Expected Shortfall (ES) is a risk measure computed by averting potential losses above a certain level given by the Value at Risk (VaR). It can be shown that the Expected Shortfall at the confidence level $p$ coincides with the Tail Value at Risk (TVaR) defined as the average of losses suffered in the worst $(1 - p)\%$ of events. This chapter presents the concept of coherent risk measure, including Expected Shortfall and Tail Value at Risk (TVaR), together with experiments based on financial data sets.

### 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 $V_X^p$. The next figure illustrates the limitations of Value at Risk, namely its inability to capture the properties of a probability distribution beyond $V_X^p$.†

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

$$TV^p_X := \frac{1}{1-p} \int_p^1 V^q_X dq. \quad (4.1)$$

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

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

**Example: exponential distribution**

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

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

$$= -\frac{1}{\lambda(1-p)} \int_p^1 \log(1-q) dq$$

$$= -\frac{1}{\lambda(1-p)} \int_0^{1-p} (\log q) dq$$

$$= 1 - p + (1-p) \log \frac{1}{1-p}$$

$$= \frac{1}{\lambda} + \frac{1}{\lambda} \log \frac{1}{1-p}$$

$$= \mathbb{E}[X] \left( 1 + \log \frac{1}{1-p} \right)$$
Expected Shortfall

\[
= \mathbb{E}[X] + V^p_X.
\]

4.2 Conditional tail expectation (CTE)

Recall that by Lemma 10.12, 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

\[
\mathbb{P}(\{k\}) = \frac{1}{6}, \quad k = 1, 3, -1, -2, 5, 7,
\]

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

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

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

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

\[\bigodot\]
The use of the strict inequality “>” in the definition of the Conditional Tail Expectation is motivated by the necessity to avoid dependence on \( P(X = V^p_X) \) and to consider risky values strictly beyond \( V^p_X \).

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

```r
library(quantmod)
gtSymbs("^HSI",from="2013-06-01",to="2014-10-01",src="yahoo")
returns <- as.vector(diff(log(Ad(HSI)/gtSymbs("^HSI","2013-06-01","2014-10-01","yahoo"))))
library(PerformanceAnalytics)
var=VaR(returns, p=.95, method="historical")
cte=mean(returns[returns<as.numeric(var)],na.rm=TRUE)
```

The next proposition shows by which amount the Conditional Tail Expectation exceeds the Value at Risk.

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

\[
\text{CTE}^p_X = \mathbb{E} [X \mid X > V^p_X] = V^p_X + \mathbb{E} [(X - V^p_X)^+ \mid X > V^p_X].
\]

*Proof.* We have

\[
\mathbb{E} [X \mid X > V^p_X] = \frac{1}{P(X > V^p_X)} \mathbb{E} [X \mathbb{1}_{\{X > V^p_X\}}]
\]

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

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

\[
= V^p_X + \frac{1}{P(X > V^p_X)} \mathbb{E} [(X - V^p_X)^+] = V^p_X + \mathbb{E} [(X - V^p_X)^+ \mid X > V^p_X].
\]

Next, we check that when \( P(X = V^p_X) = 0 \), the Conditional Tail Expectation coincides with the Tail Value at Risk.

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

\[
\text{CTE}^p_X = \mathbb{E} [X \mid X > V^p_X] = \mathbb{E} [X \mid X \geq V^p_X] = \frac{1}{1 - p} \int_p^1 V^q_X dq = \text{TV}^p_X.
\]
Expected Shortfall

Proof. By Lemma 3.11 we construct $X$ as $X = V^U_X$ where $U$ is uniformly distributed on $[0, 1]$, with

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

and

$$X > V^p_X \implies V^U_X > V^p_X \implies U > p.$$ 

Since $\mathbb{P}(X = V^p_X) = 0$ we find that, with probability 1,

$$U \geq p \iff U > p \iff V^U_X \geq V^p_X \iff X \geq V^p_X \iff X > V^p_X,$$

hence

$$\text{CTE}^p_X = \mathbb{E} \left[ X \mid X > V^p_X \right]$$

$$= \mathbb{E} \left[ V^U_X \mid V^U_X > V^p_X \right]$$

$$= \mathbb{E} \left[ V^U_X \mid U \geq p \right]$$

$$= \frac{1}{\mathbb{P}(U \geq p)} \mathbb{E} \left[ V^U_X \mathbb{1}_{\{U \geq p\}} \right]$$

$$= \frac{1}{1 - p} \int_p^1 V^q_X dq.$$ 

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.](image_url)

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

(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.

\textbf{Proof.} Using the relation \( \mathbb{P}(X > V_{X}^p) = 1 - p \), cf. Proposition 3.7, we have
\[
\begin{align*}
\text{CTE}_X^p &= \text{TV}_X^p \\
&= \mathbb{E}[X | X > V_{X}^p] \\
&= \frac{1}{\mathbb{P}(X > V_{X}^p)} \mathbb{E} \left[ X 1 \{X > V_{X}^p\} \right] \\
&= \frac{1}{1 - p} \int_{V_{X}^p}^{\infty} x e^{-(x-\mu)^2/(2\sigma^2)} \frac{dx}{\sqrt{2\pi}\sigma^2} \\
&= \frac{\mu}{1 - p} \int_{V_{X}^p}^{\infty} e^{-(x-\mu)^2/(2\sigma^2)} \frac{dx}{\sqrt{2\pi}\sigma^2} + \frac{1}{1 - p} \int_{V_{X}^p}^{\infty} (x - \mu) e^{-(x-\mu)^2/(2\sigma^2)} \frac{dx}{\sqrt{2\pi}\sigma^2} \\
&= \frac{\mu}{1 - p} \mathbb{P}(X \geq V_{X}^p) + \frac{\sigma^2}{(1-p)\sqrt{2\pi}\sigma^2} \left[ e^{-(x-\mu)^2/(2\sigma^2)} \right]_{V_{X}^p}^{\infty} \\
&= \mu + \frac{\sigma^2}{(1-p)\sqrt{2\pi}} e^{-(V_{X}^p-\mu)/\sigma^2}/2 \\
&= \mu + \frac{\sigma}{(1-p)\sqrt{2\pi}} e^{-(V_{Z}^p)^2/2} \\
&= \mu + \frac{\sigma}{(1-p)} \phi(V_{Z}^p),
\end{align*}
\]
due to the rescaling relation \( V_{X}^p = \mu + \sigma q_{Z}^p \), cf. (3.7). \qed

4.3 Expected Shortfall (ES)

There are several variants for the definition of the Expected Shortfall \( \text{ES}_X^p \). Next is a frequently used definition.

\textbf{Definition 4.6.} The Expected Shortfall \( \text{ES}_X^p \) of a random variable \( X \) at the level \( p \in (0,1) \) is defined by
\[
\text{ES}_X^p := \frac{1}{1 - p} \mathbb{E} \left[ X 1 \{X > V_{X}^p\} \right] + \frac{V_{X}^p}{1 - p} (1 - \mathbb{P}(X \geq V_{X}^p)).
\]

(4.4)

We also have
Expected Shortfall

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

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 have

\[
\text{ES}^p_X = \mathbb{E} \left[ X | X > V^p_X \right] = \text{TV}^p_X.
\]

**Proof.** By Proposition 3.7 we have

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

hence

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

by Proposition 4.4. \( \square \)

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

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

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

\[
\text{ES}^p_X = \frac{1}{1-p} \mathbb{E}[X f_X(X)] = \frac{1}{1-p} \int_0^1 V^q_X f_X(V^q_X) dq, \quad (4.5)
\]

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

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

**Proof.** We have
ES\textsubscript{p} \text{X} = \frac{1}{1 - p} \mathbb{E} \left[ X \mathbb{1}_{\{X \geq V\text{p}_X\}} \right] + \frac{V\text{p}_X}{1 - p} \frac{1 - p - \mathbb{P}(X \geq V\text{p}_X)}{\mathbb{P}(X = V\text{p}_X)} \mathbb{1}_{\{X = V\text{p}_X\}}.

\[ \text{The distortion function } f_X \text{ is a non-decreasing } [0,1] \text{-valued function that satisfies} \]
\[ \mathbb{E}[f_X(X)] = \mathbb{E} \left[ \mathbb{1}_{\{X > V\text{p}_X\}} + \mathbb{1}_{\{X = V\text{p}_X\}} \frac{1 - p - \mathbb{P}(X > V\text{p}_X)}{\mathbb{P}(X = V\text{p}_X)} \mathbb{1}_{\{X = V\text{p}_X\}} \right] \]
\[ = \mathbb{E} \left[ \mathbb{1}_{\{X > V\text{p}_X\}} + 1 - p - \mathbb{P}(X > V\text{p}_X) \right] \]
\[ = \mathbb{P}(X > V\text{p}_X) + 1 - p - \mathbb{P}(X \geq V\text{p}_X) \]
\[ = 1 - p, \quad x \in \mathbb{R}. \quad (4.6) \]

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 \( TV\text{p}_X \).

**Proposition 4.9.** The Expected Shortfall \( ES\text{p}_X \) coincides with the Tail Value at Risk \( TV\text{p}_X \) for any \( p \in (0,1) \), i.e. we have
\[ ES\text{p}_X = TV\text{p}_X = \frac{1}{1 - p} \int_p^1 V\text{p}_X dq. \]

**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
\[ U \geq p \implies V^U_X \geq V\text{p}_X \implies X \geq V\text{p}_X \]
and
\[ (U < p \text{ and } X \geq V\text{p}_X) \implies (V^U_X \leq V\text{p}_X \text{ and } X \geq V\text{p}_X) \]
\[ \implies (X \leq V\text{p}_X \text{ and } X \geq V\text{p}_X) \]
\[ \implies X = V\text{p}_X. \]

Hence by (4.4) and the relations
\[ 1 - p = \mathbb{E} \left[ \mathbb{1}_{\{U \geq p\}} \right] \text{ and } \mathbb{P}(X \geq V\text{p}_X) = \mathbb{E} \left[ \mathbb{1}_{\{X \geq V\text{p}_X\}} \right], \]
we have
Expected Shortfall

\[ V_X^p (1 - p - \mathbb{P} (X \geq V_X^p)) = -V_X^p \mathbb{E} \left[ \mathbb{I}_{\{X \geq V_X^p\}} - \mathbb{I}_{\{U \geq p\}} \right] \]
\[ = -V_X^p \mathbb{E} \left[ \mathbb{I}_{\{X \geq V_X^p\}} \mathbb{I}_{\{U \geq p\}} \right] \]
\[ = -V_X^p \mathbb{E} \left[ \mathbb{I}_{\{X \geq V_X^p\}} \mathbb{I}_{\{U < p\}} \right] \]
\[ = -\mathbb{E} \left[ X \mathbb{I}_{\{X \geq V_X^p\}} \mathbb{I}_{\{U < p\}} \right], \]

hence

\[ \text{ES}_X^p = \frac{1}{1 - p} \mathbb{E} [X \mathbb{I}_{\{X \geq V_X^p\}}] + \frac{V_X^p}{1 - p} (1 - p - \mathbb{P} (X \geq V_X^p)) \]
\[ = \frac{1}{1 - p} \mathbb{E} [X \mathbb{I}_{\{X \geq V_X^p\}}] - \frac{1}{1 - p} \mathbb{E} [X \mathbb{I}_{\{X \geq V_X^p\}} \mathbb{I}_{\{U < p\}}] \]
\[ = \frac{1}{1 - p} \mathbb{E} [V_X^U \mathbb{I}_{\{V_X^U \geq V_X^p\}}] - \frac{1}{1 - p} \mathbb{E} [V_X^U \mathbb{I}_{\{V_X^U \geq V_X^p\}} \mathbb{I}_{\{U < p\}}] \]
\[ = \frac{1}{1 - p} \mathbb{E} [V_X^U \mathbb{I}_{\{V_X^U \geq V_X^p\}} \mathbb{I}_{\{U \geq p\}}] \]
\[ = \frac{1}{1 - p} \int_p^1 V_X^q dq, \]

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

\[ \text{ES}_X^p = TV_X^p \]
\[ = \frac{1}{1 - p} \int_p^1 V_X^q dq \]
\[ \leq \frac{1}{1 - p} \int_p^1 V_Y^q dq \]
\[ = TV_Y^p \]
\[ \leq \text{ES}_Y^p. \]
for all $p \in (0, 1)$.

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

$$
\text{ES}_{\mu+\lambda X}^p = T\text{V}_{\mu+\lambda X}^p
= \frac{1}{1-p} \int_0^1 V_{\mu+\lambda X}^q dq
= \frac{1}{1-p} \int_0^1 (\mu + \lambda V_X^q) dq
= \mu + \lambda \int_0^1 V_X^q dq
= \mu + \lambda \text{TV}_X^p
\leq \mu + \lambda \text{ES}_Y^p
$$

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

(iii) Sub-additivity. We have

$$(1-p) \left( \text{ES}_{X+Y}^p - \text{ES}_X^p - \text{ES}_Y^p \right)
= \mathbb{E}[(X + Y) f_{X+Y}(X + Y)] - \mathbb{E}[X f_X(X)] - \mathbb{E}[Y f_Y(Y)]
= \mathbb{E}[X(f_{X+Y}(X + Y) - f_X(X))] - \mathbb{E}[Y(f_{X+Y}(X + Y) - f_Y(Y))]
\geq V_X^p \mathbb{E}[f_{X+Y}(X + Y) - f_X(X)] - V_Y^p \mathbb{E}[f_{X+Y}(X + Y) - f_Y(Y)]
= V_X^p (1 - p - (1-p)) - V_Y^p (1 - p - (1-p))
= 0,$$

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

$$(1-p)f_{X+Y}(x+y) - f_X(x) = \mathbb{I}_{x+y>V_{X+Y}^p} - \mathbb{I}_{x>V_X^p}
+ \mathbb{I}_{\mathbb{P}(x+y>V_{X+Y}^p)>0} \frac{1-p - \mathbb{P}(X + Y > V_{X+Y}^p)}{\mathbb{P}(X = V_X^p)} \mathbb{I}_{x=V_X^p}
- \mathbb{I}_{\mathbb{P}(x>V_X^p)>0} \frac{1-p - \mathbb{P}(X > V_X^p)}{\mathbb{P}(X = V_X^p)} \mathbb{I}_{x=V_X^p}
= \mathbb{I}_{x+y>V_{X+Y}^p} + \mathbb{I}_{\mathbb{P}(x+y>V_{X+Y}^p)>0} \frac{1-p - \mathbb{P}(X + Y > V_{X+Y}^p)}{\mathbb{P}(X + Y = V_{X+Y}^p)} \mathbb{I}_{x+y>V_{X+Y}^p}
\geq 0, \quad x < V_X^p,$$

and, for $x > V_X^p$,

$$(1-p)f_{X+Y}(x+y) - f_X(x) = \mathbb{I}_{x+y>V_{X+Y}^p} - \mathbb{I}_{x>V_X^p}$$
Expected Shortfall

\[
+ \mathbb{I}\{P(X+Y=V^P_X+Y) > 0\} \frac{1 - p - P(X + Y > V^P_X+Y)}{P(X + Y = V^P_X+Y)} \mathbb{I}\{x+y=V^P_X+Y\} \\
- \mathbb{I}\{P(X=V^P_X) > 0\} \frac{1 - p - P(X > V^P_X)}{P(X = V^P_X)} \mathbb{I}\{x=V^P_X\} \\
= \mathbb{I}\{x+y>V^P_X+Y\} - \mathbb{I}\{x>V^P_X\} \\
+ \mathbb{I}\{P(X+Y=V^P_X+Y) > 0\} \frac{1 - p - P(X + Y > V^P_X+Y)}{P(X + Y = V^P_X+Y)} \mathbb{I}\{x+y=V^P_X+Y\} \\
\leq \mathbb{I}\{x+y>V^P_X+Y\} - \mathbb{I}\{x>V^P_X\} \\
+ \mathbb{I}\{x+y=V^P_X+Y\} \\
= \mathbb{I}\{x+y>V^P_X+Y\} - \mathbb{I}\{x>V^P_X\} \\
\geq 0, \quad x < V^P_X.
\]

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 \(ES^{95\%}_X = -0.02087832\). 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

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

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
CTE90= -0.04536057
ES90= -0.04464267
Historical VaR90= -0.03258733
Gaussian VaR90= -0.03190036
Historical ES90= -0.04446267
Gaussian ES90= -0.04025347
```

![Graph showing data samples, frequency histogram, and cumulative distribution function](image)

**Fig. 4.3:** Value at Risk and Expected Shortfall.

**Value at Risk vs Expected Shortfall**

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

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. $\mathbb{P}(|X| \leq 3\sigma) = 0.9973002$. 

Expected Shortfall

Fig. 4.4: Value at Risk vs Expected Shortfall.

Fig. 4.5: Market returns vs 2610 normalized Gaussian returns.
This mismatch can be further illustrated by the empirical probability density plot in Figure 4.6, which is obtained from the following R code.

```R
x <- seq(-0.25, 0.25, length=100); qx <- dnorm(x,mean=m,sd=s)
stock.dens=density(stock.rtn,na.rm=TRUE)
plot(stock.dens, xlab = 'Sample Quantiles', lwd=2, col="red", ylab = '
', main = 'Empirical Cumulative Distribution', panel.first = abline(h = 0, col="grey", lwd =0.2))
lines(x, qx, type="l", lty=2, lwd=2, col="blue",xlab="x value",ylab="Density",
main="Gaussian cdf")
legend("topleft", legend=c("Empirical cdf", "Gaussian cdf"),col=c("red", "blue"), lty=1:2,
cex=0.8)
```

Clearly, the Gaussian Value at Risk will be strictly lower than the empirical Value at Risk for large return values.

![Empirical Cumulative Distribution](image)

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

```r
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

- $p1$
  
  ```text
  [1] -0.184717249 -0.001591433 0.001385017
  ```

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

which yields a rational fraction of the form

$$x \mapsto \frac{0.001385017 - 0.001591433 \times x - 0.184717249 \times x^2}{1.314672 \times 10^{-5} - 6.460948 \times 10^{-4} \times x + x^2} \approx -0.184717249 - \frac{0.001591433}{x} + \frac{0.001385017}{x^2},$$

which approximates the empirical probability density of DJIA returns in the least squares sense.
The next Figure 4.8 uses the above R code to compare the historical and Gaussian values at risk.

![Graph showing comparison between historical and Gaussian Value at Risk](image)

**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.

![Graph showing comparison between historical and Gaussian Expected Shortfall](image)

**Fig. 4.9:** Quantile function.

In Table 4.1 we summarize some properties of risk measures.

<table>
<thead>
<tr>
<th>Table 4.1: Summary of Risk Measures</th>
<th></th>
<th></th>
</tr>
</thead>
</table>

This version: September 13, 2020
https://www.ntu.edu.sg/home/nprivault/index.html
Expected Shortfall

<table>
<thead>
<tr>
<th>Risk Measure</th>
<th>Additivity</th>
<th>Homogeneity</th>
<th>Subadditivity</th>
<th>Coherence</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_X^p$</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>$CTE_X^p$</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>$TV_X^p$</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$ES_X^p$</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Note that Value at Risk $V_X^p$ is \textit{coherent} on Gaussian random variables according to Remark 3.10. Similarly, the Conditional Tail Expectation $CTE_X^p$ is \textit{coherent} on random variables having a continuous CDF by Propositions 4.4 and 4.10.

Exercises

Exercise 4.1 Consider the following data set.

Find the Value at Risk $\text{VaR}_X^p$ and the Conditional Tail Expectation $\text{CTE}_X^p = \mathbb{E} [X \mid X > \text{VaR}_X^p]$ 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) $\text{VaR}_{90\%}^X$,
b) $\mathbb{E} \left[ X \mathbb{1}_{\{X > V_{90\%}^X\}} \right]$,
c) $\mathbb{P}(X > V_{90\%}^X)$,
d) $\text{CTE}_{90\%}^X = \mathbb{E} \left[ X \mid X > V_{90\%}^X \right] = \mathbb{E} \left[ X \mathbb{1}_{\{X > V_{90\%}^X\}} \right] / \mathbb{P}(X > V_{90\%}^X)$,
e) $\mathbb{P} \left( X \geq V_{90\%}^X \right)$,
f) $\mathbb{P}(X \geq V_{90\%}^X)$,
g) $\text{ES}_{90\%}^X = \frac{1}{1-p} \left( \mathbb{E} \left[ X \mathbb{1}_{\{X > V_{90\%}^X\}} \right] + V_{90\%}^X \left( 1 - p - \mathbb{P}(X \geq V_{90\%}^X) \right) \right)$,
h) $\text{TV}_{90\%}^X = \frac{1}{1-p} \int_0^1 V_q^X dq$

and mark the values of $\text{VaR}_{90\%}^X$, $\text{CTE}_{90\%}^X$, $\text{ES}_{90\%}^X$, $\text{TV}_{90\%}^X$ on the above graph.

Exercise 4.3 Consider a random variable $X \in \{10, 100, 150\}$ with the distribution

$\mathbb{P}(X = 10) = 96\%$, $\mathbb{P}(X = 100) = 3\%$, $\mathbb{P}(X = 150) = 1\%$.

Compute

a) the Value at Risk $V_{98\%}^X$,
b) the Tail Value at Risk $\text{TV}_{98\%}^X$,
c) the Conditional Tail Expectation $\mathbb{E} \left[ X \mid X > V_{98\%}^X \right]$,
d) the Expected Shortfall $E_{98\%}^X$.

Exercise 4.4 Consider two independent random variables $X$ and $Y$ with same distribution given by

$\mathbb{P}(X = 0) = \mathbb{P}(Y = 0) = 90\%$ and $\mathbb{P}(X = 100) = \mathbb{P}(Y = 100) = 10\%$.
a) Plot the cumulative distribution function of $X$ on the following graph:

![Cumulative distribution function of X](image1)

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

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

![Cumulative distribution function of X+Y](image2)

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