As before, the following inequality holds:

$$P(h(X, 0.7, 10, 15) \leq 7.48) = P\{.7X_1 + .3X_2 \leq 7.48)\} \geq 0.9.$$

This means that our, unequally weighted, two-sided risk of missing the rainfall interval $[10, 15]$ is less than 7.48 with probability 0.9. Similarly,

$$P(h(X, 0.3, 10, 15) \leq 3.21) = P\{.3X_1 + .7X_2 \leq 3.21)\} \geq 0.9.$$

We can see that, in this instance, the VaR drops sharply as we place less weight on the lower risk X_1.

This shows that the weight γ plays an important role. However, its influence is interconnected with the size and location of the interval $[\nu_1, \nu_2]$ in the domain of the density function of this asymmetric distribution. From Figure 7.4 we see that VaR can both decrease or increase as γ

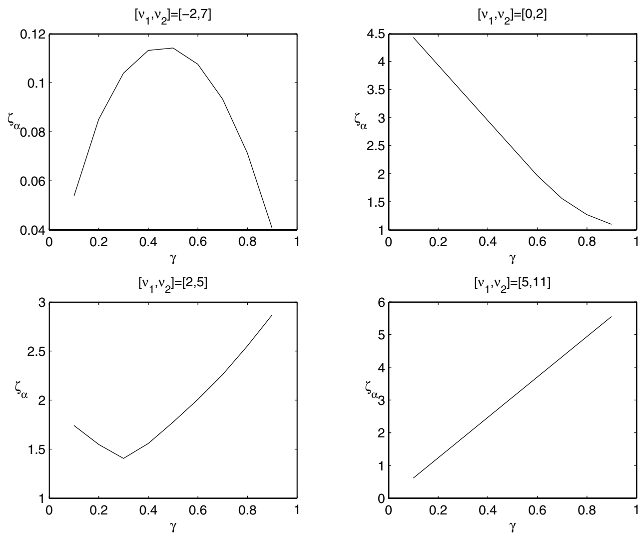

Figure 7.4. Relationship between ζ_α for fixed α and γ for different intervals.

increases, depending on the exact specification of the riskless interval. For instance, as a function of γ, VaR could be concave, convex, linear, or nonlinear.

Of course, with γ and the riskless interval held fixed, VaR and CVaR exhibit the usual dependence on the percentile parameter α. For instance, in the above, with fixed $\gamma = 0.3, \nu_1 = 10, \nu_2 = 15$, we observe the relationship between ζ_α, ϕ_α and α that is displayed in Figure 7.5.

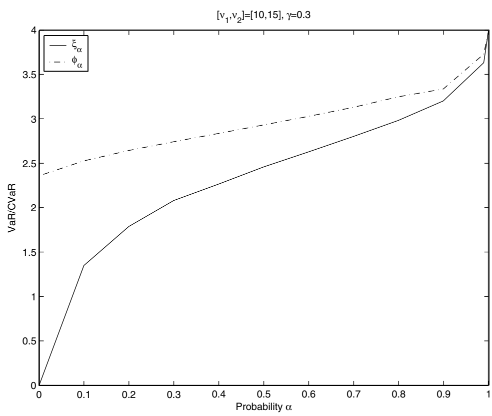

Figure 7.5. Relationship between ζ_α, ϕ_α and α.

2. Two types of risk

Recall that, X_1 was defined as the part of X that is lower than ν_1, and X_2 as the part of X that exceeds ν_2. In the environmental problems that motivated the preceding section it was natural to aim to minimize X_1 and X_2 simultaneously. Hence, a convex combination of X_1 and X_2 was a good choice for that purpose.

However, there are some applications (e.g., the standard financial return) where we have a different requirement with respect to X_1 and X_2. For example, we may want X_1 (a loss below ν_1 threshold) to be small and X_2 (a gain above ν_2 threshold) to be large. In such a case, we require a different analysis of the two tails of the underlying probability distribution. In what follows, we discuss this problem in the special case where $\nu_1 = \nu_2 = 0$. The analysis in the general, $\nu_1 \neq \nu_2$, case can be performed in an analogous manner.

Thus, as before, we begin by considering $X_1 = \max\{-X, 0\}$ and $X_2 = \max\{X, 0\}$, where X_1, X_2 are negative part and positive part of X respectively. In a typical financial market, X_1, X_2 will, respectively, represent the loss and the gain resulting from an investment. However, in this case, we clearly want X_1 to be small whereas X_2 to be large.

Unlike the discussion in Section 1, it will be convenient to deal with the above as two separate, yet interrelated, aspects of the underlying portfolio optimization problem. The essential observation is that in many (most?) situations, investments that increase a probability of a large gain may also increase a probability of large loss. In this sense, the problem is reminiscent of the classical problem of Type I and II errors in Statistics.

Type I and Type II risk

Following the above motivation we define the risk associated with large values of X_1 as *Type I risk*, and the risk associated with small values of X_2 as *Type II risk*.

We note that the above formulation of Type I risk is similar to already standard concepts (e.g., see [16]). In particular, we now briefly recall definitions of VaR and CVaR on X_1. More detailed discussion together with some financial applications will be given in Section 3.

Mathematically, we treat above random variables (r.v.'s) as functions $X : \Omega \to \mathbb{R}$ that belong to the linear space $\mathcal{L}^2 = \mathcal{L}^2(\Omega, \mathcal{F}, P)$, that is, (measurable) functions for which the mean and variance exist.

We denote by $\Psi_1(\cdot)$ on $\mathcal{R}$ the distribution function of X_1 as follows:

$$\Psi_1(\zeta) = P\{X_1 \leq \zeta\}.$$

DEFINITION 7.1 *The value-at-risk (VaR) of the loss X_1 associated with a confidence level α is the functional $\zeta_\alpha : \mathcal{L}^2 \to (-\infty, \infty)$:*

$$\zeta_\alpha(X) := \inf\{\zeta | P\{X_1 \leq \zeta\} \geq \alpha\} = \inf\{\zeta | \Psi_1(\zeta) \geq \alpha\},$$

which shows the maximal loss the investor will face with the confidence level α. That is, $\zeta_\alpha(X)$ is the maximal amount of loss that will be incurred with probability at least α. However, with probability $1-\alpha$, the loss will be greater than $\zeta_\alpha(X)$, so we will define:

DEFINITION 7.2 *Conditional value-at-risk (CVaR) is the functional $\phi_\alpha : \mathcal{L}^2 \to (-\infty, \infty)$:*

$$\phi_\alpha(X) = \textit{mean of the } \alpha - \textit{tail distribution of } X_1,$$

where the distribution in question is the one with distribution function $\Psi_{1,\alpha}(\zeta)$ defined by

$$\Psi_{1,\alpha}(\zeta) = \begin{cases} 0, & \zeta < \zeta_\alpha(X), \\ (\Psi_1(\zeta) - \alpha)/(1-\alpha), & \zeta \geq \zeta_\alpha(X). \end{cases}$$

Since $X_1 = \max\{-X, 0\}$ is a convex function of X, $\phi_\alpha(X)$ defined above is a convex function of X as well.

Similarly, let us denote by $\Psi_2(\cdot)$ on $\mathcal{R}$ the distribution function of X_2 as follows:

$$\Psi_2(\xi) = P\{X_2 \leq \xi\}.$$

DEFINITION 7.3 *The value-of-gain (VoG) of the gain X_2 associated with a assurance level β is the functional $\xi_\beta : \mathcal{L}^2 \to (-\infty, \infty)$:*

$$\xi_\beta(X) = \sup\{\xi | P\{X_2 > \xi\} \geq \beta\} = \sup\{\xi | 1 - \Psi_2(\xi) \geq \beta\},$$

which shows the minimum gain the investor can achieve with a specified assurance level β. That is, $\xi_\beta(X)$ is the minimal amount of gain that will be incurred with probability at least β. However, with probability $1-\beta$, the gain will be less than $\xi_\beta(X)$, so we will define:

DEFINITION 7.4 *Conditional value-of-gain (CVoG) is the functional $\psi_\beta : \mathcal{L}^2 \to (-\infty, \infty)$:*

$$\psi_\beta(X) = \textit{mean of the } \beta - \textit{left tail distribution of } X_2,$$

where the distribution in question is the one with distribution function $\Psi_{2,\beta}(\cdot)$ defined by

$$\Psi_{2,\beta}(x) = \begin{cases} \Psi(x)/(1-\beta), & x \leq \xi_\beta(X), \\ 1, & x > \xi_\beta(X). \end{cases}$$

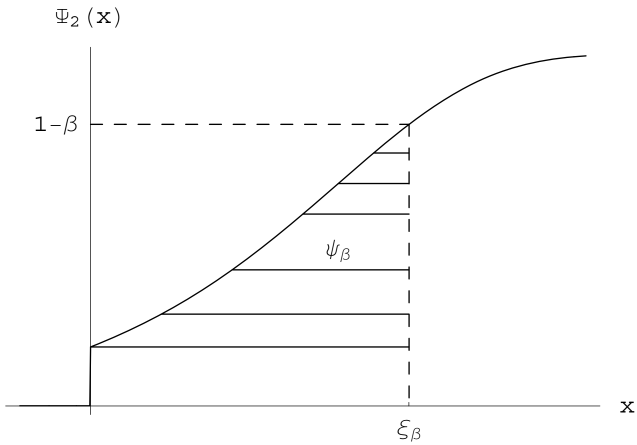

Figure 7.6. $\beta-$VoG (ξ_β) and $\beta-$CVoG (ψ_β).

Figure 7.6 portrays the essence of these concepts and you could see the differences between CVoG and CVaR.

One of the important properties of conditional value-of-gain, $\psi_\beta(X)$, concave in X, will be proved in Section 3 together with some financial applications.

Two problems and properties of parameters

Investors who want to minimize Type I risk will try to minimize CVaR, and those who want to minimize Type II risk will try to maximize CVoG. However, one of these types of risk will tend to stay high when the other one is minimized. In addition, some parties may want to minimize a combination of Type I and Type II risks. Our discussion below indicates one reasonable approach to these important and difficult problems.

Basically, we are assuming that by choosing a "portfolio" (defined formally in the next section) an investor can select the r.v. X from a family of r.v.'s with known probability distributions. Hence, an "optimal portfolio" may involve solving the following two problems:

Problem I

$$\min_X \phi_\alpha(X) \ \text{(minimize the CVaR loss)}$$

Subject to:

$$\psi_\beta(X) \geq \tau. \ \text{(guarantee a CVoG gain level of } \tau\text{)}$$

Since Conditional Value of Gain, $\psi_\beta(X)$ is a concave function of X, for any real number τ the set $\{X : \psi_\beta(X) \geq \tau\}$ is a convex set. Hence, the above optimization problem is a convex problem, that is, in principle, suitable for fast numerical solution.

Problem II

$$\max_X \ \psi_\beta(X), \ \text{(maximize the CVoG gain)}$$

Subject to:

$$\phi_\alpha(X) \leq v. \ \text{(tolerable CVaR risk level } v\text{)}$$

Since $-\psi_\beta(X)$ is a convex function of the decision variable x and the set $\{X : \phi_\alpha(X) \leq v\}$ is a convex set, above problem is also a convex problem.

In above problems, we have four parameters in total. Those are *confidence level* α, *assurance level* β, *Type I risk tolerance* v and *gain target* τ. Selection of values of these parameters constitutes a characterization of the investor's attitudes towards the "loss versus gain dilemma". However, an intelligent investor will want to select these values on the basis of their interrelationship that are, ultimately, influenced by the probability distribution function of the asset X. The analysis below, should enable such an investor to make an informed decision.

Firstly, we assume that we have chosen and fixed α and β, and in this case, we want to choose τ and v so that above two problems are meaningful and interesting. We shall require following notations:

$$\tau^*(\beta) := \max_X \psi_\beta(X), \quad v_*(\alpha) := \min_X \phi_\alpha(X),$$

and we shall denote the optimal objective function value of Problems I and II by

$$Z_1(\alpha, \beta, \tau), \quad Z_2(\alpha, \beta, v),$$

respectively. Select $X_*(\alpha) \in \text{argmin}\phi_\alpha(X)$ and let $\psi_\beta(X_*(\alpha)) = \tau_*(\beta, \alpha)$, then we have following lemma.

Ideally, an investor would want a portfolio that is an optimal solution to both Problems I and II. However, in order to achieve this, some adjustments to the target τ (respectively, tolerance level v) maybe needed.

LEMMA 7.5 (1) *If* $(\alpha, \beta) \in \{(\alpha, \beta) | \tau_*(\beta, \alpha)\} = \tau^*(\beta)\}$, *then we can choose* $\tau \leq \tau^*(\beta)$, *and for any such* $\tau, Z_1(\alpha, \beta, \tau) = v_*(\alpha)$.

(2) *If* $(\alpha, \beta) \in \{(\alpha, \beta) | \tau_*(\beta, \alpha) < \tau^*(\beta)\}$, *then a choice of* $\tau \in (\tau_*(\beta, \alpha), \tau^*(\beta)]$ *yields* $Z_1(\alpha, \beta, \tau) > v_*(\alpha)$.

The first case in the Lemma corresponds to the ideal situation since we obtain the maximum gain while at the same time we minimize our risk of loss. However, when the second case occurs, namely the strict inequality holds, perhaps, the best we can do is to choose our gain target level τ in that kind of interval. Of course, we will face a greater risk of loss when we do this.

For Problem II, we can obtain similar conditions for choosing the risk tolerance v. As before, we define the notation: $X^*(\beta) \in \mathrm{argmax}\psi_\beta(X)$, and let $v^*(\alpha,\beta) = \phi_\alpha(X^*(\beta))$, then the following lemma follows immediately:

LEMMA 7.6 (1) *If* $(\alpha,\beta) \in \{(\alpha,\beta)|v_*(\alpha) = v^*(\alpha,\beta)\}$, *then we can choose* $v \geq v_*(\alpha)$, *and for any such* $v, Z_2(\alpha,\beta,v) = \tau^*(\beta)$.

(2) *If* $(\alpha,\beta) \in \{(\alpha,\beta)|v_*(\alpha) < v^*(\alpha,\beta)\}$, *then a choice of* $v \in [v_*(\alpha), v^*(\alpha,\beta))$ *yields* $Z_2(\alpha,\beta,v) < \tau^*(\beta)$.

The first case of this lemma corresponds to the ideal situation where we attain minimum risk of loss while maximizing our gain. But, when the second case happens, that means the risk of loss has not been minimized. We can improve it while trying to maximize our gain but, of course, we will sacrifice part of the gain.

In fact, combining the analysis of Problem I with that of II, we obtain the following equation:

$$\{(\alpha,\beta)|\tau_*(\beta,\alpha) = \tau^*(\beta)\} =$$

$$\{(\alpha,\beta)|\,\{\mathrm{argmin}_X\phi_\alpha(X)\}\bigcap\{\mathrm{argmax}_X\psi_\beta(X)\} \neq \emptyset\} =$$

$$\{(\alpha,\beta)|v_*(\alpha) = v^*(\alpha,\beta)\}.$$

What we are interested in now is the set

$$\{(\alpha,\beta)|\,\{\mathrm{argmin}_X\phi_\alpha(X)\}\bigcap\{\mathrm{argmax}_X\psi_\beta(X)\} \neq \emptyset\}.$$

From an investor's point of view, the larger this set is, the better. We experimented with different distributions of X and obtained a range of results. For a symmetric distribution, e.g. normal distribution, it is easy to find parameters μ, σ that permit the above set to be large. However, for asymmetric distributions, it is harder to do so.

In fact, generally speaking, if we denote the distribution function of X by $F(x)$, then the distributions of X_1, X_2 are $(1 - F(-x))I_{x\geq 0}$ and $F(x)I_{x\geq 0}$, respectively. We observe their typical graphs in Figure 7.7.

The next lemma shows that the non-overlapping feature in the left panel of Figure 7.7 always occurs in the case of a symmetric underlying distribution of X.

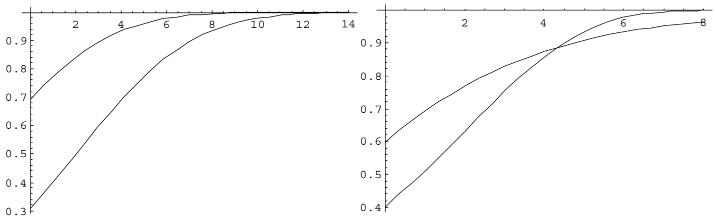

Figure 7.7. Distribution function of the loss and gain, when X is a normal distribution on the left and an extreme distribution on the right.

LEMMA 7.7 *(symmetric property) Assume X is a symmetric random variable with the distribution function $F(x)$. By symmetric with respect to μ, $F(\mu + x) + F(\mu - x) = 1$. Then if $\mu \neq 0$, $F(x) \neq 1 - F(-x)$, in fact, $F(x) < 1 - F(-x)$, when $\mu > 0$; $F(x) > 1 - F(-x)$, when $\mu < 0$ and $F(x) = 1 - F(-x)$, when $\mu = 0$.*

However, after calculating some examples, we found that for asymmetric distributions it is hard to find α, β such that $\{\text{argmin}_X \phi_\alpha(X)\} \cap \{\text{argmax}_X \psi_\beta(X)\} \neq \emptyset$.

Remark: We note that definitions of optimality for Problems I and II could be generalised to "ε-optimality", $\varepsilon > 0$ and (typically) very small. This is because, in practice, investors would be satisfied with portfolios that are only slightly sub-optimal. All of the previous analysis generalizes to this situation in a natural way. For details we refer the reader to Boda's thesis [8].

3. Financial interpretation

In this section, we explicitly apply the analysis of Type I and Type II risk to the portfolio optimization problem and interpret the results. Therefore we concentrate on financial analysis and related optimization algorithms.

Loss function, gain function and Type I risk

We now apply the general concepts of two types of risk to a specific portfolio optimization problem and derive methods to optimize and balance Type I and Type II risks.

Let vector $Y = (Y_1, \cdots, Y_m)^T$ be the random return on m stocks. We define a *portfolio* to be an m-vector $x = (x_1, \cdots, x_m)^T$ such that $x^T e = 1, x \geq 0$. We also define the random *loss function* and a *gain* *function*, induced by the portfolio $x = (x_1, \cdots, x_m)^T$ as follows:

$$l(x, Y) = \max\{-x^T Y, 0\}, \ g(x, Y) = \max\{x^T Y, 0\}.$$

Note that we are not assuming that the distribution of Y_j's is symmetric. For a portfolio, we believe it is a *loss* if it is negative, otherwise it is a *gain.* For the loss function $l(x, Y)$, we define *value-at-risk (VaR)* similarly to Definition 7.1 or, equivalently, [15].

Namely, for each x, we denote by $L(x, \cdot)$ on $\mathcal{R}$ the distribution function of $l(x, Y)$ as follows:

$$L(x, \zeta) = P_Y\{l(x, Y) \leq \zeta\}.$$

Next, choose and fix a confidence level $\alpha \in [0, 1]$. The α-VaR of the loss associated with a portfolio x, and the loss function $l(x, Y)$ is the value:

$$\zeta_\alpha(x) = \min\{\zeta | L(x, \zeta) \geq \alpha\},$$

which shows the maximal loss the investor will face with the confidence level α.

Further, we recall that the conditional value-at-risk (CVaR) was defined as:

$$\phi_\alpha(x) = \text{mean of the } \alpha - \text{tail distribution of } Z = l(x, Y),$$

where the distribution in question is the one with distribution function $L_\alpha(x, \cdot)$ defined by

$$L_\alpha(x, \zeta) = \begin{cases} 0, & \zeta < \zeta_\alpha(x), \\ (L(x, \zeta) - \alpha)/(1 - \alpha), & \zeta \geq \zeta_\alpha(x). \end{cases}$$

Analogously to the analysis in [15] we use $l(x, Y) = \max\{-x^T Y, 0\}$ in place of $f(x, Y)$ to define VaR and CVaR. Note that $l(x, Y)$ is convex with x, so if we let

$$F_\alpha(x, \zeta) = \zeta + \frac{1}{1 - \alpha} E\{[l(x, Y) - \zeta]^+\},$$

then following conclusions will hold, by the same arguments as those given in [15].

THEOREM 7.8 *As a function of $\zeta \in R$, $F_\alpha(x, \zeta)$ is finite and convex (hence continuous), with*

$$\phi_\alpha(x) = \min_\zeta F_\alpha(x, \zeta),$$

and moreover,

$$\zeta_\alpha(x) \in argmin_\zeta F_\alpha(x, \zeta).$$

In particular, one always has:

$$\zeta_\alpha(x) \in argmin_\zeta F_\alpha(x, \zeta),\ \phi_\alpha(x) = F_\alpha(x, \zeta_\alpha(x)),$$

COROLLARY 7.9 *The conditional value-at-risk, $\phi_\alpha(x)$, is convex with respect to x. Indeed, in this case $F_\alpha(x, \zeta)$ is jointly convex in (x, ζ).*

THEOREM 7.10 *Minimizing $\phi_a(x)$ with respect to $x \in X$ is equivalent to minimizing $F_\alpha(x, \zeta)$ over all $(x, \zeta) \in X \times R$, in the sense that*

$$\min_{x \in X} \phi_\alpha(x) = \min_{(x,\zeta) \in X \times R} F_\alpha(x, \zeta)$$

where moreover,

$$(x^*, \zeta^*) \in argmin_{(x,\zeta) \in X \times R} F_\alpha(x, \zeta) \Longleftrightarrow$$

$$x^* \in argmin_{x \in X} \phi_\alpha(x),\ \zeta^* \in argmin_{\zeta \in R} F_\alpha(x^*, \zeta).$$

One kind of approximation of $F_\alpha(x, \zeta)$ obtained by sampling the probability distribution of Y. So a sample set $y_1, \cdots, y_N$ of observations of Y yields the approximation function

$$\tilde{F}_\alpha(x, \zeta) = \zeta + \frac{1}{N(1-\alpha)} \sum_{k=1}^{N} \max\{\max\{-x^T y_k, 0\} - \zeta, 0\}.$$

Because here $l(x, y_k) = \max\{-x^T y_k, 0\}$ is a non-smooth function of x, the formulation of the problem $\min_{(x,\zeta)} \tilde{F}_\alpha(x, \zeta)$ in [15] should be changed to the following linear programming problem:

$$\min\ \zeta + \frac{1}{N(1-\alpha)} \sum_{k=1}^{N} u_k$$

Subject to:

$$x \geq 0, x^T \mathbf{e} = 1;$$

$$l_k \geq 0, l_k + x^T y_k \geq 0, k = 1, \cdots, N;$$

$$u_k \geq 0, l_k - \zeta - u_k \leq 0, k = 1, \cdots, N.$$

Similarly to arguments in [14], VaR and CVaR corresponding to $l(x, Y)$ can be approximated by the optimizer and the optimal objective function value of the above linear programming problem. Of course, the quality of this approximation increases with the sample size N.

Type II risk

Whereas the preceding analysis of Type I risk was completely analogous to that in [14], when considering properties of Type II risk a few, natural, adjustments need to be made when considering the problem of minimizing the risk associated with the new gain function $g(x, Y) = \max\{x^T Y, 0\}$ failing to take sufficiently large values.

Concept of conditional value-of-gain. For each x, the distribution function of $g(x, Y)$ is defined by: $G(x, \xi) = P_Y\{g(x, Y) \leq \xi\}$. Choose and fix $\beta \in [0, 1]$, the investor's *assurance level.*

Definition 7.11 *The value-of-gain (VoG) associated with a portfolio x and $g(x, Y)$ is the value:*

$$\xi_\beta(x) = \sup\{\xi | P_Y\{g(x, Y) > \xi\} \geq \beta\} = \sup\{\xi | 1 - G(x, \xi) \geq \beta\},$$

which shows the minimum gain the investor can achieve with a specified assurance level β.

However, with probability $1 - \beta$, the gain will be less than $\xi_\beta(x)$, so the following definition is now natural.

Definition 7.12 *Conditional value-of-gain (CVoG):*

$$\psi_\beta(x) = \textit{mean of the } \beta - \textit{left tail distribution of } Z = g(x, Y),$$

where the distribution in question is the one with distribution function $G_\beta(x, \cdot)$ defined by

$$G_\beta(x, \xi) = \begin{cases} G(x, \xi)/(1 - \beta), & \xi \leq \xi_\beta(x), \\ 1, & \xi > \xi_\beta(x). \end{cases}$$

The fact that the distribution function of $Z = g(x, Y)$ need not be continuous necessitates the following two additional definitions.

Definition 7.13 *The $\beta-$ CVoG$^+$ ("upper" $\beta-$ CVoG) of the gain associated with a decision x is the value:*

$$\psi_\beta^+(x) = E\{g(x, Y) | g(x, Y) \leq \xi_\beta(x)\},$$

whereas the $\beta-$ CVoG$^-$ ("lower" $\beta-$ CVoG) of the gain is the value:

$$\psi_\beta^-(x) = E\{g(x, Y) | g(x, Y) < \xi_\beta(x)\}.$$

It is important to differentiate between the cases where the upper and lower conditional values-of-gain coincide, or differ. This is done in the following proposition that is proved in Section 4.

PROPOSITION 7.14 *(Basic CVoG relations). If there is no probability atom at $\xi_\beta(x)$, one simply has:*

$$\psi_\beta^-(x) = \psi_\beta(x) = \psi_\beta^+(x).$$

If a probability atom does exist at $\xi_\beta(x)$, one has:

$$\psi_\beta^-(x) < \psi_\beta(x) = \psi_\beta^+(x), \textit{when}\ G(x, \xi_\beta(x)) = 1 - \beta,$$

or on the other hand,

$$\psi_\beta(x) = \psi_\beta^+(x), \textit{when}\ G(x, \xi_\beta(x)) = 0,$$

(with $\psi_\beta^-(x)$ then being ill defined). But in all the remaining cases, we have

$$0 < G(x, \xi_\beta(x)) < 1 - \beta,$$

and one has the strict inequality

$$\psi_\beta^-(x) < \psi_\beta(x) < \psi_\beta^+(x).$$

The next proposition (also proved in Section 4) shows that $\psi_\beta(x) =$ mean of the $\beta-$left tail distribution of $Z = g(x, Y)$ can be expressed as convex combination of value-of-gain and the upper conditional value-of-gain.

PROPOSITION 7.15 *(CVoG as a weighted average). Let $\lambda_\beta(x)$ be the probability assigned to the gain amount $z = \xi_\beta(x)$ by the $\beta-$ left tail distribution, namely*

$$\lambda_\beta(x) = G(x, \xi_\beta(x))/(1 - \beta) \in [0, 1].$$

If $G(x, \xi_\beta(x)) > 0$, so there is a positive probability of a gain less than $\xi_\beta(x)$, then

$$\psi_\beta(x) = \lambda_\beta(x)\psi_\beta^+(x) + [1 - \lambda_\beta(x)]\xi_\beta(x),$$

with $\lambda_\beta(x) < 1$. However, if $G(x, \xi_\beta(x)) = 0$, $\xi_\beta(x)$ is the lowest gain that can occur (and thus $\lambda_\beta(x) = 0$ but $\psi_\beta^-(x)$ is ill defined), then

$$\psi_\beta(x) = \xi_\beta(x).$$

For a gain function in finance, following [15], we can easily derive some useful properties of CVoG as a measure of risk with significant advantages over VoG. For a discrete distribution and a “scenario case”, the following results will illustrate a method to estimate VoG and CVoG from historical data.

PROPOSITION 7.16 *(CVoG for scenario models). Suppose the probability measure P is concentrated on finitely many points y_k of Y, so that for each $x \in X$ the distribution of the gain $Z = g(x, Y)$ is likewise concentrated on finitely many points, and $G(x, \cdot)$ is a step function with jumps at those points. Fixing x, let those corresponding gain points $z_k := g(x, y_k)$ be ordered as $z_1 < z_2 < \cdots < z_N$, with the probability of z_k being $p_k > 0$. For any fixed assurance level $\beta \in [0, 1]$, let k_β be the unique index such that*

$$\sum_{k=1}^{k_\beta} p_k \leq 1 - \beta < \sum_{k=1}^{k_\beta + 1} p_k.$$

The $\beta-$ VoG of the gain is given by

$$\xi_\beta(x) = z_{k_\beta},$$

whereas the $\beta-$CVoG of the gain is given by

$$\psi_\beta(x) = \frac{1}{1-\beta}\left[\sum_{k=1}^{k_\beta} p_k z_k + \left(1 - \beta - \sum_{k=1}^{k_\beta} p_k\right) z_{k_\beta}\right].$$

Furthermore, in this situation the weight from Proposition 7.15 is given by

$$\lambda_\beta(x) = \frac{1}{1-\beta}\sum_{k=1}^{k_\beta} p_k.$$

COROLLARY 7.17 *(Lowest gain). In the notation of Proposition 7.16, if z_1 is the lowest point with probability $p_1 > 1 - \beta$, then $\psi_\beta(x) = \xi_\beta(x) = z_1$.*

Maximization rule and coherence. We can define the function

$$H_\beta(x, \xi) := \xi - \frac{1}{1-\beta} E\{[g(x, Y) - \xi]^-\}$$

that enables us to determine CVoG and VoG from solutions of an appropriate optimization problem formulated in the next theorem. The proof (given in Section 4) is inspired by the line of argument used to prove somewhat similar results in [14, 15].

THEOREM 7.18 *As a function of* $\xi \in R$, $H_\beta(x,\xi)$ *is finite and concave (hence continuous), with*

$$\psi_\beta(x) = \max_\xi H_\beta(x,\xi),$$

and moreover,

$$\xi_\beta(x) \in argmax_\xi H_\beta(x,\xi).$$

In particular, one always has:

$$\psi_\beta(x) = H_\beta(x, \xi_\beta(x)).$$

COROLLARY 7.19 *(Concavity of CVoG) If* $g(x,y)$ *is convex with respect to* x*, then* $\psi_\beta(x)$ *is concave with respect to* x*. Indeed, in this case* $H_\beta(x,\xi)$ *is jointly concave in* (x,ξ)*.*

THEOREM 7.20 *Maximizing* $\psi_\beta(x)$ *with respect to* $x \in X$ *is equivalent to maximizing* $H_\beta(x,\xi)$ *over all* $(x,\xi) \in X \times R$*, in the sense that*

$$\max_{x\in X} \psi_\beta(x) = \max_{(x,\xi)\in X\times R} H_\beta(x,\xi).$$

Moreover,

$$(x^*,\xi^*) \in argmax_{(x,\xi)\in X\times R} H_\beta(x,\xi) \Longleftrightarrow$$
$$x^* \in argmax_{x\in X}\psi_\beta(x),\ \zeta^* \in argmax_{\zeta\in R} H_\beta(x^*,\xi).$$

It is also possible, and interesting, to optimize an arbitrary portfolio performance function subject to a number of gain-assurance level constraints. This is summarized in the next theorem.

THEOREM 7.21 *(Gain-shaping with CVoG) Let* g *be any objective function chosen on* X*. For any selection of assurance levels* β_i *and corresponding target levels* $\tau_i, i = 1,\cdots,l$*, the problem*

$$minimize\ \ g(x)\ \ over\ \ x \in X\ \ satisfying\ \ \psi_{\beta_i}(x) \geq \tau_i\ \ for\ \ i = 1,\cdots,l,$$

is equivalent to the problem

$$minimize\ \ g(x)\ \ over\ \ (x,\xi_1,\cdots,\xi_l) \in X \times \mathcal{R} \times \cdots \times \mathcal{R}$$
$$satisfying\ \ H_{\beta_i}(x,\xi_i) \geq \tau_i,\ \ for\ \ i = 1,\cdots,l.$$

Indeed, $(x^*, \xi_1^*, \cdots, \xi_l^*)$ *solves the second problem if and only if* x^* *solves the first problem and the inequalities* $H_{\beta_i}(x^*, \xi_i^*) \geq \tau_i$, *hold for* $i = 1, \cdots, l$.

Moreover one then has $\psi_{\beta_i}(x^*) \geq \tau_i$ *for every* i, *and actually* $\psi_{\beta_i}(x^*) = \tau_i$ *for each* i *such that* $H_{\beta_i}(x^*, \xi_i^*) = \tau_i$ *(i.e., those that correspond to active CVoG constraints).*

Based on Theorem 7.18, we can construct the following approximating algorithm for maximizing the CVoG. We assume that Y_k's are i.i.d distributed according to $p(y)$, and a sample of observations from $p(y)$ is denoted by $y_1, y_2, \cdots, y_N$. Maximizing CVoG can then be approximated by the following mathematical programming problem:

$$\max\ \xi - \frac{1}{N(1-\beta)} \sum_{k=1}^{N} u_k$$

subject to

$$x \geq 0,\ x^T \mathbf{e} = 1;$$

$$u_k \geq 0,\ g(x, y_k) - \xi + u_k \geq 0,\ k = 1, \cdots, N.$$

However, since $g(x, y_k) = \max\{x^T y_k, 0\}$ is not a smooth function of x, the above mathematical programming problem is not in a directly tractable form. Hence, we introduce $0 - 1$ integer variables for each sample point to change the above mathematical programming problem to one that can be solved using mixed integer programming method as follows:

$$\max\ \xi - \frac{1}{N(1-\beta)} \sum_{k=1}^{N} u_k$$

subject to

$$x \geq 0,\ x^T \mathbf{e} = 1;$$

$$n_k \in \{0, 1\}, 0 \leq g_k \leq n_k M, 0 \leq l_k \leq (1 - n_k) M,\ k = 1, \cdots, N;$$

$$g_k - l_k = x^T y,\ k = 1, \cdots, N;$$

$$u_k \geq 0,\ g_k - \xi + u_k \geq 0,\ k = 1, \cdots, N,$$

where, M is a suitably chosen large number. Of course, when the sample size N is large, this means that the computational effort required to solve the above optimality can be prohibitively difficult. Nonetheless, there are now many heuristics for solving large scale integer programming problems.

Table 7.1. Portfolio mean return m

Instrument	Mean return
S&P	0.0101110
Gov.bond	0.0043532
Small cap	0.0137058

Table 7.2. Portfolio variance-covariance matrix V

	S&P	Gov.bond	Small cap
S&P	0.00324625	0.00022983	0.00420395
Gov.bond	0.00022983	0.00049937	0.00019247
Small cap	0.00420395	0.00019247	0.00764097

Examples. Here we use the data from [14] to calculate VaR, CVaR and VoG, CVoG using our new loss and gain function. The data are as follows:

We assume the return of three stocks satisfy a multivariate normal distribution $N(m, V)$. The mean vector and variance-covariance matrix are shown in Table 7.1 and Table 7.2 respectively. We can use these parameters to generate samples that satisfy multivariate normal distribution and then use the samples and the above constructions and mathematical programs to calculate VaR, VoG and optimize CVaR and CVoG.

The calculated VaR, optimized CVaR and the optimal portfolio corresponding to different confidence levels α and the sample size of $10,000$ are shown in Table 7.3. The calculated VoG, optimized CVoG and the optimal portfolio corresponding to different assurance levels β and the sample size of $1,000$ are shown in Table 7.4.

Table 7.3. Results of Minimizing Type I risk (minimizing CVaR) using the new loss function

α	Sample Size	S&P	Gov.bond	Small cap	VaR	CVaR
0.90	10000	0.0784	0.9215	0	0.0193	0.0233
0.95	10000	0.0741	0.9258	0	0.0321	0.0413
0.99	10000	0.0924	0.9075	0	0.0479	0.0532

The results are as might be expected. In particular, we note that when optimizing the portfolio with respect Type I risk, we find that zero weight is allocated to “Small cap”, a small weight to “S&P” and a large weight

Table 7.4. Results of Minimizing Type II risk (maximizing CVoG) using the new gain function

β	Sample Size	S&P	Gov.bond	Small cap	VoG	CVoG
0.33	1000	0	0	1	0.0919	0.0210
0.4	1000	0	0	1	0.0575	0.0098
0.6	1000	0.0951	0	0.9048	0.0219	0.0024

to "Gov. bond". This merely reflects the fact that government bonds have a variance that is very close to zero, followed by "S&P", followed by "Small cap". Correspondingly, when optimizing the portfolio with respect to Type II risk, we find that zero weight is allocated to "Gov. bond", and very large weights are allocated to "Small cap". This reflects the fact that the mean return of "Small cap" is the highest while the government bonds have the lowest mean return.

Balancing two types of risks

Returning to the discussion of Section 2, we will continue to consider two problems associated with our loss and gain functions. Since we have seen that an investor who focuses on just Type I risk will obtain very different results to those that would be obtained if Type II risk were of the main concern. However, most investors will be sensitive - albeit, in varying degrees - to both types of risk. Hence, the challenge is to formulate a portfolio optimization problem that captures these dual concerns. Here, we will give two possible formulations of this "risk balancing problem" and illustrate them with an example based on the above data.

Two problems. Basically, we want to solve the following problems:

Problem I

$$\min_x \phi_\alpha(x), \quad \text{(minimize the CVaR loss)}$$

Subject to:

$$\psi_\beta(x) \geq \tau. \quad \text{(guarantee a CVoG target level of } \tau\text{).}$$

Since conditional value-of-gain, $\psi_\beta(x)$ is a concave function of x, the set $\{x : \psi_\beta(x) \geq \tau\}$ for any real number τ is a convex set. Hence, the above Problem I is a convex programming problem that is, in principle, suitable for fast numerical solution.

Problem II

$$\max_x \ \ \psi_\beta(x), \ \ \text{(maximize the CVoG gain)}$$

Subject to:

$$\phi_\alpha(x) \le v. \ \ \text{(tolerable CVaR risk level } v\text{)}$$

Since $-\psi_\beta(x)$ is a convex function of the decision variable x and the set $\{x : \phi_\alpha(x) \le v\}$ is a convex set, the above problem is also a convex programming problem.

According to our Theorem 7.21 and Theorem 16 in [15], if we have observations of returns $y_k, k = 1, \cdots, N$ generated by the distribution $p(y)$, then we can change the above two problems to the following Mixed Integer Programming (MIP) problems:

Problem I'

$$\min \ \zeta + \frac{1}{N(1-\alpha)} \sum_{k=1}^{N} \eta_k$$

Subject to:

$$x \ge 0, \ x^T \mathbf{e} = 1;$$

$$n_k \in \{0, 1\}, \ 0 \le g_k \le n_k M, \ 0 \le f_k \le (1 - n_k)M, \ k = 1, \cdots, N;$$

$$g_k - f_k - x^T y_k = 0, \ k = 1, \cdots, N;$$

$$\eta_k \ge 0, \ f_k - \zeta - \eta_k \le 0, \ k = 1, \cdots, N;$$

$$u_k \ge 0, \ g_k - \xi + u_k \ge 0, \ k = 1, \cdots, N;$$

$$\xi - \frac{1}{N(1-\beta)} \sum_{k=1}^{N} u_k \ge \tau;$$

where, M is a suitably chosen large number.

Problem II'

$$\max \ \xi - \frac{1}{N(1-\beta)} \sum_{k=1}^{N} u_k$$

Subject to:

$$x \ge 0, \ x^T \mathbf{e} = 1;$$

$$n_k \in \{0, 1\}, \ 0 \le g_k \le n_k M, \ 0 \le f_k \le (1 - n_k)M, \ k = 1, \cdots, N;$$

$$g_k - f_k - x^T y_k = 0, \ k = 1, \cdots, N;$$

$$u_k \ge 0, \ g_k - \xi + u_k \ge 0, \ k = 1, \cdots, N;$$

$$\eta_k \ge 0, \ f_k - \zeta - \eta_k \le 0, \ k = 1, \cdots, N;$$

$$\zeta + \frac{1}{N(1-\alpha)} \sum_{k=1}^{N} \eta_k \leq v;$$

where, M is a suitably chosen large number.

From above, it is easy to see that if we chose α, β, τ and v, we can use the above two optimization problems to calculate two optimized portfolios that capture a given investor's attitude to Type I and Type II risks.

Examples. We still use the data in Table 7.1 and Table 7.2. We let $\alpha = 0.9, \beta = 0.6$ be fixed. Using a sample size of 1000, we first calculate the maximal conditional value-of-gain, $\tau^*(\beta) = \tau^*(0.6) = 0.0024$ and the minimal conditional value-at-risk $v_*(\alpha) = v_*(0.9) = 0.0233$. Then, we choose $\tau \leq \tau^*(0.6)$ and $v \geq v_*(0.9)$ and solve the preceding two mixed integer programming problems. We obtain the following results.

Table 7.5. Results of Problem I

τ	Sample Size	S&P	Gov.bond	Small cap	$Z_1(0.9, 0.6, \tau)$
0.001	1000	0.0809	0.9190	0	0.0249
0.0015	1000	0.0902	0.9097	0	0.0257
0.002	1000	0.2160	0.5468	0.2371	0.0626
0.0023	1000	0.3597	0.2655	0.3746	0.0955

From Table 7.5 we can see that, $Z_1(\alpha, \beta, \tau)$ is very close to $v_*(\alpha)$ when our target level τ is very small, that means low requirement for the gain will yield low risk and vice versa.

Table 7.6. Results of Problem II

v	Sample Size	S&P	Gov.bond	Small cap	$Z_2(0.9, 0.6, v)$
0.024	1000	0.0697	0.9302	0	0.0003
0.030	1000	0.0576	0.8566	0.0857	0.0016
0.060	1000	0.2046	0.5691	0.2262	0.0019
0.1	1000	0.3791	0.2276	0.3931	0.0023

We can see that $Z_2(\alpha, \beta, v)$ is close to $\tau^*(\beta)$ when our risk tolerance v is large from Table 7.6, that shows the relaxation of the risk requirement will lead to a larger gain and vice versa.

4. Proofs

In this section, we will provide proofs of each lemma, proposition, theorem and corollary in above sections.

Proof of Lemma 7.5: If $(\alpha,\beta) \in \{(\alpha,\beta)|\tau_*(\beta,\alpha)\} = \tau^*(\beta)\}$, then $\psi_\beta(X_*(\alpha)) = \tau_*(\beta,\alpha) = \tau^*(\beta) = \max_X \psi_\beta(X)$. So $X_*(\alpha) \in \text{argmax}_X$ $\psi_\beta(X)$, that means for any $X \in \mathcal{L}^2$, we can't find one such that $\psi_\beta(X) > \tau^*(\beta)$, so we should choose $\tau \leq \tau^*(\beta)$, however based on the above relationship, there exists at least one X that will minimize $\phi_\alpha(X)$ and maximize $\psi_\beta(X)$ simultaneously, so in this case, $Z_1(\alpha,\beta,\tau) = v_*(\alpha)$.

In another case when $(\alpha,\beta) \in \{(\alpha,\beta)|\tau_*(\beta,\alpha) < \tau^*(\beta)\}$, that means we will obtain a lower gain $\tau_*(\beta,\alpha)$ when we try to minimize the risk $\phi_\alpha(X)$, but we can't obtain a gain that will exceed $\tau^*(\beta)$, reasonably, we will choose $\tau \in (\tau_*(\beta,\alpha),\tau^*(\beta)]$, however for this τ, there won't exist $X \in \text{argmin}_X \phi_\alpha(X)$ such that $\psi_\beta(X) \geq \tau$, this will yield a higher risk, that is $Z_1(\alpha,\beta,\tau) > v_*(\alpha)$. □

Proof of Lemma 7.6: Similarly to the proof of Lemma 7.5, we can prove Lemma 7.6. □

Proof of Lemma 7.7: Since $F(\mu+x)+F(\mu-x) = 1$, so $F(x) = F(\mu-\mu+x) = F(\mu-(\mu-x)) = 1-F(\mu+(\mu-x)) = 1-F(2\mu-x)$. It is easy to use the above relationship to prove the lemma together with the monotonicity of the distribution function $F(x)$. □

Proof of Proposition 7.14: We define:

$$\beta^-(x) = G(x,\xi_\beta(x)^-),\ \beta^+(x) = G(x,\xi_\beta(x)^+).$$

In comparison with the definition of $\psi_\beta(x)$ in Definition 7.12, $\psi_\beta^-(x)$ is the mean of the gain distribution associated with

$$G_\beta^-(x,\xi) = \begin{cases} G(x,\xi)/(1-\beta^-(x)), & \xi \leq \xi_\beta(x), \\ 1, & \xi > \xi_\beta(x), \end{cases}$$

whereas the $\psi_\beta^+(x)$ value is the mean of the gain distribution associated with

$$G_\beta^+(x,\xi) = \begin{cases} G(x,\xi)/(1-\beta^+(x)), & \xi \leq \xi_\beta(x), \\ 1, & \xi > \xi_\beta(x). \end{cases}$$

It is easy to see that $\beta^-(x)$ and $\beta^+(x)$ mark the bottom and top of the vertical gap at $\xi_\beta(x)$ for the original distribution function $G(x,\cdot)$ (if a jump occurs there).

The case of there being no probability atom at $\xi_\beta(x)$ corresponds to having $\beta^-(x) = \beta^+(x) = \beta \in (0,1)$. Then the first equation holds because the distribution functions $G_\beta^-(x,\xi), G_\beta(x,\xi)$ and $G_\beta^+(x,\xi)$ are identical.

When a probability atom exists but $\beta = \beta^+(x)$, we have: $\beta^-(x) < \beta^+(x) < 1$ and thus the second relations. If $\beta^+(x) = 0$, we can nevertheless get the third one since $\beta^-(x) < \beta^+(x) < 1$. Under the alternative of $0 < G(x, \xi_\beta(x)) < 1$, the strict inequalities in the fifth prevail. $\square$

Proof of Proposition 7.15: According to the definition of CVoG and when $G(x, \xi_\beta(x)) > 0$, we can calculate the mean in the definition directly as follows: for a fixed portfolio x,

$$\begin{aligned}
\psi_\beta(x) &= \int_0^{\xi_\beta(x)} d(G(x,\xi)/(1-\beta)) + (1 - G(x, \xi_\beta(x))/(1-\beta))\xi_\beta(x) \\
&= \tfrac{G(x,\xi_\beta(x))}{1-\beta} \times \tfrac{\int_0^{\xi_\beta(x)} dG(x,\xi)}{G(x,\xi_\beta(x))} + \left(1 - \tfrac{G(x,\xi_\beta(x))}{1-\beta}\right) \xi_\beta(x) \\
&= \tfrac{G(x,\xi_\beta(x))}{1-\beta} E\{g(x,Y) | g(x,Y) \le \xi_\beta(x)\} + \left(1 - \tfrac{G(x,\xi_\beta(x))}{1-\beta}\right) \xi_\beta(x) \\
&= \tfrac{G(x,\xi_\beta(x))}{1-\beta} \psi_\beta^+(x) + \left(1 - \tfrac{G(x,\xi_\beta(x))}{1-\beta}\right) \xi_\beta(x),
\end{aligned}$$

so we can obtain the following equation after defining $\lambda_\beta(x) = \frac{G(x,\xi_\beta(x))}{(1-\beta)}$,

$$\psi_\beta(x) = \lambda_\beta(x)\psi_\beta^+(x) + [1 - \lambda_\beta(x)]\xi_\beta(x).$$

We know $\lambda_\beta(x) \in [0,1]$ since $0 \le G_\beta(x, \xi_\beta(x)) \le 1 - \beta$. If $G(x, \xi_\beta(x)) = 0$, $\xi_\beta(x)$ is the lowest gain that can occur (and thus $\lambda_\beta(x) = 0$ but $\psi_\beta^-(x)$ is ill defined), then $\psi_\beta(x) = \xi_\beta(x)$. $\square$

Proof of Proposition 7.16: According to the following relationship:

$$\sum_{k=1}^{k_\beta} p_k \le 1 - \beta < \sum_{k=1}^{k_\beta+1} p_k,$$

we have

$$G(x, \xi_\beta(x)) = \sum_{k=1}^{k_\beta} p_k, \quad G(x, \xi_\beta(x)^-) = \sum_{k=1}^{k_\beta-1} p_k,$$

$$G(x, \xi_\beta(x)) - G(x, \xi_\beta(x)^-) = p_{k_\beta}.$$

The assertions then follow from Definition 7.12 and Proposition 7.15. $\square$

Proof of Corollary 7.17: This amounts to the special case in Proposition 7.16 with $k_\beta = 0$, then we know $\psi_\beta(x) = \xi_\beta(x) = z_1$. $\square$

Proof of Theorem 7.18: Firstly, we will prove the Theorem in the case that the distribution function $G(x, \xi)$ of the gain $g(x, Y)$ for fixed x is everywhere continuous with respect to ξ. We also assume the random return Y has desity function $p(y)$. Before proceeding main steps, we will give a lemma for preparation.

LEMMA 7.22 *With x fixed, let $Q(\xi) = \int_{y \in R^n} q(\xi, y)p(y)dy$, where $q(\xi, y) = [g(x, y) - \xi]^-$. Then Q is a convex continuously differentiable function with derivative*

$$Q'(\xi) = G(x, \xi).$$

Proof: This lemma follows from Proposition 2.1 of Shapiro and wardi (1994) in [20].

Now, let's prove the Theorem in this particular case. In view of the defining formula for $H_\beta(x, \xi)$,

$$H_\beta(x, \xi) = \xi - \frac{1}{1-\beta} E\{[g(x, Y) - \xi]^-\},$$

it is immediate from Lemma 7.22 and the fact that linear function is a concave function that $H_\beta(x, \xi)$ is concave and continuously differentiable with derivative

$$\frac{\partial}{\partial \xi} H_\beta(x, \xi) = 1 - \frac{1}{1-\beta} G(x, \xi).$$

Therefore, the values of ξ that furnish the maximum of $H_\beta(x, \xi)$ are precisely those for which $G(x, \xi) = 1 - \beta$. They form a nonempty closed interval, inasmuch as $G(x, \xi)$ is continuous and nondecreasing in ξ with limit 1 as $\xi \to \infty$ and limit 0 as $\xi \to -\infty$. This further yields the validity of the formula $\xi_\beta(x) \in \text{argmax}_\xi H_\beta(x, \xi)$. In particular, then, we have

$$\max_{\xi \in R} H_\beta(x, \xi) = H_\beta(x, \xi_\beta(x)) = \xi_\beta(x) - \frac{1}{1-\beta} \int_{y \in R} [g(x, y) - \xi_\beta(x)]^- p(y)dy.$$

But the integral here equals

$$\int_{g(x,y) \le \xi_\beta(x)} [\xi_\beta(x) - g(x, y)]p(y)dy = \xi_\beta(x) \int_{g(x,y) \le \xi_\beta(x)} p(y)dy -$$

$$\int_{g(x,y) \le \xi_\beta(x)} g(x, y)p(y)dy,$$

where the first integral on the right is by definition $G(x, \xi_\beta(x)) = 1 - \beta$ and the second is $(1 - \beta)\psi_\beta(x)$. Thus,

$$\max_{\xi \in R} H_\beta(x, \xi) = \xi_\beta(x) - \frac{1}{1-\beta} [(1 - \beta)\xi_\beta(x) - (1 - \beta)\psi_\beta(x)] = \psi_\beta(x).$$

This confirms the formula for $\beta-$ CVoG, $\psi_\beta(x) = \max_\xi H_\beta(x, \xi)$, and completes the proof of the Theorem in this special case.

In the following, I'll prove it in a more general sense, including the discreteness of the distribution.

The finiteness of $H_\beta(x,\xi)$ is a consequence of our assumption that $E\{|g(x,y)|\} < \infty$ for each $x \in X$. It's concave follows at once from the convexity of $[g(x,y)-\xi]^-$ with respect to ξ. Similar to convex function, a finite concave function, $H_\beta(x,\xi)$ has finite right and left derivatives at any ξ. The following approach of proving the rest of the assertions in the theorem will rely on first establishing for these one-sided derivatives, the formulas,

$$\frac{\partial^+ H_\beta}{\partial \xi}(x,\xi) = \frac{1-\beta-G(x,\xi)}{1-\beta}, \quad \frac{\partial^- H_\beta}{\partial \xi}(x,\xi) = \frac{1-\beta-G(x,\xi^-)}{1-\beta}. \tag{7.1}$$

We start by observing that

$$\frac{H_\beta(x,\xi')-H_\beta(x,\xi)}{\xi'-\xi} = 1 - \frac{1}{1-\beta} E\left\{\frac{[g(x,Y)-\xi']^- - [g(x,Y)-\xi]^-}{\xi'-\xi}\right\}.$$

When $\xi' > \xi$, we have:

$$\frac{[g(x,Y)-\xi']^- - [g(x,Y)-\xi]^-}{\xi'-\xi} \begin{cases} = 0 & \text{if } g(x,Y) \geq \xi' \\ = 1 & \text{if } g(x,Y) \leq \xi \\ \in (0,1) & \text{if } \xi < g(x,Y) < \xi' \end{cases}$$

Since $P_Y\{\xi < g(x,Y) \leq \xi'\} = G(x,\xi') - G(x,\xi)$, this yields the existence of a value $\rho(\xi,\xi') \in [0,1]$ for which

$$E\left\{\frac{[g(x,Y)-\xi']^- - [g(x,Y)-\xi]^-}{\xi'-\xi}\right\} = G(x,\xi) + \rho(\xi,\xi')[G(x,\xi')-G(x,\xi)].$$

Since furthermore $G(x,\xi') \searrow G(x,\xi)$ as $\xi' \searrow \xi$, it follows that

$$\lim_{\xi' \searrow \xi} E\left\{\frac{[g(x,Y)-\xi']^- - [g(x,Y)-\xi]^-}{\xi'-\xi}\right\} = G(x,\xi).$$

So, we obtain

$$\lim_{\xi' \searrow \xi} \frac{H_\beta(x,\xi')-H_\beta(x,\xi)}{\xi'-\xi} = 1 - \frac{1}{1-\beta} G(x,\xi) = \frac{1-\beta-G(x,\xi)}{1-\beta},$$

thereby verifying the first formula in (7.1). For the second formula in (7.1), we argue similarly that when $\xi' < \xi$ we have

$$\frac{[g(x,Y)-\xi']^- - [g(x,Y)-\xi]^-}{\xi'-\xi} \begin{cases} = 0 & \text{if } g(x,Y) \geq \xi \\ = 1 & \text{if } g(x,Y) \leq \xi' \\ \in (0,1) & \text{if } \xi' < g(x,Y) < \xi \end{cases}$$

where $P_Y\{\xi' < g(x,Y) < \xi\} = G(x,\xi^-) - G(x,\xi')$. Since $G(x,\xi') \nearrow G(x,\xi^-)$ as $\xi' \nearrow \xi$ we obtain

$$\lim_{\xi' \nearrow \xi} E\left\{\frac{[g(x,Y)-\xi']^- - [g(x,Y)-\xi]^-}{\xi'-\xi}\right\} = G(x,\xi^-),$$

and then

$$\lim_{\xi' \nearrow \xi} \frac{H_\beta(x,\xi') - H_\beta(x,\xi)}{\xi'-\xi} = 1 - \frac{1}{1-\beta}G(x,\xi^-) = \frac{1-\beta-G(x,\xi^-)}{1-\beta}.$$

That gives the second formula in (7.1).

Because of concavity, the one-sided derivatives in (7.1) are non-increasing with respect to ξ, with the formulas assuring that

$$\lim_{\xi\to\infty} \frac{\partial^+ H_\beta}{\partial\xi}(x,\xi) = \lim_{\xi\to\infty} \frac{\partial^- H_\beta}{\partial\xi}(x,\xi) = -\frac{\beta}{1-\beta}$$

and on the other hand,

$$\lim_{\xi\to-\infty} \frac{\partial^+ H_\beta}{\partial\xi}(x,\xi) = \lim_{\xi\to-\infty} \frac{\partial^- H_\beta}{\partial\xi}(x,\xi) = 1.$$

On the basis of these limits, we know that the level set $\{\xi | H_\beta(x,\xi) \geq c\}$ are bounded (for any choice of $c \in R$) and therefore that the maximum in the theorem is attained, with the argmax set being a closed, bounded interval. The values of ξ in that set are characterized as the ones such that

$$\frac{\partial^+ H_\beta}{\partial\xi}(x,\xi) \leq 0 \leq \frac{\partial^- H_\beta}{\partial\xi}(x,\xi).$$

According to the formulas in (7.1), they are the values of ξ satisfying $G(x,\xi^-) \leq 1-\beta \leq G(x,\xi)$. The rest of the Theorem is a direct conclusion of the above results. □

Proof of Corollary 7.19: The joint concavity of $H_\beta(x,\xi)$ in (x,ξ) is an elementary consequence of the definition of this function, the relationship between convexity and concavity and the convexity of the function $(x,\xi) \to [g(x,y)-\xi]^-$ when $g(x,y)$ is convex in x. The concave of $\psi_\beta(x)$ in x follows immediately then from the maximization formula in Theorem 7.18. (In convex analysis, when a convex function of two vector variables is minimized with respect to one of them, the residual is a convex function of the other: see Rockafellar [13]. Here, we can obtain the result just simply applying the above theory to the convex function $-H_\beta(x,\xi)$.) □

Proof of Theorem 7.20: This rests on the principle in optimization that maximization with respect to (x,ξ) can be carried out by maximizing with respect to ξ for each x and then maximizing the residual with respect to x. In the situation at hand, we invoke Theorem 7.18 and in particular, in order to get the equivalence in the second formula in the Theorem, the fact there that the maximum of $H_\beta(x, \xi)$ in ξ (for fixed x) is always attained. $\square$

Proof of Theorem 7.21: This relies on the maximization formula in Theorem 7.18 and the assured attainment of the maximum there. The arguments are very much like that for Theorem 7.20. Because $\psi_\beta(x) = \max_\xi H_\beta(x, \xi)$, we have $\psi_{\beta_i}(x) \geq \tau_i$, if and only if there exists ξ_i such that $H_{\beta_i}(x, \xi_i) \geq \tau_i$. $\square$

References

[1] C. Acerbi, C. Nordio and C. Sirtori, Expected Shortfall as a Tool for Financial Risk Management, working paper, 2001.

[2] C. Acerbi and D. Tasche, On the coherence of Expected Shortfall, *Journal of Banking and Finance*, **26 (7)**, 1487–1503, 2002.

[3] P. Artzner, F. Delbaen, J. M. Eber and D. Heath, Coherent Measures of Risk, *Mathematical Finance*, **9**, 203–227, 1999.

[4] S. Basak and A. Shapiro, Value-at-Risk Based Risk Management: Optimal Policies and Asset Prices, *The Review of Financial Studies*, **14(2)**, 371–405, 2001.

[5] D. Bertsimas, G. J. Lauprete and A. Samarov, Shortfall as a Risk Measure: Properties Optimization and Application. *Journal of Economic Dynamics and Control*, **28**, 1227–1480, 2004.

[6] G. A. Holton, Value-at-risk: theory and practice. Amsterdam; Boston: Academic Press, 2003.

[7] P. Kall and S. W. Wallace, Stochastic Programming, Wiley, Chichester, 1994.

[8] Boda Kang, Measures of risk: issues of time consistency and surrogate processes, PhD thesis, University of South Australia, 2006.

[9] P. Krokhmal, J. Palmquist and S. Uryasev, Portfolio Optimization with Conditional Value-At-Risk Objective and Constraints, *The Journal of Risk*, **4 (2)**, 11-27, 2002.

[10] N. Larsen, H. Mausser and S. Uryasev, Algorithms for Optimization of value-at-risk, *P. Pardalos and V.K. Tsitsiringos, (Eds.) Financial Engineering, e-Commerce and Supply Chain*, Kluwer Academic Publishers, 129–157, 2002.

[11] H. Markowitz, Portfolio Selection: Efficient Diversification of Investment, J.Wiley and Sons., 1959.

[12] G. C. Pflug, Some Remarks on the Value-at-Risk and the Conditional Value-at-Risk, *Probabilistic Constrained Optimization: Methodology and Applications (S.P. Uryasev, ed.)*, Kluwer, 278–287, 2000.

[13] R. T. Rockafellar, Convex Analysis, Princeton University Press, available since 1997 in paperback in the series Princeton Landmarks in Mathematics and Physics, 1997.

[14] R. T. Rockafellar and S. Uryasev, Optimization of conditional value-at-risk, *Journal of Risk*, **2**, 21–42, 2000.

[15] R. T. Rockafellar and S. Uryasev, Conditional value-at-risk for general loss distributions, *Journal of Banking* & *Finance*, **26**, 1443–1471, 2002.

[16] R.T. Rockafellar and S. Uryasev, Deviation Measures in Risk Analysis and Optimization, *Research Report 2002-7*, Dept. of Industrial and Systems Engineering, University of Florida, 2002.

[17] R.T. Rockafellar, S. Uryasev and M. Zabarankin, Generalized Deviations in Risk Analysis. To appear on *Finance and Stochastics*, **9**, 2005.

[18] R.T. Rockafellar, S. Uryasev and M. Zabarankin, Optimality Conditions in Portfolio Analysis with Generalized Deviation Measures, *Mathematical Programming*, accepted for publication, 2005.

[19] R.T. Rockafellar, S. Uryasev and M. Zabarankin, Master Funds in Portfolio Analysis with General Deviation Measures, *The Journal of Banking and Finance*, Vol.**29**, 2005.

[20] A. Shapiro, Y. Wardi, Nondifferentiability of the steady-state function in discrete event dynamic systems, *IEEE Transactions on Automatic Control*, **39**, 1707-1711, 1994.