Chapter 4
Expected Shortfall

This chapter deals coherent risk measures such as Expected Shortfall and Tail Value at Risk (TVaR).

4.1 Tail Value at Risk (TVaR)

A natural shortcoming of Value at Risk is to fail to provide information on the behavior of probability distribution tails beyond $V_p^X$. The next figure illustrates the limitations of Value at Risk, namely its inability to capture the properties of a probability distribution beyond $V_p^X$.†

† “Value at Risk is like an airbag that works all the time, except when you have a car accident”. - David Einhorn, hedge fund manager.

Fig. 4.1: Two distributions having the same Value at Risk $V_{X}^{95\%} = 2.145$. 
The Tail Value at Risk aims at providing a solution to the tail distribution problem observed with Value at Risk at the level $p$, by averaging over confidence levels ranging from $p$ to 1.

**Definition 4.1.** The Tail Value at Risk of a random variable $X$ at the level $p \in (0, 1)$ is defined by the average

\[
TV^p_X := \frac{1}{1-p} \int_p^1 V^p_X dq.
\]  

(4.1)

Note that since the function $p \mapsto V^p_X$ is non-decreasing, we always have

\[
TV^p_X = \frac{1}{1-p} \int_p^1 V^p_X dq \geq \frac{1}{1-p} \int_p^1 V^p_X dq = V^p_X.
\]

**Example: Exponential distribution**

For example, if $X$ has an exponential distribution with parameter $\lambda > 0$ and mean $1/\lambda$, we have

\[
TV^p_X = \frac{1}{1-p} \int_p^1 V^p_X dq = \frac{1}{\lambda(1-p)} \int_p^1 \log(1-q) dq = \frac{1}{\lambda(1-p)} \int_0^{1-p} \log q dq = 1-p + (1-p) \log \frac{1}{1-p} = \frac{1}{\lambda} + \frac{1}{\lambda} \log \frac{1}{1-p} = \mathbb{E}[X] + V^p_X.
\]

### 4.2 Conditional tail expectation (CTE)

Recall that by Lemma 10.11, given an event $A$ such that $\mathbb{P}(A) > 0$, the conditional expectation of $X : \Omega \rightarrow \mathbb{N}$ given the event $A$ satisfies

\[
\mathbb{E}[X \mid A] = \frac{1}{\mathbb{P}(A)} \mathbb{E}[X 1_A].
\]

For example, consider $\Omega = \{1, 3, -1, -2, 5, 7\}$ with the uniform probability measure given by
Expected Shortfall

$$\mathbb{P}(\{k\}) = 1/6, \quad k = 1, 3, -1, -2, 5, 7,$$

and the random variable

$$X : \Omega \rightarrow \mathbb{Z}$$

given by

$$X(k) = k, \quad k = 1, 3, -1, -2, 5, 7.$$  

Then $\mathbb{E}[X \mid X > 1]$ denotes the expected value of $X$ given

$$A = \{X > 1\} = \{3, 5, 7\} \subset \Omega,$$

i.e. the mean value of $X$ given that $X$ is strictly positive. This conditional expectation can be computed as

$$\mathbb{E}[X \mid X > 1] = \frac{3 + 5 + 7}{6} \times \frac{3}{6} = \frac{1}{\mathbb{P}(X > 1)} \mathbb{E} \left[ X \mathbb{1}_{\{X > 1\}} \right],$$

where $\mathbb{P}(X > 1) = 3/6$ and the truncated expectation $\mathbb{E} \left[ X \mathbb{1}_{\{X > 1\}} \right]$ is given by $\mathbb{E} \left[ X \mathbb{1}_{\{X > 1\}} \right] = (3 + 5 + 7)/6.$

For example, when $X$ has a geometric distribution we can estimate $\mathbb{E}[X \mid X > 10]$ as follows using $R$:

```r
geo_samples <- rgeom(100000, prob = 1/4)
mean(geo_samples)
mean(geo_samples[geo_samples>10])
```

**Definition 4.2.** The Conditional Tail Expectation of a random variable $X$ at the level $p \in (0, 1)$ is the quantity

$$\text{CTE}_X^p := \mathbb{E} \left[ X \mid X > V_X^p \right] = \frac{1}{\mathbb{P}(X > V_X^p)} \mathbb{E} \left[ X \mathbb{1}_{\{X > V_X^p\}} \right].$$

The use of the strict inequality “$>$” in the definition of the Conditional Tail Expectation is motivated by the necessity to avoid dependence on $\mathbb{P}(X = V_X^p)$ and to consider risky values strictly beyond $V_X^p$.

Examples of Conditional Tail Expectations can be computed as in the following code.

1 geo_samples <- rgeom(100000, prob = 1/4)
2 mean(geo_samples)
3 mean(geo_samples[geo_samples>10])
The next proposition shows by which amount the Conditional Tail Expectation exceeds the Value at Risk.

**Proposition 4.3.** In general we have \( \text{CTE}_X^p \geq \text{V}_X^p \) and, more precisely,

\[
\text{CTE}_X^p = \mathbb{E}[X | X > V_X^p] = V_X^p + \mathbb{E}[(X - V_X^p)^+ | X > V_X^p].
\]

**Proof.** We have

\[
\begin{align*}
\mathbb{E}[X | X > V_X^p] &= \frac{1}{\mathbb{P}(X > V_X^p)} \mathbb{E}[X 1_{\{X > V_X^p\}}] \\
&= \frac{1}{\mathbb{P}(X > V_X^p)} \left( \mathbb{E}[(X - V_X^p) 1_{\{X > V_X^p\}}] + V_X^p \mathbb{E}[1_{\{X > V_X^p\}}] \right) \\
&= \frac{1}{\mathbb{P}(X > V_X^p)} \left( \mathbb{E}[(X - V_X^p)^+] + V_X^p \mathbb{P}(X > V_X^p) \right) \\
&= V_X^p + \frac{1}{\mathbb{P}(X > V_X^p)} \mathbb{E}[(X - V_X^p)^+] \\
&= V_X^p + \mathbb{E}[(X - V_X^p)^+ | X > V_X^p].
\end{align*}
\]

\( \square \)

Next, we check that when \( \mathbb{P}(X = V_X^p) = 0 \), the Conditional Tail Expectation coincides with the Tail Value at Risk.

**Proposition 4.4.** Assume that \( \mathbb{P}(X = V_X^p) = 0 \). Then we have \( \text{CTE}_X^p = \text{TV}_X^p \), i.e.

\[
\text{CTE}_X^p = \mathbb{E}[X | X > V_X^p] = \mathbb{E}[X | X \geq V_X^p] = \frac{1}{1 - p} \int_p^1 V_X^q dq = \text{TV}_X^p.
\]

**Proof.** By Lemma 3.11 we construct \( X = V_X^U \) where \( U \) is uniformly distributed on \([0, 1] \), with

\[
U \geq p \implies V_X^U \geq V_X^p \implies X \geq V_X^p,
\]

and

\[
X > V_X^p \implies V_X^U > V_X^p \implies U > p.
\]

Since \( \mathbb{P}(X = V_X^p) = 0 \) we find that, with probability 1,
Expected Shortfall

\[ U \geq p \iff U > p \iff V_X^U \geq V_X^p \iff X \geq V_X^p \iff X > V_X^p, \]

hence

\[
\begin{align*}
\text{CTE}_X^p &= \mathbb{E}[X \mid X > V_X^p] \\
&= \mathbb{E}[V_X^U \mid V_X^U > V_X^p] \\
&= \mathbb{E}[V_X^U \mid U > p] \\
&= \frac{1}{\mathbb{P}(U > p)} \mathbb{E}[V_X^U \mathbb{1}_{\{U > p\}}] \\
&= \frac{1}{1 - p} \int_p^1 V_X^q dq.
\end{align*}
\]

The next figure shows the location of Value at Risk and Conditional Tail Expectation on a data set. Note that the sign of the data has been changed according to Proposition 3.8.

Fig. 4.2: Value at Risk and Conditional Tail Expectation.

The Conditional Tail Expectation of a Gaussian \( \mathcal{N}(\mu, \sigma^2) \) random variable is computed in the next proposition.

**Proposition 4.5.** Gaussian CTE. Given \( X \simeq \mathcal{N}(\mu, \sigma^2) \) we have

\[
\text{CTE}_X^p = \mu + \frac{\sigma}{(1 - p)} \phi(V_Z^p) = \mu + \frac{\sigma}{(1 - p)\sqrt{2\pi}} e^{-(V_Z^p)^2/2}, \tag{4.3}
\]

where \( V_Z^p \) is the Value at Risk of \( Z \simeq \mathcal{N}(0, 1) \) at the level \( p \) and

\[
\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad x \in \mathbb{R},
\]

is the standard normal probability density function.
Proof. Using the relation $\mathbb{P}(X > V^p_X) = 1 - p$, cf. Proposition 3.7, we have

\[
\text{CTE}^p_X = TV^p_X = \mathbb{E} \left[ X \mid X > V^p_X \right] = \frac{1}{\mathbb{P}(X > V^p_X)} \mathbb{E} \left[ X \mathbb{1}_{\{X > V^p_X\}} \right] = \frac{1}{1 - p} \int_{V^p_X}^{\infty} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} \frac{dx}{\sqrt{2\pi\sigma^2}}.
\]

\[
= \frac{\mu}{1 - p} \int_{V^p_X}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \frac{dx}{\sqrt{2\pi\sigma^2}} + \frac{1}{1 - p} \int_{V^p_X}^{\infty} (x - \mu) e^{-\frac{(x-\mu)^2}{2\sigma^2}} \frac{dx}{\sqrt{2\pi\sigma^2}}.
\]

\[
= \frac{\mu}{1 - p} \mathbb{P}(X \geq V^p_X) + \frac{\sigma^2}{(1-p)\sqrt{2\pi\sigma^2}} \left[ -e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right]_{V^p_X}^{\infty} = \mu + \frac{\sigma^2}{(1-p)\sqrt{2\pi\sigma^2}} e^{-\frac{((V^p_X-\mu)/\sigma)^2}{2}} - e^{-\frac{(V^p_X-\mu)^2}{2\sigma^2}}.
\]

\[
= \mu + \frac{\sigma^2}{(1-p)\sqrt{2\pi\sigma^2}} e^{-\frac{(V^p_X-\mu)^2}{2\sigma^2}} - e^{-\frac{(V^p_Z)^2}{2\sigma_Z}} = \mu + \frac{\sigma}{(1-p)\sqrt{2\pi}} e^{-\frac{(V^p_Z)^2}{2\sigma_Z}},
\]

due to the rescaling relation $V^p_X = \mu + \sigma q^p_Z$, cf. (3.7). □

4.3 Expected Shortfall (ES)

There are several variants for the definition of the Expected Shortfall $\text{ES}^p_X$. Next is a frequently used definition.

**Definition 4.6.** The Expected Shortfall $\text{ES}^p_X$ of a random variable $X$ at the level $p \in (0, 1)$ is defined by

\[
\text{ES}^p_X := \frac{1}{1 - p} \mathbb{E} \left[ X \mathbb{1}_{\{X \geq V^p_X\}} \right] + \frac{V^p_X}{1 - p} (1 - \mathbb{P}(X \geq V^p_X)).
\] (4.4)

We also have

\[
\text{ES}^p_X = \begin{cases} 
\frac{1}{1 - p} \mathbb{E} \left[ X \mathbb{1}_{\{X > V^p_X\}} \right] = \mathbb{E} \left[ X \mid X > V^p_X \right] = TV^p_X & \text{if } \mathbb{P}(X = V^p_X) = 0, \\
\frac{1}{1 - p} \mathbb{E} \left[ X \mathbb{1}_{\{X > V^p_X\}} \right] + \frac{V^p_X}{1 - p} (1 - \mathbb{P}(X \geq V^p_X)) & \text{if } \mathbb{P}(X = V^p_X) > 0,
\end{cases}
\]

as shown in the next proposition.

**Proposition 4.7.** When $\mathbb{P}(X = V^p_X) = 0$ the Expected Shortfall coincides with the Conditional Tail Expectation and the Tail Value at Risk, i.e. we...
Expected Shortfall

have

\[ \text{ES}_X^p = \mathbb{E} \left[ X \mid X > V_X^p \right] = \text{TV}_X^p. \]

**Proof.** By Proposition 3.7 we have

\[ p = \mathbb{P}(X \leq V_X^p) \quad \text{and} \quad 1 - p = \mathbb{P}(X > V_X^p) = \mathbb{P}(X \geq V_X^p), \]

hence

\[
\text{ES}_X^p = \frac{1}{1 - p} \mathbb{E} \left[ X \mathbb{1}_{\{X > V_X^p\}} \right] = \frac{1}{\mathbb{P}(X > V_X^p)} \mathbb{E} \left[ X \mathbb{1}_{\{X > V_X^p\}} \right] = \mathbb{E} \left[ X \mid X > V_X^p \right] = \text{TV}_X^p,
\]

by Proposition 4.4.

When \( \mathbb{P}(X = V_X^p) = 0 \), we also have

\[
\text{ES}_X^p = \frac{1}{1 - p} \mathbb{E} \left[ X \mathbb{1}_{\{X \geq V_X^p\}} \right] + V_X^p - \frac{V_X^p}{1 - p} \mathbb{P}(X \geq V_X^p) = \frac{1}{1 - p} \mathbb{E} \left[ X \mathbb{1}_{\{X > V_X^p\}} \right] + \frac{V_X^p}{1 - p} \left( 1 - p - \mathbb{P}(X > V_X^p) \right).
\]

**Proposition 4.8.** The Expected Shortfall \( \text{ES}_X^p \) can be written as the distorted risk measure

\[
\text{ES}_X^p = \frac{1}{1 - p} \mathbb{E} \left[ X f_X(X) \right] = \frac{1}{1 - p} \int_0^1 V_X^q f_X(V_X^q) dq,
\]

where \( f_X \) is the distortion function defined by

\[
f_X(x) := \mathbb{1}_{\{x > V_X^p\}} + \mathbb{1}_{\{\mathbb{P}(X = V_X^p) > 0\}} \frac{1 - p - \mathbb{P}(X > V_X^p)}{\mathbb{P}(X = V_X^p)} \mathbb{1}_{\{x = V_X^p\}}, \quad x \in \mathbb{R}.
\]

**Proof.** We have

\[
\text{ES}_X^p = \frac{1}{1 - p} \mathbb{E} \left[ X \mathbb{1}_{\{X > V_X^p\}} \right] + \frac{V_X^p}{1 - p} \left( 1 - p - \mathbb{P}(X \geq V_X^p) \right) = \frac{1}{1 - p} \mathbb{E} \left[ X \mathbb{1}_{\{X \geq V_X^p\}} \right] + \mathbb{1}_{\{\mathbb{P}(X = V_X^p) > 0\}} \frac{V_X^p}{1 - p} \left( 1 - p - \mathbb{P}(X > V_X^p) \right) = \frac{1}{1 - p} \mathbb{E} \left[ X \left( \mathbb{1}_{\{X > V_X^p\}} + \mathbb{1}_{\{\mathbb{P}(X = V_X^p) > 0\}} \frac{1 - p - \mathbb{P}(X > V_X^p)}{\mathbb{P}(X = V_X^p)} \mathbb{1}_{\{x = V_X^p\}} \right) \right].
\]

This version: December 23, 2019

https://www.ntu.edu.sg/home/nprivault/indext.html
The distortion function $f_X$ is a non-decreasing $[0,1]$-valued function that satisfies

$$
\mathbb{E}[f_X(X)] = \mathbb{E}
\left[
\mathbf{1}_{\{X > V^p_X\}} + \mathbf{1}_{\{P(X = V^p_X) > 0\}} \frac{1 - p - \mathbb{P}(X > V^p_X)}{\mathbb{P}(X = V^p_X)} \mathbf{1}_{\{X = V^p_X\}}
\right]
$$

$$
= \mathbb{E}
\left[
\mathbf{1}_{\{X > V^p_X\}} + 1 - p - \mathbb{P}(X > V^p_X)
\right]
$$

$$
= \mathbb{P}(X \geq V^p_X) + 1 - p - \mathbb{P}(X \geq V^p_X)
$$

$$
= 1 - p, \quad x \in \mathbb{R}.
$$

(4.6)

The following proposition, cf. Acerbi and Tasche (2001), shows that in general, the Expected Shortfall at the level $p \in (0,1)$ coincides with the Tail Value at Risk $\text{TV}^p_X$.

**Proposition 4.9.** The Expected Shortfall $\text{ES}^p_X$ coincides with the Tail Value at Risk $\text{TV}^p_X$ for any $p \in (0,1)$, i.e. we have

$$
\text{ES}^p_X = \text{TV}^p_X = \frac{1}{1 - p} \int_p^1 V^q_X dq.
$$

Proof. Constructing $X$ as $X = V^U_X$ where $U$ is uniformly distributed on $[0,1]$ as in Lemma 3.11, by Proposition 3.6 we have

$$
U \geq p \implies V^U_X \geq V^p_X \implies X \geq V^p_X
$$

and

$$
(U < p \text{ and } X \geq V^p_X) \implies (V^U_X \leq V^p_X \text{ and } X \geq V^p_X)
$$

$$
\implies (X \leq V^p_X \text{ and } X \geq V^p_X)
$$

$$
\implies X = V^p_X.
$$

Hence by (4.4) and the relations

$$
1 - p = \mathbb{E}\left[ \mathbf{1}_{\{U \geq p\}} \right] \quad \text{and} \quad \mathbb{P}(X \geq V^p_X) = \mathbb{E}\left[ \mathbf{1}_{\{X \geq V^p_X\}} \right],
$$

we have

$$
V^p_X\left(1 - p - \mathbb{P}(X \geq V^p_X)\right) = -V^p_X \mathbb{E}\left[ \mathbf{1}_{\{X \geq V^p_X\}} - \mathbf{1}_{\{U \geq p\}} \right]
$$

$$
= -V^p_X \mathbb{E}\left[ \mathbf{1}_{\{X \geq V^p_X\} \setminus \{U \geq p\}} \right]
$$

$$
= -V^p_X \mathbb{E}\left[ \mathbf{1}_{\{X \geq V^p_X\} \cap \{U < p\}} \right]
$$

$$
= -\mathbb{E}\left[ X \mathbf{1}_{\{X \geq V^p_X\} \cap \{U < p\}} \right],
$$

hence
Expected Shortfall

\[ \text{ES}_X^p = \frac{1}{1-p} \mathbb{E} \left[ X \mathbb{1}_{\{X \geq V_X^p\}} \right] + \frac{V_X^p}{1-p} \left( 1 - p - \mathbb{P}(X \geq V_X^p) \right) \]

\[ = \frac{1}{1-p} \mathbb{E} \left[ X \mathbb{1}_{\{X \geq V_X^p\}} \right] - \frac{1}{1-p} \mathbb{E} \left[ X \mathbb{1}_{\{X \geq V_X^p\} \cap \{ U < p \}} \right] \]

\[ = \frac{1}{1-p} \mathbb{E} \left[ V_X^U \mathbb{1}_{\{V_X^U \geq V_X^p\}} \right] - \frac{1}{1-p} \mathbb{E} \left[ V_X^U \mathbb{1}_{\{V_X^U \geq V_X^p\} \cap \{ U < p \}} \right] \]

\[ = \frac{1}{1-p} \mathbb{E} \left[ V_X^U \mathbb{1}_{\{V_X^U \geq V_X^p\} \cap \{ U \geq p \}} \right] \]

\[ = \frac{1}{1-p} \int_p^1 V_X^U \, dq, \]

which is the Tail Value at Risk \( TV_X^p \).

As a consequence of (4.3), Proposition 4.4 and Proposition 4.9, the Gaussian Expected Shortfall at the level \( p \) is also given by

\[ \text{ES}_X^p = \mu + \frac{\sigma}{(1-p)} \phi(V_Z^p). \]

**Proposition 4.10.** The Expected Shortfall \( \text{ES}_X^p \) and the Tail Value at Risk \( TV_X^p \) are coherent risk measures.

**Proof.** As \( \text{ES}_X^p \) coincides with \( TV_X^p \) for all \( p \in (0, 1) \) from Proposition 4.9, we can use either Relation (4.4) in Definition 4.6 or Relation (4.1) in Definition 4.1.

(i) Monotonicity. If \( X \leq Y \), since Value at Risk is monotone we have

\[ \text{ES}_X^p = TV_X^p \]

\[ = \frac{1}{1-p} \int_p^1 V_X^U \, dq \]

\[ \leq \frac{1}{1-p} \int_p^1 V_Y^U \, dq \]

\[ = TV_Y^p \]

\[ \leq \text{ES}_Y^p. \]

for all \( p \in (0, 1) \).

(ii) Homogeneity and translation invariance. Similarly, since Value at Risk satisfies the homogeneity and translation invariance properties, for all \( \mu \in \mathbb{R} \) and \( \lambda > 0 \) we have monotone we have

\[ \text{ES}_{\mu + \lambda X}^p = TV_{\mu + \lambda X}^p \]

\[ = \frac{1}{1-p} \int_p^1 V_{\mu + \lambda X}^U \, dq \]

\[ \leq \text{ES}_{\mu + \lambda X}^p. \]
\[ \frac{1}{1-p} \int_p^1 (\mu + \lambda V_Y^p) dq \]

\[ = \mu + \lambda \frac{1}{1-p} \int_p^1 V_Y^p dq \]

\[ = \mu + \lambda TV_Y^p \]

\[ \leq \mu + \lambda ES_Y^p \]

for all \( p \in (0, 1) \).

(iii) Sub-additivity. We have

\[ (1-p) \left( \mathbb{E}_{X,Y}^p \mathbb{E}_X^p - \mathbb{E}_X^p - \mathbb{E}_X^p \right) \]

\[ = \mathbb{E}[X+Y] f_{X,Y} - \mathbb{E}[X f_X] - \mathbb{E}[Y f_Y] \]

\[ = \mathbb{E}[X f_{X,Y} (X+Y)] - \mathbb{E}[Y f_{X,Y} (Y)] \]

\[ = \mathbb{E}[X (f_{X,Y} (X+Y) - f_X (X))] - \mathbb{E}[Y (f_{X,Y} (X+Y) - f_Y (Y))] \]

\[ \geq V_X^p \mathbb{E}[f_{X,Y} (X+Y) - f_X (X)] - V_Y^p \mathbb{E}[f_{X,Y} (X+Y) - f_Y (Y)] \]

\[ = V_X^p (1-p - (1-p)) - V_Y^p (1-p - (1-p)) \]

\[ = 0, \]

where we have used (4.6) and the facts that, for \( x < V_X^p \),

\[ (1-p)(f_{X,Y}(x+y) - f_X(x)) = \mathbb{I}\{x+y>V_X^p\} - \mathbb{I}\{x>V_X^p\} \]

\[ + \mathbb{I}\{p(x+y=V_X^p)\} \frac{1-p - \mathbb{P}(X+Y > V_X^p)}{\mathbb{P}(X+Y = V_X^p)} \mathbb{I}\{x+y=V_X^p\} \]

\[ - \mathbb{I}\{p(x=V_X^p)\} \frac{1-p - \mathbb{P}(X > V_X^p)}{\mathbb{P}(X = V_X^p)} \mathbb{I}\{x=V_X^p\} \]

\[ = \mathbb{I}\{x+y>V_X^p\} + \mathbb{I}\{p(x+y=V_X^p)\} \frac{1-p - \mathbb{P}(X+Y > V_X^p)}{\mathbb{P}(X+Y = V_X^p)} \mathbb{I}\{x+y=V_X^p\} \]

\[ \geq 0, \quad x < V_X^p, \]

and, for \( x > V_X^p \),

\[ (1-p)(f_{X,Y}(x+y) - f_X(x)) = \mathbb{I}\{x+y>V_X^p\} - \mathbb{I}\{x>V_X^p\} \]

\[ + \mathbb{I}\{p(x+y=V_X^p)\} \frac{1-p - \mathbb{P}(X+Y > V_X^p)}{\mathbb{P}(X+Y = V_X^p)} \mathbb{I}\{x+y=V_X^p\} \]

\[ - \mathbb{I}\{p(x=V_X^p)\} \frac{1-p - \mathbb{P}(X > V_X^p)}{\mathbb{P}(X = V_X^p)} \mathbb{I}\{x=V_X^p\} \]

\[ = \mathbb{I}\{x+y>V_X^p\} - \mathbb{I}\{x>V_X^p\} \]

\[ + \mathbb{I}\{p(x+y=V_X^p)\} \frac{1-p - \mathbb{P}(X+Y > V_X^p)}{\mathbb{P}(X+Y = V_X^p)} \mathbb{I}\{x+y=V_X^p\} \]
Expected Shortfall

\[
\mathbb{1}_{\{x+y>V_p^X+Y\}} - \mathbb{1}_{\{x>V_p^X\}} + \mathbb{1}_{\{x+y=V_p^X+Y\}} = \mathbb{1}_{\{x+y\geq V_p^X+Y\}} - \mathbb{1}_{\{x>V_p^X\}} \geq 0, \quad x < V_p^X.
\]

Note that in general, the Conditional Tail Expectation is not a coherent risk measure when \( P(X = V_p^X) > 0 \).

Performance analytics in R - Expected Shortfall (ES)

```r
library(PerformanceAnalytics)

ES(returns, p=.95, method="historical")
```

The 95% Expected Shortfall is \( \text{ES}^{95\%}_X = -0.02087832 \). The historical Expected Shortfall can be exactly recovered by the empirical Conditional Tail Expectation (CTE) as

```r
mean(returns[returns<(VaR(returns, p=.95, method="historical")[1]),na.rm=TRUE])
```

The Gaussian Expected Shortfall is given as \( -0.0191359 \) by

```r
ES(returns, p=.95, method="gaussian")
```

and it can be recovered from (4.3) as

\[
\text{ES}^p_x = \mu + \frac{\sigma}{(1-p)}\phi(V^p_Z) = \mu + \frac{\sigma}{(1-p)\sqrt{2\pi}}e^{-(V^p_Z)^2/2}
\]

with \( \mu = \mathbb{E}[X] \) and \( \sigma^2 = \text{Var}[X] \), i.e.

```r
q=qnorm(.95, mean=0, sd=1)
m=mean(returns,na.rm=TRUE)
s=sd(returns,na.rm=TRUE)
m-s*dnorm(q)/0.05
```

with output \( -0.01916536 \).

The attached **R code** computes the Expected Shortfall and compares its output to the that of the PerformanceAnalytics package, as illustrated in the next Figure 4.3.

```r
> source("comparison.R")
```

Number of samples= 266
VaR90= -0.03279002, Threshold= 0.05263158

This version: December 23, 2019

https://www.ntu.edu.sg/home/nprivault/index.html
CTE90 = -0.04536057
ES90 = -0.04507702
Historical VaR90 = -0.03258733
Gaussian VaR90 = -0.03190036
Historical ES90 = -0.04446267
Gaussian ES90 = -0.04025347

Fig. 4.3: Value at Risk and Expected Shortfall.

Value at Risk vs Expected Shortfall

```r
chart.VaRSensitivity(ts(returns), methods=c("HistoricalVaR", "HistoricalES"), colorset=bluefocus, lwd=2)
```

Fig. 4.4: Value at Risk vs Expected Shortfall.
4.4 Gaussian Measures of Risk vs Market Returns

Market returns vs Gaussian and Power Tails

Consider for example the following DJIA returns data obtained with Quantmod.

![Graph showing market returns vs 2610 normalized Gaussian returns.](figure)

Figure 4.5 shows the mismatch between the distributional properties of market vs standardized Gaussian returns which tends to underestimate the probabilities of extreme events. Note that when $X \sim \mathcal{N}(0, \sigma^2)$, 99.73% of samples of $X$ are falling within the interval $[-3\sigma, +3\sigma]$, i.e. $P(|X| \leq 3\sigma) = 0.9973002$.

```r
library(quantmod)
getSymbols("DJIA", from="1990-01-03", to="2015-02-01", src="FRED")
stock.rtn = diff(log(DJIA)); returns <- as.vector(stock.rtn)
m = mean(returns, na.rm = TRUE); s = sd(returns, na.rm = TRUE); times = index(stock.rtn)
n = sum(is.na(returns)) + sum(!is.na(returns)); x = seq(1,n); y = rnorm(n, mean = m, sd = s)
plot(times, returns, pch = 19, xaxs="i", cex = 0.03, col = "blue", ylab = "X", xlab = "n", main = "")
segments(x0 = times, x1 = times, y0 = 0, y1 = returns, col = "blue")
points(times, y, pch = 19, cex = 0.03, col = "red", ylab = "X", xlab = "n", main = "")
abline(h = 3*s, col = "black", lwd =1); abline(h = -3*s, col = "black", lwd =1)
```

This mismatch can be further illustrated by the empirical density plot in Figure 4.6, which is obtained from the following code.
Clearly, the Gaussian Value at Risk will be strictly lower than the empirical Value at Risk for large return values.

![Empirical Cumulative Distribution](image1)

**Fig. 4.6:** Empirical density vs normalized Gaussian density.

From Figure 4.7 we note that power tail (e.g. power tail) densities with a decay in $x \mapsto 1/|x|^\alpha$ can provide a better fit of empirical densities.

![Empirical density vs power density](image2)

**Fig. 4.7:** Empirical density vs power density.

The above fitting of empirical density is using a power probability density defined by a rational fraction obtained by the following R script.
Expected Shortfall

install.packages("pracma")
x <- seq(-0.25, 0.25, length=1000)
stock.dens=density(returns,na.rm=TRUE, from = -0.1, to = 0.1, n = 1000)
library(pracma)
a<-rationalfit(stock.dens$x, stock.dens$y, d1=2, d2=2)

plot(stock.dens$x,stock.dens$y, type = "l",xlab = "x value", col="red",ylab = "Density", main = "")
lines(x,(a$p1[3]+a$p1[2]*x+a$p1[1]*x^2)/(a$p2[3]+a$p2[2]*x+a$p2[1]*x^2),type="l",lty=2,
col="blue",xlab="x value",ylab="Density",main="")
legend("topleft", legend=c("Empirical density", "power density"),col=c("red", "blue"),
lty=1:2, cex=0.8)
The output of the command rationalfit is

$p1
[1] -0.184717249 -0.001591433 0.001385017

$p2
[1] 1.000000e+00 -6.460948e-04 1.314672e-05

which yields a rational fraction of the form

\[
x \mapsto \frac{0.001385017 - 0.001591433 x - 0.184717249 x^2}{1.314672 \times 10^{-5} - 6.460948 \times 10^{-4} x + x^2}
\]

\[
\approx -0.184717249 - \frac{0.001591433}{x} + \frac{0.001385017}{x^2},
\]

which approximates the empirical density of DJIA returns in the least squares sense.

chart.VaRSensitivity(stock.rtn[,1,drop=FALSE],methods=c("HistoricalVaR","GaussianVaR"),
colorset=bluefocus, lwd=2)
The next Figure 4.8 uses the above code to compare the historical and Gaussian values at risk.

Fig. 4.8: Historical vs Gaussian estimates of Value at Risk.
In the next Figure 4.9 we compare the Gaussian and historical estimates of Expected Shortfall.

![Risk Confidence Sensitivity of 1800.HK.Adjusted](image)

Fig. 4.9: Quantile function.

In Table 4.1 we summarize some properties of risk measures.

<table>
<thead>
<tr>
<th>Risk Measure</th>
<th>Additivity</th>
<th>Homogeneity</th>
<th>Subadditivity</th>
<th>Coherence</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V^p_X$</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>$CTE^p_X$</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>$TV^p_X$</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$ES^p_X$</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Note that Value at Risk $V^p_X$ is coherent on Gaussian random variables according to Remark 3.10. Similarly, the Conditional Tail Expectation $CTE^p_X$ is coherent on random variables having a continuous CDF by Propositions 4.4 and 4.10.
Expected Shortfall

Exercises

Exercise 4.1 Consider the following data set.

Find the Value at Risk VaR$^p_X$ and the Conditional Tail Expectation CTE$^p_X = \mathbb{E} [X \mid X > \text{VaR}^p_X]$ and mark their values on the graph in the following cases.

a) $p = 0.9$.
b) $p = 0.8$.

Exercise 4.2

Let $p = 0.9$. For the above data set represented by the uniformly distributed random variable $X$, compute the numerical values of the following quantities.

a) $\text{VaR}_{X}^{90}$,
b) $\mathbb{E} [X 1_{\{X > \text{VaR}_{X}^{90}\}}]$,
c) $\mathbb{P}(X > \text{VaR}_{X}^{90})$,
d) $\text{CTE}_{X}^{90} = \mathbb{E} [X \mid X > \text{VaR}_{X}^{90}] = \mathbb{E} [X 1_{\{X > \text{VaR}_{X}^{90}\}}] / \mathbb{P}(X > \text{VaR}_{X}^{90})$,
e) $\mathbb{E} [X 1_{\{X \geq \text{VaR}_{X}^{90}\}}]$,
f) $\mathbb{P}(X \geq \text{VaR}_{X}^{90})$,
Exercise 4.3 Consider a random variable $X \in \{10, 100, 150\}$ with the distribution
\[
\mathbb{P}(X = 10) = 96\%, \quad \mathbb{P}(X = 100) = 3\%, \quad \mathbb{P}(X = 150) = 1\%.
\]
Compute
a) the Value at Risk $\text{VaR}_X^{98\%}$,
b) the Tail Value at Risk $\text{TV}_X^{98\%}$,
c) the Conditional Tail Expectation $\mathbb{E}[X | X > \text{VaR}_X^{98\%}]$, and
d) the Expected Shortfall $\text{ES}_X^{98\%}$.

Exercise 4.4 Consider two independent random variables $X$ and $Y$ with the same distribution given by
\[
\mathbb{P}(X = 0) = \mathbb{P}(Y = 0) = 90\% \quad \text{and} \quad \mathbb{P}(X = 100) = \mathbb{P}(Y = 100) = 10\%.
\]
a) Plot the cumulative distribution function of $X$ on the following graph:

![Cumulative distribution function of X](image)

Fig. 4.10: Cumulative distribution function of $X$.

b) Plot the cumulative distribution function of $X + Y$ on the following graph:
Expected Shortfall

Fig. 4.11: Cumulative distribution function of $X + Y$.

c) Give the values at risk $V_{X+Y}^{99\%}$, $V_{X+Y}^{95\%}$, $V_{X+Y}^{90\%}$.

d) Compute the Tail Value at Risk

$$TV_X^{90\%} := \frac{1}{1 - p} \int_{p}^{1} V_X^q dq$$

at the level $p = 90\%$.

e) Compute the Tail Value at Risk

$$TV_{X+Y}^{p} := \frac{1}{1 - p} \int_{p}^{1} V_{X+Y}^q dq$$

at the levels $p = 90\%$ and $p = 80\%$.

Exercise 4.5 (Exercise 3.2 continued).

a) Compute the Tail Value at Risk

$$TV_X^{p} := \frac{1}{1 - p} \int_{p}^{1} V_X^q dq$$

for all $p$ in the interval $[0.99, 1]$, and give the value of $TV_X^{99\%}$.

b) Taking $p = 0.98$, compute the Conditional Tail Expectation

$$CTE_X^{98\%} = \mathbb{E} \left[ X \mid X > V_X^{98\%} \right] = \frac{1}{\mathbb{P}(X > V_X^{p})} \mathbb{E} \left[ X \mathbb{1}_{\{X > V_X^{p}\}} \right].$$