Chapter 7
Expected Shortfall

This chapter presents the construction of Tail Value at Risk (TVaR) and the Expected Shortfall (ES), which, unlike Value at Risk, are coherent risk measures. The Tail Value at Risk at the confidence level $p \in (0, 1)$ is defined as the average of losses suffered in the worst $(1 - p)\%$ of events. Expected Shortfall provides an alternative computation of Tail Value at Risk (TVaR) by averaging potential losses above the VaR level. Experiments based on financial data sets are also included.

7.1 Tail Value at Risk (TVaR)

A basic shortcoming of Value at Risk is failing 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". - D. Einhorn, hedge fund manager.

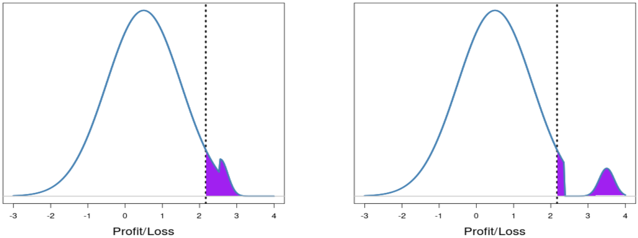


Fig. 7.1: Two distributions having the same Value at Risk $V_X^{95\%} = 2.145$.

The Tail Value at Risk (or Conditional Value at Risk) aims at providing a solution to the tail distribution problem observed with Value at Risk at the level $p \in (0,1)$ by averaging over confidence levels ranging from p to 1.

Definition 7.1. *The* Tail Value at Risk *(TVaR) of a random variable X at the level $p \in (0,1)$ is defined by the average*

$$\mathrm{TV}_X^p := \frac{1}{1-p} \int_p^1 V_X^q dq. \tag{7.1}$$

Note that since the function $p \longmapsto V_X^p$ is non-decreasing, see Proposition 6.11, we always have

$$\mathrm{TV}_X^p = \frac{1}{1-p} \int_p^1 V_X^q dq \geqslant \frac{1}{1-p} \int_p^1 V_X^p dq = V_X^p.$$

7.2 Conditional Tail Expectation (CTE)

Recall that by Lemma A.15, given an event A such that $\mathbb{P}(A) > 0$, the *conditional expectation* of $X : \Omega \longrightarrow \mathbb{N}$ given the event A satisfies

$$\mathbb{E}[X \mid A] = \frac{1}{\mathbb{P}(A)} \mathbb{E}\left[X \mathbb{1}_A\right],$$

see Section 6.1 for an example.

Definition 7.2. *Consider a random variable X such that $\mathbb{P}\big(X > V_X^p\big) > 0$. The* Conditional Tail Expectation *of X at the level $p \in (0,1)$ is the quantity*

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

This version: January 10, 2024
https://personal.ntu.edu.sg/nprivault/indext.html

The *Conditional Tail Expectation* CTE_X^p at the level $p \in (0,1)$ can be written as the *distortion risk measure*

$$\mathrm{CTE}_X^p := \mathbb{E}[X f_X(X)], \tag{7.2}$$

where the function f_X defined by

$$f_X(x) := \frac{1}{\mathbb{P}(X > V_X^p)} \mathbb{1}_{\{x > V_X^p\}}, \qquad x \in \mathbb{R},$$

is a *distortion function* according to Definition 6.4.

The use of the strict inequality ">" in the definition of the Conditional Tail Expectation allows us to avoid any dependence on $\mathbb{P}(X = V_X^p)$, and to consider risky values strictly beyond V_X^p. The Conditional Tail Expectation is also called Conditional Value at Risk (CVaR).

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

```
library(quantmod)
getSymbols("^HSI",from="2013-06-01",to="2014-10-01",src="yahoo")
returns <- as.vector(diff(log(Ad(`HSI`))))
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 7.3. *Let X be a random variable X such that $\mathbb{P}\big(X > V_X^p\big) > 0$. For any $p \in (0,1]$ we have $\mathrm{CTE}_X^p > \mathbb{E}[X]$ and $\mathrm{CTE}_X^p > V_X^p$ with, more precisely,*

$$\mathrm{CTE}_X^p = \mathbb{E}\big[X \mid X > V_X^p\big] = V_X^p + \mathbb{E}\big[\big(X - V_X^p\big)^+ \mid X > V_X^p\big].$$

Proof. We have

$$\begin{aligned}
\mathbb{E}\big[X \mid X > V_X^p\big] &= \frac{1}{\mathbb{P}\big(X > V_X^p\big)} \mathbb{E}\big[X \mathbb{1}_{\{X > V_X^p\}}\big] \\
&= \frac{1}{\mathbb{P}\big(X > V_X^p\big)} \Big(\mathbb{E}\big[\big(X - V_X^p\big)\mathbb{1}_{\{X > V_X^p\}}\big] + V_X^p \mathbb{E}\big[\mathbb{1}_{\{X > V_X^p\}}\big]\Big) \\
&= \frac{1}{\mathbb{P}\big(X > V_X^p\big)} \Big(\mathbb{E}\big[\big(X - V_X^p\big)^+\big] + V_X^p \mathbb{P}\big(X > V_X^p\big)\Big) \\
&= V_X^p + \frac{1}{\mathbb{P}\big(X > V_X^p\big)} \mathbb{E}\big[\big(X - V_X^p\big)^+\big]
\end{aligned}$$

$$= V_X^p + \mathbb{E}\big[\big(X - V_X^p\big)^+ \,\big|\, X > V_X^p\big].$$

See Exercise 7.2-(d) for a proof of $\mathrm{CTE}_X^p > \mathbb{E}[X]$. □

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

$$\mathbb{P}\big(X > V_X^p\big) = 1 - p > 0$$

by (6.8) in Proposition 6.12.

Proposition 7.4. *Assume that* $\mathbb{P}\big(X = V_X^p\big) = 0$. *Then we have*

$$\mathrm{CTE}_X^p = \mathrm{TV}_X^p,$$

i.e.

$$\mathrm{CTE}_X^p = \mathbb{E}\big[X \mid X > V_X^p\big] = \mathbb{E}\big[X \mid X \geqslant V_X^p\big] = \frac{1}{1-p}\int_p^1 V_X^q dq = \mathrm{TV}_X^p. \tag{7.3}$$

Proof. By Lemma 6.17 we construct X as $X = V_X^U$ where U is uniformly distributed on $[0,1]$, with

$$U \geqslant p \Longrightarrow V_X^U \geqslant V_X^p \Longleftrightarrow X \geqslant V_X^p,$$

and

$$U \leqslant p \Longrightarrow V_X^U \leqslant V_X^p \Longleftrightarrow X \leqslant V_X^p,$$

hence

$$X > V_X^p \Longleftrightarrow V_X^U > V_X^p \Longrightarrow U > p.$$

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

$$U \geqslant p \Longleftrightarrow U > p \Longleftrightarrow V_X^U \geqslant V_X^p \Longleftrightarrow X \geqslant V_X^p \Longleftrightarrow X > V_X^p,$$

hence

$$\begin{aligned}
\mathrm{CTE}_X^p &= \mathbb{E}\big[X \mid X > V_X^p\big] \\
&= \mathbb{E}\big[V_X^U \mid V_X^U > V_X^p\big] \\
&= \mathbb{E}\big[V_X^U \mid U \geqslant p\big] \\
&= \frac{1}{\mathbb{P}(U \geqslant p)}\mathbb{E}\big[V_X^U \mathbb{1}_{\{U \geqslant p\}}\big] \\
&= \frac{1}{1-p}\int_p^1 V_X^q dq.
\end{aligned}$$

This version: January 10, 2024
https://personal.ntu.edu.sg/nprivault/indext.html

□

Figure 7.2 shows the locations of Value at Risk and Conditional Tail Expectation on a given data set. Note that here, the computation is done on sign-changed data according to Proposition 6.14, i.e. the results are computed according to the "practitioner" point of view.

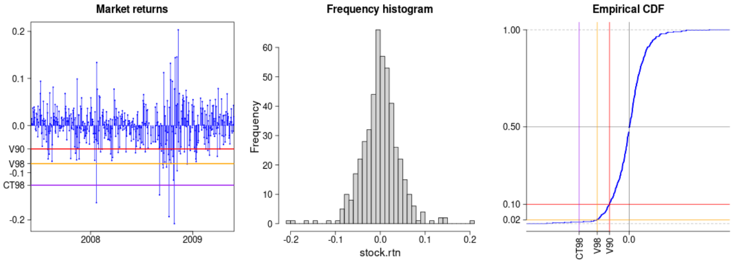


Fig. 7.2: Value at Risk and Conditional Tail Expectation.

The Conditional Tail Expectation of a Gaussian $\mathcal{N}(\mu, \sigma^2)$ random variable is computed in the next proposition, see also Proposition 6.15.

Proposition 7.5. Gaussian CTE. *Given* $X \simeq \mathcal{N}(\mu_X, \sigma_X^2)$ *we have*

$$\mathrm{CTE}_X^p = \mu_X + \frac{\sigma_X}{1-p}\phi(q_Z^p) = \mu_X + \frac{\sigma_X}{(1-p)\sqrt{2\pi}}\mathrm{e}^{-(q_Z^p)^2/2}, \tag{7.4}$$

where $q_Z^p = \Phi^{-1}(p)$ *is the Gaussian quantile of* $Z \simeq \mathcal{N}(0,1)$ *at the level* $p \in (0,1)$ *and*

$$\phi(z) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-z^2/2}, \qquad z \in \mathbb{R},$$

is the standard normal probability density function.

Proof. Using the relation $\mathbb{P}(X \geqslant V_X^p) = \mathbb{P}(X > V_X^p) = 1-p$, cf. Proposition 6.13, we have

$$\begin{aligned}
\mathrm{CTE}_X^p &= \mathrm{TV}_X^p \\
&= \mathbb{E}\big[X \mid X > V_X^p\big] \\
&= \frac{1}{\mathbb{P}(X > V_X^p)}\mathbb{E}\big[X \mathbb{1}_{\{X > V_X^p\}}\big] \\
&= \frac{1}{1-p}\int_{V_X^p}^\infty x\mathrm{e}^{-(x-\mu_X)^2/(2\sigma_X^2)}\frac{dx}{\sqrt{2\pi\sigma_X^2}}
\end{aligned}$$

$$= \frac{\mu_X}{1-p}\int_{V_X^p}^\infty \mathrm{e}^{-(x-\mu_X)^2/(2\sigma_X^2)}\frac{dx}{\sqrt{2\pi\sigma_X^2}} + \frac{1}{1-p}\int_{V_X^p}^\infty (x-\mu_X)\mathrm{e}^{-(x-\mu_X)^2/(2\sigma_X^2)}\frac{dx}{\sqrt{2\pi\sigma_X^2}}$$

$$= \frac{\mu_X}{1-p}\mathbb{P}(X \geqslant V_X^p) + \frac{\sigma_X^2}{(1-p)\sqrt{2\pi\sigma_X^2}}\left[-\mathrm{e}^{-(x-\mu_X)^2/(2\sigma_X^2)}\right]_{V_X^p}^\infty$$

$$= \mu_X + \frac{\sigma_X^2}{(1-p)\sqrt{2\pi\sigma_X^2}}\mathrm{e}^{-((V_X^p-\mu_X)/\sigma_X)^2/2}$$

$$= \mu_X + \frac{\sigma_X}{(1-p)\sqrt{2\pi}}\mathrm{e}^{-(q_Z^p)^2/2}$$

$$= \mu_X + \frac{\sigma_X}{1-p}\phi(q_Z^p),$$

due to the rescaling relation $V_X^p = \mu_X + \sigma_X q_Z^p$, cf. (6.12). □

7.3 Expected Shortfall (ES)

There are several variants for the definition of the Expected Shortfall ES_X^p. Next is a frequently used definition.

Definition 7.6. *The Expected Shortfall* ES_X^p *of a random variable* X *at the level* $p \in (0,1)$ *is defined by*

$$\mathrm{ES}_X^p := V_X^p + \frac{1}{1-p}\mathbb{E}[(X-V_X^p)^+]. \tag{7.5}$$

From Propositions 7.7 and 7.8, we deduce that

$$\mathrm{ES}_X^p = \begin{cases} \dfrac{1}{1-p}\mathbb{E}\big[X\mathbb{1}_{\{X\geqslant V_X^p\}}\big] = \mathbb{E}\big[X \mid X > V_X^p\big] = \mathrm{TV}_X^p & \text{if } \mathbb{P}\big(X = V_X^p\big) = 0, \\ \dfrac{1}{1-p}\mathbb{E}\big[X\mathbb{1}_{\{X\geqslant V_X^p\}}\big] + \dfrac{V_X^p}{1-p}\big(1-p-\mathbb{P}\big(X \geqslant V_X^p\big)\big) & \text{if } \mathbb{P}\big(X = V_X^p\big) \geqslant 0. \end{cases}$$

Proposition 7.7. *The Expected Shortfall of* X *at the level* $p \in (0,1)$ *can be written as*

$$\mathrm{ES}_X^p = \frac{1}{1-p}\mathbb{E}\big[X\mathbb{1}_{\{X\geqslant V_X^p\}}\big] + \frac{V_X^p}{1-p}\big(1-p-\mathbb{P}\big(X \geqslant V_X^p\big)\big).$$

Proof. By Proposition 6.2, we have

$$\mathrm{ES}_X^p = V_X^p + \frac{1}{1-p}\mathbb{E}[(X-V_X^p)^+]$$

This version: January 10, 2024
https://personal.ntu.edu.sg/nprivault/indext.html

$$
\begin{aligned}
&= V_X^p + \frac{1}{1-p}\mathbb{E}\big[(X - V_X^p)\mathbb{1}_{\{X \geqslant V_X^p\}}\big] \\
&= V_X^p + \frac{\mathbb{P}(X \geqslant V_X^p)}{1-p}\mathbb{E}\big[X - V_X^p \mid X \geqslant V_X^p\big] \\
&= V_X^p + \frac{1}{1-p}\big(\mathbb{E}\big[X\mathbb{1}_{\{X \geqslant V_X^p\}}\big] - V_X^p\mathbb{E}\big[\mathbb{1}_{\{X \geqslant V_X^p\}}\big]\big) \\
&= V_X^p + \frac{1}{1-p}\big(\mathbb{E}\big[X\mathbb{1}_{\{X \geqslant V_X^p\}}\big] - V_X^p\mathbb{P}(X \geqslant V_X^p)\big) \\
&= \frac{1}{1-p}\mathbb{E}\big[X\mathbb{1}_{\{X \geqslant V_X^p\}}\big] + \frac{V_X^p}{1-p}\big(1 - p - \mathbb{P}(X \geqslant V_X^p)\big).
\end{aligned}
$$

□

When $\mathbb{P}(X = V_X^p) = 0$, Proposition 7.7 also yields

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

Proposition 7.8. *When $\mathbb{P}(X = V_X^p) = 0$ the Expected Shortfall ES_X^p coincides with the* Conditional Tail Expectation CTE_X^p *and with the Tail Value at Risk TV_X^p, i.e., we have*

$$
\mathrm{ES}_X^p = \mathbb{E}\big[X \mid X > V_X^p\big] = \mathbb{E}\big[X \mid X \geqslant V_X^p\big] = \mathrm{TV}_X^p.
$$

Proof. By Relation (6.9) in Proposition 6.12, when $\mathbb{P}(X = V_X^p) = 0$ we have

$$
p = \mathbb{P}(X \leqslant V_X^p) \quad \text{and} \quad 1 - p = \mathbb{P}(X > V_X^p) = \mathbb{P}(X \geqslant V_X^p),
$$

hence

$$
\begin{aligned}
\mathrm{ES}_X^p &= \frac{1}{1-p}\mathbb{E}\big[X\mathbb{1}_{\{X \geqslant V_X^p\}}\big] \\
&= \frac{1}{1-p}\mathbb{E}\big[X\mathbb{1}_{\{X > V_X^p\}}\big] \\
&= \frac{1}{\mathbb{P}(X > V_X^p)}\mathbb{E}\big[X\mathbb{1}_{\{X > V_X^p\}}\big] \\
&= \mathbb{E}\big[X \mid X > V_X^p\big] \\
&= \mathrm{TV}_X^p,
\end{aligned}
$$

by Proposition 7.4. □

In particular, by Propositions 7.5 and 7.8, the Gaussian Expected Shortfall of $X \simeq \mathcal{N}(\mu, \sigma^2)$ at the level $p \in (0, 1)$ is also given by

$$\mathrm{ES}_X^p = \mathrm{CTE}_X^p = \mu + \frac{\sigma}{1-p}\phi(\Phi^{-1}(p)) = \mu + \frac{\sigma}{(1-p)\sqrt{2\pi}}\mathrm{e}^{-(\Phi^{-1}(p))^2/2}.$$

Proposition 7.9. *The Expected Shortfall* ES_X^p *at the level* $p \in (0,1)$ *can be written as the* distortion risk measure

$$\mathrm{ES}_X^p = \mathbb{E}[X f_X(X)], \tag{7.6}$$

where the function f_X *defined by*

$$f_X(x) := \frac{1}{1-p}\mathbb{1}_{\{x > V_X^p\}} + \mathbb{1}_{\{\mathbb{P}(X=V_X^p)>0\}}\frac{1-p-\mathbb{P}\big(X > V_X^p\big)}{(1-p)\mathbb{P}\big(X = V_X^p\big)}\mathbb{1}_{\{x=V_X^p\}},$$

$x \in \mathbb{R}$, *is a* distortion function.

Proof. By Proposition 7.7, we have

$$\begin{aligned}
\mathrm{ES}_X^p &= \frac{1}{1-p}\mathbb{E}\big[X\mathbb{1}_{\{X \geqslant V_X^p\}}\big] + \frac{V_X^p}{1-p}\big(1-p-\mathbb{P}\big(X \geqslant V_X^p\big)\big) \\
&= \frac{1}{1-p}\mathbb{E}\big[X\mathbb{1}_{\{X \geqslant V_X^p\}}\big] + \mathbb{1}_{\{\mathbb{P}(X=V_X^p)>0\}}\frac{V_X^p}{1-p}\big(1-p-\mathbb{P}\big(X \geqslant V_X^p\big)\big) \\
&= \frac{1}{1-p}\mathbb{E}\left[\left(\mathbb{1}_{\{X \geqslant V_X^p\}} + \mathbb{1}_{\{\mathbb{P}(X=V_X^p)>0\}}\frac{1-p-\mathbb{P}\big(X \geqslant V_X^p\big)}{\mathbb{P}\big(X = V_X^p\big)}\mathbb{1}_{\{X=V_X^p\}}\right)X\right] \\
&= \frac{1}{1-p}\mathbb{E}\Bigg[\Big(\mathbb{1}_{\{X \geqslant V_X^p\}} - \mathbb{1}_{\{\mathbb{P}(X=V_X^p)>0\}}\mathbb{1}_{\{X=V_X^p\}} \\
&\qquad + \mathbb{1}_{\{\mathbb{P}(X=V_X^p)>0\}}\frac{1-p-\mathbb{P}\big(X > V_X^p\big)}{\mathbb{P}\big(X = V_X^p\big)}\mathbb{1}_{\{X=V_X^p\}}\Big)X\Bigg] \\
&= \frac{1}{1-p}\mathbb{E}\left[\left(\mathbb{1}_{\{X > V_X^p\}} + \mathbb{1}_{\{\mathbb{P}(X=V_X^p)>0\}}\frac{1-p-\mathbb{P}\big(X > V_X^p\big)}{\mathbb{P}\big(X = V_X^p\big)}\mathbb{1}_{\{X=V_X^p\}}\right)X\right].
\end{aligned}$$

In order to show that f_X is a distortion function according to Definition 6.4, we check that

$$\begin{aligned}
\mathbb{E}[f_X(X)] &= \frac{1}{1-p}\mathbb{E}\left[\mathbb{1}_{\{X > V_X^p\}} + \mathbb{1}_{\{\mathbb{P}(X=V_X^p)>0\}}\frac{1-p-\mathbb{P}\big(X > V_X^p\big)}{\mathbb{P}\big(X = V_X^p\big)}\mathbb{1}_{\{X=V_X^p\}}\right] \\
&= \frac{1}{1-p}\left(\mathbb{E}\big[\mathbb{1}_{\{X > V_X^p\}}\big] + 1-p-\mathbb{P}\big(X > V_X^p\big)\right) \\
&= \frac{1}{1-p}\left(\mathbb{P}\big(X > V_X^p\big) + 1-p-\mathbb{P}\big(X > V_X^p\big)\right) \\
&= 1.
\end{aligned} \tag{7.7}$$

□

By Lemma 6.17 and Proposition 7.9, we also have

$$\mathrm{ES}_X^p = \int_0^1 V_X^q f_X(V_X^q) dq,$$

and we check that the distortion function f_X of Proposition 7.9 is a non-decreasing function that satisfies

$$f_X(x) \leqslant \frac{1}{1-p}, \qquad x \in \mathbb{R},$$

by (6.8). The following proposition, see 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_X^p.

Theorem 7.10. *The Expected Shortfall* ES_X^p *coincides with the Tail Value at Risk* TV_X^p *for any* $p \in (0,1)$*, i.e. we have*

$$\mathrm{ES}_X^p = \mathrm{TV}_X^p = \frac{1}{1-p} \int_p^1 V_X^q dq.$$

Proof. Constructing X as $X = V_X^U$ where U is uniformly distributed on $[0,1]$ as in Lemma 6.17, by Proposition 6.11 we have

$$U \geqslant p \Longrightarrow V_X^U \geqslant V_X^p \Longrightarrow X \geqslant V_X^p$$

and

$$\begin{aligned}
\big(U < p \text{ and } X \geqslant V_X^p\big) &\Longrightarrow \big(V_X^U \leqslant V_X^p \text{ and } X \geqslant V_X^p\big) \\
&\Longrightarrow \big(X \leqslant V_X^p \text{ and } X \geqslant V_X^p\big) \\
&\Longrightarrow X = V_X^p.
\end{aligned}$$

Hence by (7.5) and the relations

$$1 - p = \mathbb{E}\big[\mathbb{1}_{\{U \geqslant p\}}\big] \quad \text{and} \quad \mathbb{P}\big(X \geqslant V_X^p\big) = \mathbb{E}\big[\mathbb{1}_{\{X \geqslant V_X^p\}}\big],$$

we have

$$\begin{aligned}
V_X^p\big(1 - p - \mathbb{P}\big(X \geqslant V_X^p\big)\big) &= -V_X^p \mathbb{E}\big[\mathbb{1}_{\{X \geqslant V_X^p\}} - \mathbb{1}_{\{U \geqslant p\}}\big] \\
&= -V_X^p \mathbb{E}\big[\mathbb{1}_{\{X \geqslant V_X^p\} \setminus \{U \geqslant p\}}\big] \\
&= -V_X^p \mathbb{E}\big[\mathbb{1}_{\{X \geqslant V_X^p\} \cap \{U < p\}}\big] \\
&= -\mathbb{E}\big[X \mathbb{1}_{\{X \geqslant V_X^p\} \cap \{U < p\}}\big],
\end{aligned}$$

hence

$$\begin{aligned}
\mathrm{ES}_X^p &= \frac{1}{1-p}\mathbb{E}\big[X\mathbb{1}_{\{X\geqslant V_X^p\}}\big] + \frac{V_X^p}{1-p}\big(1-p-\mathbb{P}(X\geqslant V_X^p)\big) \\
&= \frac{1}{1-p}\mathbb{E}\big[X\mathbb{1}_{\{X\geqslant V_X^p\}}\big] - \frac{1}{1-p}\mathbb{E}\big[X\mathbb{1}_{\{X\geqslant V_X^p\}\cap\{U<p\}}\big] \\
&= \frac{1}{1-p}\mathbb{E}\big[V_X^U\mathbb{1}_{\{V_X^U\geqslant V_X^p\}}\big] - \frac{1}{1-p}\mathbb{E}\big[V_X^U\mathbb{1}_{\{V_X^U\geqslant V_X^p\}}\mathbb{1}_{\{U<p\}}\big] \\
&= \frac{1}{1-p}\mathbb{E}\big[V_X^U\mathbb{1}_{\{V_X^U\geqslant V_X^p\}}\mathbb{1}_{\{U\geqslant p\}}\big] \\
&= \frac{1}{1-p}\mathbb{E}\big[V_X^U\mathbb{1}_{\{U\geqslant p\}}\big] \\
&= \frac{1}{1-p}\int_p^1 V_X^q dq,
\end{aligned}$$

which is the Tail Value at Risk TV_X^p. □

Theorem 7.11. *Expected Shortfall* ES_X^p *and Tail Value at Risk* TV_X^p *are coherent risk measures.*

Proof. As ES_X^p coincides with TV_X^p for all $p\in(0,1)$ from Theorem 7.10, we can use either Relation (7.5) in Definition 7.6 or Relation (7.1) in Definition 7.1.

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

$$\begin{aligned}
\mathrm{ES}_X^p &= \mathrm{TV}_X^p \\
&= \frac{1}{1-p}\int_p^1 V_X^q dq \\
&\leqslant \frac{1}{1-p}\int_p^1 V_Y^q dq \\
&= \mathrm{TV}_Y^p \\
&= \mathrm{ES}_Y^p
\end{aligned}$$

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

$$\begin{aligned}
\mathrm{ES}_{\mu+\lambda X}^p &= \mathrm{TV}_{\mu+\lambda X}^p \\
&= \frac{1}{1-p}\int_p^1 V_{\mu+\lambda X}^q dq
\end{aligned}$$

This version: January 10, 2024
https://personal.ntu.edu.sg/nprivault/indext.html

$$
\begin{aligned}
&= \frac{1}{1-p}\int_p^1 (\mu + \lambda V_X^q) dq \\
&= \mu + \lambda \frac{1}{1-p}\int_p^1 V_X^q dq \\
&= \mu + \lambda \mathrm{TV}_Y^p \\
&= \mu + \lambda \mathrm{ES}_Y^p
\end{aligned}
$$

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

(iii) Sub-additivity. By Proposition 7.9, we have

$$
\begin{aligned}
&(1-p)\left(\mathrm{ES}_{X+Y}^p - \mathrm{ES}_X^p - \mathrm{ES}_Y^p\right) \\
&= \mathbb{E}[(X+Y)f_{X+Y}(X+Y)] - \mathbb{E}[Xf_X(X)] - \mathbb{E}[Yf_Y(Y)] \\
&= \mathbb{E}[X(f_{X+Y}(X+Y) - f_X(X))] + \mathbb{E}[Y(f_{X+Y}(X+Y) - f_Y(Y))] \\
&= V_X^p \mathbb{E}[f_{X+Y}(X+Y) - f_X(X)] + \mathbb{E}[(X - V_X^p)(f_{X+Y}(X+Y) - f_X(X))] \\
&\quad + V_Y^p \mathbb{E}[f_{X+Y}(X+Y) - f_Y(Y)] + \mathbb{E}[(Y - V_Y^p)(f_{X+Y}(X+Y) - f_Y(Y))] \\
&= (1-1)V_X^p + \mathbb{E}[(X - V_X^p)(f_{X+Y}(X+Y) - f_X(X))] \\
&\quad + (1-1)V_Y^p + \mathbb{E}[(Y - V_Y^p)(f_{X+Y}(X+Y) - f_Y(Y))] \\
&\leqslant 0,
\end{aligned}
$$

where we have used (7.7) and the following facts.

- When $x - V_X^p < 0$, we have

$$
\begin{aligned}
(1-p)(f_{X+Y}(x+y) - f_X(x)) &= \mathbb{1}_{\{x+y > V_{X+Y}^p\}} - \mathbb{1}_{\{x > V_X^p\}} \\
&\quad + \mathbb{1}_{\{\mathbb{P}(X+Y = V_{X+Y}^p) > 0\}} \frac{1 - p - \mathbb{P}(X+Y > V_{X+Y}^p)}{\mathbb{P}(X+Y = V_{X+Y}^p)} \mathbb{1}_{\{x+y = V_{X+Y}^p\}} \\
&\quad - \mathbb{1}_{\{\mathbb{P}(X = V_X^p) > 0\}} \frac{1 - p - \mathbb{P}\big(X > V_X^p\big)}{\mathbb{P}\big(X = V_X^p\big)} \mathbb{1}_{\{x = V_X^p\}} \\
&= \mathbb{1}_{\{x+y > V_{X+Y}^p\}} + \mathbb{1}_{\{\mathbb{P}(X+Y = V_{X+Y}^p) > 0\}} \frac{1 - p - \mathbb{P}(X+Y > V_{X+Y}^p)}{\mathbb{P}(X+Y = V_{X+Y}^p)} \mathbb{1}_{\{x+y = V_{X+Y}^p\}} \\
&\geqslant 0, \qquad x < V_X^p,
\end{aligned}
$$

  where we applied (6.8).
- When $x - V_X^p > 0$, we have

$$
\begin{aligned}
(1-p)(f_{X+Y}(x+y) - f_X(x)) &= \mathbb{1}_{\{x+y > V_{X+Y}^p\}} - \mathbb{1}_{\{x > V_X^p\}} \\
&\quad + \mathbb{1}_{\{\mathbb{P}(X+Y = V_{X+Y}^p) > 0\}} \frac{1 - p - \mathbb{P}(X+Y > V_{X+Y}^p)}{\mathbb{P}(X+Y = V_{X+Y}^p)} \mathbb{1}_{\{x+y = V_{X+Y}^p\}}
\end{aligned}
$$

$$
\begin{aligned}
&-\mathbb{1}_{\{\mathbb{P}(X=V_X^p)>0\}}\frac{1-p-\mathbb{P}\big(X>V_X^p\big)}{\mathbb{P}\big(X=V_X^p\big)}\mathbb{1}_{\{x=V_X^p\}} \\
&= \mathbb{1}_{\{x+y>V_{X+Y}^p\}} - \mathbb{1}_{\{x>V_X^p\}} \\
&\quad +\mathbb{1}_{\{\mathbb{P}(X+Y=V_{X+Y}^p)>0\}}\frac{1-p-\mathbb{P}(X+Y>V_{X+Y}^p)}{\mathbb{P}(X+Y=V_{X+Y}^p)}\mathbb{1}_{\{x+y=V_{X+Y}^p\}} \\
&\leqslant \mathbb{1}_{\{x+y>V_{X+Y}^p\}} - \mathbb{1}_{\{x>V_X^p\}} + \mathbb{1}_{\{x+y=V_{X+Y}^p\}} \\
&= \mathbb{1}_{\{x+y\geqslant V_{X+Y}^p\}} - \mathbb{1}_{\{x>V_X^p\}} \\
&\leqslant 0, \qquad x > V_X^p,
\end{aligned}
$$

where we applied (6.8).

□

Note that in general, the Conditional Tail Expectation is not a coherent risk measure when $\mathbb{P}\big(X = V_X^p\big) > 0$.

7.4 Numerical Estimates

We are using the PerformanceAnalytics R package, see also § 6.1.1 of Mina and Xiao (2001). In case we care about negative return values, Definitions 7.2 and 7.6 are replaced with

$$
\overline{\mathrm{CTE}}_X^p := \mathbb{E}\big[X \mid X < \overline{V}_X^p\big] = \frac{1}{\mathbb{P}\big(X < \overline{V}_X^p\big)}\mathbb{E}\big[X\mathbb{1}_{\{X<\overline{V}_X^p\}}\big]
$$

and

$$
\overline{\mathrm{ES}}_X^p := \overline{V}_X^p + \frac{1}{1-p}\mathbb{E}\big[\big(X-\overline{V}_X^p\big)\mathbb{1}_{\{X\leqslant\overline{V}_X^p\}}\big].
$$

From Proposition 6.14, when the cumulative distribution function F_X is continuous and strictly increasing we have

$$
\begin{aligned}
\overline{\mathrm{CTE}}_X^p &= \mathrm{ES}_X^p \\
&= \mathbb{E}\big[X \mid X < \overline{V}_X^p\big] \\
&= \mathbb{E}\big[X \mid X < -V_X^{1-p}\big] \\
&= \frac{1}{\mathbb{P}\big(X < -V_X^{1-p}\big)}\mathbb{E}\big[X\mathbb{1}_{\{X<-V_X^{1-p}\}}\big] \\
&= -\frac{1}{\mathbb{P}\big(-X > V_X^{1-p}\big)}\mathbb{E}\big[-X\mathbb{1}_{\{-X>V_X^{1-p}\}}\big]
\end{aligned}
$$

This version: January 10, 2024
https://personal.ntu.edu.sg/nprivault/indext.html

$$
\begin{aligned}
&= -\mathbb{E}\big[-X \mid -X > V_X^{1-p}\big] \\
&= -\overline{\mathrm{CTE}}_{-X}^{1-p} \\
&= -\mathrm{ES}_{-X}^{1-p}.
\end{aligned}
$$

```
library(PerformanceAnalytics)
ES(returns, p=.95, method="historical",invert="TRUE"))
ES(returns, p=.95, method="historical",invert="FALSE")
```

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

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

The Gaussian Expected Shortfall is given as -0.0191359 by

```
ES(returns, p=.95, method="gaussian",invert="FALSE")
ES(returns, p=.95, method="gaussian",invert="TRUE")
```

It can be recovered from (7.4) (after sign inversion) as

$$
\begin{aligned}
\mathrm{ES}_X^p &= -\mathrm{ES}_{-X}^{1-p} \\
&= -\left(-\mu + \frac{\sigma}{1-p}\phi(V_Z^{1-p})\right) \\
&= \mu - \frac{\sigma}{1-p}\phi(V_Z^p) \\
&= \mu - \frac{\sigma}{(1-p)\sqrt{2\pi}}\mathrm{e}^{-(V_Z^p)^2/2},
\end{aligned}
$$

i.e.

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

with output -0.01916536.

The attached **code 1** and **code 2** compute the Expected Shortfall from the practitioner and academic points of views, and compare their outputs to the that of the PerformanceAnalytics package, as illustrated in the next Figure 7.3.

```
> source("comparison.R")
Number of samples= 265
```

VaR 95 = -0.03420879 , Threshold= 0.9433962
CTE 95 = -0.04646176
ES 95 = -0.04623058
Historical VaR 95 0= -0.03316604
Gaussian VaR 95 = -0.03209374
Historical ES 95 = -0.04552403
Gaussian ES 95 = -0.04043227

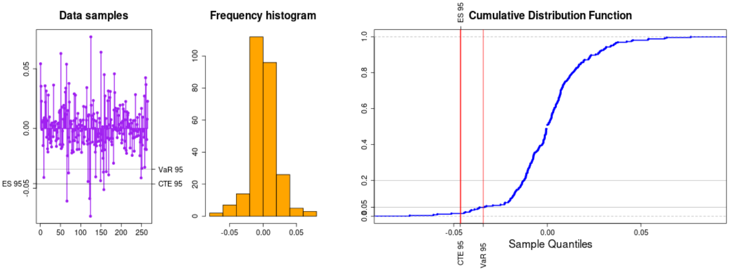


Fig. 7.3: Value at Risk and Expected Shortfall.

Value at Risk *vs.* Expected Shortfall

```
dev.new(width=16,height=8)
chart.VaRSensitivity(ts(returns),methods=c("HistoricalVaR","HistoricalES"),
      colorset=bluefocus, lwd=4)
```

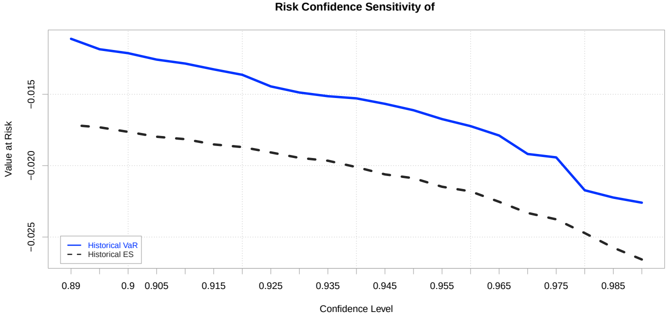


Fig. 7.4: Value at Risk *vs.* Expected Shortfall.

Historical *vs.* Gaussian risk measures

```
dev.new(width=16,height=8)
chart.VaRSensitivity(ts(returns),methods=c("HistoricalVaR","GaussianVaR"),
      colorset=bluefocus, lwd=4)
```

The next Figure 7.5 uses the above R code to compare the historical and Gaussian values at risk.

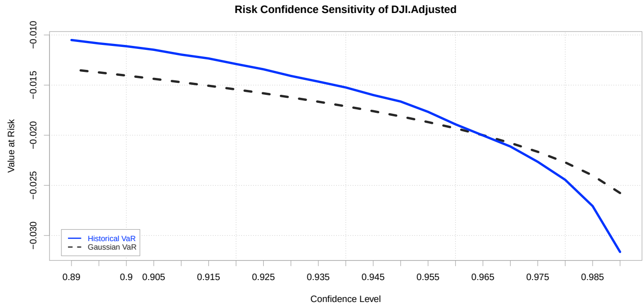


Fig. 7.5: Historical *vs.* Gaussian estimates of Value at Risk.

```
dev.new(width=16,height=8)
chart.VaRSensitivity(ts(returns),methods=c("HistoricalES","GaussianES"), colorset=bluefocus,
      lwd=4)
```

In the next Figure 7.6 we compare the Gaussian and historical estimates of Expected Shortfall.

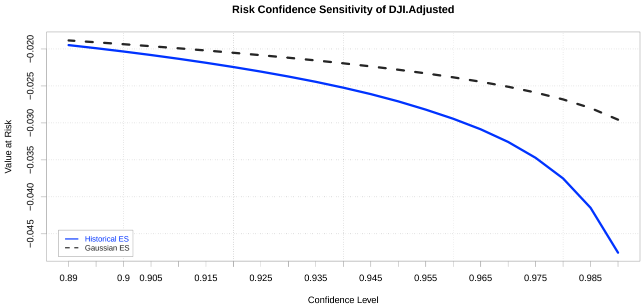


Fig. 7.6: Quantile function.

In Table 7.1 we summarize some properties of risk measures.

Risk Measure	Additivity	Homogeneity	Subadditivity	Coherence
V_X	✓	✓	✗	✗
CTE_X	✓	✓	✗	✗
TV_X	✓	✓	✓	✓
ES_X	✓	✓	✓	✓

Table 7.1: Summary of Risk Measures.

Note that Value at Risk V_X^p is *coherent* on Gaussian random variables according to Remark 6.16. Similarly, the Conditional Tail Expectation CTE_X^p

This version: January 10, 2024
https://personal.ntu.edu.sg/nprivault/indext.html

is *coherent* on random variables having a continuous CDF by Proposition 7.4 and Theorem 7.11.

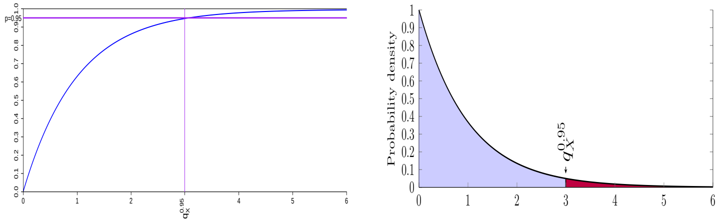
Exercises

Exercise 7.1 Let X denote an exponentially distributed random variable with parameter $\lambda > 0$, *i.e.* the distribution of X has the cumulative distribution function (CDF)

$$F_X(x) = \mathbb{P}(X \leqslant x) = 1 - \mathrm{e}^{-\lambda x}, \qquad x \geqslant 0,$$

and the probability density function (PDF)

$$f_X(x) = F_X'(x) = \lambda \mathrm{e}^{-\lambda x}, \qquad x \geqslant 0.$$



(a) Exponential quantile and CDF. (b) Exponential PDF.

a) Compute the conditional tail expectation

$$\mathbb{E}[X \mid X > \mathrm{VaR}_X^p] = \frac{1}{\mathbb{P}(X > \mathrm{VaR}_X^p)} \int_{\mathrm{VaR}_X^p}^{\infty} x f_X(x) dx.$$

b) Compute the tail value at risk

$$\mathrm{TV}_X^p = \frac{1}{1-p} \int_p^1 V_X^q dq.$$

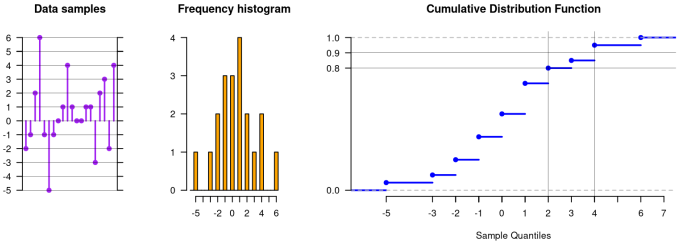
Exercise 7.2 Consider X an (integrable) random variable and $z \in \mathbb{R}$ such that $\mathbb{P}(X > z) > 0$.

a) Show that $\mathbb{E}[X \mid X > z] > z$.
b) Show that $\mathbb{E}[X \mid X > z] \geqslant \mathbb{E}[X]$.

This version: January 10, 2024
https://personal.ntu.edu.sg/nprivault/indext.html

c) Show that $\mathbb{E}[X \mid X > z] > \mathbb{E}[X]$ if $\mathbb{P}(X \leqslant z) > 0$.
d) Show that $\mathrm{CTE}_X^p > \mathbb{E}[X]$.

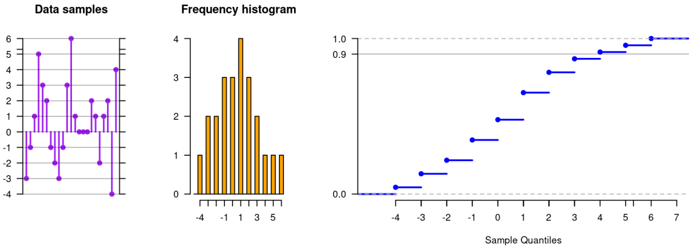
Exercise 7.3 Consider the following data set.



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

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

Exercise 7.4



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

a) VaR_X^{90},
b) $\mathbb{E}\big[X \mathbb{1}_{\{X > V_X^{90}\}}\big]$,
c) $\mathbb{P}\big(X > V_X^{90}\big)$,
d) $\mathrm{CTE}_X^{90} = \mathbb{E}\big[X \mid X > V_X^{90}\big] = \mathbb{E}\big[X \mathbb{1}_{\{X > V_X^{90}\}}\big] / \mathbb{P}\big(X > V_X^{90}\big)$,
e) $\mathbb{E}\big[X \mathbb{1}_{\{X \geqslant V_X^{90}\}}\big]$,

f) $\mathbb{P}(X \geqslant V_X^{90})$,

g) $\mathrm{ES}_X^{90} = \dfrac{1}{1-p}\big(\mathbb{E}\big[X \mathbb{1}_{\{X \geqslant V_X^{90}\}}\big] + V_X^{90}\big(1 - p - \mathbb{P}(X \geqslant V_X^{90})\big)\big)$,

h) $\mathrm{TV}_X^{90} = \dfrac{1}{1-p}\displaystyle\int_p^1 V_X^q dq$,

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

Exercise 7.5 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 $V_X^{98\%}$,

b) the Tail Value at Risk $\mathrm{TV}_X^{98\%}$,

c) the Conditional Tail Expectation $\mathbb{E}\big[X \mid X > V_X^{98\%}\big]$, and

d) the Expected Shortfall $E_X^{98\%}$.

Exercise 7.6 Consider two independent random variables X and Y with 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:

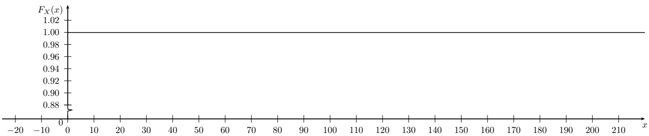


Fig. 7.8: Cumulative distribution function of X.

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

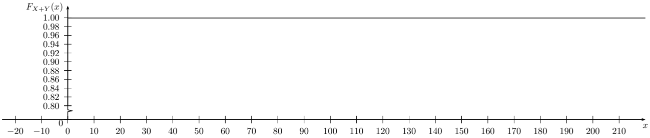


Fig. 7.9: Cumulative distribution function of $X+Y$.

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*

$$\mathrm{TV}_X^{90\%} := \frac{1}{1-p}\int_p^1 V_X^q dq$$

at the level $p = 90\%$.

e) Compute the *Tail Value at Risk*

$$\mathrm{TV}_{X+Y}^p := \frac{1}{1-p}\int_p^1 V_{X+Y}^q dq$$

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

Exercise 7.7 (Exercise 6.2 continued).

a) Compute the Tail Value at Risk

$$\mathrm{TV}_X^p := \frac{1}{1-p}\int_p^1 V_X^q dq$$

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

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

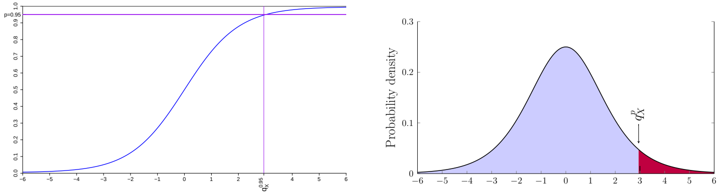
$$\mathrm{CTE}_X^{98\%} = \mathbb{E}\big[X \mid X > V_X^{98\%}\big] = \frac{1}{\mathbb{P}\big(X > V_X^p\big)}\mathbb{E}\left[X\mathbb{1}_{\{X>V_X^p\}}\right].$$

Exercise 7.8 We assume that the payoff X of a portfolio follows the standard logistic distribution with cumulative distribution function (CDF)

$$F_X(x) = \mathbb{P}(X \leqslant x) = \frac{1}{1+\mathrm{e}^{-x}}, \qquad x \in \mathbb{R},$$

and the probability density function (PDF)

$$f_X(x) = F_X'(x) = \frac{e^{-x}}{(1+e^{-x})^2}, \qquad x \in \mathbb{R}.$$



(a) Logistic quantile and CDF.

(b) Logistic PDF.

a) Compute the quantile $q_X^p = \mathrm{VaR}_X^p$ of X at any level $p \in [0,1]$, defined by the relation

$$F_X(q_X^p) = \mathbb{P}(X \leqslant \mathrm{VaR}_X^p) = p.$$

b) Compute the conditional tail expectation

$$\mathbb{E}[X \mid X > \mathrm{VaR}_X^p] = \frac{1}{\mathbb{P}(X > \mathrm{VaR}_X^p)} \int_{\mathrm{VaR}_X^p}^{\infty} x f_X(x) dx.$$

Hint. We have

$$\int_a^{\infty} \frac{x e^{-x}}{(1+e^{-x})^2} dx = \log(1+e^a) - \frac{a e^a}{1+e^a}, \qquad a \in \mathbb{R}.$$

c) Compute the tail value at risk

$$\mathrm{TV}_X^p = \frac{1}{1-p} \int_p^1 V_X^q dq.$$

Hint. We have $\int_p^1 \log q dq = p - 1 - p \log p$, $p \in (0,1)$.

Exercise 7.9

a) Show that for any random variable Z with probability density function $f_Z : \mathbb{R} \to \mathbb{R}_+$ we have

$$q\mathbb{P}(Z \geqslant q) \leqslant \mathbb{E}\big[Z \mathbb{1}_{\{Z \geqslant q\}}\big] = \int_q^{\infty} x f_Z(x) dx, \qquad q \geqslant 0. \tag{7.8}$$

b) Compute the left hand side and right hand side of the inequality (7.8) when $Z \simeq \mathcal{N}(0,1)$ has the standard normal distribution and q is the quantile q_Z^p of Z at the level $p \in [0,1]$.

c) Given $X \simeq \mathcal{N}(\mu_X, \sigma_X^2)$ a Gaussian random variable with mean μ_X and variance σ_X^2, show that the Gaussian Value at Risk

$$V_X^p = \mu_X + \sigma_X q_Z^p$$

is upper bounded by the Gaussian conditional tail expectation

$$\mathrm{CTE}_X^p = \mu_X + \frac{\sigma_X}{1-p}\phi(q_Z^p)$$

where

$$\phi(z) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-z^2/2}, \qquad z \in \mathbb{R},$$

is the standard normal probability density function.