Risk Analytics

Machine Learning and Optimization
for Data-Driven Decision Making

Fernando S. Oliveira

Draft version — April 22, 2026

Chapter 6
Tail Risk under Unknown Loss Distributions

Chapter 6

Tail Risk under Unknown Loss Distributions

6.1 Introduction

Chapters 4 and 5 developed downside-risk analysis when the loss distribution is known. Chapter 4 introduced VaR and CVaR in continuous settings and showed how they can be embedded into optimization problems. Chapter 5 then showed that, in discrete settings, additional care is required because probability mass may lie directly at the VaR threshold. The present chapter takes the next natural step and considers situations in which the loss distribution is not known in closed form.

This is the empirically relevant case in many practical applications. Managers rarely observe a fully specified probability law for losses. Instead, they have historical data, simulated outcomes, partial structural knowledge, forecasts, and judgment. Risk analysis must therefore move from a purely analytical exercise to a data-driven one in which the analyst constructs and defends a representation of uncertainty before measuring or optimizing risk [2, 13, 14, 8].

Unknown-distribution problems arise in procurement, inventory, energy systems, finance, disaster-risk management, and climate-related risk. A refinery may evaluate procurement policies using scenario-based price and margin simulations [9]. Strategic reserve management and energy trading problems rely on uncertain future prices, demand, weather, and operational disruptions rather than on a known distribution [12, 10, 11]. In disaster and climate settings, historical observations are often sparse precisely where the most important tail losses occur, so scenario design and stress analysis become central [4, 5].

The main message of this chapter is that VaR, TCE, and CVaR remain conceptually meaningful when the loss distribution is unknown, but their implementation becomes empirical, scenario-based, and subject to estimation uncertainty. More importantly for practice, once the analyst works with a finite set of scenarios, CVaR can be reformulated through auxiliary variables and scenario-wise constraints, so that the resulting problem can be solved as a linear program whenever the loss function is linear, or can be linearized, in the decision variables [13, 15, 1]. This is the practical computational approach adopted in much of the empirical literature and the one emphasized in this chapter.

The chapter proceeds as follows. Section 6.2 develops empirical VaR and CVaR for unknown loss distributions and shows how empirical CVaR is connected to a finite-scenario linear-programming representation. Section 6.3 presents the corresponding data-driven tail-risk optimization model. Section 6.4 explains how scenarios are generated for this purpose. Section 6.5 illustrates the framework with a newsvendor example solved using the same methodology. Section 6.6 concludes the chapter with the main lessons and key takeaways.

6.2 Empirical VaR and CVaR under Unknown Loss Distributions

Let

$$L(q, \xi)$$

denote the loss generated by decision q when uncertainty ξ is realized. In Chapters 4 and 5, the distribution of ξ, and therefore of $L(q, \xi)$, was assumed known. Here that assumption is relaxed. The analyst now treats the distribution of ξ as unknown and approximates it using data or finite scenarios.

The population definitions of the tail-risk measures remain unchanged. For a confidence level $\alpha \in (0, 1)$,

$$\mathrm{VaR}_\alpha(L) = \inf\{\ell \in \mathbb{R} : P(L \le \ell) \ge \alpha\}.$$

VaR identifies the threshold beyond which only the worst $1 - \alpha$ fraction of outcomes remain. CVaR then measures the average severity of losses in that adverse tail [14, 8]. What changes in the present chapter is that these quantities must be estimated from finite information.

The most direct approach is historical simulation. For a fixed decision q, suppose the analyst observes

$$\xi_1, \xi_2, \ldots, \xi_N$$

and computes the induced losses

$$L_1(q), L_2(q), \dots, L_N(q), \qquad L_i(q) = L(q, \xi_i).$$

If the losses are sorted in increasing order,

$$L_{(1)}(q) \le \cdots \le L_{(N)}(q),$$

a natural empirical estimator of VaR is the empirical quantile

$$\widehat{\mathrm{VaR}}_\alpha(q) = L_{(\lceil N\alpha \rceil)}(q).$$

This estimator is intuitive, but it also reveals a crucial point: the empirical distribution is discrete by construction, even if the unknown underlying distribution is continuous. For that reason, the discrete-tail logic of Chapter 5 remains directly relevant. In particular, TCE and exact CVaR need not coincide in finite samples whenever there is probability mass at the empirical VaR threshold [14, 8].

A first empirical tail measure is the tail conditional expectation,

$$\widehat{\mathrm{TCE}}_\alpha(q) = \frac{\sum_{i: L_i(q) \ge \widehat{\mathrm{VaR}}_\alpha(q)} L_i(q)}{\#\{i : L_i(q) \ge \widehat{\mathrm{VaR}}_\alpha(q)\}}.$$

However, exactly as in Chapter 5, this may average over more than the intended worst $1-\alpha$ probability mass if many observations lie exactly at the VaR point. To correct for this, exact empirical CVaR can be written as

$$\widehat{\mathrm{CVaR}}_\alpha(q) = \frac{\frac{1}{N}\sum_{i=1}^{N} L_i(q)\mathbf{1}\{L_i(q) > v_\alpha(q)\} + v_\alpha(q)\big(\widehat{F}_q(v_\alpha(q)) - \alpha\big)}{1-\alpha},$$

where $\mathbf{1}\{A\}$ denotes the indicator function, equal to 1 if the event A is true and 0 otherwise, and

$$v_\alpha(q) = \widehat{\mathrm{VaR}}_\alpha(q), \qquad \widehat{F}_q(\ell) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}\{L_i(q) \le \ell\}.$$

All losses strictly above the VaR threshold are included in full, and only the fraction of probability mass needed at the threshold is retained. This is the exact empirical analogue of the discrete CVaR formula developed in Chapter 5.

In many applications the scenarios are not equally weighted. Historical observations may be reweighted, recent periods may receive greater importance, or simulated scenarios may come with assigned probabilities. If the analyst has losses

$$L_1(q), \dots, L_N(q)$$

with probabilities

$$p_1, \dots, p_N, \qquad \sum_{n=1}^{N} p_n = 1,$$

then the weighted empirical distribution is

$$\widehat{F}_q(\ell) = \sum_{n=1}^{N} p_n \mathbf{1}\{L_n(q) \leq \ell\},$$

and the weighted empirical VaR is

$$\widehat{\mathrm{VaR}}_\alpha(q) = \inf \left\{ \ell \in \mathbb{R} : \sum_{n=1}^{N} p_n \mathbf{1}\{L_n(q) \leq \ell\} \geq \alpha \right\}.$$

This weighted setting is especially useful because it unifies equally weighted historical simulation, recency-weighted historical simulation, and model-based finite-scenario approximations within one framework [17].

For optimization, the most important representation is the Rockafellar–Uryasev formula

$$\mathrm{CVaR}_\alpha(L) = \min_{\eta \in \mathbb{R}} \left\{ \eta + \frac{1}{1-\alpha} \mathbb{E}\big[(L-\eta)^+\big] \right\}.$$

When the distribution is unknown and replaced by a finite scenario set, the expectation is replaced by a scenario-weighted average, yielding

$$\widehat{\mathrm{CVaR}}_\alpha(q) = \min_{\eta \in \mathbb{R}} \left\{ \eta + \frac{1}{1-\alpha} \sum_{n=1}^{N} p_n (L_n(q) - \eta)^+ \right\}.$$

For practical implementation, this is usually converted into a linear-programming representation by introducing nonnegative auxiliary variables u_n satisfying

$$u_n \geq L_n(q) - \eta, \qquad u_n \geq 0, \qquad n = 1, \dots, N.$$

Then empirical CVaR can be written as

$$\widehat{\mathrm{CVaR}}_\alpha(q) = \min_{\eta, u_1, \dots, u_N} \left\{ \eta + \frac{1}{1-\alpha} \sum_{n=1}^{N} p_n u_n \right\}$$

subject to the scenario constraints above. This representation is operationally central because it transforms empirical CVaR minimization into a linear or convex optimization problem whenever the loss function itself can be expressed linearly, or piecewise linearly, in the decision variables [13, 14, 15].

Because the distribution is unknown, the tail estimators themselves are uncertain. Empirical VaR and CVaR can vary substantially across samples, especially when α is high and only a few observations determine the tail. This makes bootstrap analysis useful as a diagnostic tool. By resampling the observed data and recomputing the tail measures or the resulting optimal policy, the analyst can assess not only outcome uncertainty but also the stability of the estimated risk measures themselves [7, 17]. The practical conclusion is that unknown-distribution tail-risk analysis is not a single technique, but a family of closely related empirical methods: historical simulation, weighted empirical estimation, finite-scenario reformulation, and bootstrap-based stability assessment.

6.3 Data-Driven Tail-Risk Optimization

Under an unknown loss distribution, the decision maker still faces the same basic problem as in Chapter 4: choose a feasible decision before uncertainty resolves. Let $q \in \mathcal{Q}$ denote the decision variable, and let $L(q, \xi)$ denote the induced loss. A natural downside-sensitive objective is

$$q^*_{\lambda,\alpha} \in \arg\min_{q \in \mathcal{Q}} \left\{ \mathbb{E}[L(q,\xi)] + \lambda \, \mathrm{CVaR}_\alpha(L(q,\xi)) \right\}, \qquad \lambda \geq 0.$$

This mean–CVaR structure preserves a transparent interpretation. The first term measures average performance; the second penalizes severe upper-tail losses. The parameter λ governs how strongly the decision maker values tail protection relative to expected performance [13, 15].

When the distribution of ξ is unknown, the expectation and CVaR must be approximated from finite scenarios. Suppose the analyst has scenarios

$$\xi^1, \ldots, \xi^N$$

with probabilities

$$p_1, \ldots, p_N, \qquad \sum_{n=1}^{N} p_n = 1.$$

Let

$$L_n(q) = L(q, \xi^n).$$

Then the data-driven mean–CVaR problem becomes

$$q^*_{\lambda,\alpha} \in \arg\min_{q \in \mathcal{Q}} \left\{ \sum_{n=1}^{N} p_n L_n(q) + \lambda \, \widehat{\mathrm{CVaR}}_\alpha(q) \right\}.$$

In practical empirical work, this is not usually solved by directly evaluating the plus function inside the objective. Instead, the problem is reformulated as a mathematical program with scenario-specific auxiliary variables. Introducing the threshold variable η and excess-loss variables u_n, the empirical mean–CVaR problem becomes

$$(q^*_{\lambda,\alpha}, \eta^*, u^*) \in \arg\min_{q,\eta,u} \left\{ \sum_{n=1}^{N} p_n L_n(q) + \lambda \left[\eta + \frac{1}{1-\alpha} \sum_{n=1}^{N} p_n u_n \right] \right\},$$

subject to

$$u_n \geq L_n(q) - \eta, \qquad u_n \geq 0, \qquad n = 1, \ldots, N,$$

together with the feasibility constraints defining $q \in \mathcal{Q}$.

This is the practical formulation most often implemented in empirical CVaR optimization. The key advantage is that the tail measure is handled through a set of scenario constraints rather than through direct manipulation of order statistics. When the loss function $L_n(q)$ is linear, or can be linearized through auxiliary variables, the full model becomes a linear program. This is exactly the type of formulation used in empirical and scenario-based applications of risk-averse stochastic programming, including the papers motivating this chapter [9, 10, 11].

The same idea remains useful even when the original economic problem is nonlinear. In many applications, the nonlinearities can be isolated and represented through auxiliary variables, piecewise-linear approximations, or other standard reformulations, while the CVaR component still enters through linear scenario constraints. This is one of the main reasons the Rockafellar–Uryasev representation is so important operationally: it converts downside-risk control into a constraint-based optimization structure that is much easier to solve than direct VaR minimization [13, 1].

The resulting optimizer is not the exact solution to the unknown-distribution problem. It is the solution to a finite approximation of that problem. This is precisely the sample-average approximation perspective developed in stochastic programming: the analyst solves an empirical optimization problem and studies how its solution depends on the available sample or scenario set [7, 18, 17]. In practice, this also makes clear why scenario quality matters so much. A data-driven optimum is only as meaningful as the empirical approximation on which it is based.

In some applications the analyst may not wish to trust a single empirical distribution completely. Data may be sparse, the environment may be changing, or severe tail scenarios may be underrepresented. In such cases, an ambiguity-aware extension is natural:

$$q^* \in \arg\min_{q \in \mathcal{Q}} \sup_{P \in \mathcal{P}} \left\{ \mathbb{E}_P[L(q, \xi)] + \lambda \ \mathrm{CVaR}_\alpha^P(L(q, \xi)) \right\},$$

where $\mathcal{P}$ is a family of plausible distributions around the empirical model. This robust extension is useful when direct reliance on the raw empirical distribution would be unjustified [6, 16].

The framework applies far beyond finance. In operations, procurement, energy, and reserve management, decision makers frequently combine expected performance with explicit tail control in exactly this way [9, 12, 10, 11]. The general lesson is that once the loss distribution is unknown, tail-risk optimization becomes inseparable from empirical approximation, scenario design, and the practical construction of a scenario-based mathematical program.

6.4 Scenario Generation

Once the loss distribution is unknown, scenario generation becomes the operational bridge between abstract uncertainty and usable decision inputs. Historical simulation is the natural baseline: past realizations of the uncertainty drivers are treated as scenarios, and the resulting loss sample is used to compute empirical VaR, empirical CVaR, and the mean–CVaR objective. This approach is transparent and nonparametric, but it is limited by the representativeness of the historical record and by the fact that it cannot generate losses more extreme than those already observed [8, 17].

A useful refinement is weighted historical simulation, in which some observations receive more weight than others. This is appropriate when recent periods are judged more informative, when volatility has changed, or when domain knowledge suggests that some observations should matter more than others. Weighted historical simulation remains within the empirical framework of this chapter while allowing the analyst to incorporate judgment explicitly through scenario probabilities.

A second approach is model-based simulation. Here the analyst estimates a probabilistic model for the uncertainty drivers and simulates a finite scenario set from that model. This can generate a richer scenario set than the historical sample alone, especially when the analyst wants to represent dependence, trend, or changing volatility. Its strength is flexibility; its weakness is model risk. If the model fits the tail poorly, the resulting empirical CVaR will also be poor [8, 17].

Bootstrap methods are useful because they address not only future uncertainty but also estimation uncertainty. By repeatedly resampling the available data and recomputing the tail measures or optimal policy, the analyst can assess whether the recommendation is stable or whether it changes sharply across nearby pseudo-samples. Stress augmentation plays a complementary role. Historical and model-based scenarios often underrepresent severe but plausible losses, so the analyst may deliberately add or overweight adverse scenarios to ensure that economically important tail events remain visible in the empirical distribution. When rare events dominate the tail, tail-focused refinements such as importance sampling can also improve the numerical quality of empirical tail estimation [7, 3].

For the purposes of this chapter, the key point is simple. Scenario generation is not a separate theoretical topic here. It is the practical stage at which an unknown distribution becomes a finite object that can support empirical VaR, empirical CVaR, and one-period tail-risk optimization. In implementation terms, it is the stage at which scenarios, probabilities, and scenario-wise losses are generated so that the linear-programming formulation of Section 6.3 can be solved.

6.5 Numerical Illustration and Sensitivity Analysis: A Newsvendor with Unknown Demand Distribution

The newsvendor problem provides a natural setting for illustrating tail-risk optimization under unknown loss distributions. Its economic structure is simple, but the trade-off it captures is fundamental: ordering too little exposes the firm to shortage losses in high-demand states, while ordering too much creates overage losses in ordinary states. Chapters 4 and 5 analyzed this trade-off under known continuous and discrete demand distributions. In the present chapter, the same structure is retained, but demand is represented empirically rather than through a known probability law.

Consider a retailer that must choose an order quantity q of portable battery backup kits before a peak outage season. The loss function is

$$L(q, D) = C_o(q - D)^+ + C_u(D - q)^+,$$

where D is uncertain demand, C_o is the overage loss per excess unit, and C_u is the underage loss per unmet unit. This piecewise-linear specification is stylized, but it is economically transparent: each unsold unit generates overage cost, while each unmet unit generates shortage cost.

Worked Example: Portable Battery Backup Kits

A retailer must choose an order quantity q for portable battery backup kits before the outage season begins. To preserve continuity with Chapter 4, the economic parameters are

$$C_o = 230, \qquad C_u = 360.$$

Unlike Chapter 4, demand is not assumed to follow a known parametric distribution. Instead, the analyst works with an empirical sample of 1000 demand scenarios constructed to reflect two market regimes: a regular season with moderate demand and a surge season with much higher demand, together with a small proportion of extreme stress scenarios. The resulting empirical demand distribution is therefore asymmetric and bimodal.

The purpose of the example is to show how empirical tail-risk optimization works when the data themselves define the loss distribution. The baseline preference parameters are

$$(\lambda, \alpha) = (0.6, 0.90).$$

Figure 6.1 shows the empirical demand sample used in the analysis. Most observations correspond to ordinary seasons, a second concentration reflects surge-demand periods, and a thinner right tail captures stress events. This is precisely the kind of setting in which a single standard parametric distribution would be too restrictive and an empirical scenario-based representation becomes more informative.

Given the empirical demand sample $\{d_1, \ldots, d_N\}$ with $N = 1000$, the analyst solves a scenario-based optimization model in which the order quantity q, the threshold η, and the excess-loss variables u_i are chosen jointly:

$$(q^*_{\lambda,\alpha}, \eta^*, u^*) \in \arg\min_{q,\eta,u} \left\{ \frac{1}{N} \sum_{i=1}^{N} L(q, d_i) + \lambda \left[\eta + \frac{1}{(1-\alpha)N} \sum_{i=1}^{N} u_i \right] \right\},$$

subject to

$$u_i \geq L(q, d_i) - \eta, \qquad u_i \geq 0, \qquad i = 1, \ldots, N.$$

This is the empirical linear-programming formulation of mean–CVaR optimization specialized to the newsvendor setting.

Table 6.1 reports empirical expected loss, empirical VaR, exact empirical CVaR, and the baseline mean–CVaR objective for three benchmark order

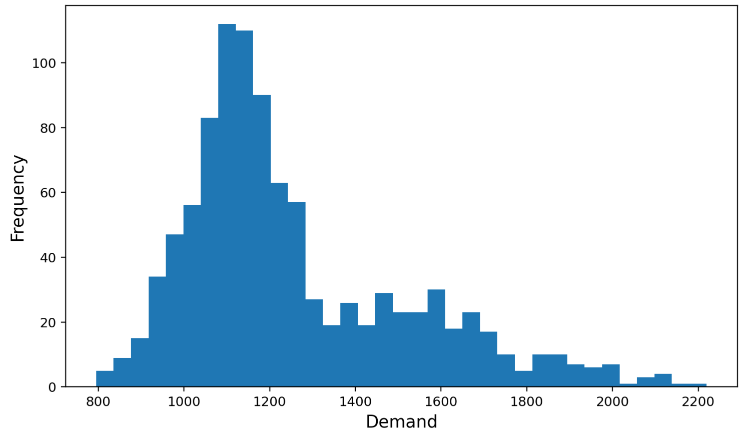

Figure 6.1: Empirical demand sample used in the newsvendor illustration. The sample contains 1000 observations and is asymmetric and bimodal, reflecting regular seasons, surge seasons, and a small stress tail.

quantities, each associated with a different decision criterion. The first,

$$q_E^* = 1230.00,$$

minimizes empirical expected loss alone. The second,

$$q_{0.6,0.90}^* = 1417.80,$$

is the exact LP-based solution to the baseline mean–CVaR problem at $(\lambda, \alpha) = (0.6, 0.90)$. The third,

$$q_{\text{CVaR},0.90}^* = 1506.37,$$

minimizes empirical $\text{CVaR}_{0.90}$ alone.

Table 6.1: Illustrative empirical performance measures for benchmark order quantities in the unknown-demand newsvendor example.

q	$\widehat{\mathbb{E}}[L(q,D)]$	$\text{VaR}_{0.90}$	$\text{CVaR}_{0.90}$	$\widehat{\mathbb{E}}[L] + 0.6\,\text{CVaR}$
1230.00	61,232.84	152,280.00	214,315.20	189,821.96
1417.80	71,204.11	113,113.22	155,077.26	164,250.46
1506.37	80,225.88	126,585.76	150,252.86	170,377.60

The logic of the table is clear. The risk-neutral solution $q_E^* = 1230.00$ achieves the lowest expected loss, but it performs poorly in the tail, with substantially larger VaR and CVaR. At the other extreme, the CVaR-minimizing solution $q^*_{\text{CVaR},0.90} = 1506.37$ reduces tail severity, but at the cost of higher average loss. The baseline mean–CVaR solution $q^*_{0.6,0.90} = 1417.80$ lies between these two extremes and therefore represents a compromise between average efficiency and tail protection.

Figure 6.2 and Figure 6.3 establish this baseline case visually. Figure 6.2 shows the empirical loss distribution, together with the corresponding empirical $\text{VaR}_{0.90}$ and $\text{CVaR}_{0.90}$. Figure 6.3 then shows the baseline mean–CVaR objective as a function of q at $(\lambda, \alpha) = (0.6, 0.90)$. Together, the figure and the table show that the preferred solution lies in an interior region: ordering too little leaves the retailer exposed to severe shortage losses, while ordering too much increases routine overage costs.

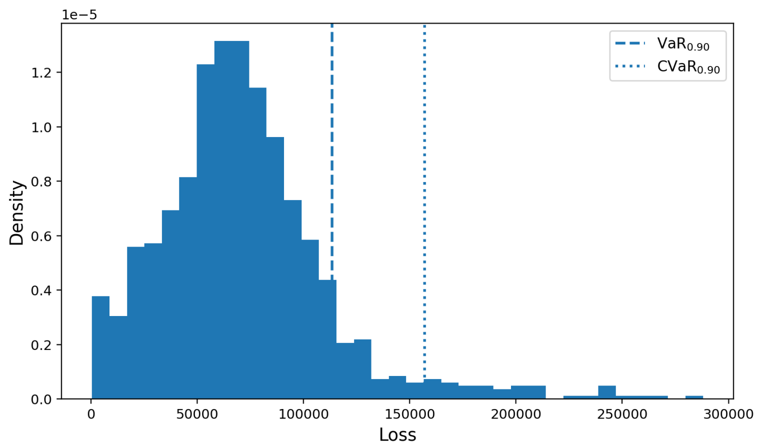

Figure 6.2: Empirical loss distribution for the unknown-demand newsvendor. The dashed line marks $\text{VaR}_{0.90}(L(q, D))$, while the dotted line marks $\text{CVaR}_{0.90}(L(q, D))$.

The next question is how the recommended policy changes when the decision maker becomes more or less tail-sensitive. Sensitivity to λ answers that question directly. Figure 6.4 shows how the mean–CVaR objective changes as larger weights are placed on downside risk, while Figure 6.5 shows the corresponding effect on the optimal order quantity q^*. When $\lambda = 0$, the problem is risk-neutral and the solution minimizes expected loss alone. As λ increases, the optimization places more weight on severe shortage scenarios, and the recommended order quantity rises. Economically, the decision maker accepts somewhat larger routine overage losses in exchange for lower exposure to adverse tail outcomes.

Sensitivity to α addresses a different dimension of tail preference. Instead

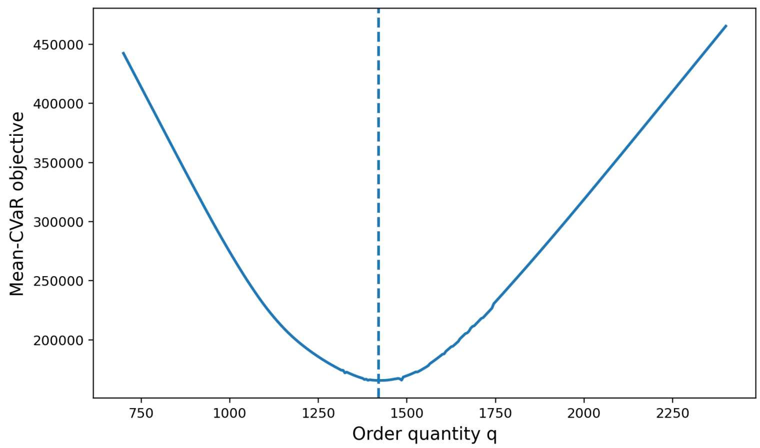

Figure 6.3: Mean–CVaR objective as a function of q in the empirical unknown-demand newsvendor example, evaluated at $(\lambda, \alpha) = (0.6, 0.90)$.

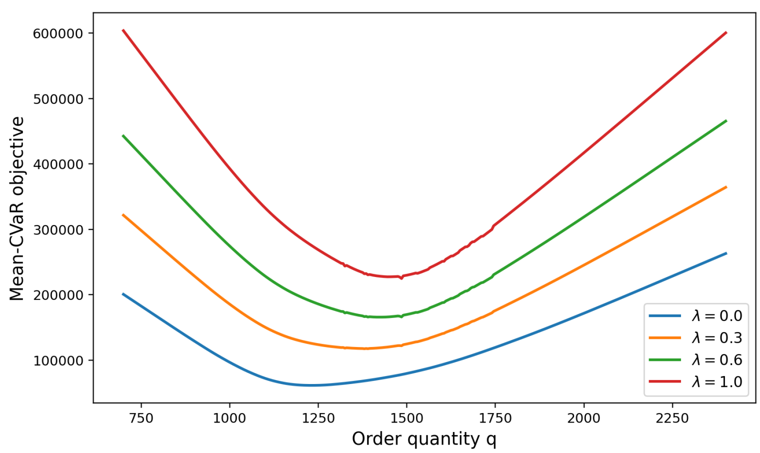

Figure 6.4: Mean–CVaR objective as a function of q for different values of λ, holding $\alpha = 0.90$ fixed.

of changing the weight placed on CVaR, it changes the depth of the tail being emphasized. Figure 6.6 shows that larger values of α place greater emphasis on increasingly extreme losses in the objective. Figure 6.7 then shows how the optimal order quantity responds. As α increases, the optimization focuses more heavily on the worst empirical scenarios, and the recommended order quantity rises accordingly. The curve in Figure 6.7 should therefore be

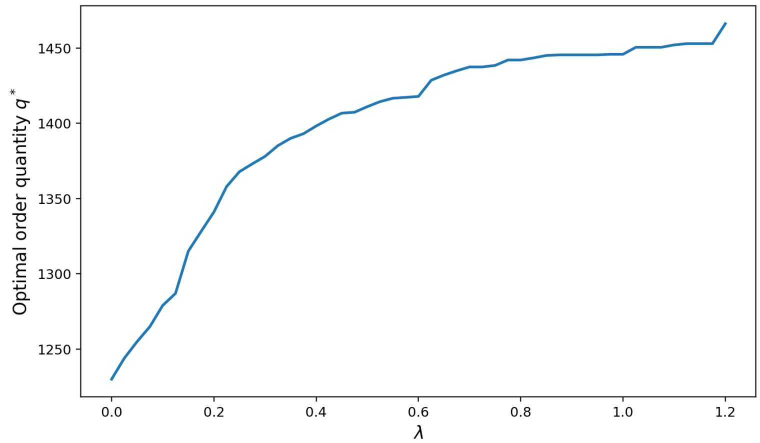

Figure 6.5: Optimal order quantity q^* as a function of λ, holding $\alpha = 0.90$ fixed.

interpreted as a finite-sample sensitivity result rather than as a deterministic comparative-statics law.

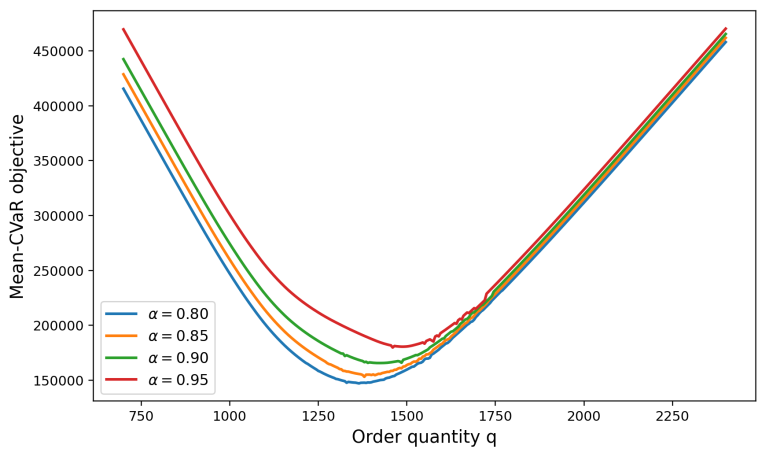

Figure 6.6: Mean–CVaR objective as a function of q for different values of α, holding $\lambda = 0.6$ fixed.

Finally, Table 6.2 examines how the *baseline mean–CVaR solution* changes with sample size. For each value of N, the table reports the LP-based

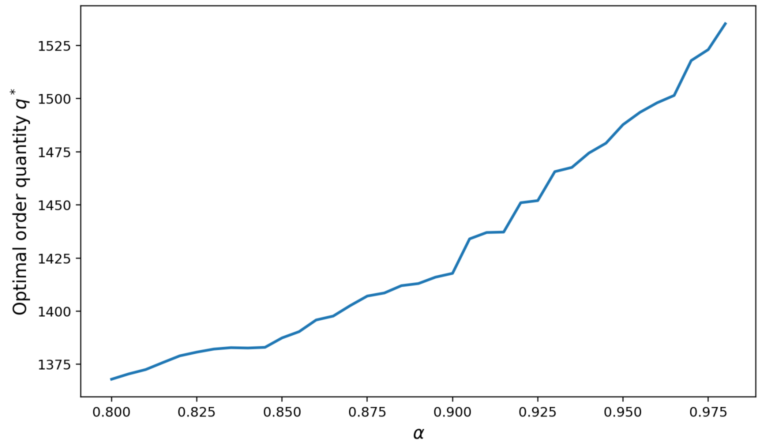

Figure 6.7: Optimal order quantity q^* as a function of α, holding $\lambda = 0.6$ fixed.

optimum

$$q^*_{0.6,0.90} \in \arg\min_q \left\{ \widehat{\mathbb{E}}[L(q,D)] + 0.6\,\widehat{\mathrm{CVaR}}_{0.90}(L(q,D)) \right\},$$

together with the corresponding empirical expected loss and empirical $\mathrm{CVaR}_{0.90}$ evaluated at that optimum. The table does not compare different optimization criteria. Instead, it shows how the estimated baseline tail-sensitive policy changes when the amount of available empirical information changes.

Table 6.2: Sensitivity of the baseline mean–CVaR optimum to effective sample size. For each sample size N, $q^*_{0.6,0.90}$ denotes the LP solution to the empirical objective $\widehat{\mathbb{E}}[L(q,D)] + 0.6\,\widehat{\mathrm{CVaR}}_{0.90}(L(q,D))$.

Sample size N	$q^*_{0.6,0.90}$	$\widehat{\mathbb{E}}[L(q^*_{0.6,0.90}, D)]$	$\mathrm{CVaR}_{0.90}(L(q^*_{0.6,0.90}, D))$
50	1432.61	66,382.94	155,348.34
100	1397.42	64,999.96	150,177.46
250	1412.68	67,609.25	151,440.33
1000	1417.80	71,204.11	157,339.52

The purpose of Table 6.2 is not to demonstrate monotonic convergence. Rather, it shows that the estimated baseline policy remains sample-dependent when the available data are limited. The broader lesson is that, under unknown demand distributions, the analyst is not only choosing a risk measure but also deciding how much trust to place in finite empirical information. When the sample is limited, this sample dependence becomes more pronounced, which is why robustness checks and resampling-based analysis are valuable.

6.6 Summary and Key Takeaways

This chapter extended the downside-risk framework of Chapters 4 and 5 to situations in which the loss distribution is not known parametrically. Its central message is that VaR, TCE, and CVaR remain conceptually meaningful under distributional uncertainty, but their practical implementation becomes empirical, scenario-based, and subject to estimation uncertainty.

The chapter first showed how empirical VaR and exact empirical CVaR can be constructed from finite samples of losses. It emphasized that once the distribution is represented empirically, the resulting loss distribution is discrete by construction, so the distinction between TCE and exact CVaR developed in Chapter 5 remains directly relevant. The chapter then reformulated downside-sensitive decision-making as a data-driven optimization problem and showed that, in finite-scenario settings, empirical CVaR can be handled through auxiliary variables and scenario-wise constraints. This yields a practical LP representation that makes empirical mean–CVaR optimization computationally tractable in many applications.

The chapter next explained that scenario generation is not a secondary technical detail, but part of the risk model itself. Historical simulation, weighted empirical scenarios, model-based simulation, bootstrap methods, and stress augmentation were presented as alternative ways of constructing the uncertainty representation on which empirical VaR and CVaR depend and from which the scenario-based optimization model is built. The newsvendor illustration then showed how the recommended decision changes once demand is treated empirically rather than through a known closed-form law, and how the resulting policy depends not only on the preference parameters λ and α, but also on the amount of information available for tail estimation.

From a managerial perspective, the main lesson is that unknown loss distributions are the rule rather than the exception. In practice, tail-risk measurement is inseparable from the way scenarios are constructed, weighted, and stress-tested. Sensitivity analysis therefore plays a substantive role: it helps reveal whether a recommended policy is robust to plausible changes in tail emphasis and empirical information, or whether it is fragile because it depends too heavily on limited data.

More broadly, the chapter completes the downside-risk block of Part II by moving from known distributions to empirical, scenario-based, and computationally implementable risk analysis. It also prepares the transition to Part III. In Part II, the main question has been how to evaluate and optimize downside risk once a loss function is specified. In Part III, the focus shifts toward predicting which cases, units, or situations are risky before losses occur. The empirical and data-driven mindset developed here therefore provides a natural bridge from loss-based risk evaluation to predictive risk analytics.

Key Takeaways

- VaR and CVaR remain meaningful when the loss distribution is unknown, but they must be estimated empirically from data or scenarios.
- The empirical loss distribution is discrete by construction, so the distinction between TCE and exact CVaR remains important in finite samples.
- In finite-scenario settings, empirical CVaR can be represented using auxiliary variables and scenario-wise constraints.
- Mean–CVaR optimization can therefore be implemented as a linear or piecewise-linear mathematical program in many practical applications.
- Scenario generation is part of the risk model itself, because empirical tail measures and the associated optimization model depend directly on how uncertainty is represented.
- Bootstrap methods and stress testing are useful not only for refinement, but also for revealing instability in tail estimates and policy recommendations.
- Sensitivity analysis under unknown distributions must address informational robustness as well as preference variation.
- Chapter 6 completes the move from known distributions to empirical, scenario-based, and computationally implementable downside-risk analysis.

Bibliography

[1] Gordon J. Alexander, Thomas F. Coleman, and Yuying Li. Minimizing CVaR and VaR for a portfolio of derivatives. *Journal of Banking & Finance*, 30(2):583–605, 2006.

[2] Terje Aven. Risk assessment and risk management: Review of recent advances on their foundation. *European Journal of Operational Research*, 253(1):1–13, 2016.

[3] Olivier Bardou, Noufel Frikha, and Gilles Pagès. Computing var and cvar using stochastic approximation and adaptive unconstrained importance sampling. *Monte Carlo Methods and Applications*, 15(3):173–210, 2009.

[4] Olivier de Bandt, Laura-Chloé Kuntz, Nora Pankratz, Fulvio Pegoraro, Haakon Solheim, Greg Sutton, Azusa Takeyama, and Dora Xia. The effects of climate change-related risks on banks: A literature review. Working Paper 40, Basel Committee on Banking Supervision, Bank for International Settlements, 2023.

[5] Adele Fontana, Barbara Jarmulska, Benedikt Scheid, Christopher Scheins, and Claudia Schwarz. From flood to fire: Is physical climate risk taken into account in banks' residential mortgage rates? Working Paper Series 3036, European Central Bank, 2024.

[6] Joel Goh and Melvyn Sim. Distributionally robust optimization and its tractable approximations. *Operations Research*, 58(4-part-1):902–917, 2010.

[7] Anton J. Kleywegt, Alexander Shapiro, and Tito Homem-de Mello. The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, 12(2):479–502, 2002.

[8] Alexander J. McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton University Press, Princeton, NJ, 2005.

[9] Fernando S. Oliveira. Procurement risk management in a petroleum refinery. *Decision Sciences*, 54(3):277–296, 2023.

[10] Fernando S. Oliveira and Carlos Ruiz. Analysis of futures and spot electricity markets under risk aversion. *European Journal of Operational Research*, 291(3):1132–1148, 2021.

[11] Fernando S. Oliveira and Carlos Ruiz. Risk management in solar power plants with storage: a comparative study. *International Journal of Production Research*, 63(21):8074–8090, 2025.

[12] Fernando S. Oliveira, N. B. Zahur, and F. Wu. Analysis of the optimal policy for managing strategic petroleum reserves under long-term uncertainty: The ASEAN case. *Computers & Industrial Engineering*, 175:108834, 2023.

[13] R. Tyrrell Rockafellar and Stanislav Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2(3):21–42, 2000.

[14] R. Tyrrell Rockafellar and Stanislav Uryasev. Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26(7):1443–1471, 2002.

[15] Sergiy Sarykalin, Gian Paolo Serraino, and Stanislav Uryasev. Value-at-risk vs. conditional value-at-risk in risk management and optimization. *Tutorials in Operations Research*, pages 270–294, 2008.

[16] Chuanhao See and Melvyn Sim. Robust approximation to multistage stochastic optimization. *Operations Research*, 58(3):865–879, 2010.

[17] Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczyński. *Lectures on Stochastic Programming: Modeling and Theory*. SIAM and MOS, Philadelphia, PA, 2nd edition, 2009.

[18] Alexander Shapiro and Tito Homem-de Mello. On the rate of convergence of optimal solutions of monte carlo approximations of stochastic programs. *SIAM Journal on Optimization*, 11(1):70–86, 2000.

