Chapter 5

Statistical Inference

Alexander Shapiro

5.1 Statistical Properties of Sample Average Approximation Estimators

Consider the following stochastic programming problem:

$$\operatorname*{Min}_{x \in X} \big\{ f(x) := \mathbb{E}[F(x, \xi)] \big\}. \tag{5.1}$$

Here X is a nonempty closed subset of $\mathbb{R}^n$, ξ is a random vector whose probability distribution P is supported on a set $\Xi \subset \mathbb{R}^d$, and $F : X \times \Xi \to \mathbb{R}$. In the framework of two-stage stochastic programming, the objective function $F(x, \xi)$ is given by the optimal value of the corresponding second-stage problem. Unless stated otherwise, we assume in this chapter that the expectation function $f(x)$ is well defined and *finite valued* for all $x \in X$. This implies, of course, that for every $x \in X$ the value $F(x, \xi)$ is finite for a.e. $\xi \in \Xi$. In particular, for two-stage programming this implies that the recourse is relatively complete.

Suppose that we have a sample $\xi^1, \ldots, \xi^N$ of N realizations of the random vector ξ. This random sample can be viewed as historical data of N observations of ξ, or it can be generated in the computer by Monte Carlo sampling techniques. For any $x \in X$ we can estimate the expected value $f(x)$ by averaging values $F(x, \xi^j)$, $j = 1, \ldots, N$. This leads to the so-called sample average approximation (SAA)

$$\operatorname*{Min}_{x \in X} \left\{ \hat{f}_N(x) := \frac{1}{N} \sum_{j=1}^{N} F(x, \xi^j) \right\} \tag{5.2}$$

of the "true" problem (5.1). Let us observe that we can write the sample average function as the expectation

$$\hat{f}_N(x) = \mathbb{E}_{P_N}[F(x, \xi)] \tag{5.3}$$

taken with respect to the *empirical distribution*[18] (measure) $P_N := N^{-1} \sum_{j=1}^{N} \Delta(\xi^j)$. Therefore, for a given sample, the SAA problem (5.2) can be considered as a stochastic programming problem with respective scenarios $\xi^1, \ldots, \xi^N$, each taken with probability $1/N$.

As with data vector ξ, the sample $\xi^1, \ldots, \xi^N$ can be considered from two points of view: as a sequence of random vectors or as a particular realization of that sequence. Which of these two meanings will be used in a particular situation will be clear from the context. The SAA problem is a function of the considered sample and in that sense is random. For a particular realization of the random sample, the corresponding SAA problem is a stochastic programming problem with respective scenarios $\xi^1, \ldots, \xi^N$ each taken with probability $1/N$. We always assume that each random vector ξ^j in the sample has the same (marginal) distribution P as the data vector ξ. If, moreover, each ξ^j, $j = 1, \ldots, N$, is distributed independently of other sample vectors, we say that the sample is *independently identically distributed* (iid).

By the Law of Large Numbers we have that, under some regularity conditions, $\hat{f}_N(x)$ converges pointwise w.p. 1 to $f(x)$ as $N \to \infty$. In particular, by the classical LLN this holds if the sample is iid. Moreover, under mild additional conditions the convergence is uniform (see section 7.2.5). We also have that $\mathbb{E}[\hat{f}_N(x)] = f(x)$, i.e., $\hat{f}_N(x)$ is an *unbiased* estimator of $f(x)$. Therefore, it is natural to expect that the optimal value and optimal solutions of the SAA problem (5.2) converge to their counterparts of the true problem (5.1) as $N \to \infty$. We denote by ϑ^* and S the optimal value and the set of optimal solutions, respectively, of the true problem (5.1) and by $\hat{\vartheta}_N$ and $\hat{S}_N$ the optimal value and the set of optimal solutions, respectively, of the SAA problem (5.2).

We can view the sample average functions $\hat{f}_N(x)$ as defined on a common probability space $(\Omega, \mathcal{F}, P)$. For example, in the case of the iid sample, a standard construction is to consider the set $\Omega := \Xi^\infty$ of sequences $\{(\xi_1, \ldots)\}_{\xi_i \in \Xi, i \in \mathbb{N}}$, equipped with the product of the corresponding probability measures. Assume that $F(x, \xi)$ is a *Carathéodory function*, i.e., continuous in x and measurable in ξ. Then $\hat{f}_N(x) = \hat{f}_N(x, \omega)$ is also a Carathéodory function and hence is a random lower semicontinuous function. It follows (see section 7.2.3 and Theorem 7.37 in particular) that $\hat{\vartheta}_N = \hat{\vartheta}_N(\omega)$ and $\hat{S}_N = \hat{S}_N(\omega)$ are measurable. We also consider a particular optimal solution $\hat{x}_N$ of the SAA problem and view it as a measurable selection $\hat{x}_N(\omega) \in \hat{S}_N(\omega)$. Existence of such measurable selection is ensured by the measurable selection theorem (Theorem 7.34). This takes care of the measurability questions.

Next we discuss statistical properties of the SAA estimators $\hat{\vartheta}_N$ and $\hat{S}_N$. Let us make the following useful observation.

Proposition 5.1. *Let $f : X \to \mathbb{R}$ and $f_N : X \to \mathbb{R}$ be a sequence of (deterministic) real valued functions. Then the following two properties are equivalent:* (i) *for any $\bar{x} \in X$ and any sequence $\{x_N\} \subset X$ converging to $\bar{x}$ it follows that $f_N(x_N)$ converges to $f(\bar{x})$, and* (ii) *the function $f(\cdot)$ is continuous on X and $f_N(\cdot)$ converges to $f(\cdot)$ uniformly on any compact subset of X.*

[18] Recall that $\Delta(\xi)$ denotes measure of mass one at the point ξ.

Proof. Suppose that property (i) holds. Consider a point $\bar{x} \in X$, a sequence $\{x_N\} \subset X$ converging to $\bar{x}$ and a number $\varepsilon > 0$. By taking a sequence with each element equal x_1, we have by (i) that $f_N(x_1) \to f(x_1)$. Therefore, there exists N_1 such that $|f_{N_1}(x_1) - f(x_1)| < \varepsilon/2$. Similarly, there exists $N_2 > N_1$ such that $|f_{N_2}(x_2) - f(x_2)| < \varepsilon/2$, and so on. Consider now a sequence, denoted x'_N, constructed as follows: $x'_i = x_1$, $i = 1, \ldots, N_1$, $x'_i = x_2$, $i = N_1 + 1, \ldots, N_2$, and so on. We have that this sequence x'_N converges to $\bar{x}$ and hence $|f_N(x'_N) - f(\bar{x})| < \varepsilon/2$ for all N large enough. We also have that $|f_{N_k}(x'_{N_k}) - f(x_k)| < \varepsilon/2$, and hence $|f(x_k) - f(\bar{x}| < \varepsilon$ for all k large enough. This shows that $f(x_k) \to f(\bar{x})$ and hence $f(\cdot)$ is continuous at $\bar{x}$.

Now let C be a compact subset of X. Arguing by contradiction, suppose that $f_N(\cdot)$ does not converge to $f(\cdot)$ uniformly on C. Then there exists a sequence $\{x_N\} \subset C$ and $\varepsilon > 0$ such that $|f_N(x_N) - f(x_N)| \geq \varepsilon$ for all N. Since C is compact, we can assume that $\{x_N\}$ converges to a point $\bar{x} \in C$. We have

$$|f_N(x_N) - f(x_N)| \leq |f_N(x_N) - f(\bar{x})| + |f(x_N) - f(\bar{x})|. \tag{5.4}$$

The first term in the right-hand side of (5.4) tends to zero by (i) and the second term tends to zero since $f(\cdot)$ is continuous, and hence these terms are less that $\varepsilon/2$ for N large enough. This gives a designed contradiction.

Conversely, suppose that property (ii) holds. Consider a sequence $\{x_N\} \subset X$ converging to a point $\bar{x} \in X$. We can assume that this sequence is contained in a compact subset of X. By employing the inequality

$$|f_N(x_N) - f(\bar{x})| \leq |f_N(x_N) - f(x_N)| + |f(x_N) - f(\bar{x})| \tag{5.5}$$

and noting that the first term in the right-hand side of this inequality tends to zero because of the uniform convergence of f_N to f and the second term tends to zero by continuity of f, we obtain that property (i) holds. □

5.1.1 Consistency of SAA Estimators

In this section we discuss convergence properties of the SAA estimators $\hat{\vartheta}_N$ and $\hat{S}_N$. It is said that an estimator $\hat{\theta}_N$ of a parameter θ is *consistent* if $\hat{\theta}_N$ converges w.p. 1 to θ as $N \to \infty$. Let us consider first consistency of the SAA estimator of the optimal value. We have that for any fixed $x \in X$, $\hat{\vartheta}_N \leq \hat{f}_N(x)$, and hence if the pointwise LLN holds, then

$$\limsup_{N\to\infty} \hat{\vartheta}_N \leq \lim_{N\to\infty} \hat{f}_N(x) = f(x) \ \text{ w.p. 1.}$$

It follows that if the pointwise LLN holds, then

$$\limsup_{N\to\infty} \hat{\vartheta}_N \leq \vartheta^* \ \text{ w.p. 1.} \tag{5.6}$$

Without some additional conditions, the inequality in (5.6) can be strict.

Proposition 5.2. *Suppose that $\hat{f}_N(x)$ converges to $f(x)$ w.p. 1, as $N \to \infty$, uniformly on X. Then $\hat{\vartheta}_N$ converges to ϑ^* w.p. 1 as $N \to \infty$.*

Proof. The uniform convergence w.p. 1 of $\hat{f}_N(x) = \hat{f}_N(x, \omega)$ to $f(x)$ means that for any $\varepsilon > 0$ and a.e. $\omega \in \Omega$ there is $N^* = N^*(\varepsilon, \omega)$ such that the following inequality holds for all $N \geq N^*$:

$$\sup_{x \in X} \left| \hat{f}_N(x, \omega) - f(x) \right| \leq \varepsilon. \tag{5.7}$$

It follows then that $|\hat{\vartheta}_N(\omega) - \vartheta^*| \leq \varepsilon$ for all $N \geq N^*$, which completes the proof. □

In order to establish consistency of the SAA estimators of optimal solutions, we need slightly stronger conditions. Recall that $\mathbb{D}(A, B)$ denotes the deviation of set A from set B. (See (7.4) for the corresponding definition.)

Theorem 5.3. *Suppose that there exists a compact set $C \subset \mathbb{R}^n$ such that:* (i) *the set S of optimal solutions of the true problem is nonempty and is contained in C,* (ii) *the function $f(x)$ is finite valued and continuous on C,* (iii) *$\hat{f}_N(x)$ converges to $f(x)$ w.p. 1, as $N \to \infty$, uniformly in $x \in C$, and* (iv) *w.p. 1 for N large enough the set $\hat{S}_N$ is nonempty and $\hat{S}_N \subset C$. Then $\hat{\vartheta}_N \to \vartheta^*$ and $\mathbb{D}(\hat{S}_N, S) \to 0$ w.p. 1 as $N \to \infty$.*

Proof. Assumptions (i) and (iv) imply that both the true and the SAA problem can be restricted to the set $X \cap C$. Therefore we can assume without loss of generality that the set X is compact. The assertion that $\hat{\vartheta}_N \to \vartheta^*$ w.p. 1 follows by Proposition 5.2. It suffices to show now that $\mathbb{D}(\hat{S}_N(\omega), S) \to 0$ for every $\omega \in \Omega$ such that $\hat{\vartheta}_N(\omega) \to \vartheta^*$ and assumptions (iii) and (iv) hold. This is basically a deterministic result; therefore, we omit ω for the sake of notational convenience.

We argue now by a contradiction. Suppose that $\mathbb{D}(\hat{S}_N, S) \not\to 0$. Since X is compact, by passing to a subsequence if necessary, we can assume that there exists $\hat{x}_N \in \hat{S}_N$ such that $\operatorname{dist}(\hat{x}_N, S) \geq \varepsilon$ for some $\varepsilon > 0$ and that $\hat{x}_N$ tends to a point $x^* \in X$. It follows that $x^* \notin S$ and hence $f(x^*) > \vartheta^*$. Moreover, $\hat{\vartheta}_N = \hat{f}_N(\hat{x}_N)$ and

$$\hat{f}_N(\hat{x}_N) - f(x^*) = [\hat{f}_N(\hat{x}_N) - f(\hat{x}_N)] + [f(\hat{x}_N) - f(x^*)]. \tag{5.8}$$

The first term in the right-hand side of (5.8) tends to zero by assumption (iii) and the second term by continuity of $f(x)$. That is, we obtain that $\hat{\vartheta}_N$ tends to $f(x^*) > \vartheta^*$, a contradiction. □

Recall that by Proposition 5.1, assumptions (ii) and (iii) in the above theorem are equivalent to the condition that for any sequence $\{x_N\} \subset C$ converging to a point $\bar{x}$ it follows that $\hat{f}_N(x_N) \to f(\bar{x})$ w.p. 1. Assumption (iv) in the above theorem holds, in particular, if the feasible set X is closed, the functions $\hat{f}_N(x)$ are lower semicontinuous, and for some $\alpha > \vartheta^*$ the level sets $\left\{x \in X : \hat{f}_N(x) \leq \alpha\right\}$ are uniformly bounded w.p. 1. This condition is often referred to as the *inf-compactness condition*. Conditions ensuring the uniform convergence of $\hat{f}_N(x)$ to $f(x)$ (assumption (iii)) are given in Theorems 7.48 and 7.50, for example.

The assertion that $\mathbb{D}(\hat{S}_N, S) \to 0$ w.p. 1 means that for any (measurable) selection $\hat{x}_N \in \hat{S}_N$, of an optimal solution of the SAA problem, it holds that $\operatorname{dist}(\hat{x}_N, S) \to 0$ w.p. 1. If, moreover, $S = \{\bar{x}\}$ is a singleton, i.e., the true problem has unique optimal solution $\bar{x}$,

then this means that $\hat{x}_N \to \bar{x}$ w.p. 1. The inf-compactness condition ensures that $\hat{x}_N$ cannot escape to infinity as N increases.

If the problem is convex, it is possible to relax the required regularity conditions. In the following theorem we assume that the integrand function $F(x, \xi)$ is an extended real valued function, i.e., can also take values $\pm\infty$. Denote

$$\bar{F}(x, \xi) := F(x, \xi) + \mathbb{I}_X(x), \ \bar{f}(x) := f(x) + \mathbb{I}_X(x), \ \tilde{f}_N(x) := \hat{f}_N(x) + \mathbb{I}_X(x), \quad (5.9)$$

i.e., $\bar{f}(x) = f(x)$ if $x \in X$ and $\bar{f}(x) = +\infty$ if $x \notin X$, and similarly for functions $F(\cdot, \xi)$ and $\hat{f}_N(\cdot)$. Clearly $\bar{f}(x) = \mathbb{E}[\bar{F}(x, \xi)]$ and $\tilde{f}_N(x) = N^{-1} \sum_{j=1}^N \bar{F}(x, \xi^j)$. Note that if the set X is convex, then the above penalization operation preserves convexity of respective functions.

Theorem 5.4. *Suppose that:* (i) *the integrand function F is random lower semicontinuous,* (ii) *for almost every $\xi \in \Xi$ the function $F(\cdot, \xi)$ is convex,* (iii) *the set X is closed and convex,* (iv) *the expected value function f is lower semicontinuous and there exists a point $\bar{x} \in X$ such that $f(x) < +\infty$ for all x in a neighborhood of $\bar{x}$,* (v) *the set S of optimal solutions of the true problem is nonempty and bounded, and* (vi) *the LLN holds pointwise. Then $\hat{\vartheta}_N \to \vartheta^*$ and $\mathbb{D}(\hat{S}_N, S) \to 0$ w.p. 1 as $N \to \infty$.*

Proof. Clearly we can restrict both the true and the SAA problem to the affine space generated by the convex set X. Relative to that affine space, the set X has a nonempty interior. Therefore, without loss of generality we can assume that the set X has a nonempty interior. Since it is assumed that $f(x)$ possesses an optimal solution, we have that ϑ^* is finite and hence $f(x) \geq \vartheta^* > -\infty$ for all $x \in X$. Since $f(x)$ is convex and is greater than $-\infty$ on an open set (e.g., interior of X), it follows that $f(\cdot)$ is subdifferentiable at any point $x \in \mathrm{int}(X)$ such that $f(x)$ is finite. Consequently $f(x) > -\infty$ for all $x \in \mathbb{R}^n$, and hence f is proper.

Observe that the pointwise LLN for $F(x, \xi)$ (assumption (vi)) implies the corresponding pointwise LLN for $\bar{F}(x, \xi)$. Since X is convex and closed, it follows that $\bar{f}$ is convex and lower semicontinuous. Moreover, because of the assumption (iv) and since the interior of X is nonempty, we have that $\mathrm{dom}\, \bar{f}$ has a nonempty interior. By Theorem 7.49 it follows then that $\tilde{f}_N \overset{e}{\to} \bar{f}$ w.p. 1. Consider a compact set K with a nonempty interior and such that it does not contain a boundary point of $\mathrm{dom}\, \bar{f}$, and $\bar{f}(x)$ is finite valued on K. Since $\mathrm{dom}\, \bar{f}$ has a nonempty interior, such a set exists. Then it follows from $\tilde{f}_N \overset{e}{\to} \bar{f}$ that $\tilde{f}_N(\cdot)$ converge to $\bar{f}(\cdot)$ uniformly on K, all w.p. 1 (see Theorem 7.27). It follows that w.p. 1 for N large enough the functions $\tilde{f}_N(x)$ are finite valued on K and hence are proper.

Now let C be a compact subset of $\mathbb{R}^n$ such that the set S is contained in the interior of C. Such set exists since it is assumed that the set S is bounded. Consider the set $\widetilde{S}_N$ of minimizers of $\tilde{f}_N(x)$ over C. Since C is nonempty and compact and $\tilde{f}_N(x)$ is lower semicontinuous and proper for N large enough, and because by the pointwise LLN we have that for any $x \in S$, $\tilde{f}_N(x)$ is finite w.p. 1 for N large enough, the set $\widetilde{S}_N$ is nonempty w.p. 1 for N large enough. Let us show that $\mathbb{D}(\widetilde{S}_N, S) \to 0$ w.p. 1. Let $\omega \in \Omega$ be such that $\tilde{f}_N(\cdot, \omega) \overset{e}{\to} \bar{f}(\cdot)$. We have that this happens for a.e. $\omega \in \Omega$. We argue now by a contradiction. Suppose that there exists a minimizer $\tilde{x}_N = \tilde{x}_N(\omega)$ of $\tilde{f}_N(x, \omega)$ over C such that $\mathrm{dist}(\tilde{x}_N, S) \geq \varepsilon$ for some $\varepsilon > 0$. Since C is compact, by passing to a subsequence if necessary, we can assume that $\tilde{x}_N$ tends to a point $x^* \in C$. It follows that $x^* \notin S$. On the other hand, we have

by Proposition 7.26 that $x^* \in \arg\min_{x\in C} \bar{f}(x)$. Since $\arg\min_{x\in C} \bar{f}(x) = S$, we obtain a contradiction.

Now because of the convexity assumptions, any minimizer of $\tilde{f}_N(x)$ over C which lies inside the interior of C is also an optimal solution of the SAA problem (5.2). Therefore, w.p. 1 for N large enough we have that $\widetilde{S}_N = \hat{S}_N$. Consequently, we can restrict both the true and the SAA optimization problems to the compact set C, and hence the assertions of the above theorem follow. □

Let us make the following observations. Lower semicontinuity of $f(\cdot)$ follows from lower semicontinuity $F(\cdot, \xi)$, provided that $F(x, \cdot)$ is bounded from below by an integrable function. (See Theorem 7.42 for a precise formulation of this result.) It was assumed in the above theorem that the LLN holds pointwise for all $x \in \mathbb{R}^n$. Actually, it suffices to assume that this holds for all x in some neighborhood of the set S. Under the assumptions of the above theorem we have that $f(x) > -\infty$ for every $x \in \mathbb{R}^n$. The above assumptions do not prevent, however, $f(x)$ from taking value $+\infty$ at some points $x \in X$. Nevertheless, it was possible to push the proof through because in the considered convex case local optimality implies global optimality. There are two possible reasons $f(x)$ can be $+\infty$. Namely, it can be that $F(x, \cdot)$ is finite valued but grows sufficiently fast so that its integral is $+\infty$, or it can be that $F(x, \cdot)$ is equal $+\infty$ on a set of positive measure. Of course, it can be both. For example, in the case of two-stage programming it may happen that for some $x \in X$ the corresponding second stage problem is infeasible with a positive probability p. Then w.p. 1 for N large enough, for at least one of the sample points ξ^j the corresponding second-stage problem will be infeasible, and hence $\hat{f}_N(x) = +\infty$. Of course, if the probability p is very small, then the required sample size for such event to happen could be very large.

We assumed so far that the feasible set X of the SAA problem is fixed, i.e., independent of the sample. However, in some situations it also should be estimated. Then the corresponding SAA problem takes the form

$$\operatorname*{Min}_{x\in X_N} \hat{f}_N(x), \tag{5.10}$$

where X_N is a subset of $\mathbb{R}^n$ depending on the sample and therefore is random. As before we denote by $\hat{\vartheta}_N$ and $\hat{S}_N$ the optimal value and the set of optimal solutions, respectively, of the SAA problem (5.10).

Theorem 5.5. *Suppose that in addition to the assumptions of Theorem* 5.3 *the following conditions hold:*

(a) *If $x_N \in X_N$ and x_N converges w.p.* 1 *to a point x, then $x \in X$.*

(b) *For some point $x \in S$ there exists a sequence $x_N \in X_N$ such that $x_N \to x$ w.p.* 1.

Then $\hat{\vartheta}_N \to \vartheta^*$ and $\mathbb{D}(\hat{S}_N, S) \to 0$ w.p. 1 *as $N \to \infty$.*

Proof. Consider an $\hat{x}_N \in \hat{S}_N$. By compactness arguments we can assume that $\hat{x}_N$ converges w.p. 1 to a point $x^* \in \mathbb{R}^n$. Since $\hat{S}_N \subset X_N$, we have that $\hat{x}_N \in X_N$, and hence it follows by condition (a) that $x^* \in X$. We also have (see Proposition 5.1) that $\hat{\vartheta}_N = \hat{f}_N(\hat{x}_N)$ tends w.p. 1 to $f(x^*)$, and hence $\liminf_{N\to\infty} \hat{\vartheta}_N \geq \vartheta^*$ w.p. 1. On the other hand, by condition (b), there exists a sequence $x_N \in X_N$ converging to a point $x \in S$ w.p. 1. Consequently, $\hat{\vartheta}_N \leq \hat{f}_N(\hat{x}_N) \to f(x) = \vartheta^*$ w.p. 1, and hence $\limsup_{N\to\infty} \hat{\vartheta}_N \leq \vartheta^*$. It follows that

$\hat{\vartheta}_N \to \vartheta^*$ w.p. 1. The remainder of the proof can be completed by the same arguments as in the proof of Theorem 5.3. □

The SAA problem (5.10) is convex if the functions $\hat{f}_N(\cdot)$ and the sets X_N are convex w.p. 1. It is also possible to show consistency of the SAA estimators of problem (5.10) under the assumptions of Theorem 5.4 together with conditions (a) and (b) of the above Theorem 5.5, and convexity of the set X_N.

Suppose, for example, that the set X is defined by the constraints

$$X := \{x \in X_0 : g_i(x) \leq 0,\ i = 1, \ldots, p\}, \tag{5.11}$$

where X_0 is a nonempty closed subset of $\mathbb{R}^n$ and the constraint functions are given as the expected value functions

$$g_i(x) := \mathbb{E}[G_i(x, \xi)], \quad i = 1, \ldots, p, \tag{5.12}$$

with $G_i(x, \xi)$, $i = 1, \ldots, p$, being random lower semicontinuous functions. Then the set X can be estimated by

$$X_N := \big\{x \in X_0 : \hat{g}_{iN}(x) \leq 0,\ i = 1, \ldots, p\big\}, \tag{5.13}$$

where

$$\hat{g}_{iN}(x) := \frac{1}{N} \sum_{j=1}^{N} G_i(x, \xi^j).$$

If for a given point $x \in X_0$, every function $\hat{g}_{iN}$ converges uniformly to g_i w.p. 1 on a neighborhood of x and the functions g_i are continuous, then condition (a) of Theorem 5.5 holds.

Remark 5. Let us note that the samples used in construction of the SAA functions $\hat{f}_N$ and $\hat{g}_{iN}$, $i = 1, \ldots, p$, can be the same or can be different, independent of each other. That is, for random samples $\xi^{i1}, \ldots, \xi^{iN_i}$, possibly of different sample sizes N_i, $i = 1, \ldots, p$, and independent of each other and of the random sample used in $\hat{f}_N$, the corresponding SAA functions are

$$\hat{g}_{iN_i}(x) := \frac{1}{N_i} \sum_{j=1}^{N_i} G_i(x, \xi^{ij}), \quad i = 1, \ldots, p.$$

The question of how to generate the respective random samples is especially relevant for Monte Carlo sampling methods discussed later. For consistency type results we only need to verify convergence w.p. 1 of the involved SAA functions to their true (expected value) counterparts, and this holds under appropriate regularity conditions in both cases—of the same and independent samples. However, from a variability point of view, it is advantageous to use independent samples (see Remark 9 on page 173).

In order to ensure condition (b) of Theorem 5.5, one needs to impose a constraint qualification (on the true problem). Consider, for example, $X := \{x \in \mathbb{R} : g(x) \leq 0\}$ with $g(x) := x^2$. Clearly $X = \{0\}$, while an arbitrary small perturbation of the function $g(\cdot)$ can result in the corresponding set X_N being empty. It is possible to show that if a constraint

qualification for the true problem is satisfied at x, then condition (b) follows. For instance, if the set X_0 is convex and for every $\xi \in \Xi$ the functions $G_i(\cdot, \xi)$ are convex, and hence the corresponding expected value functions $g_i(\cdot)$, $i = 1, \ldots, p$, are also convex, then such a simple constraint qualification is the Slater condition. Recall that it is said that the *Slater condition* holds if there exists a point $x^* \in X_0$ such that $g_i(x^*) < 0$, $i = 1, \ldots, p$.

As another example, suppose that the feasible set is given by probabilistic (chance) constraints in the form

$$X = \left\{x \in \mathbb{R}^n : \Pr\big(C_i(x, \xi) \leq 0\big) \geq 1 - \alpha_i,\ i = 1, \ldots, p\right\}, \tag{5.14}$$

where $\alpha_i \in (0, 1)$ and $C_i : \mathbb{R}^n \times \Xi \to \mathbb{R}$, $i = 1, \ldots, p$, are Carathéodory functions. Of course, we have that[19]

$$\Pr\big(C_i(x, \xi) \leq 0\big) = \mathbb{E}\left[\mathbf{1}_{(-\infty,0]}\big(C_i(x, \xi)\big)\right]. \tag{5.15}$$

Consequently, we can write the above set X in the form (5.11)–(5.12) with $X_0 := \mathbb{R}^n$ and

$$G_i(x, \xi) := 1 - \alpha_i - \mathbf{1}_{(-\infty,0]}\big(C_i(x, \xi)\big). \tag{5.16}$$

The corresponding set X_N can be written as

$$X_N = \left\{x \in \mathbb{R}^n : \frac{1}{N}\sum_{j=1}^{N} \mathbf{1}_{(-\infty,0]}\big(C_i(x, \xi^j)\big) \geq 1 - \alpha_i,\ i = 1, \ldots, p\right\}. \tag{5.17}$$

Note that $\sum_{j=1}^{N} \mathbf{1}_{(-\infty,0]}\big(C_i(x, \xi^j)\big)$, in the above formula, counts the number of times that the event "$C_i(x, \xi^j) \leq 0$", $j = 1, \ldots, N$, happens. The additional difficulty here is that the (step) function $\mathbf{1}_{(-\infty,0]}(t)$ is discontinuous at $t = 0$. Nevertheless, suppose that the sample is iid and for every x in a neighborhood of the set X and $i = 1, \ldots, p$, the event "$C_i(x, \xi) = 0$" happens with probability zero, and hence $G_i(\cdot, \xi)$ is continuous at x for a.e. ξ. By Theorem 7.48 this implies that the expectation function $g_i(x)$ is continuous and $\hat{g}_{iN}(x)$ converge uniformly w.p. 1 on compact neighborhoods to $g_i(x)$, and hence condition (a) of Theorem 5.5 holds. Condition (b) could be verified by ad hoc methods.

Remark 6. As pointed out in Remark 5, it is possible to use different, independent of each other, random samples $\xi^{i1}, \ldots, \xi^{iN_i}$, possibly of different sample sizes N_i, $i = 1, \ldots, p$, for constructing the corresponding SAA functions. That is, constraints $\Pr\big(C_i(x, \xi) > 0\big) \leq \alpha_i$ are approximated by

$$\frac{1}{N_i}\sum_{j=1}^{N_i} \mathbf{1}_{(0,\infty)}\big(C_i(x, \xi^{ij})\big) \leq \alpha_i,\ i = 1, \ldots, p. \tag{5.18}$$

From the point of view of reducing variability of the respective SAA estimators, it could be preferable to use this approach of independent, rather than the same, samples.

[19] Recall that $\mathbf{1}_{(-\infty,0]}(t) = 1$ if $t \leq 0$ and $\mathbf{1}_{(-\infty,0]}(t) = 0$ if $t > 0$.

5.1.2 Asymptotics of the SAA Optimal Value

Consistency of the SAA estimators gives a certain assurance that the error of the estimation approaches zero in the limit as the sample size grows to infinity. Although important conceptually, this does not give any indication of the magnitude of the error for a given sample. Suppose for the moment that the sample is iid and let us fix a point $x \in X$. Then we have that the sample average estimator $\hat{f}_N(x)$, of $f(x)$, is unbiased and has variance $\sigma^2(x)/N$, where $\sigma^2(x) := \mathbb{V}\text{ar}[F(x,\xi)]$ is supposed to be finite. Moreover, by the CLT we have that

$$N^{1/2}\left[\hat{f}_N(x) - f(x)\right] \xrightarrow{\mathcal{D}} Y_x, \tag{5.19}$$

where $\xrightarrow{\mathcal{D}}$ denotes convergence in *distribution* and Y_x has a normal distribution with mean 0 and variance $\sigma^2(x)$, written $Y_x \sim \mathcal{N}\left(0, \sigma^2(x)\right)$. That is, $\hat{f}_N(x)$ has *asymptotically normal* distribution, i.e., for large N, $\hat{f}_N(x)$ has approximately normal distribution with mean $f(x)$ and variance $\sigma^2(x)/N$.

This leads to the following (approximate) $100(1-\alpha)\%$ confidence interval for $f(x)$:

$$\left[\hat{f}_N(x) - \frac{z_{\alpha/2}\hat{\sigma}(x)}{\sqrt{N}},\ \hat{f}_N(x) + \frac{z_{\alpha/2}\hat{\sigma}(x)}{\sqrt{N}}\right], \tag{5.20}$$

where $z_{\alpha/2} := \Phi^{-1}(1-\alpha/2)$ and[20]

$$\hat{\sigma}^2(x) := \frac{1}{N-1}\sum_{j=1}^{N}\left[F(x,\xi^j) - \hat{f}_N(x)\right]^2 \tag{5.21}$$

is the sample variance estimate of $\sigma^2(x)$. That is, the error of estimation of $f(x)$ is (stochastically) of order $O_p(N^{-1/2})$.

Consider now the optimal value $\hat{\vartheta}_N$ of the SAA problem (5.2). Clearly we have that for any $x' \in X$ the inequality $\hat{f}_N(x') \geq \inf_{x\in X}\hat{f}_N(x)$ holds. By taking the expected value of both sides of this inequality and minimizing the left-hand side over all $x' \in X$, we obtain

$$\inf_{x\in X}\mathbb{E}\left[\hat{f}_N(x)\right] \geq \mathbb{E}\left[\inf_{x\in X}\hat{f}_N(x)\right]. \tag{5.22}$$

Note that the inequality (5.22) holds even if $f(x) = +\infty$ or $f(x) = -\infty$ for some $x \in X$. Since $\mathbb{E}[\hat{f}_N(x)] = f(x)$, it follows that $\vartheta^* \geq \mathbb{E}[\hat{\vartheta}_N]$. In fact, typically, $\mathbb{E}[\hat{\vartheta}_N]$ is strictly less than ϑ^*, i.e., $\hat{\vartheta}_N$ is a downward *biased* estimator of ϑ^*. As the following result shows, this bias decreases monotonically with increase of the sample size N.

Proposition 5.6. *Let $\hat{\vartheta}_N$ be the optimal value of the SAA problem* (5.2)*, and suppose that the sample is* iid*. Then $\mathbb{E}[\hat{\vartheta}_N] \leq \mathbb{E}[\hat{\vartheta}_{N+1}] \leq \vartheta^*$ for any $N \in \mathbb{N}$.*

[20] Here $\Phi(\cdot)$ denotes the cdf of the standard normal distribution. For example, to 95% confidence intervals corresponds $z_{0.025} = 1.96$.

Proof. It was already shown above that $\mathbb{E}[\hat{\vartheta}_N] \leq \vartheta^*$ for any $N \in \mathbb{N}$. We can write

$$\hat{f}_{N+1}(x) = \frac{1}{N+1} \sum_{i=1}^{N+1} \left[\frac{1}{N} \sum_{j \neq i} F(x, \xi^j) \right].$$

Moreover, since the sample is iid we have

$$\begin{aligned}
\mathbb{E}[\hat{\vartheta}_{N+1}] &= \mathbb{E}\left[\inf_{x \in X} \hat{f}_{N+1}(x)\right] \\
&= \mathbb{E}\left[\inf_{x \in X} \tfrac{1}{N+1} \textstyle\sum_{i=1}^{N+1} \left(\frac{1}{N} \sum_{j \neq i} F(x, \xi^j) \right)\right] \\
&\geq \mathbb{E}\left[\tfrac{1}{N+1} \textstyle\sum_{i=1}^{N+1} \left(\inf_{x \in X} \frac{1}{N} \sum_{j \neq i} F(x, \xi^j) \right)\right] \\
&= \tfrac{1}{N+1} \textstyle\sum_{i=1}^{N+1} \mathbb{E}\left[\inf_{x \in X} \frac{1}{N} \sum_{j \neq i} F(x, \xi^j)\right] \\
&= \tfrac{1}{N+1} \textstyle\sum_{i=1}^{N+1} \mathbb{E}[\hat{\vartheta}_N] = \mathbb{E}[\hat{\vartheta}_N],
\end{aligned}$$

which completes the proof. □

First Order Asymptotics of the SAA Optimal Value

We use the following assumptions about the integrand F:

(A1) For some point $\tilde{x} \in X$ the expectation $\mathbb{E}[F(\tilde{x}, \xi)^2]$ is finite.

(A2) There exists a measurable function $C : \Xi \to \mathbb{R}_+$ such that $\mathbb{E}[C(\xi)^2]$ is finite and

$$\left|F(x, \xi) - F(x', \xi)\right| \leq C(\xi)\|x - x'\| \tag{5.23}$$

for all $x, x' \in X$ and a.e. $\xi \in \Xi$.

The above assumptions imply that the expected value $f(x)$ and variance $\sigma^2(x)$ are finite valued for all $x \in X$. Moreover, it follows from (5.23) that

$$|f(x) - f(x')| \leq \kappa \|x - x'\|, \quad \forall x, x' \in X,$$

where $\kappa := \mathbb{E}[C(\xi)]$, and hence $f(x)$ is Lipschitz continuous on X. If X is compact, we have then that the set S, of minimizers of $f(x)$ over X, is nonempty.

Let Y_x be random variables defined in (5.19). These variables depend on $x \in X$ and we also use notation $Y(x) = Y_x$. By the (multivariate) CLT we have that for any finite set $\{x_1, \ldots, x_m\} \subset X$, the random vector $(Y(x_1), \ldots, Y(x_m))$ has a multivariate normal distribution with zero mean and the same covariance matrix as the covariance matrix of $(F(x_1, \xi), \ldots, F(x_m, \xi))$. Moreover, by assumptions (A1) and (A2), compactness of X, and since the sample is iid, we have that $N^{1/2}(\hat{f}_N - f)$ converges in distribution to Y, viewed as a *random element*[21] of $C(X)$. This is a so-called functional CLT (see, e.g., Araujo and Giné [4, Corollary 7.17]).

[21] Recall that $C(X)$ denotes the space of continuous functions equipped with the sup-norm. A random element of $C(X)$ is a mapping $Y : \Omega \to C(X)$ from a probability space $(\Omega, \mathcal{F}, P)$ into $C(X)$ which is measurable with respect to the Borel sigma algebra of $C(X)$, i.e., $Y(x) = Y(x, \omega)$ can be viewed as a random function.

Theorem 5.7. *Let $\hat{\vartheta}_N$ be the optimal value of the SAA problem* (5.2)*. Suppose that the sample is iid, the set X is compact, and assumptions* (A1) *and* (A2) *are satisfied. Then the following holds:*

$$\hat{\vartheta}_N = \inf_{x \in S} \hat{f}_N(x) + o_p(N^{-1/2}), \tag{5.24}$$

$$N^{1/2}\left(\hat{\vartheta}_N - \vartheta^*\right) \xrightarrow{\mathcal{D}} \inf_{x \in S} Y(x). \tag{5.25}$$

If, moreover, $S = \{\bar{x}\}$ is a singleton, then

$$N^{1/2}\left(\hat{\vartheta}_N - \vartheta^*\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(\bar{x})). \tag{5.26}$$

Proof. Proof is based on the functional CLT and the Delta theorem (Theorem 7.59). Consider Banach space $C(X)$ of continuous functions $\psi : X \to \mathbb{R}$ equipped with the sup-norm $\|\psi\| := \sup_{x \in X} |\psi(x)|$. Define the min-value function $V(\psi) := \inf_{x \in X} \psi(x)$. Since X is compact, the function $V : C(X) \to \mathbb{R}$ is real valued and measurable (with respect to the Borel sigma algebra of $C(X)$). Moreover, it is not difficult to see that $|V(\psi_1) - V(\psi_2)| \leq \|\psi_1 - \psi_2\|$ for any $\psi_1, \psi_2 \in C(X)$, i.e., $V(\cdot)$ is Lipschitz continuous with Lipschitz constant one. By the Danskin theorem (Theorem 7.21), $V(\cdot)$ is directionally differentiable at any $\mu \in C(X)$ and

$$V'_\mu(\delta) = \inf_{x \in \bar{X}(\mu)} \delta(x), \quad \forall \delta \in C(X), \tag{5.27}$$

where $\bar{X}(\mu) := \arg\min_{x \in X} \mu(x)$. Since $V(\cdot)$ is Lipschitz continuous, directional differentiability in the Hadamard sense follows (see Proposition 7.57). As discussed above, we also have here under assumptions (A1) and (A2) and since the sample is iid that $N^{1/2}(\hat{f}_N - f)$ converges in distribution to the random element Y of $C(X)$. Noting that $\hat{\vartheta}_N = V(\hat{f}_N)$, $\vartheta^* = V(f)$, and $\bar{X}(f) = S$, and by applying the Delta theorem to the min-function $V(\cdot)$ at $\mu := f$ and using (5.27), we obtain (5.25) and that

$$\hat{\vartheta}_N - \vartheta^* = \inf_{x \in S}\left[\hat{f}_N(x) - f(x)\right] + o_p(N^{-1/2}). \tag{5.28}$$

Since $f(x) = \vartheta^*$ for any $x \in S$, we have that assertions (5.24) and (5.28) are equivalent. Finally, (5.26) follows from (5.25). □

Under mild additional conditions (see Remark 32 on page 382), it follows from (5.25) that $N^{1/2}\mathbb{E}\left[\hat{\vartheta}_N - \vartheta^*\right]$ tends to $\mathbb{E}\left[\inf_{x \in S} Y(x)\right]$ as $N \to \infty$, that is,

$$\mathbb{E}[\hat{\vartheta}_N] - \vartheta^* = N^{-1/2}\mathbb{E}\left[\inf_{x \in S} Y(x)\right] + o(N^{-1/2}). \tag{5.29}$$

In particular, if $S = \{\bar{x}\}$ is a singleton, then by (5.26) the SAA optimal value $\hat{\vartheta}_N$ has asymptotically normal distribution and, since $\mathbb{E}[Y(\bar{x})] = 0$, we obtain that in this case the bias $\mathbb{E}[\hat{\vartheta}_N] - \vartheta^*$ is of order $o(N^{-1/2})$. On the other hand, if the true problem has more than one optimal solution, then the right-hand side of (5.25) is given by the minimum of a number of random variables. Although each $Y(x)$ has mean zero, their minimum $\inf_{x \in S} Y(x)$ typically has a negative mean if the set S has more than one element. Therefore, if S is not a singleton, then the bias $\mathbb{E}[\hat{\vartheta}_N] - \vartheta^*$ typically is strictly less than zero and is of order $O(N^{-1/2})$. Moreover, the bias tends to be bigger the larger the set S is. For a further discussion of the bias issue, see Remark 7 on page 168.

5.1.3 Second Order Asymptotics

Formula (5.24) gives a first order expansion of the SAA optimal value $\hat{\vartheta}_N$. In this section we discuss a second order term in an expansion of $\hat{\vartheta}_N$. It turns out that the second order analysis of $\hat{\vartheta}_N$ is closely related to deriving (first order) asymptotics of optimal solutions of the SAA problem. We assume in this section that the true (expected value) problem (5.1) has unique optimal solution $\bar{x}$ and denote by $\hat{x}_N$ an optimal solution of the corresponding SAA problem. In order to proceed with the second order analysis we need to impose considerably stronger assumptions.

Our analysis is based on the second order Delta theorem, Theorem 7.62, and second order perturbation analysis of section 7.1.5. As in section 7.1.5, we consider a convex compact set $U \subset \mathbb{R}^n$ such that $X \subset \mathrm{int}(U)$, and we work with the space $W^{1,\infty}(U)$ of Lipschitz continuous functions $\psi : U \to \mathbb{R}$ equipped with the norm

$$\|\psi\|_{1,U} := \sup_{x\in U} |\psi(x)| + \sup_{x\in U'} \|\nabla\psi(x)\|, \tag{5.30}$$

where $U' \subset \mathrm{int}(U)$ is the set of points where $\psi(\cdot)$ is differentiable.

We make the following assumptions about the true problem:

(S1) The function $f(x)$ is Lipschitz continuous on U, has unique minimizer $\bar{x}$ over $x \in X$, and is twice continuously differentiable at $\bar{x}$.

(S2) The set X is second order regular at $\bar{x}$.

(S3) The quadratic growth condition (7.70) holds at $\bar{x}$.

Let $\mathcal{K}$ be the subset of $W^{1,\infty}(U)$ formed by differentiable at $\bar{x}$ functions. Note that the set $\mathcal{K}$ forms a closed (in the norm topology) linear subspace of $W^{1,\infty}(U)$. Assumption (S1) ensures that $f \in \mathcal{K}$. In order to ensure that $\hat{f}_N \in \mathcal{K}$ w.p. 1, we make the following assumption:

(S4) Function $F(\cdot,\xi)$ is Lipschitz continuous on U and differentiable at $\bar{x}$ for a.e. $\xi \in \Xi$.

We view $\hat{f}_N$ as a random element of $W^{1,\infty}(U)$, and assume, further, that $N^{1/2}(\hat{f}_N - f)$ converges in distribution to a random element Y of $W^{1,\infty}(U)$.

Consider the min-function $V : W^{1,\infty}(U) \to \mathbb{R}$ defined as

$$V(\psi) := \inf_{x\in X} \psi(x), \quad \psi \in W^{1,\infty}(U).$$

By Theorem 7.23, under assumptions (S1)–(S3), the min-function $V(\cdot)$ is second order Hadamard directionally differentiable at f tangentially to the set $\mathcal{K}$ and we have the following formula for the second order directional derivative in a direction $\delta \in \mathcal{K}$:

$$V'_f(\delta) = \inf_{h\in C(\bar{x})} \left\{2h^{\mathrm{T}}\nabla\delta(\bar{x}) + h^{\mathrm{T}}\nabla^2 f(\bar{x})h - \mathrm{s}\big(-\nabla f(\bar{x}), \mathcal{T}^2_X(\bar{x},h)\big)\right\}. \tag{5.31}$$

Here $C(\bar{x})$ is the critical cone of the true problem, $\mathcal{T}^2_X(\bar{x},h)$ is the second order tangent set to X at $\bar{x}$ and $\mathrm{s}(\cdot, A)$ denotes the support function of set A. (See page 386 for the definition of second order directional derivatives.)

Moreover, suppose that the set X is given in the form

$$X := \{x \in \mathbb{R}^n : G(x) \in K\}, \tag{5.32}$$

where $G : \mathbb{R}^n \to \mathbb{R}^m$ is a twice continuously differentiable mapping and $K \subset \mathbb{R}^m$ is a closed convex cone. Then, under Robinson constraint qualification, the optimal value of the right-hand side of (5.31) can be written in a dual form (compare with (7.84)), which results in the following formula for the second order directional derivative in a direction $\delta \in \mathcal{K}$:

$$V''_f(\delta) = \inf_{h \in C(\bar{x})} \sup_{\lambda \in \Lambda(\bar{x})} \left\{2h^{\mathrm{T}}\nabla\delta(\bar{x}) + h^{\mathrm{T}}\nabla^2_{xx}L(\bar{x}, \lambda)h - \mathrm{s}\big(\lambda, \mathfrak{T}(h)\big)\right\}. \tag{5.33}$$

Here

$$\mathfrak{T}(h) := \mathcal{T}^2_K\big(G(\bar{x}), [\nabla G(\bar{x})]h\big), \tag{5.34}$$

and $L(x, \lambda)$ is the Lagrangian and $\Lambda(\bar{x})$ is the set of Lagrange multipliers of the true problem.

Theorem 5.8. *Suppose that the assumptions* (S1)–(S4) *hold and* $N^{1/2}(\hat{f}_N - f)$ *converges in distribution to a random element* Y *of* $W^{1,\infty}(U)$. *Then*

$$\hat{\vartheta}_N = \hat{f}_N(\bar{x}) + \tfrac{1}{2}V''_f(\hat{f}_N - f) + o_p(N^{-1}), \tag{5.35}$$

and

$$N\big[\hat{\vartheta}_N - \hat{f}_N(\bar{x})\big] \xrightarrow{\mathcal{D}} \tfrac{1}{2}V''_f(Y). \tag{5.36}$$

Moreover, suppose that for every $\delta \in \mathcal{K}$ *the problem in the right-hand side of* (5.31) *has unique optimal solution* $\bar{h} = \bar{h}(\delta)$. *Then*

$$N^{1/2}\big(\hat{x}_N - \bar{x}\big) \xrightarrow{\mathcal{D}} \bar{h}(Y). \tag{5.37}$$

Proof. By the second order Delta theorem, Theorem 7.62, we have that

$$\hat{\vartheta}_N = \vartheta^* + V'_f(\hat{f}_N - f) + \tfrac{1}{2}V''_f(\hat{f}_N - f) + o_p(N^{-1})$$

and

$$N\big[\hat{\vartheta}_N - \vartheta^* - V'_f(\hat{f}_N - f)\big] \xrightarrow{\mathcal{D}} \tfrac{1}{2}V''_f(Y).$$

We also have (compare with formula (5.27)) that

$$V'_f(\hat{f}_N - f) = \hat{f}_N(\bar{x}) - f(\bar{x}) = \hat{f}_N(\bar{x}) - \vartheta^*,$$

and hence (5.35) and (5.36) follow.

Now consider a (measurable) mapping $\mathfrak{x} : W^{1,\infty}(U) \to \mathbb{R}^n$ such that

$$\mathfrak{x}(\psi) \in \arg\min_{x \in X} \psi(x), \quad \psi \in W^{1,\infty}(U).$$

We have that $\mathfrak{x}(f) = \bar{x}$, and by (7.82) of Theorem 7.23 we have that $\mathfrak{x}(\cdot)$ is Hadamard directionally differentiable at f tangentially to $\mathcal{K}$, and for $\delta \in \mathcal{K}$ the directional derivative $\mathfrak{x}'(f, \delta)$ is equal to the optimal solution in the right-hand side of (5.31), provided

that it is unique. By applying the Delta theorem, Theorem 7.61, this completes the proof of (5.37). □

One of the difficulties in applying the above theorem is verification of convergence in distribution of $N^{1/2}(\hat{f}_N - f)$ in the space $W^{1,\infty}(X)$. Actually, it could be easier to prove asymptotic results (5.35)–(5.37) by direct methods. Note that formulas (5.31) and (5.33), for the second order directional derivatives $V''_f(\hat{f}_N - f)$, involve statistical properties of $\hat{f}_N(x)$ only at the (fixed) point $\bar{x}$. Note also that by the (finite dimensional) CLT we have that $N^{1/2}\big[\nabla \hat{f}_N(\bar{x}) - \nabla f(\bar{x})\big]$ converges in distribution to normal $\mathcal{N}(0, \Sigma)$ with the covariance matrix

$$\Sigma = \mathbb{E}\big[\big(\nabla F(\bar{x}, \xi) - \nabla f(\bar{x})\big)\big(\nabla F(\bar{x}, \xi) - \nabla f(\bar{x})\big)^{\mathsf{T}}\big], \tag{5.38}$$

provided that this covariance matrix is well defined and $\mathbb{E}[\nabla F(\bar{x}, \xi)] = \nabla f(\bar{x})$, i.e., the differentiation and expectation operators can be interchanged (see Theorem 7.44).

Let Z be a random vector having normal distribution, $Z \sim \mathcal{N}(0, \Sigma)$, with covariance matrix Σ defined in (5.38), and let the set X be given in the form (5.32). Then by the above discussion and formula (5.33), we have that under appropriate regularity conditions,

$$N\big[\hat{\vartheta}_N - \hat{f}_N(\bar{x})\big] \xrightarrow{\mathcal{D}} \tfrac{1}{2}\mathfrak{v}(Z), \tag{5.39}$$

where $\mathfrak{v}(Z)$ is the optimal value of the problem

$$\operatorname*{Min}_{h \in C(\bar{x})} \sup_{\lambda \in \Lambda(\bar{x})} \left\{2h^{\mathsf{T}}Z + h^{\mathsf{T}}\nabla^2_{xx}L(\bar{x}, \lambda)h - \mathrm{s}\big(\lambda, \mathfrak{T}(h)\big)\right\}, \tag{5.40}$$

with $\mathfrak{T}(h)$ being the second order tangent set defined in (5.34). Moreover, if for all Z, problem (5.40) possesses unique optimal solution $\bar{h} = \mathfrak{h}(Z)$, then

$$N^{1/2}\big(\hat{x}_N - \bar{x}\big) \xrightarrow{\mathcal{D}} \mathfrak{h}(Z). \tag{5.41}$$

Recall also that if the cone K is polyhedral, then the curvature term $\mathrm{s}\big(\lambda, \mathfrak{T}(h)\big)$ vanishes.

Remark 7. Note that $\mathbb{E}\big[\hat{f}_N(\bar{x})\big] = f(\bar{x}) = \vartheta^*$. Therefore, under the respective regularity conditions, in particular under the assumption that the true problem has unique optimal solution $\bar{x}$, we have by (5.39) that the expected value of the term $\frac{1}{2}N^{-1}\mathfrak{v}(Z)$ can be viewed as the asymptotic bias of $\hat{\vartheta}_N$. This asymptotic bias is of order $O(N^{-1})$. This can be compared with formula (5.29) for the asymptotic bias of order $O(N^{-1/2})$ when the set of optimal solutions of the true problem is not a singleton. Note also that $\mathfrak{v}(\cdot)$ is nonpositive; to see this, just take $h = 0$ in (5.40).

As an example, consider the case where the set X is defined by a finite number of constraints:

$$X := \big\{x \in \mathbb{R}^n : g_i(x) = 0,\ i = 1, \ldots, q,\ g_i(x) \leq 0,\ i = q+1, \ldots, p\big\} \tag{5.42}$$

with the functions $g_i(x)$, $i = 1, \ldots, p$, being twice continuously differentiable. This is a particular form of (5.32) with $G(x) := (g_1(x), \ldots, g_p(x))$ and $K := \{0_q\} \times \mathbb{R}^{p-q}_-$. Denote

$$\mathcal{I}(\bar{x}) := \{i : g_i(\bar{x}) = 0,\ i = q+1, \ldots, p\}$$

the index set of active at $\bar{x}$ inequality constraints. Suppose that the linear independence constraint qualification (LICQ) holds at $\bar{x}$, i.e., the gradient vectors $\nabla g_i(\bar{x})$, $i \in \{1, \ldots, q\} \cup \mathcal{I}(\bar{x})$, are linearly independent. Then the corresponding set of Lagrange multipliers is a singleton, $\Lambda(\bar{x}) = \{\bar{\lambda}\}$. In that case

$$C(\bar{x}) = \left\{h : h^{\mathrm{T}} \nabla g_i(\bar{x}) = 0,\ i \in \{1, \ldots, q\} \cup \mathcal{I}_+(\bar{\lambda}),\ h^{\mathrm{T}} \nabla g_i(\bar{x}) \leq 0, i \in \mathcal{I}_0(\bar{\lambda})\right\},$$

where

$$\mathcal{I}_0(\bar{\lambda}) := \left\{i \in \mathcal{I}(\bar{x}) : \bar{\lambda}_i = 0\right\} \quad \text{and} \quad \mathcal{I}_+(\bar{\lambda}) := \left\{i \in \mathcal{I}(\bar{x}) : \bar{\lambda}_i > 0\right\}.$$

Consequently problem (5.40) takes the form

$$\begin{array}{ll} \underset{h \in \mathbb{R}^n}{\mathrm{Min}} & 2h^{\mathrm{T}} Z + h^{\mathrm{T}} \nabla^2_{xx} L(\bar{x}, \bar{\lambda}) h \\ \text{s.t.} & h^{\mathrm{T}} \nabla g_i(\bar{x}) = 0,\ i \in \{1, \ldots, q\} \cup \mathcal{I}_+(\bar{\lambda}),\ h^{\mathrm{T}} \nabla g_i(\bar{x}) \leq 0, i \in \mathcal{I}_0(\bar{\lambda}). \end{array} \tag{5.43}$$

This is a quadratic programming problem. The linear independence constraint qualification implies that problem (5.43) has a unique vector $\alpha(Z)$ of Lagrange multipliers and that it has a unique optimal solution $\mathfrak{h}(Z)$ if the Hessian matrix $H := \nabla^2_{xx} L(\bar{x}, \bar{\lambda})$ is positive definite over the linear space defined by the first $q + |\mathcal{I}_+(\bar{\lambda})|$ (equality) linear constraints in (5.43).

If, furthermore, the strict complementarity condition holds, i.e., $\bar{\lambda}_i > 0$ for all $i \in \mathcal{I}_+(\bar{\lambda})$, or in other words $\mathcal{I}_0(\bar{\lambda}) = \emptyset$, then $h = \mathfrak{h}(Z)$ and $\alpha = \alpha(Z)$ can be obtained as solutions of the following system of linear equations

$$\begin{bmatrix} H & A \\ A^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} h \\ \alpha \end{bmatrix} = \begin{bmatrix} Z \\ 0 \end{bmatrix}. \tag{5.44}$$

Here $H = \nabla^2_{xx} L(\bar{x}, \bar{\lambda})$ and A is the $n \times (q + |\mathcal{I}(\bar{x})|)$ matrix whose columns are formed by vectors $\nabla g_i(\bar{x})$, $i \in \{1, \ldots, q\} \cup \mathcal{I}(\bar{x})$. Then

$$N^{1/2} \begin{bmatrix} \hat{x}_N - \bar{x} \\ \hat{\lambda}_N - \bar{\lambda} \end{bmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}\big(0,\, J^{-1} \Upsilon J^{-1}\big), \tag{5.45}$$

where

$$J := \begin{bmatrix} H & A \\ A^{\mathrm{T}} & 0 \end{bmatrix} \quad \text{and} \quad \Upsilon := \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix},$$

provided that the matrix J is nonsingular.

Under the linear independence constraint qualification and strict complementarity condition, we have by the second order necessary conditions that the Hessian matrix $H = \nabla^2_{xx} L(\bar{x}, \bar{\lambda})$ is positive semidefinite over the linear space $\{h : A^{\mathrm{T}} h = 0\}$. Note that this linear space coincides here with the critical cone $C(\bar{x})$. It follows that the matrix J is nonsingular iff H is positive definite over this linear space. That is, here the nonsingularity of the matrix J is equivalent to the second order sufficient conditions at $\bar{x}$.

Remark 8. As mentioned earlier, the curvature term $\mathrm{s}\big(\lambda, \mathfrak{T}(h)\big)$ in the auxiliary problem (5.40) vanishes if the cone K is polyhedral. In particular, this happens if $K = \{0_q\} \times \mathbb{R}^{p-q}_-$, and hence the feasible set X is given in the form (5.42). This curvature term can also be written in an explicit form for some nonpolyhedral cones, in particular for the cone of positive semidefinite matrices (see [22, section 5.3.6]).

5.1.4 Minimax Stochastic Programs

Sometimes it is worthwhile to consider minimax stochastic programs of the form

$$\operatorname*{Min}_{x\in X} \sup_{y\in Y} \big\{ f(x, y) := \mathbb{E}[F(x, y, \xi)] \big\}, \tag{5.46}$$

where $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$ are closed sets, $F : X \times Y \times \Xi \to \mathbb{R}$ and $\xi = \xi(\omega)$ is a random vector whose probability distribution is supported on set $\Xi \subset \mathbb{R}^d$. The corresponding SAA problem is obtained by using the sample average as an approximation of the expectation $f(x, y)$, that is,

$$\operatorname*{Min}_{x\in X} \sup_{y\in Y} \left\{ \hat{f}_N(x, y) := \frac{1}{N} \sum_{j=1}^{N} F(x, y, \xi^j) \right\}. \tag{5.47}$$

As before, denote by, ϑ^* and $\hat{\vartheta}_N$ the optimal values of (5.46) and (5.47), respectively, and by $S_x \subset X$ and $\hat{S}_{x,N} \subset X$ the respective sets of optimal solutions. Recall that $F(x, y, \xi)$ is said to be a *Carathéodory function* if $F(x, y, \xi(\cdot))$ is measurable for every (x, y) and $F(\cdot, \cdot, \xi)$ is continuous for a.e. $\xi \in \Xi$. We make the following assumptions:

(A′1) $F(x, y, \xi)$ is a Carathéodory function.

(A′2) The sets X and Y are nonempty and compact.

(A′3) $F(x, y, \xi)$ is dominated by an integrable function, i.e., there is an open set $N \subset \mathbb{R}^{n+m}$ containing the set $X \times Y$ and an integrable, with respect to the probability distribution of the random vector ξ, function $h(\xi)$ such that $|F(x, y, \xi)| \leq h(\xi)$ for all $(x, y) \in N$ and a.e. $\xi \in \Xi$.

By Theorem 7.43 it follows that the expected value function $f(x, y)$ is continuous on $X \times Y$. Since Y is compact, this implies that the max-function

$$\phi(x) := \sup_{y\in Y} f(x, y)$$

is continuous on X. It also follows that the function $\hat{f}_N(x, y) = \hat{f}_N(x, y, \omega)$ is a Carathéodory function. Consequently, the sample average max-function

$$\hat{\phi}_N(x, \omega) := \sup_{y\in Y} \hat{f}_N(x, y, \omega)$$

is a Carathéodory function. Since $\hat{\vartheta}_N = \hat{\vartheta}_N(\omega)$ is given by the minimum of the Carathéodory function $\hat{\phi}_N(x, \omega)$, it follows that it is measurable.

Theorem 5.9. *Suppose that assumptions* (A′1)–(A′3) *hold and the sample is* iid. *Then* $\hat{\vartheta}_N \to \vartheta^*$ *and* $\mathbb{D}(\hat{S}_{x,N}, S_x) \to 0$ *w.p.* 1 *as* $N \to \infty$.

Proof. By Theorem 7.48 we have that under the specified assumptions, $\hat{f}_N(x, y)$ converges to $f(x, y)$ w.p. 1 uniformly on $X \times Y$. That is, $\Delta_N \to 0$ w.p. 1 as $N \to \infty$, where

$$\Delta_N := \sup_{(x,y)\in X\times Y} \left| \hat{f}_N(x, y) - f(x, y) \right|.$$

Consider $\hat{\phi}_N(x) := \sup_{y\in Y} \hat{f}_N(x, y)$ and $\phi(x) := \sup_{y\in Y} f(x, y)$. We have that

$$\sup_{x\in X} \left|\hat{\phi}_N(x) - \phi(x)\right| \le \Delta_N,$$

and hence $\left|\hat{\vartheta}_N - \vartheta^*\right| \le \Delta_N$. It follows that $\hat{\vartheta}_N \to \vartheta^*$ w.p. 1.

The function $\phi(x)$ is continuous and $\hat{\phi}_N(x)$ is continuous w.p. 1. Consequently, the set S_x is nonempty and $\hat{S}_{x,N}$ is nonempty w.p. 1. Now to prove that $\mathbb{D}(\hat{S}_{x,N}, S_x) \to 0$ w.p. 1, one can proceed exactly in the same way as in the proof of Theorem 5.3. □

We discuss now asymptotics of $\hat{\vartheta}_N$ in the convex–concave case. We make the following additional assumptions:

(A′4) The sets X and Y are convex, and or a.e. $\xi \in \Xi$ the function $F(\cdot, \cdot, \xi)$ is convex–concave on $X \times Y$, i.e., the function $F(\cdot, y, \xi)$ is convex on X for every $y \in Y$, and the function $F(x, \cdot, \xi)$ is concave on Y for every $x \in X$.

It follows that the expected value function $f(x, y)$ is convex concave and continuous on $X \times Y$. Consequently, problem (5.46) and its dual

$$\operatorname*{Max}_{y\in Y} \inf_{x\in X} f(x, y) \tag{5.48}$$

have nonempty and bounded sets of optimal solutions $S_x \subset X$ and $S_y \subset Y$, respectively. Moreover, the optimal values of problems (5.46) and (5.48) are equal to each other and $S_x \times S_y$ forms the set of saddle points of these problems.

(A′5) For some point $(x, y) \in X \times Y$, the expectation $\mathbb{E}[F(x, y, \xi)^2]$ is finite, and there exists a measurable function $C : \Xi \to \mathbb{R}_+$ such that $\mathbb{E}[C(\xi)^2]$ is finite and the inequality

$$|F(x, y, \xi) - F(x', y', \xi)| \le C(\xi)\big(\|x - x'\| + \|y - y'\|\big) \tag{5.49}$$

holds for all $(x, y), (x', y') \in X \times Y$ and a.e. $\xi \in \Xi$.

The above assumption implies that $f(x, y)$ is Lipschitz continuous on $X \times Y$ with Lipschitz constant $\kappa = \mathbb{E}[C(\xi)]$.

Theorem 5.10. *Consider the minimax stochastic problem* (5.46) *and the SAA problem* (5.47) *based on an iid sample. Suppose that assumptions* (A′1)–(A′2) *and* (A′4)–(A′5) *hold. Then*

$$\hat{\vartheta}_N = \inf_{x\in S_x} \sup_{y\in S_y} \hat{f}_N(x, y) + o_p(N^{-1/2}). \tag{5.50}$$

Moreover, if the sets $S_x = \{\bar{x}\}$ and $S_y = \{\bar{y}\}$ are singletons, then $N^{1/2}(\hat{\vartheta}_N - \vartheta^*)$ converges in distribution to normal with zero mean and variance $\sigma^2 = \mathbb{V}\mathrm{ar}[F(\bar{x}, \bar{y}, \xi)]$.

Proof. Consider the space $C(X, Y)$ of continuous functions $\psi : X \times Y \to \mathbb{R}$ equipped with the sup-norm $\|\psi\| = \sup_{x\in X, y\in Y} |\psi(x, y)|$, and set $\mathcal{K} \subset C(X, Y)$ formed by convex–concave on $X \times Y$ functions. It is not difficult to see that the set $\mathcal{K}$ is a closed (in the

norm topology of $C(X, Y)$) and convex cone. Consider the optimal value function $V : C(X, Y) \to \mathbb{R}$ defined as

$$V(\psi) := \inf_{x \in X} \sup_{y \in Y} \psi(x, y) \quad \text{for } \psi \in C(X, Y). \tag{5.51}$$

Recall that it is said that $V(\cdot)$ is Hadamard directionally differentiable at $f \in \mathcal{K}$, tangentially to the set $\mathcal{K}$, if the following limit exists for any $\gamma \in \mathcal{T}_{\mathcal{K}}(f)$:

$$V'_f(\gamma) := \lim_{\substack{t \downarrow 0, \eta \to \gamma \\ f + t\eta \in \mathcal{K}}} \frac{V(f + t\eta) - V(f)}{t}. \tag{5.52}$$

By Theorem 7.24 we have that the optimal value function $V(\cdot)$ is Hadamard directionally differentiable at f tangentially to the set $\mathcal{K}$ and

$$V'_f(\gamma) = \inf_{x \in S_x} \sup_{y \in S_y} \gamma(x, y) \tag{5.53}$$

for any $\gamma \in \mathcal{T}_{\mathcal{K}}(f)$.

By the assumption (A′5) we have that $N^{1/2}(\hat{f}_N - f)$, considered as a sequence of random elements of $C(X, Y)$, converges in distribution to a random element of $C(X, Y)$. Then by noting that $\vartheta^* = f(x^*, y^*)$ for any $(x^*, y^*) \in S_x \times S_y$ and using Hadamard directional differentiability of the optimal value function, tangentially to the set $\mathcal{K}$, together with formula (5.53) and a version of the Delta method given in Theorem 7.61, we can complete the proof. □

Suppose now that the feasible set X is defined by constraints in the form (5.11). The Lagrangian function of the true problem is

$$L(x, \lambda) := f(x) + \sum_{i=1}^{p} \lambda_i g_i(x).$$

Suppose also that the problem is *convex*, that is, the set X_0 is convex and for all $\xi \in \Xi$ the functions $F(\cdot, \xi)$ and $G_i(\cdot, \xi), i = 1, \ldots, p$, are convex. Suppose, further, that the functions $f(x)$ and $g_i(x)$ are finite valued on a neighborhood of the set S (of optimal solutions of the true problem) and the Slater condition holds. Then with every optimal solution $\bar{x} \in S$ is associated a nonempty and bounded set Λ of Lagrange multipliers vectors $\lambda = (\lambda_1, \ldots, \lambda_p)$ satisfying the optimality conditions

$$\bar{x} \in \arg\min_{x \in X_0} L(x, \lambda), \quad \lambda_i \geq 0 \ \text{ and } \ \lambda_i g_i(\bar{x}) = 0, \ i = 1, \ldots, p. \tag{5.54}$$

The set Λ coincides with the set of optimal solutions of the dual of the true problem and therefore is the same for any optimal solution $\bar{x} \in S$.

Let $\hat{\vartheta}_N$ be the optimal value of the SAA problem (5.10) with X_N given in the form (5.13). That is, $\hat{\vartheta}_N$ is the optimal value of the problem

$$\operatorname*{Min}_{x \in X_0} \hat{f}_N(x) \ \text{ subject to } \ \hat{g}_{iN}(x) \leq 0, \ i = 1, \ldots, p, \tag{5.55}$$

with $\hat{f}_N(x)$ and $\hat{g}_{iN}(x)$ being the SAA functions of the respective integrands $F(x, \xi)$ and $G_i(x, \xi)$, $i = 1, \ldots, p$. Assume that conditions (A1) and (A2), formulated on page 164, are satisfied for the integrands F and $G_i, i = 1, \ldots, p$, i.e., finiteness of the corresponding

second order moments and the Lipschitz continuity condition of assumption (A2) hold for each function. It follows that the corresponding expected value functions $f(x)$ and $g_i(x)$ are finite valued and continuous on X. As in Theorem 5.7, we denote by $Y(x)$ random variables which are normally distributed and have the same covariance structure as $F(x, \xi)$. We also denote by $Y_i(x)$ random variables which are normally distributed and have the same covariance structure as $G_i(x, \xi)$, $i = 1, \ldots, p$.

Theorem 5.11. *Let $\hat{\vartheta}_N$ be the optimal value of the SAA problem* (5.55)*. Suppose that the sample is iid, the problem is convex, and the following conditions are satisfied:* (i) *the set S, of optimal solutions of the true problem, is nonempty and bounded,* (ii) *the functions $f(x)$ and $g_i(x)$ are finite valued on a neighborhood of S,* (iii) *the Slater condition for the true problem holds, and* (iv) *the assumptions* (A1) *and* (A2) *hold for the integrands F and G_i, $i = 1, \ldots, p$. Then*

$$N^{1/2}\left(\hat{\vartheta}_N - \vartheta^*\right) \xrightarrow{\mathcal{D}} \inf_{x \in S} \sup_{\lambda \in \Lambda} \left[Y(x) + \sum_{i=1}^{p} \lambda_i Y_i(x) \right]. \tag{5.56}$$

If, moreover, $S = \{\bar{x}\}$ and $\Lambda = \{\bar{\lambda}\}$ are singletons, then

$$N^{1/2}\left(\hat{\vartheta}_N - \vartheta^*\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2) \tag{5.57}$$

with

$$\sigma^2 := \mathbb{V}\mathrm{ar}\left[F(\bar{x}, \xi) + \sum_{i=1}^{p} \bar{\lambda}_i G_i(\bar{x}, \xi) \right]. \tag{5.58}$$

Proof. Since the problem is convex and the Slater condition (for the true problem) holds, we have that ϑ^* is equal to the optimal value of the (Lagrangian) dual

$$\operatorname*{Max}_{\lambda \geq 0} \inf_{x \in X_0} L(x, \lambda), \tag{5.59}$$

and the set of optimal solutions of (5.59) is nonempty and compact and coincides with the set of Lagrange multipliers Λ. Since the problem is convex and S is nonempty and bounded, the problem can be considered on a bounded neighborhood of S, i.e., without loss of generality it can be assumed that the set X is compact. The proof can now be completed by applying Theorem 5.10. □

Remark 9. There are two possible approaches to generating random samples in construction of SAA problems of the form (5.55) by Monte Carlo sampling techniques. One is to use the same sample $\xi^1, \ldots, \xi^N$ for estimating the functions $f(x)$ and $g_i(x)$, $i = 1, \ldots, p$, by their SAA counterparts. The other is to use independent samples, possibly of different sizes, for each of these functions (see Remark 5 on page 161). The asymptotic results of Theorem 5.11 are for the case of the same sample. The (asymptotic) variance σ^2, given in (5.58), is equal to the sum of the variances of $F(\bar{x}, \xi)$ and $\bar{\lambda}_i G_i(\bar{x}, \xi)$, $i = 1, \ldots, p$, and all their covariances. If we use the independent samples construction, then a similar result holds but without the corresponding covariance terms. Since in the case of the same sample these covariance terms could be expected to be positive, it would be advantageous to use the independent, rather than the same, samples approach in order to reduce variability of the SAA estimates.

5.2 Stochastic Generalized Equations

In this section we discuss the following so-called *stochastic generalized equations*. Consider a random vector ξ whose distribution is supported on a set $\Xi \subset \mathbb{R}^d$, a mapping $\Phi : \mathbb{R}^n \times \Xi \to \mathbb{R}^n$, and a multifunction $\Gamma : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$. Suppose that the expectation $\phi(x) := \mathbb{E}[\Phi(x, \xi)]$ is well defined and finite valued. We refer to

$$\phi(x) \in \Gamma(x) \tag{5.60}$$

as true, or expected value, generalized equation and say that a point $\bar{x} \in \mathbb{R}^n$ is a solution of (5.60) if $\phi(\bar{x}) \in \Gamma(\bar{x})$.

The above abstract setting includes the following cases. If $\Gamma(x) = \{0\}$ for every $x \in \mathbb{R}^n$, then (5.60) becomes the ordinary equation $\phi(x) = 0$. As another example, let $\Gamma(\cdot) := \mathcal{N}_X(\cdot)$, where X is a nonempty closed convex subset of $\mathbb{R}^n$ and $\mathcal{N}_X(x)$ denotes the (outward) normal cone to X at x. Recall that, by the definition, $\mathcal{N}_X(x) = \emptyset$ if $x \notin X$. In that case $\bar{x}$ is a solution of (5.60) iff $\bar{x} \in X$ and the following so-called variational inequality holds:

$$(x - \bar{x})^{\mathrm{T}} \phi(\bar{x}) \leq 0, \quad \forall x \in X. \tag{5.61}$$

Since the mapping $\phi(x)$ is given in the form of the expectation, we refer to such variational inequalities as *stochastic variational inequalities*. Note that if $X = \mathbb{R}^n$, then $\mathcal{N}_X(x) = \{0\}$ for any $x \in \mathbb{R}^n$, and hence in that case the above variational inequality is reduced to the equation $\phi(x) = 0$. Let us also remark that if $\Phi(x, \xi) := -\nabla_x F(x, \xi)$ for some real valued function $F(x, \xi)$, and the interchangeability formula $\mathbb{E}[\nabla_x F(x, \xi)] = \nabla f(x)$ holds, i.e., $\phi(x) = -\nabla f(x)$, where $f(x) := \mathbb{E}[F(x, \xi)]$, then (5.61) represents first order necessary, and if $f(x)$ is convex, sufficient conditions for $\bar{x}$ to be an optimal solution for the optimization problem (5.1).

If the feasible set X of the optimization problem (5.1) is defined by constraints in the form

$$X := \big\{x \in \mathbb{R}^n : g_i(x) = 0,\ i = 1, \ldots, q,\ g_i(x) \leq 0,\ i = q + 1, \ldots, p\big\} \tag{5.62}$$

with $g_i(x) := \mathbb{E}[G_i(x, \xi)]$, $i = 1, \ldots, p$, then the corresponding first-order Karush–Kuhn–Tucker (KKT) optimality conditions can be written in a form of variational inequality. That is, let $z := (x, \lambda) \in \mathbb{R}^{n+p}$ and

$$\begin{array}{rcl} L(z, \xi) & := & F(x, \xi) + \sum_{i=1}^p \lambda_i G_i(x, \xi), \\ \ell(z) & := & \mathbb{E}[L(z, \xi)] = f(x) + \sum_{i=1}^p \lambda_i g_i(x) \end{array}$$

be the corresponding Lagrangians. Define

$$\Phi(z, \xi) := \left[\begin{array}{c} \nabla_x L(z, \xi) \\ G_1(x, \xi) \\ \cdots \\ G_p(x, \xi) \end{array}\right] \text{ and } \Gamma(z) := \mathcal{N}_K(z), \tag{5.63}$$

where $K := \mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R}^{p-q}_+ \subset \mathbb{R}^{n+p}$. Note that if $z \in K$, then

$$\mathcal{N}_K(z) = \left\{(v, \gamma) \in \mathbb{R}^{n+p} : \begin{array}{l} v = 0 \text{ and } \gamma_i = 0,\ i = 1, \ldots, q, \\ \gamma_i = 0,\ i \in \mathcal{I}_+(\lambda),\ \gamma_i \leq 0,\ i \in \mathcal{I}_0(\lambda) \end{array}\right\}, \tag{5.64}$$

where

$$\begin{array}{ll} \mathcal{I}_0(\lambda) & := \{i : \lambda_i = 0,\ i = q+1, \ldots, p\}, \\ \mathcal{I}_+(\lambda) & := \{i : \lambda_i > 0,\ i = q+1, \ldots, p\}, \end{array} \tag{5.65}$$

and $\mathcal{N}_K(z) = \emptyset$ if $z \notin K$. Consequently, assuming that the interchangeability formula holds, and hence $\mathbb{E}[\nabla_x L(z,\xi)] = \nabla \ell_x(z) = \nabla f(x) + \sum_{i=1}^p \lambda_i \nabla g_i(x)$, we have that

$$\phi(z) := \mathbb{E}[\Phi(z,\xi)] = \begin{bmatrix} \nabla_x \ell(z) \\ g_1(x) \\ \cdots \\ g_p(x) \end{bmatrix}, \tag{5.66}$$

and variational inequality $\phi(z) \in \mathcal{N}_K(z)$ represents the KKT optimality conditions for the true optimization problem.

We make the following assumption about the multifunction $\Gamma(x)$:

(E1) The multifunction $\Gamma(x)$ is *closed*, that is, the following holds: if $x_k \to x$, $y_k \in \Gamma(x_k)$ and $y_k \to y$, then $y \in \Gamma(x)$.

The above assumption implies that the multifunction $\Gamma(x)$ is closed valued, i.e., for any $x \in \mathbb{R}^n$ the set $\Gamma(x)$ is closed. For variational inequalities, assumption (E1) always holds, i.e., the multifunction $x \mapsto \mathcal{N}_X(x)$ is closed.

Now let $\xi^1, \ldots, \xi^N$ be a random sample of N realizations of the random vector ξ and let $\hat{\phi}_N(x) := N^{-1} \sum_{j=1}^N \Phi(x, \xi^j)$ be the corresponding sample average estimate of $\phi(x)$. We refer to

$$\hat{\phi}_N(x) \in \Gamma(x) \tag{5.67}$$

as the SAA generalized equation. There are standard numerical algorithms for solving nonlinear equations which can be applied to (5.67) in the case $\Gamma(x) \equiv \{0\}$, i.e., when (5.67) is reduced to the ordinary equation $\hat{\phi}_N(x) = 0$. There are also numerical procedures for solving variational inequalities. We are not going to discuss such numerical algorithms but rather concentrate on statistical properties of solutions of SAA equations. We denote by S and $\hat{S}_N$ the sets of (all) solutions of the true (5.60) and SAA (5.67) generalized equations, respectively.

5.2.1 Consistency of Solutions of the SAA Generalized Equations

In this section we discuss convergence properties of the SAA solutions.

Theorem 5.12. *Let C be a compact subset of $\mathbb{R}^n$ such that $S \subset C$. Suppose that:* (i) *the multifunction $\Gamma(x)$ is closed (assumption* (E1)*),* (ii) *the mapping $\phi(x)$ is continuous on C,* (iii) *w.p. 1 for N large enough the set $\hat{S}_N$ is nonempty and $\hat{S}_N \subset C$, and* (iv) *$\hat{\phi}_N(x)$ converges to $\phi(x)$ w.p. 1 uniformly on C as $N \to \infty$. Then $\mathbb{D}(\hat{S}_N, S) \to 0$ w.p. 1 as $N \to \infty$.*

Proof. The above result basically is deterministic in the sense that if we view $\hat{\phi}_N(x) = \hat{\phi}_N(x,\omega)$ as defined on a common probability space, then it should be verified for a.e. ω. Therefore we omit saying "w.p. 1." Consider a sequence $\hat{x}_N \in \hat{S}_N$. Because of assumption (iii), by passing to a subsequence if necessary, we need to show only that if $\hat{x}_N$ converges to a point x^*, then $x^* \in S$ (compare with the proof of Theorem 5.3). Now since it is

assumed that $\phi(\cdot)$ is continuous and $\hat{\phi}_N(x)$ converges to $\phi(x)$ uniformly, it follows that $\hat{\phi}_N(\hat{x}_N) \to \phi(x^*)$ (see Proposition 5.1). Since $\hat{\phi}_N(\hat{x}_N) \in \Gamma(\hat{x}_N)$, it follows by assumption (E1) that $\phi(x^*) \in \Gamma(x^*)$, which completes the proof. □

A few remarks about the assumptions involved in the above consistency result are now in order. By Theorem 7.48 we have that, in the case of iid sampling, the assumptions (ii) and (iv) of the above proposition are satisfied for any compact set C if the following assumption holds:

(E2) For every $\xi \in \Xi$ the function $\Phi(\cdot, \xi)$ is continuous on C and $\|\Phi(x, \xi)\|_{x \in C}$ is dominated by an integrable function.

There are two parts to assumption (iii) of Theorem 5.12, namely, that the SAA generalized equations do not have a solution which escapes to infinity, and that they possess at least one solution w.p. 1 for N large enough. The first of these assumptions often can be verified by ad hoc methods. The second assumption is more subtle. We will discuss it next. The following concept of strong regularity is due to Robinson [170].

Definition 5.13. *Suppose that the mapping $\phi(x)$ is continuously differentiable. We say that a solution $\bar{x} \in S$ is* strongly regular *if there exist neighborhoods $\mathcal{N}_1$ and $\mathcal{N}_2$ of $0 \in \mathbb{R}^n$ and $\bar{x}$, respectively, such that for every $\delta \in \mathcal{N}_1$ the (linearized) generalized equation*

$$\delta + \phi(\bar{x}) + \nabla\phi(\bar{x})(x - \bar{x}) \in \Gamma(x) \tag{5.68}$$

has a unique solution in $\mathcal{N}_2$, denoted $\tilde{x} = \tilde{x}(\delta)$, and $\tilde{x}(\cdot)$ is Lipschitz continuous on $\mathcal{N}_1$.

Note that it follows from the above conditions that $\tilde{x}(0) = \bar{x}$. In the case $\Gamma(x) \equiv \{0\}$, strong regularity simply means that the $n \times n$ Jacobian matrix $J := \nabla\phi(\bar{x})$ is invertible or, in other words, nonsingular. Also in the case of variational inequalities, the strong regularity condition was investigated extensively, we discuss this later.

Let $\mathcal{V}$ be a convex compact neighborhood of $\bar{x}$, i.e., $\bar{x} \in \mathrm{int}(\mathcal{V})$. Consider the space $C^1(\mathcal{V}, \mathbb{R}^n)$ of continuously differentiable mappings $\psi : \mathcal{V} \to \mathbb{R}^n$ equipped with the norm

$$\|\psi\|_{1,\mathcal{V}} := \sup_{x \in \mathcal{V}} \|\phi(x)\| + \sup_{x \in \mathcal{V}} \|\nabla\phi(x)\|.$$

The following (deterministic) result is essentially due to Robinson [171].

Suppose that $\phi(x)$ is continuously differentiable on $\mathcal{V}$, i.e., $\phi \in C^1(\mathcal{V}, \mathbb{R}^n)$. Let $\bar{x}$ be a strongly regular solution of the generalized equation (5.60). Then there exists $\varepsilon > 0$ such that for any $u \in C^1(\mathcal{V}, \mathbb{R}^n)$ satisfying $\|u - \phi\|_{1,\mathcal{V}} \leq \varepsilon$, the generalized equation $u(x) \in \Gamma(x)$ has a unique solution $\hat{x} = \hat{x}(u)$ in a neighborhood of $\bar{x}$, such that $\hat{x}(\cdot)$ is Lipschitz continuous (with respect the norm $\|\cdot\|_{1,\mathcal{V}}$), and

$$\hat{x}(u) = \tilde{x}\big(u(\bar{x}) - \phi(\bar{x})\big) + o(\|u - \phi\|_{1,\mathcal{V}}). \tag{5.69}$$

Clearly, we have that $\hat{x}(\phi) = \bar{x}$ and $\hat{x}\big(\hat{\phi}_N\big)$ is a solution, in a neighborhood of $\bar{x}$, of the SAA generalized equation provided that $\|\hat{\phi}_N - \phi\|_{1,\mathcal{V}} \leq \varepsilon$. Therefore, by employing the above results for the mapping $u(\cdot) := \hat{\phi}_N(\cdot)$ we immediately obtain the following.

Theorem 5.14. *Let $\bar{x}$ be a strongly regular solution of the true generalized equation* (5.60)*, and suppose that $\phi(x)$ and $\hat{\phi}_N(x)$ are continuously differentiable in a neighborhood $\mathcal{V}$ of $\bar{x}$ and $\|\hat{\phi}_N - \phi\|_{1,\mathcal{V}} \to 0$ w.p. 1 as $N \to \infty$. Then w.p. 1 for N large enough the SAA generalized equation* (5.67) *possesses a unique solution $\hat{x}_N$ in a neighborhood of $\bar{x}$, and $\hat{x}_N \to \bar{x}$ w.p. 1 as $N \to \infty$.*

The assumption that $\|\hat{\phi}_N - \phi\|_{1,\mathcal{V}} \to 0$ w.p. 1, in the above theorem, means that $\hat{\phi}_N(x)$ and $\nabla\hat{\phi}_N(x)$ converge w.p. 1 to $\phi(x)$ and $\nabla\phi(x)$, respectively, uniformly on $\mathcal{V}$. By Theorem 7.48, in the case of iid sampling this is ensured by the following assumption:

(E3) For a.e. ξ the mapping $\Phi(\cdot, \xi)$ is continuously differentiable on $\mathcal{V}$, and $\|\Phi(x, \xi)\|_{x\in\mathcal{V}}$ and $\|\nabla_x\Phi(x, \xi)\|_{x\in\mathcal{V}}$ are dominated by an integrable function.

Note that the assumption that $\Phi(\cdot, \xi)$ is continuously differentiable on a neighborhood of $\bar{x}$ is essential in the above analysis. By combining Theorems 5.12 and 5.14 we obtain the following result.

Theorem 5.15. *Let C be a compact subset of $\mathbb{R}^n$ and let $\bar{x}$ be a unique in C solution of the true generalized equation* (5.60)*. Suppose that:* (i) *the multifunction $\Gamma(x)$ is closed (assumption* (E1)*),* (ii) *for a.e. ξ the mapping $\Phi(\cdot, \xi)$ is continuously differentiable on C, and $\|\Phi(x, \xi)\|_{x\in C}$ and $\|\nabla_x\Phi(x, \xi)\|_{x\in C}$ are dominated by an integrable function,* (iii) *the solution $\bar{x}$ is strongly regular, and* (iv) *$\hat{\phi}_N(x)$ and $\nabla\hat{\phi}_N(x)$ converge w.p. 1 to $\phi(x)$ and $\nabla\phi(x)$, respectively, uniformly on C. Then w.p. 1 for N large enough the SAA generalized equation possesses unique in C solution $\hat{x}_N$ converging to $\bar{x}$ w.p. 1 as $N \to \infty$.*

Note again that if the sample is iid, then assumption (iv) in the above theorem is implied by assumption (ii) and hence is redundant.

5.2.2 Asymptotics of SAA Generalized Equations Estimators

By using the first order approximation (5.69) it is also possible to derive asymptotics of $\hat{x}_N$. Suppose for the moment that $\Gamma(x) \equiv \{0\}$. Then strong regularity means that the Jacobian matrix $J := \nabla\phi(\bar{x})$ is nonsingular and $\tilde{x}(\delta)$ is the solution of the corresponding linear equations and hence can be written in the form

$$\tilde{x}(\delta) = \bar{x} - J^{-1}\delta. \tag{5.70}$$

By using (5.70) and (5.69) with $u(\cdot) := \hat{\phi}_N(\cdot)$, we obtain under certain regularity conditions, which ensure that the remainder in (5.69) is of order $o_p(N^{-1/2})$, that

$$N^{1/2}(\hat{x}_N - \bar{x}) = -J^{-1}Y_N + o_p(1), \tag{5.71}$$

where $Y_N := N^{1/2}\left[\hat{\phi}_N(\bar{x}) - \phi(\bar{x})\right]$. Moreover, in the case of iid sample, we have by the CLT that $Y_N \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma)$, where Σ is the covariance matrix of the random vector $\Phi(\bar{x}, \xi)$. Consequently, $\hat{x}_N$ has asymptotically normal distribution with mean vector $\bar{x}$ and the covariance matrix $N^{-1}J^{-1}\Sigma J^{-1}$.

Suppose now that $\Gamma(\cdot) := \mathcal{N}_X(\cdot)$ with the set X being nonempty closed convex and *polyhedral*, and let $\bar{x}$ be a strongly regular solution of (5.60). Let $\tilde{x}(\delta)$ be the (unique) solution, of the corresponding linearized variational inequality (5.68), in a neighborhood of $\bar{x}$. Consider the cone

$$\mathcal{C}_X(\bar{x}) := \left\{ y \in \mathcal{T}_X(\bar{x}) : y^{\mathrm{T}} \phi(\bar{x}) = 0 \right\}, \tag{5.72}$$

called the *critical cone*, and the Jacobian matrix $J := \nabla\phi(\bar{x})$. Then for all δ sufficiently close to $0 \in \mathbb{R}^n$, we have that $\tilde{x}(\delta) - \bar{x}$ coincides with the solution $\tilde{d}(\delta)$ of the variational inequality

$$\delta + Jd \in \mathcal{N}_{\mathcal{C}_X(\bar{x})}(d). \tag{5.73}$$

Note that the mapping $\tilde{d}(\cdot)$ is positively homogeneous, i.e., for any $\delta \in \mathbb{R}^n$ and $t \geq 0$, it follows that $\tilde{d}(t\delta) = t\tilde{d}(\delta)$. Consequently, under the assumption that the solution $\bar{x}$ is strongly regular, we obtain by (5.69) that $\tilde{d}(\cdot)$ is the directional derivative of $\hat{x}(u)$, at $u = \phi$, in the Hadamard sense. Therefore, under appropriate regularity conditions ensuring functional CLT for $N^{1/2}(\hat{\phi}_N - \phi)$ in the space $C^1(\mathcal{V}, \mathbb{R}^n)$, it follows by the Delta theorem that

$$N^{1/2}(\hat{x}_N - \bar{x}) \xrightarrow{\mathcal{D}} \tilde{d}(Y), \tag{5.74}$$

where $Y \sim \mathcal{N}(0, \Sigma)$ and Σ is the covariance matrix of $\Phi(\bar{x}, \xi)$. Consequently, $\hat{x}_N$ is asymptotically normal iff the mapping $\tilde{d}(\cdot)$ is linear. This, in turn, holds if the cone $\mathcal{C}_X(\bar{x})$ is a linear space.

In the case $\Gamma(\cdot) := \mathcal{N}_X(\cdot)$, with the set X being nonempty closed convex and polyhedral, there is a complete characterization of the strong regularity in terms of the so-called *coherent orientation* associated with the matrix (mapping) $J := \nabla\phi(\bar{x})$ and the critical cone $\mathcal{C}_X(\bar{x})$. The interested reader is referred to [172], [79] for a discussion of this topic. Let us just remark that if $\mathcal{C}_X(\bar{x})$ is a linear subspace of $\mathbb{R}^n$, then the variational inequality (5.73) can be written in the form

$$P\delta + PJd = 0, \tag{5.75}$$

where P denotes the orthogonal projection matrix onto the linear space $\mathcal{C}_X(\bar{x})$. Then $\bar{x}$ is strongly regular iff the matrix (mapping) PJ restricted to the linear space $\mathcal{C}_X(\bar{x})$ is invertible or, in other words, nonsingular.

Suppose now that $S = \{\bar{x}\}$ is such that $\phi(\bar{x})$ belongs to the interior of the set $\Gamma(\bar{x})$. Then, since $\hat{\phi}_N(\bar{x})$ converges w.p. 1 to $\phi(\bar{x})$, it follows that the event "$\hat{\phi}_N(\bar{x}) \in \Gamma(\bar{x})$" happens w.p. 1 for N large enough. Moreover, by the LD principle (see (7.191)) we have that this event happens with probability approaching one exponentially fast. Of course, $\hat{\phi}_N(\bar{x}) \in \Gamma(\bar{x})$ means that $\hat{x}_N = \bar{x}$ is a solution of the SAA generalized equation (5.67). Therefore, in such case one may compute an exact solution of the true problem (5.60) by solving the SAA problem, with probability approaching one exponentially fast with increase of the sample size. Note that if $\Gamma(\cdot) := \mathcal{N}_X(\cdot)$ and $\bar{x} \in S$, then $\phi(\bar{x}) \in \operatorname{int}\Gamma(\bar{x})$ iff the critical cone $\mathcal{C}_X(\bar{x})$ is equal to $\{0\}$. In that case, the variational inequality (5.73) has solution $\bar{d} = 0$ for any δ, i.e., $\tilde{d}(\delta) \equiv 0$.

The above asymptotics can be applied, in particular, to the generalized equation (variational inequality) $\phi(z) \in \mathcal{N}_K(z)$, where $K := \mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R}^{p-q}_+$ and $\mathcal{N}_K(z)$ and $\phi(z)$ are given in (5.64) and (5.66), respectively. Recall that this variational inequality represents the KKT optimality conditions of the expected value optimization problem (5.1) with the

feasible set X given in the form (5.62). (We assume that the expectation functions $f(x)$ and $g_i(x)$, $i = 1, \ldots, p$, are continuously differentiable.) Let $\bar{x}$ be an optimal solution of the (expected value) problem (5.1). It is said that the LICQ holds at the point $\bar{x}$ if the gradient vectors $\nabla g_i(\bar{x})$, $i \in \{i : g_i(\bar{x}) = 0,\ i = 1, \ldots, p\}$, (of active at $\bar{x}$ constraints) are linearly independent. Under the LICQ, to $\bar{x}$ corresponds a unique vector $\bar{\lambda}$ of Lagrange multipliers, satisfying the KKT optimality conditions. Let $\bar{z} = (\bar{x}, \bar{\lambda})$ and $\mathit{1}_0(\lambda)$ and $\mathit{1}_+(\lambda)$ be the index sets defined in (5.65). Then

$$\mathcal{T}_K(\bar{z}) = \mathbb{R}^n \times \mathbb{R}^q \times \left\{\gamma \in \mathbb{R}^{p-q} : \gamma_i \geq 0,\ i \in \mathit{1}_0(\bar{\lambda})\right\}. \tag{5.76}$$

In order to simplify notation, let us assume that *all* constraints are active at $\bar{x}$, i.e., $g_i(\bar{x}) = 0$, $i = 1, \ldots, p$. Since for sufficiently small perturbations of x inactive constraints remain inactive, we do not lose generality in the asymptotic analysis by considering only active at $\bar{x}$ constraints. Then $\phi(\bar{z}) = 0$, and hence $\mathcal{C}_K(\bar{z}) = \mathcal{T}_K(\bar{z})$.

Assuming, further, that $f(x)$ and $g_i(x)$, $i = 1, \ldots, p$, are twice continuously differentiable, we have that the following second order necessary conditions hold at $\bar{x}$:

$$h^{\mathrm{T}} \nabla^2_{xx} \ell(\bar{z}) h \geq 0, \quad \forall h \in C_X(\bar{x}), \tag{5.77}$$

where

$$C_X(\bar{x}) := \left\{h : h^{\mathrm{T}} \nabla g_i(\bar{x}) = 0,\ i \in \{1, \ldots, q\} \cup \mathit{1}_+(\bar{\lambda}),\ h^{\mathrm{T}} \nabla g_i(\bar{x}) \leq 0,\ i \in \mathit{1}_0(\bar{\lambda})\right\}.$$

The corresponding second order sufficient conditions are

$$h^{\mathrm{T}} \nabla^2_{xx} \ell(\bar{z}) h > 0, \quad \forall h \in C_X(\bar{x}) \setminus \{0\}. \tag{5.78}$$

Moreover, $\bar{z}$ is a strongly regular solution of the corresponding generalized equation iff the LICQ holds at $\bar{x}$ and the following (strong) form of second order sufficient conditions is satisfied:

$$h^{\mathrm{T}} \nabla^2_{xx} \ell(\bar{z}) h > 0, \quad \forall h \in \mathrm{lin}(C_X(\bar{x})) \setminus \{0\}, \tag{5.79}$$

where

$$\mathrm{lin}(C_X(\bar{x})) := \left\{h : h^{\mathrm{T}} \nabla g_i(\bar{x}) = 0,\ i \in \{1, \ldots, q\} \cup \mathit{1}_+(\bar{\lambda})\right\}. \tag{5.80}$$

Under the LICQ, the set defined in the right-hand side of (5.80) is, indeed, the linear space generated by the cone $C_X(\bar{x})$. We also have here

$$J := \nabla \phi(\bar{z}) = \begin{bmatrix} H & A \\ A^{\mathrm{T}} & 0 \end{bmatrix}, \tag{5.81}$$

where $H := \nabla^2_{xx} \ell(\bar{z})$ and $A := \left[\nabla g_1(\bar{x}), \ldots, \nabla g_p(\bar{x})\right]$.

It is said that the *strict complementarity condition* holds at $\bar{x}$ if the index set $\mathit{1}_0(\bar{\lambda})$ is empty, i.e., all Lagrange multipliers corresponding to active at $\bar{x}$ inequality constraints are strictly positive. We have here that $\mathcal{C}_K(\bar{z})$ is a linear space, and hence the SAA estimator $\hat{z}_N = [\hat{x}_N, \hat{\lambda}_N]$ is asymptotically normal iff the strict complementarity condition holds. If the strict complementarity condition holds, then $\mathcal{C}_K(\bar{z}) = \mathbb{R}^{n+p}$ (recall that it is assumed that all constraints are active at $\bar{x}$), and hence the normal cone to $\mathcal{C}_K(\bar{z})$, at every point, is $\{0\}$. Consequently, the corresponding variational inequality (5.73) takes the form

$\delta + Jd = 0$. Under the strict complementarity condition, $\bar{z}$ is strongly regular iff the matrix J is nonsingular. It follows that under the above assumptions together with the strict complementarity condition, the following asymptotics hold (compare with (5.45)):

$$N^{1/2}\big(\hat{z}_N - \bar{z}\big) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, J^{-1}\Sigma J^{-1}\right), \tag{5.82}$$

where Σ is the covariance matrix of the random vector $\Phi(\bar{z}, \xi)$ defined in (5.63).

5.3 Monte Carlo Sampling Methods

In this section we assume that a random sample $\xi^1, \ldots, \xi^N$ of N realizations of the random vector ξ can be generated in the computer. In the Monte Carlo sampling method this is accomplished by generating a sequence $U^1, U^2, \ldots$ of independent random (or rather pseudorandom) numbers uniformly distributed on the interval [0,1], and then constructing the sample by an appropriate transformation. In that way we can consider the sequence $\omega := \{U^1, U^2, \ldots\}$ as an element of the probability space equipped with the corresponding product probability measure, and the sample $\xi^j = \xi^j(\omega)$, $i = 1, 2, \ldots$, as a function of ω. Since computer is a finite deterministic machine, sooner or later the generated sample will start to repeat itself. However, modern random numbers generators have a very large cycle period, and this method was tested in numerous applications. We view now the corresponding SAA problem (5.2) as a way of *approximating* the true problem (5.1) while drastically reducing the number of generated scenarios. For a statistical analysis of the constructed SAA problems, a particular numerical algorithm applied to solve these problems is irrelevant.

Let us also remark that values of the sample average function $\hat{f}_N(x)$ can be computed in two somewhat different ways. The generated sample $\xi^1, \ldots, \xi^N$ can be stored in the computer memory and called every time a new value (at a different point x) of the sample average function should be computed. Alternatively, the same sample can be generated by using a common seed number in an employed pseudorandom numbers generator. (This is why this approach is called the *common random number generation* method.)

The idea of common random number generation is well known in simulation. That is, suppose that we want to compare values of the objective function at two points $x_1, x_2 \in X$. In that case we are interested in the difference $f(x_1) - f(x_2)$ rather than in the individual values $f(x_1)$ and $f(x_2)$. If we use sample average estimates $\hat{f}_N(x_1)$ and $\hat{f}_N(x_2)$ based on *independent* samples, both of size N, then $\hat{f}_N(x_1)$ and $\hat{f}_N(x_2)$ are uncorrelated and

$$\mathbb{V}\mathrm{ar}\big[\hat{f}_N(x_1) - \hat{f}_N(x_2)\big] = \mathbb{V}\mathrm{ar}\big[\hat{f}_N(x_1)\big] + \mathbb{V}\mathrm{ar}\big[\hat{f}_N(x_2)\big]. \tag{5.83}$$

On the other hand, if we use the *same* sample for the estimators $\hat{f}_N(x_1)$ and $\hat{f}_N(x_2)$, then

$$\mathbb{V}\mathrm{ar}\big[\hat{f}_N(x_1) - \hat{f}_N(x_2)\big] = \mathbb{V}\mathrm{ar}\big[\hat{f}_N(x_1)\big] + \mathbb{V}\mathrm{ar}\big[\hat{f}_N(x_2)\big] - 2\mathbb{C}\mathrm{ov}\big(\hat{f}_N(x_1), \hat{f}_N(x_2)\big). \tag{5.84}$$

In both cases, $\hat{f}_N(x_1) - \hat{f}_N(x_2)$ is an unbiased estimator of $f(x_1) - f(x_2)$. However, in the case of the same sample, the estimators $\hat{f}_N(x_1)$ and $\hat{f}_N(x_2)$ tend to be positively correlated with each other, in which case the variance in (5.84) is smaller than the one in

(5.83). The difference between the independent and the common random number generated estimators of $f(x_1) - f(x_2)$ can be especially dramatic when the points x_1 and x_2 are close to each other and hence the common random number generated estimators are highly positively correlated.

By the results of section 5.1.1 we have that under mild regularity conditions, the optimal value and optimal solutions of the SAA problem (5.2) converge w.p. 1, as the sample size increases, to their true counterparts. These results, however, do not give any indication of quality of solutions for a given sample of size N. In the next section we discuss *exponential* rates of convergence of optimal and nearly optimal solutions of the SAA problem (5.2). This allows us to give an estimate of the sample size which is required to solve the true problem with a given accuracy by solving the SAA problem. Although such estimates of the sample size typically are *too conservative for a practical use*, they give insight into the *complexity* of solving the true (expected value) problem.

Unless stated otherwise, we assume in this section that the random sample $\xi^1, \ldots, \xi^N$ is iid, and make the following assumption:

(M1) The expectation function $f(x)$ is well defined and finite valued for all $x \in X$.

For $\varepsilon \geq 0$ we denote by

$$S^\varepsilon := \{x \in X : f(x) \leq \vartheta^* + \varepsilon\} \text{ and } \hat{S}_N^\varepsilon := \{x \in X : \hat{f}_N(x) \leq \hat{\vartheta}_N + \varepsilon\}$$

the sets of ε-optimal solutions of the true and the SAA problems, respectively.

5.3.1 Exponential Rates of Convergence and Sample Size Estimates in the Case of a Finite Feasible Set

In this section we assume that the feasible set X is finite, although its cardinality $|X|$ can be very large. Since X is finite, the sets S^ε and $\hat{S}_N^\varepsilon$ are nonempty and finite. For parameters $\varepsilon \geq 0$ and $\delta \in [0, \varepsilon]$, consider the event $\{\hat{S}_N^\delta \subset S^\varepsilon\}$. This event means that any δ-optimal solution of the SAA problem is an ε-optimal solution of the true problem. We estimate now the probability of that event.

We can write

$$\left\{\hat{S}_N^\delta \not\subset S^\varepsilon\right\} = \bigcup_{x \in X \setminus S^\varepsilon} \bigcap_{y \in X} \left\{\hat{f}_N(x) \leq \hat{f}_N(y) + \delta\right\}, \tag{5.85}$$

and hence

$$\mathsf{Pr}\left(\hat{S}_N^\delta \not\subset S^\varepsilon\right) \leq \sum_{x \in X \setminus S^\varepsilon} \mathsf{Pr}\left(\bigcap_{y \in X} \left\{\hat{f}_N(x) \leq \hat{f}_N(y) + \delta\right\}\right). \tag{5.86}$$

Consider a mapping $u : X \setminus S^\varepsilon \to X$. If the set $X \setminus S^\varepsilon$ is empty, then any feasible point $x \in X$ is an ε-optimal solution of the true problem. Therefore we assume that this set is nonempty. It follows from (5.86) that

$$\mathsf{Pr}\left(\hat{S}_N^\delta \not\subset S^\varepsilon\right) \leq \sum_{x \in X \setminus S^\varepsilon} \mathsf{Pr}\left\{\hat{f}_N(x) - \hat{f}_N(u(x)) \leq \delta\right\}. \tag{5.87}$$

We assume that the mapping $u(\cdot)$ is chosen in such a way that

$$f(u(x)) \leq f(x) - \varepsilon^*, \quad \forall x \in X \setminus S^\varepsilon, \tag{5.88}$$

and for some $\varepsilon^* \geq \varepsilon$. Note that such a mapping always exists. For example, if we use a mapping $u : X \setminus S^\varepsilon \to S$, then (5.88) holds with

$$\varepsilon^* := \min_{x \in X \setminus S^\varepsilon} f(x) - \vartheta^* \tag{5.89}$$

and that $\varepsilon^* > \varepsilon$ since the set X is finite. Different choices of $u(\cdot)$ give a certain flexibility to the following derivations.

For each $x \in X \setminus S^\varepsilon$, define

$$Y(x, \xi) := F(u(x), \xi) - F(x, \xi). \tag{5.90}$$

Note that $\mathbb{E}[Y(x, \xi)] = f(u(x)) - f(x)$, and hence $\mathbb{E}[Y(x, \xi)] \leq -\varepsilon^*$ for all $x \in X \setminus S^\varepsilon$. The corresponding sample average is

$$\hat{Y}_N(x) := \frac{1}{N} \sum_{j=1}^{N} Y(x, \xi^j) = \hat{f}_N(u(x)) - \hat{f}_N(x).$$

By (5.87) we have

$$\Pr\left(\hat{S}_N^\delta \not\subset S^\varepsilon\right) \leq \sum_{x \in X \setminus S^\varepsilon} \Pr\left\{\hat{Y}_N(x) \geq -\delta\right\}. \tag{5.91}$$

Let $I_x(\cdot)$ denote the (large deviations) rate function of the random variable $Y(x, \xi)$. The inequality (5.91) together with the LD upper bound (7.173) implies

$$1 - \Pr\left(\hat{S}_N^\delta \subset S^\varepsilon\right) \leq \sum_{x \in X \setminus S^\varepsilon} e^{-N I_x(-\delta)}. \tag{5.92}$$

Note that inequality (5.92) is valid for any random sample of size N. Let us make the following assumption:

(M2) For every $x \in X \setminus S^\varepsilon$, the moment-generating function $\mathbb{E}\left[e^{tY(x,\xi)}\right]$ of the random variable $Y(x, \xi) = F(u(x), \xi) - F(x, \xi)$ is finite valued in a neighborhood of $t = 0$.

Assumption (M2) holds, for example, if the support Ξ of ξ is a bounded subset of $\mathbb{R}^d$, or if $Y(x, \cdot)$ grows at most linearly and ξ has a distribution from an exponential family.

Theorem 5.16. *Let ε and δ be nonnegative numbers. Then*

$$1 - \Pr(\hat{S}_N^\delta \subset S^\varepsilon) \leq |X|\, e^{-N\eta(\delta,\varepsilon)}, \tag{5.93}$$

where

$$\eta(\delta, \varepsilon) := \min_{x \in X \setminus S^\varepsilon} I_x(-\delta). \tag{5.94}$$

Moreover, if $\delta < \varepsilon^*$ and assumption (M2) *holds, then $\eta(\delta, \varepsilon) > 0$.*

Proof. Inequality (5.93) is an immediate consequence of inequality (5.92). If $\delta < \varepsilon^*$, then $-\delta > -\varepsilon^* \geq \mathbb{E}[Y(x,\xi)]$, and hence it follows by assumption (M2) that $I_x(-\delta) > 0$ for every $x \in X \setminus S^\varepsilon$. (See the discussion above equation (7.178).) This implies that $\eta(\delta,\varepsilon) > 0$. □

The following asymptotic result is an immediate consequence of inequality (5.93):

$$\limsup_{N\to\infty} \frac{1}{N} \ln\left[1 - \Pr(\hat{S}_N^\delta \subset S^\varepsilon)\right] \leq -\eta(\delta,\varepsilon). \tag{5.95}$$

It means that the probability of the event that any δ-optimal solution of the SAA problem provides an ε-optimal solution of the true problem approaches one *exponentially fast* as $N \to \infty$. Note that since it is possible to employ a mapping $u : X \setminus S^\varepsilon \to S$ with $\varepsilon^* > \varepsilon$ (see (5.89)), this exponential rate of convergence holds even if $\delta = \varepsilon$, and in particular if $\delta = \varepsilon = 0$. However, if $\delta = \varepsilon$ and the difference $\varepsilon^* - \varepsilon$ is small, then the constant $\eta(\delta,\varepsilon)$ could be close to zero. Indeed, for δ close to $-\mathbb{E}[Y(x,\xi)]$, we can write by (7.178) that

$$I_x(-\delta) \approx \frac{\big(-\delta - \mathbb{E}[Y(x,\xi)]\big)^2}{2\sigma_x^2} \geq \frac{(\varepsilon^* - \delta)^2}{2\sigma_x^2}, \tag{5.96}$$

where

$$\sigma_x^2 := \mathbb{V}\mathrm{ar}[Y(x,\xi)] = \mathbb{V}\mathrm{ar}[F(u(x),\xi) - F(x,\xi)]. \tag{5.97}$$

Let us make now the following assumption:

(M3) There is a constant $\sigma > 0$ such that for any $x \in X \setminus S^\varepsilon$ the moment-generating function $M_x(t)$ of the random variable $Y(x,\xi) - \mathbb{E}[Y(x,\xi)]$ satisfies

$$M_x(t) \leq \exp\left(\sigma^2 t^2/2\right), \quad \forall t \in \mathbb{R}. \tag{5.98}$$

It follows from assumption (M3) that

$$\ln \mathbb{E}\left[e^{tY(x,\xi)}\right] - t\mathbb{E}[Y(x,\xi)] = \ln M_x(t) \leq \sigma^2 t^2/2, \tag{5.99}$$

and hence the rate function $I_x(\cdot)$, of $Y(x,\xi)$, satisfies

$$I_x(z) \geq \sup_{t\in\mathbb{R}} \left\{t(z - \mathbb{E}[Y(x,\xi)]) - \sigma^2 t^2/2\right\} = \frac{\big(z - \mathbb{E}[Y(x,\xi)]\big)^2}{2\sigma^2}, \quad \forall z \in \mathbb{R}. \tag{5.100}$$

In particular, it follows that

$$I_x(-\delta) \geq \frac{\big(-\delta - \mathbb{E}[Y(x,\xi)]\big)^2}{2\sigma^2} \geq \frac{(\varepsilon^* - \delta)^2}{2\sigma^2} \geq \frac{(\varepsilon - \delta)^2}{2\sigma^2}. \tag{5.101}$$

Consequently the constant $\eta(\delta,\varepsilon)$ satisfies

$$\eta(\delta,\varepsilon) \geq \frac{(\varepsilon - \delta)^2}{2\sigma^2}, \tag{5.102}$$

and hence the bound (5.93) of Theorem 5.16 takes the form

$$1 - \Pr(\hat{S}_N^\delta \subset S^\varepsilon) \leq |X|\, e^{-N(\varepsilon-\delta)^2/(2\sigma^2)}. \tag{5.103}$$

This leads to the following result giving an estimate of the sample size which guarantees that any δ-optimal solution of the SAA problem is an ε-optimal solution of the true problem with probability at least $1-\alpha$.

Theorem 5.17. *Suppose that assumptions* (M1) *and* (M3) *hold. Then for* $\varepsilon > 0$, $0 \le \delta < \varepsilon$, *and* $\alpha \in (0,1)$, *and for the sample size* N *satisfying*

$$N \ge \frac{2\sigma^2}{(\varepsilon-\delta)^2} \ln\left(\frac{|X|}{\alpha}\right), \tag{5.104}$$

it follows that

$$\Pr(\hat{S}_N^\delta \subset S^\varepsilon) \ge 1-\alpha. \tag{5.105}$$

Proof. By setting the right-hand side of the estimate (5.103) to $\le \alpha$ and solving the obtained inequality, we obtain (5.104). □

Remark 10. A key characteristic of the estimate (5.104) is that the required sample size N depends *logarithmically* both on the size (cardinality) of the feasible set X and on the tolerance probability (significance level) α. The constant σ, postulated in assumption (M3), measures, in a sense, variability of a considered problem. If, for some $x \in X$, the random variable $Y(x,\xi)$ has a normal distribution with mean μ_x and variance σ_x^2, then its moment-generating function is equal to $\exp\left(\mu_x t + \sigma_x^2 t^2/2\right)$, and hence the moment-generating function $M_x(t)$, specified in assumption (M3), is equal to $\exp\left(\sigma_x^2 t^2/2\right)$. In that case, $\sigma^2 := \max_{x\in X\setminus S^\varepsilon} \sigma_x^2$ gives the smallest possible value for the corresponding constant in assumption (M3). If $Y(x,\xi)$ is bounded w.p. 1, i.e., there is constant $b > 0$ such that

$$\left|Y(x,\xi) - \mathbb{E}[Y(x,\xi)]\right| \le b, \quad \forall x \in X \text{ and a.e. } \xi \in \Xi,$$

then by Hoeffding inequality (see Proposition 7.63 and estimate (7.186)) we have that $M_x(t) \le \exp\left(b^2 t^2/2\right)$. In that case we can take $\sigma^2 := b^2$.

In any case for small $\varepsilon > 0$ we have by (5.96) that $I_x(-\delta)$ can be approximated from below by $(\varepsilon-\delta)^2/(2\sigma_x^2)$.

Remark 11. For, say, $\delta := \varepsilon/2$, the right-hand side of the estimate (5.104) is proportional to $(\sigma/\varepsilon)^2$. For Monte Carlo sampling based methods, such dependence on σ and ε seems to be unavoidable. In order to see that, consider a simple case when the feasible set X consists of just two elements, i.e., $X = \{x_1, x_2\}$ with $f(x_2) - f(x_1) > \varepsilon > 0$. By solving the corresponding SAA problem we make the (correct) decision that x_1 is the ε-optimal solution if $\hat{f}_N(x_2) - \hat{f}_N(x_1) > 0$. If the random variable $F(x_2,\xi) - F(x_1,\xi)$ has a normal distribution with mean $\mu = f(x_2) - f(x_1)$ and variance σ^2, then $\hat{f}_N(x_2) - \hat{f}_N(x_1) \sim \mathcal{N}(\mu, \sigma^2/N)$ and the probability of the event $\{\hat{f}_N(x_2) - \hat{f}_N(x_1) > 0\}$ (i.e., of the correct decision) is $\Phi(\mu\sqrt{N}/\sigma)$, where $\Phi(z)$ is the cumulative distribution function of $\mathcal{N}(0,1)$. We have that $\Phi(\varepsilon\sqrt{N}/\sigma) < \Phi(\mu\sqrt{N}/\sigma)$, and in order to make the probability of the incorrect decision less than α we have to take the sample size $N > z_\alpha^2\sigma^2/\varepsilon^2$, where $z_\alpha := \Phi^{-1}(1-\alpha)$. Even if $F(x_2,\xi) - F(x_1,\xi)$ is not normally distributed, the sample size of order σ^2/ε^2 could be justified asymptotically, say, by applying the CLT. It also could be mentioned that if $F(x_2,\xi) - F(x_1,\xi)$ has a normal distribution (with known variance), then the uniformly

most powerful test for testing $H_0 : \mu \leq 0$ versus $H_a : \mu > 0$ is of the form "reject H_0 if $\hat{f}_N(x_2) - \hat{f}_N(x_1)$ is bigger than a specified critical value" (this is a consequence of the Neyman–Pearson lemma). In other words, in such situations, if we only have access to a random sample, then solving the corresponding SAA problem is in a sense a best way to proceed.

Remark 12. Condition (5.98) of assumption (M3) can be replaced by a more general condition,

$$M_x(t) \leq \exp\left(\psi(t)\right), \quad \forall t \in \mathbb{R}, \tag{5.106}$$

where $\psi(t)$ is a convex even function with $\psi(0) = 0$. Then, similar to (5.100), we have

$$I_x(z) \geq \sup_{t \in \mathbb{R}} \{t(z - \mathbb{E}[Y(x,\xi)]) - \psi(t)\} = \psi^*\big(z - \mathbb{E}[Y(x,\xi)]\big), \quad \forall z \in \mathbb{R}, \tag{5.107}$$

where ψ^* is the conjugate of function ψ. Consequently, the estimate (5.93) takes the form

$$1 - \Pr(\hat{S}^\delta_N \subset S^\varepsilon) \leq |X|\, e^{-N\psi^*(\varepsilon-\delta)}, \tag{5.108}$$

and hence the estimate (5.104) takes the form

$$N \geq \frac{1}{\psi^*(\varepsilon - \delta)} \ln\left(\frac{|X|}{\alpha}\right). \tag{5.109}$$

For example, instead of assuming that condition (5.98) of assumption (M3) holds for all $t \in \mathbb{R}$, we may assume that this holds for all t in a finite interval $[-a, a]$, where $a > 0$ is a given constant. That is, we can take $\psi(t) := \sigma^2 t^2/2$ if $|t| \leq a$ and $\psi(t) := +\infty$ otherwise. In that case $\psi^*(z) = z^2/(2\sigma^2)$ for $|z| \leq a\sigma^2$ and $\psi^*(z) = a|z| - a^2\sigma^2$ for $|z| > a\sigma^2$. Consequently, the estimate (5.104) of Theorem 5.17 still holds provided that $0 < \varepsilon - \delta \leq a\sigma^2$.

5.3.2 Sample Size Estimates in the General Case

Suppose now that X is a bounded, not necessarily finite, subset of $\mathbb{R}^n$, and that $f(x)$ is finite valued for all $x \in X$. Then we can proceed in a way similar to the derivations of section 7.2.9. Let us make the following assumptions:

(M4) For any $x', x \in X$ there exists constant $\sigma_{x',x} > 0$ such that the moment-generating function $M_{x',x}(t) = \mathbb{E}[e^{tY_{x',x}}]$ of random variable $Y_{x',x} := [F(x',\xi) - f(x')] - [F(x,\xi) - f(x)]$ satisfies

$$M_{x',x}(t) \leq \exp\left(\sigma^2_{x',x} t^2/2\right), \quad \forall t \in \mathbb{R}. \tag{5.110}$$

(M5) There exists a (measurable) function $\kappa : \Xi \to \mathbb{R}_+$ such that its moment-generating function $M_\kappa(t)$ is finite valued for all t in a neighborhood of zero and

$$|F(x',\xi) - F(x,\xi)| \leq \kappa(\xi)\|x' - x\| \tag{5.111}$$

for a.e. $\xi \in \Xi$ and all $x', x \in X$.

Of course, it follows from (5.110) that

$$M_{x',x}(t) \leq \exp\left(\sigma^2 t^2/2\right), \quad \forall x', x \in X, \ \forall t \in \mathbb{R}, \tag{5.112}$$

where

$$\sigma^2 := \sup_{x',x\in X} \sigma^2_{x',x}. \tag{5.113}$$

Assumption (M4) is slightly stronger than assumption (M3), i.e., assumption (M3) follows from (M4) by taking $x' = u(x)$. Note that $\mathbb{E}[Y_{x',x}] = 0$ and recall that if $Y_{x',x}$ has a normal distribution, then equality in (5.110) holds with $\sigma^2_{x',x} := \mathbb{V}\text{ar}[Y_{x',x}]$.

The assumption (M5) implies that the expectation $\mathbb{E}[\kappa(\xi)]$ is finite and the function $f(x)$ is Lipschitz continuous on X with Lipschitz constant $L = \mathbb{E}[\kappa(\xi)]$. It follows that the optimal value ϑ^* of the true problem is finite, provided the set X is bounded. (Recall that it was assumed that X is nonempty and closed.) Moreover, by Cramér's large deviation theorem we have that for any $L' > \mathbb{E}[\kappa(\xi)]$ there exists a positive constant $\beta = \beta(L')$ such that

$$\Pr\left(\hat{\kappa}_N > L'\right) \leq \exp(-N\beta), \tag{5.114}$$

where $\hat{\kappa}_N := N^{-1}\sum_{j=1}^N \kappa(\xi^j)$. Note that it follows from (5.111) that w.p. 1

$$\left|\hat{f}_N(x') - \hat{f}_N(x)\right| \leq \hat{\kappa}_N \|x' - x\|, \quad \forall x', x \in X, \tag{5.115}$$

i.e., $\hat{f}_N(\cdot)$ is Lipschitz continuous on X with Lipschitz constant $\hat{\kappa}_N$.

By $D := \sup_{x,x'\in X} \|x' - x\|$ we denote the diameter of the set X. Of course, the set X is bounded iff its diameter is finite. We also use notation $a \vee b := \max\{a, b\}$ for numbers $a, b \in \mathbb{R}$.

Theorem 5.18. *Suppose that assumptions* (M1) *and* (M4)–(M5) *hold, with the corresponding constant* σ^2 *defined in* (5.113) *being finite, the set* X *has a finite diameter* D*, and let* $\varepsilon > 0$*,* $\delta \in [0, \varepsilon)$*,* $\alpha \in (0, 1)$*,* $L' > L := \mathbb{E}[\kappa(\xi)]$*, and* $\beta = \beta(L')$ *be the corresponding constants and* $\varrho > 0$ *be a constant specified below in* (5.118)*. Then for the sample size* N *satisfying*

$$N \geq \frac{8\sigma^2}{(\varepsilon-\delta)^2}\left[n\ln\left(\frac{8\varrho L' D}{\varepsilon-\delta}\right) + \ln\left(\frac{2}{\alpha}\right)\right] \bigvee \left[\beta^{-1}\ln\left(\frac{2}{\alpha}\right)\right], \tag{5.116}$$

it follows that

$$\Pr(\hat{S}^\delta_N \subset S^\varepsilon) \geq 1 - \alpha. \tag{5.117}$$

Proof. Let us set $\nu := (\varepsilon - \delta)/(4L')$, $\varepsilon' := \varepsilon - L'\nu$, and $\delta' := \delta + L'\nu$. Note that $\nu > 0$, $\varepsilon' = 3\varepsilon/4 + \delta/4 > 0$, $\delta' = \varepsilon/4 + 3\delta/4 > 0$ and $\varepsilon' - \delta' = (\varepsilon - \delta)/2 > 0$. Let $\bar{x}_1, \ldots, \bar{x}_M \in X$ be such that for every $x \in X$ there exists $\bar{x}_i$, $i \in \{1, \ldots, M\}$, such that $\|x - \bar{x}_i\| \leq \nu$, i.e., the set $X' := \{\bar{x}_1, \ldots, \bar{x}_M\}$ forms a ν-net in X. We can choose this net in such a way that

$$M \leq (\varrho D/\nu)^n \tag{5.118}$$

for a constant $\varrho > 0$. If the $X' \setminus S^{\varepsilon'}$ is empty, then any point of X' is an ε'-optimal solution of the true problem. Otherwise, choose a mapping $u : X' \setminus S^{\varepsilon'} \to S$ and consider the sets $\tilde{S} := \cup_{x\in X'}\{u(x)\}$ and $\tilde{X} := X' \cup \tilde{S}$. Note that $\tilde{X} \subset X$ and $|\tilde{X}| \leq (2\varrho D/\nu)^n$. Now let

us replace the set X by its subset $\tilde{X}$. We refer to the obtained true and SAA problems as respective reduced problems. We have that $\tilde{S} \subset S$, any point of the set $\tilde{S}$ is an optimal solutions of the true reduced problem and the optimal value of the true reduced problem is equal to the optimal value of the true (unreduced) problem. By Theorem 5.17 we have that with probability at least $1 - \alpha/2$ any δ'-optimal solution of the reduced SAA problem is an ε'-optimal solutions of the reduced (and hence unreduced) true problem provided that

$$N \geq \frac{8\sigma^2}{(\varepsilon - \delta)^2}\left[n \ln\left(\frac{8\varrho L' D}{\varepsilon - \delta}\right) + \ln\left(\frac{2}{\alpha}\right)\right]. \tag{5.119}$$

(Note that the right-hand side of (5.119) is greater than or equal to the estimate

$$\frac{2\sigma^2}{(\varepsilon' - \delta')^2} \ln\left(\frac{2|\tilde{X}|}{\alpha}\right)$$

required by Theorem 5.17.) We also have by (5.114) that for

$$N \geq \beta^{-1} \ln\left(\frac{2}{\alpha}\right), \tag{5.120}$$

the Lipschitz constant $\hat{\kappa}_N$ of the function $\hat{f}_N(x)$ is less than or equal to L' with probability at least $1 - \alpha/2$.

Now let $\hat{x}$ be a δ-optimal solution of the (unreduced) SAA problem. Then there is a point $x' \in \tilde{X}$ such that $\|\hat{x} - x'\| \leq \nu$, and hence $\hat{f}_N(x') \leq \hat{f}_N(\hat{x}) + L'\nu$, provided that $\hat{\kappa}_N \leq L'$. We also have that the optimal value of the (unreduced) SAA problem is smaller than or equal to the optimal value of the reduced SAA problem. It follows that x' is a δ'-optimal solution of the reduced SAA problem, provided that $\hat{\kappa}_N \leq L'$. Consequently, we have that x' is an ε'-optimal solution of the true problem with probability at least $1 - \alpha$ provided that N satisfies both inequalities (5.119) and (5.120). It follows that

$$f(\hat{x}) \leq f(x') + L\nu \leq f(x') + L'\nu \leq \vartheta^* + \varepsilon' + L'\nu = \vartheta^* + \varepsilon.$$

We obtain that if N satisfies both inequalities (5.119) and (5.120), then with probability at least $1 - \alpha$, any δ-optimal solution of the SAA problem is an ε-optimal solution of the true problem. The required estimate (5.116) follows. □

It is also possible to derive sample size estimates of the form (5.116) directly from the uniform exponential bounds derived in section 7.2.9; see Theorem 7.67 in particular.

Remark 13. If instead of assuming that condition (5.110) of assumption (M4) holds for all $t \in \mathbb{R}$, we assume that it holds for all $t \in [-a, a]$, where $a > 0$ is a given constant, then the estimate (5.116) of the above theorem still holds provided that $0 < \varepsilon - \delta \leq a\sigma^2$. (See Remark 12 on page 185.)

In a sense, the above estimate (5.116) of the sample size gives an estimate of *complexity* of solving the corresponding true problem by the SAA method. Suppose, for instance, that the true problem represents the first stage of a two-stage stochastic programming problem. For decomposition-type algorithms, the total number of iterations required to solve the SAA problem typically is independent of the sample size N (this is an empirical observation)

SPbook
2009/8/20
page 188

and the computational effort at every iteration is proportional to N. Anyway, size of the SAA problem grows linearly with increase of N. For $\delta \in [0, \varepsilon/2]$, say, the right-hand side of (5.116) is proportional to σ^2/ε^2, which suggests complexity of order σ^2/ε^2 with respect to the desirable accuracy. This is in a sharp contrast to deterministic (convex) optimization, where complexity usually is bounded in terms of $\ln(\varepsilon^{-1})$. It seems that such dependence on σ and ε is unavoidable for Monte Carlo sampling based methods. On the other hand, the estimate (5.116) is *linear* in the dimension n of the first-stage problem. It also depends linearly on $\ln(\alpha^{-1})$. This means that by increasing confidence, say, from 99% to 99.99%, we need to increase the sample size by the factor of $\ln 100 \approx 4.6$ at most. Assumption (M4) requires the probability distribution of the random variable $F(x, \xi) - F(x', \xi)$ to have sufficiently light tails. In a sense, the constant σ^2 can be viewed as a bound reflecting variability of the random variables $F(x, \xi) - F(x', \xi)$ for $x, x' \in X$. Naturally, larger variability of the data should result in more difficulty in solving the problem. (See Remark 11 on page 184.)

This suggests that by using Monte Carlo sampling techniques one can solve two-stage stochastic programs with a reasonable accuracy, say, with relative accuracy of 1% or 2%, in a reasonable time, provided that: *(a) its variability is not too large, (b) it has relatively complete recourse, and (c) the corresponding SAA problem can be solved efficiently.* Indeed, this was verified in numerical experiments with two-stage problems having a linear second-stage recourse. Of course, the estimate (5.116) of the sample size is far too conservative for actual calculations. For practical applications there are techniques which allow us to estimate (statistically) the error of a considered feasible solution $\bar{x}$ for a chosen sample size N; we will discuss this in section 5.6.

Next we discuss some modifications of the sample size estimate. It will be convenient in the following estimates to use notation $O(1)$ for a generic constant independent of the data. In that way we avoid denoting many different constants throughout the derivations.

(M6) There exists constant $\lambda > 0$ such that for any $x', x \in X$ the moment-generating function $M_{x',x}(t)$ of random variable $Y_{x',x} := [F(x', \xi) - f(x')] - [F(x, \xi) - f(x)]$ satisfies

$$M_{x',x}(t) \leq \exp\left(\lambda^2 \|x' - x\|^2 t^2/2\right), \quad \forall t \in \mathbb{R}. \tag{5.121}$$

The above assumption (M6) is a particular case of assumption (M4) with

$$\sigma^2_{x',x} = \lambda^2 \|x' - x\|^2,$$

and we can set the corresponding constant $\sigma^2 = \lambda^2 D^2$. The following corollary follows from Theorem 5.18.

Corollary 5.19. *Suppose that assumptions* (M1) *and* (M5)–(M6) *hold, the set X has a finite diameter D, and let $\varepsilon > 0$, $\delta \in [0, \varepsilon)$, $\alpha \in (0, 1)$, and $L = \mathbb{E}[\kappa(\xi)]$ be the corresponding constants. Then for the sample size N satisfying*

$$N \geq \frac{O(1)\lambda^2 D^2}{(\varepsilon - \delta)^2}\left[n \ln\left(\frac{O(1)LD}{\varepsilon - \delta}\right) + \ln\left(\frac{1}{\alpha}\right)\right], \tag{5.122}$$

it follows that

$$\Pr(\hat{S}^{\delta}_N \subset S^{\varepsilon}) \geq 1 - \alpha. \tag{5.123}$$

For example, suppose that the Lipschitz constant $\kappa(\xi)$ in assumption (M5) can be taken independent of ξ. That is, there exists a constant $L > 0$ such that

$$|F(x', \xi) - F(x, \xi)| \leq L\|x' - x\| \tag{5.124}$$

for a.e. $\xi \in \Xi$ and all $x', x \in X$. It follows that the expectation function $f(x)$ is also Lipschitz continuous on X with Lipschitz constant L, and hence the random variable $Y_{x',x}$ of assumption (M6) can be bounded as $|Y_{x',x}| \leq 2L\|x' - x\|$ w.p. 1. Moreover, we have that $\mathbb{E}[Y_{x',x}] = 0$, and hence it follows by Hoeffding's inequality (see the estimate (7.186)) that

$$M_{x',x}(t) \leq \exp\left(2L^2\|x' - x\|^2 t^2\right), \quad \forall t \in \mathbb{R}. \tag{5.125}$$

Consequently, we can take $\lambda = 2L$ in (5.121) and the estimate (5.122) takes the form

$$N \geq \left(\frac{O(1)LD}{\varepsilon - \delta}\right)^2 \left[n \ln\left(\frac{O(1)LD}{\varepsilon - \delta}\right) + \ln\left(\frac{1}{\alpha}\right)\right]. \tag{5.126}$$

Remark 14. It was assumed in Theorem 5.18 that the set X has a finite diameter, i.e., that X is bounded. For convex problems, this assumption can be relaxed. Assume that the problem is convex, the optimal value ϑ^* of the true problem is finite, and for some $a > \varepsilon$ the set S^a has a finite diameter D_a^*. (Recall that $S^a := \{x \in X : f(x) \leq \vartheta^* + a\}$.) We refer here to the respective true and SAA problems, obtained by replacing the feasible set X by its subset S^a, as reduced problems. Note that the set S^ε, of ε-optimal solutions, of the reduced and original true problems are the same. Let N^* be an integer satisfying the inequality (5.116) with D replaced by D_a^*. Then, under the assumptions of Theorem 5.18, we have that with probability at least $1 - \alpha$ all δ-optimal solutions of the reduced SAA problem are ε-optimal solutions of the true problem. Let us observe now that in this case the set of δ-optimal solutions of the reduced SAA problem coincides with the set of δ-optimal solutions of the original SAA problem. Indeed, suppose that the original SAA problem has a δ-optimal solution $x^* \in X \setminus S^a$. Let $\bar{x} \in \arg\min_{x \in S^a} \hat{f}_N(x)$, such a minimizer does exist since S^a is compact and $\hat{f}_N(x)$ is real valued convex and hence continuous. Then $\bar{x} \in S^\varepsilon$ and $\hat{f}_N(x^*) \leq \hat{f}_N(\bar{x}) + \delta$. By convexity of $\hat{f}_N(x)$ it follows that $\hat{f}_N(x) \leq \max\left\{\hat{f}_N(\bar{x}), \hat{f}_N(x^*)\right\}$ for all x on the segment joining $\bar{x}$ and x^*. This segment has a common point $\hat{x}$ with the set $S^a \setminus S^\varepsilon$. We obtain that $\hat{x} \in S^a \setminus S^\varepsilon$ is a δ-optimal solutions of the reduced SAA problem, a contradiction.

That is, with such sample size N^* we are guaranteed with probability at least $1 - \alpha$ that any δ-optimal solution of the SAA problem is an ε-optimal solution of the true problem. Also, assumptions (M4) and (M5) should be verified for x, x' in the set S^a only.

Remark 15. Suppose that the set S of optimal solutions of the true problem is nonempty. Then it follows from the proof of Theorem 5.18 that it suffices in assumption (M4) to verify condition (5.110) only for every $x \in X \setminus S^{\varepsilon'}$ and $x' := u(x)$, where $u : X \setminus S^{\varepsilon'} \to S$ and $\varepsilon' := 3/4\varepsilon + \delta/4$. If the set S is closed, we can use, for instance, a mapping $u(x)$ assigning to each $x \in X \setminus S^{\varepsilon'}$ a point of S closest to x. If, moreover, the set S is convex and the employed norm is strictly convex (e.g., the Euclidean norm), then such mapping (called metric projection onto S) is defined uniquely. If, moreover, assumption (M6) holds, then for such x and x' we have $\sigma_{x',x}^2 \leq \lambda^2 \bar{D}^2$, where $\bar{D} := \sup_{x \in X \setminus S^{\varepsilon'}} \operatorname{dist}(x, S)$. Suppose, further, that the problem is convex. Then (see Remark 14) for any $a > \varepsilon$, we can use S^a

instead of X. Therefore, if the problem is convex and the assumption (M6) holds, we can write the following estimate of the required sample size:

$$N \geq \frac{O(1)\lambda^2 \bar{D}_{a,\varepsilon}^2}{\varepsilon - \delta}\left[n \ln\left(\frac{O(1)LD_a^*}{\varepsilon - \delta}\right) + \ln\left(\frac{1}{\alpha}\right)\right], \tag{5.127}$$

where D_a^* is the diameter of S^a and $\bar{D}_{a,\varepsilon} := \sup_{x \in S^a \setminus S^{\varepsilon'}} \operatorname{dist}(x, S)$.

Corollary 5.20. *Suppose that assumptions* (M1) *and* (M5)–(M6) *hold, the problem is convex, the "true" optimal set S is nonempty, and for some $\gamma \geq 1$, $c > 0$, and $r > 0$, the following growth condition holds:*

$$f(x) \geq \vartheta^* + c\,[\operatorname{dist}(x, S)]^\gamma, \quad \forall x \in S^r. \tag{5.128}$$

Let $\alpha \in (0, 1)$, $\varepsilon \in (0, r)$, and $\delta \in [0, \varepsilon/2]$ and suppose, further, that for $a := \min\{2\varepsilon, r\}$ the diameter D_a^* of S^a is finite.

Then for the sample size N satisfying

$$N \geq \frac{O(1)\lambda^2}{c^{2/\gamma}\varepsilon^{2(\gamma-1)/\gamma}}\left[n \ln\left(\frac{O(1)LD_a^*}{\varepsilon}\right) + \ln\left(\frac{1}{\alpha}\right)\right], \tag{5.129}$$

it follows that

$$\Pr(\hat{S}_N^\delta \subset S^\varepsilon) \geq 1 - \alpha. \tag{5.130}$$

Proof. It follows from (5.128) that for any $a \leq r$ and $x \in S^a$, the inequality $\operatorname{dist}(x, S) \leq (a/c)^{1/\gamma}$ holds. Consequently, for any $\varepsilon \in (0, r)$, by taking $a := \min\{2\varepsilon, r\}$ and $\delta \in [0, \varepsilon/2]$ we obtain from (5.127) the required sample size estimate (5.129). □

Note that since $a = \min\{2\varepsilon, r\} \leq r$, we have that $S^a \subset S^r$, and if $S = \{x^*\}$ is a singleton, then it follows from (5.128) that $D_a^* \leq 2(a/c)^{1/\gamma}$. In particular, if $\gamma = 1$ and $S = \{x^*\}$ is a singleton (in that case it is said that the optimal solution x^* is *sharp*), then D_a^* can be bounded by $4c^{-1}\varepsilon$ and hence we obtain the following estimate:

$$N \geq O(1)c^{-2}\lambda^2\left[n \ln\left(O(1)c^{-1}L\right) + \ln\left(\alpha^{-1}\right)\right], \tag{5.131}$$

which does not depend on ε. For $\gamma = 2$, condition (5.128) is called the second order or quadratic growth condition. Under the quadratic growth condition, the first term in the right-hand side of (5.129) becomes of order $c^{-1}\varepsilon^{-1}\lambda^2$.

The following example shows that the estimate (5.116) of the sample size cannot be significantly improved for the class of convex stochastic programs.

Example 5.21. Consider the true problem with $F(x, \xi) := \|x\|^{2m} - 2m\,\xi^{\mathrm{T}}x$, where m is a positive constant, $\|\cdot\|$ is the Euclidean norm, and $X := \{x \in \mathbb{R}^n : \|x\| \leq 1\}$. Suppose, further, that random vector ξ has normal distribution $\mathcal{N}(0, \sigma^2 I_n)$, where σ^2 is a positive constant and I_n is the $n \times n$ identity matrix, i.e., components ξ_i of ξ are independent and $\xi_i \sim \mathcal{N}(0, \sigma^2)$, $i = 1, \ldots, n$. It follows that $f(x) = \|x\|^{2m}$, and hence for $\varepsilon \in [0, 1]$ the set of ε-optimal solutions of the true problem is given by $\{x : \|x\|^{2m} \leq \varepsilon\}$. Now let $\xi^1, \ldots, \xi^N$

be an iid random sample of ξ and $\bar{\xi}_N := (\xi^1 + \cdots + \xi^N)/N$. The corresponding sample average function is

$$\hat{f}_N(x) = \|x\|^{2m} - 2m\,\bar{\xi}_N^{\mathrm{T}} x, \tag{5.132}$$

and the optimal solution $\hat{x}_N$ of the SAA problem is $\hat{x}_N = \|\bar{\xi}_N\|^{-b}\bar{\xi}_N$, where

$$b := \begin{cases} \frac{2m-2}{2m-1} & \text{if} \quad \|\bar{\xi}_N\| \le 1, \\ 1 & \text{if} \quad \|\bar{\xi}_N\| > 1. \end{cases}$$

It follows that for $\varepsilon \in (0, 1)$, the optimal solution of the corresponding SAA problem is an ε-optimal solution of the true problem iff $\|\bar{\xi}_N\|^{\nu} \le \varepsilon$, where $\nu := \frac{2m}{2m-1}$. We have that $\bar{\xi}_N \sim \mathcal{N}(0, \sigma^2 N^{-1} I_n)$, and hence $N\|\bar{\xi}_N\|^2/\sigma^2$ has a chi-square distribution with n degrees of freedom. Consequently, the probability that $\|\bar{\xi}_N\|^{\nu} > \varepsilon$ is equal to the probability $\Pr\left(\chi_n^2 > N\varepsilon^{2/\nu}/\sigma^2\right)$. Moreover, $\mathbb{E}[\chi_n^2] = n$ and the probability $\Pr(\chi_n^2 > n)$ increases and tends to $1/2$ as n increases. Consequently, for $\alpha \in (0, 0.3)$ and $\varepsilon \in (0, 1)$, for example, the sample size N should satisfy

$$N > \frac{n\sigma^2}{\varepsilon^{2/\nu}} \tag{5.133}$$

in order to have the property, "with probability $1-\alpha$ an (exact) optimal solution of the SAA problem is an ε-optimal solution of the true problem." Compared with (5.116), the lower bound (5.133) also grows linearly in n and is proportional to $\sigma^2/\varepsilon^{2/\nu}$. It remains to note that the constant ν decreases to 1 as m increases.

Note that in this example the growth condition (5.128) holds with $\gamma = 2m$ and that the power constant of ε in the estimate (5.133) is in accordance with the estimate (5.129). Note also that here

$$[F(x', \xi) - f(x')] - [F(x, \xi) - f(x)] = 2m\,\xi^{\mathrm{T}}(x - x')$$

has normal distribution with zero mean and variance $4m^2\sigma^2\|x' - x\|^2$. Consequently, assumption (M6) holds with $\lambda^2 = 4m^2\sigma^2$.

Of course, in this example the "true" optimal solution is $\bar{x} = 0$, and one does not need sampling in order to solve this problem. Note, however, that the sample average function $\hat{f}_N(x)$ here depends on the random sample only through the data average vector $\bar{\xi}_N$. Therefore, any numerical procedure based on averaging will need a sample of size N satisfying the estimate (5.133) in order to produce an ε-optimal solution. ∎

5.3.3 Finite Exponential Convergence

We assume in this section that the problem is *convex* and the expectation function $f(x)$ is finite valued.

Definition 5.22. *It is said that $x^* \in X$ is a* sharp *(optimal) solution of the true problem* (5.1) *if there exists constant $c > 0$ such that*

$$f(x) \ge f(x^*) + c\|x - x^*\|, \quad \forall x \in X. \tag{5.134}$$

Condition (5.134) corresponds to growth condition (5.128) with the power constant $\gamma = 1$ and $S = \{x^*\}$. Since $f(\cdot)$ is convex finite valued, we have that the directional derivatives $f'(x^*, h)$ exist for all $h \in \mathbb{R}^n$, $f'(x^*, \cdot)$ is (locally Lipschitz) continuous, and formula (7.17) holds. Also, by convexity of the set X we have that the tangent cone $\mathcal{T}_X(x^*)$, to X at x^*, is given by the topological closure of the corresponding radial cone. By using these facts, it is not difficult to show that condition (5.134) is equivalent to

$$f'(x^*, h) \geq c\|h\|, \quad \forall h \in \mathcal{T}_X(x^*). \tag{5.135}$$

Since condition (5.135) is local, we have that it actually suffices to verify (5.134) for all $x \in X$ in a neighborhood of x^*.

Theorem 5.23. *Suppose that the problem is convex and assumption* (M1) *holds, and let $x^* \in X$ be a sharp optimal solution of the true problem. Then $\hat{S}_N = \{x^*\}$ w.p. 1 for N large enough. Suppose, further, that assumption* (M4) *holds. Then there exist constants $C > 0$ and $\beta > 0$ such that*

$$1 - \Pr\big(\hat{S}_N = \{x^*\}\big) \leq Ce^{-N\beta}; \tag{5.136}$$

i.e., the probability of the event that "x^* is the unique optimal solution of the SAA problem" converges to 1 *exponentially fast with the increase of the sample size N.*

Proof. By convexity of $F(\cdot, \xi)$ we have that $\hat{f}'_N(x^*, \cdot)$ converges to $f'(x^*, \cdot)$ w.p. 1 uniformly on the unit sphere (see the proof of Theorem 7.54). It follows w.p. 1 for N large enough that

$$\hat{f}'_N(x^*, h) \geq (c/2)\|h\|, \quad \forall h \in \mathcal{T}_X(x^*), \tag{5.137}$$

which implies that x^* is the sharp optimal solution of the corresponding SAA problem.

Now, under the assumptions of convexity and (M1) and (M4), we have that $\hat{f}'_N(x^*, \cdot)$ converges to $f'(x^*, \cdot)$ exponentially fast on the unit sphere. (See inequality (7.219) of Theorem 7.69.) By taking $\varepsilon := c/2$ in (7.219), we can conclude that (5.136) follows. □

It is also possible to consider the growth condition (5.128) with $\gamma = 1$ and the set S not necessarily being a singleton. That is, it is said that the set S of optimal solutions of the true problem is *sharp* if for some $c > 0$ the following condition holds:

$$f(x) \geq \vartheta^* + c\,[\mathrm{dist}(x, S)], \quad \forall x \in X. \tag{5.138}$$

Of course, if $S = \{x^*\}$ is a singleton, then conditions (5.134) and (5.138) do coincide. The set of optimal solutions of the true problem is always nonempty and sharp if its optimal value is finite and the problem is *piecewise linear* in the sense that the following conditions hold:

(P1) The set X is a convex closed polyhedron.

(P2) The support set $\Xi = \{\xi_1, \ldots, \xi_K\}$ is finite.

(P3) For every $\xi \in \Xi$ the function $F(\cdot, \xi)$ is polyhedral.

Conditions (P1)–(P3) hold in the case of two-stage linear stochastic programming problems with a finite number of scenarios.

Under conditions (P1)–(P3) the true and SAA problems are polyhedral, and hence their sets of optimal solutions are polyhedral. By using polyhedral structure and finiteness of the set Ξ, it is possible to show the following result (cf. [208]).

Theorem 5.24. *Suppose that conditions* (P1)–(P3) *hold and the set S is nonempty and bounded. Then S is polyhedral and there exist constants $C > 0$ and $\beta > 0$ such that*

$$1 - \Pr\big(\hat{S}_N \neq \emptyset \text{ and } \hat{S}_N \text{ is a face of } S\big) \leq Ce^{-N\beta}; \tag{5.139}$$

i.e., the probability of the event that "$\hat{S}_N$ is nonempty and forms a face of the set S" converges to 1 *exponentially fast with the increase of the sample size N.*

5.4 Quasi–Monte Carlo Methods

In the previous section we discussed an approach to evaluating (approximating) expectations by employing random samples generated by Monte Carlo techniques. It should be understood, however, that when dimension d (of the random data vector ξ) is small, the Monte Carlo approach may not be a best way to proceed. In this section we give a brief discussion of the so-called quasi–Monte Carlo methods. It is beyond the scope of this book to give a detailed discussion of that subject. This section is based on Niederreiter [138], to which the interested reader is referred for a further reading on that topic. Let us start our discussion by considering a one-dimensional case (of $d = 1$).

Let ξ be a real valued random variable having cdf $H(z) = \Pr(\xi \leq z)$. Suppose that we want to evaluate the expectation

$$\mathbb{E}[F(\xi)] = \int_{-\infty}^{+\infty} F(z)dH(z), \tag{5.140}$$

where $F : \mathbb{R} \to \mathbb{R}$ is a measurable function. Let $U \sim U[0, 1]$, i.e., U is a random variable uniformly distributed on $[0, 1]$. Then random variable[22] $H^{-1}(U)$ has cdf $H(\cdot)$. Therefore, by making a change of variables we can write the expectation (5.140) as

$$\mathbb{E}[\psi(U)] = \int_0^1 \psi(u)du, \tag{5.141}$$

where $\psi(u) := F(H^{-1}(u))$.

Evaluation of the above expectation by the Monte Carlo method is based on generating an iid sample $U^1, \ldots, U^N$ of N replications of $U \sim U[0, 1]$ and consequently approximating $\mathbb{E}[\psi(U)]$ by the average $\bar{\psi}_N := N^{-1} \sum_{j=1}^N \psi(U^j)$. Alternatively, one can employ the Riemann sum approximation

$$\int_0^1 \psi(u)du \approx \frac{1}{N} \sum_{j=1}^N \psi(u_j) \tag{5.142}$$

by using some points $u_j \in [(j-1)/N, j/N]$, $j = 1, \ldots, N$, e.g., taking midpoints $u_j := (2j-1)/(2N)$ of equally spaced partition intervals $[(j-1)/N, j/N]$, $j = 1, \ldots, N$.

[22] Recall that $H^{-1}(u) := \inf\{z : H(z) \geq u\}$.

If the function $\psi(u)$ is Lipschitz continuous on [0,1], then the error of the Riemann sum approximation[23] is of order $O(N^{-1})$, while the Monte Carlo sample average error is of (stochastic) order $O_p(N^{-1/2})$. An explanation of this phenomenon is rather clear, an iid sample $U^1, \ldots, U^N$ will tend to cluster in some areas while leaving other areas of the interval [0,1] uncovered.

One can argue that the Monte Carlo sampling approach has an advantage in the possibility of estimating the approximation error by calculating the sample variance,

$$s^2 := (N-1)^{-1} \sum_{j=1}^{N} \left[\psi(U^j) - \bar{\psi}_N\right]^2,$$

and consequently constructing a corresponding confidence interval. It is possible, however, to employ a similar procedure for the Riemann sums by making them random. That is, each point u_j in the right-hand side of (5.142) is generated randomly, say, uniformly distributed, on the corresponding interval $[(j-1)/N, j/N]$, independently of other points u_k, $k \neq j$. This will make the right-hand side of (5.142) a random variable. Its variance can be estimated by using several independently generated batches of such approximations.

It does not make sense to use Monte Carlo sampling methods in case of one-dimensional random data. The situation starts to change quickly with an increase of the dimension d. By making an appropriate transformation we may assume that the random data vector is distributed uniformly on the d-dimensional cube $I^d = [0,1]^d$. For $d > 1$ we denote by (bold-faced) $\boldsymbol{U}$ a random vector uniformly distributed on I^d. Suppose that we want to evaluate the expectation $\mathbb{E}[\psi(\boldsymbol{U})] = \int_{I^d} \psi(\boldsymbol{u}) d\boldsymbol{u}$, where $\psi : I^d \to \mathbb{R}$ is a measurable function. We can partition each coordinate of I^d into M equally spaced intervals, and hence partition I^d into the corresponding $N = M^d$ subintervals[24] and use a corresponding Riemann sum approximation $N^{-1} \sum_{j=1}^{N} \psi(\boldsymbol{u}_j)$. The resulting error is of order $O(M^{-1})$, provided that the function $\psi(\boldsymbol{u})$ is Lipschitz continuous. In terms of the total number N of function evaluations, this error is of order $O(N^{-1/d})$. For $d = 2$ it is still compatible with the Monte Carlo sample average approximation approach. However, for larger values of d the Riemann sums approach quickly becomes unacceptable. On the other hand, the rate of convergence (error bounds) of the Monte Carlo sample average approximation of $\mathbb{E}[\psi(\boldsymbol{U})]$ does not depend directly on dimensionality d but only on the corresponding variance $\mathbb{V}\mathrm{ar}[\psi(\boldsymbol{U})]$. Yet the problem of uneven covering of I^d by an iid sample $\boldsymbol{U}^j$, $j = 1, \ldots, N$, remains persistent.

Quasi–Monte Carlo methods employ the approximation

$$\mathbb{E}[\psi(\boldsymbol{U})] \approx \frac{1}{N} \sum_{j=1}^{N} \psi(\boldsymbol{u}_j) \tag{5.143}$$

for a carefully chosen (deterministic) sequence of points $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N \in I^d$. From the numerical point of view, it is important to be able to generate such a sequence iteratively as an infinite sequence of points $\boldsymbol{u}_j$, $j = 1, \ldots,$ in I^d. In that way, one does not need to recalculate already calculated function values $\psi(\boldsymbol{u}_j)$ with the increase of N. A basic requirement for this sequence is that the right-hand side of (5.143) converges to $\mathbb{E}[\psi(\boldsymbol{U})]$

[23] If $\psi(u)$ is continuously differentiable, then, e.g., the trapezoidal rule gives even a slightly better approximation error of order $O(N^{-2})$. Also, one should be careful in making the assumption of Lipschitz continuity of $\psi(u)$. If the distribution of ξ is supported on the whole real line, e.g., is normal, then $H^{-1}(u)$ tends to ∞ as u tends to 0 or 1. In that case, $\psi(u)$ typically will be discontinuous at $u = 0$ and $u = 1$.

[24] A set $A \subset \mathbb{R}^d$ is said to be a (d-dimensional) interval if $A = [a_1, b_1] \times \cdots \times [a_d, b_d]$.

as $N \to \infty$. It is not difficult to show that this holds (for any Riemann-integrable function $\psi(\boldsymbol{u})$) if

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} \mathbf{1}_A(\boldsymbol{u}_j) = V_d(A) \tag{5.144}$$

for any interval $A \subset I^d$. Here $V_d(A)$ denotes the d-dimensional Lebesgue measure (volume) of set $A \subset \mathbb{R}^d$.

Definition 5.25. *The* star discrepancy *of a point set* $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N\} \subset I^d$ *is defined by*

$$\mathcal{D}^*(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N) := \sup_{A\in\mathcal{A}} \left| \frac{1}{N} \sum_{j=1}^{N} \mathbf{1}_A(\boldsymbol{u}_j) - V_d(A) \right|, \tag{5.145}$$

where $\mathcal{A}$ *is the family of all subintervals of* I^d *of the form* $\prod_{i=1}^{d}[0, b_i)$.

It is possible to show that for a sequence $\boldsymbol{u}_j \in I^d$, $j = 1, \ldots,$ condition (5.144) holds iff $\lim_{N\to\infty} \mathcal{D}^*(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N) = 0$. A more important property of the star discrepancy is that it is possible to give error bounds in terms of $\mathcal{D}^*(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N)$ for quasi–Monte Carlo approximations. Let us start with the one-dimensional case. Recall that variation of a function $\psi : [0, 1] \to \mathbb{R}$ is the sup $\sum_{i=1}^{m} |\psi(t_i) - \psi(t_{i-1})|$, where the supremum is taken over all partitions $0 = t_0 < t_1 < \cdots < t_m = 1$ of the interval [0,1]. It is said that ψ has bounded variation if its variation is finite.

Theorem 5.26 (Koksma). *If* $\psi : [0, 1] \to \mathbb{R}$ *has bounded variation* $V(\psi)$*, then for any* $u_1, \ldots, u_N \in [0, 1]$ *we have*

$$\left| \frac{1}{N} \sum_{j=1}^{N} \psi(u_j) - \int_0^1 \psi(u)du \right| \le V(\psi)\mathcal{D}^*(u_1, \ldots, u_N). \tag{5.146}$$

Proof. We can assume that the sequence $u_1, \ldots, u_N$ is arranged in increasing order, and we set $u_0 = 0$ and $u_{N+1} = 1$. That is, $0 = u_0 \le u_1 \le \cdots \le u_{N+1} = 1$. Using integration by parts we have

$$\int_0^1 \psi(u)du = u\psi(u)\big|_0^1 - \int_0^1 u d\psi(u) = \psi(1) - \int_0^1 u d\psi(u),$$

and using summation by parts we have

$$\frac{1}{N} \sum_{j=1}^{N} \psi(u_j) = \psi(u_{N+1}) - \sum_{j=0}^{N} \frac{j}{N}[\psi(u_{j+1}) - \psi(u_j)];$$

we can write

$$\begin{aligned} \tfrac{1}{N}\textstyle\sum_{j=1}^{N} \psi(u_j) - \int_0^1 \psi(u)du &= -\textstyle\sum_{j=0}^{N} \frac{j}{N}[\psi(u_{j+1}) - \psi(u_j)] + \int_0^1 u d\psi(u) \\ &= \textstyle\sum_{j=0}^{N} \int_{u_j}^{u_{j+1}} \left(u - \frac{j}{N}\right) d\psi(u). \end{aligned}$$

Also for any $u \in [u_j, u_{j+1}]$, $j = 0, \ldots, N$, we have

$$\left| u - \frac{j}{N} \right| \le \mathcal{D}^*(u_1, \ldots, u_N).$$

It follows that

$$\begin{aligned} \left| \tfrac{1}{N} \textstyle\sum_{j=1}^N \psi(u_j) - \int_0^1 \psi(u) du \right| &\le \textstyle\sum_{j=0}^N \int_{u_j}^{u_{j+1}} \left| u - \tfrac{j}{N} \right| d\psi(u) \\ &\le \mathcal{D}^*(u_1, \ldots, u_N) \textstyle\sum_{j=0}^N \left| \psi(u_{j+1}) - \psi(u_j) \right|, \end{aligned}$$

and, of course, $\sum_{j=0}^N \left| \psi(u_{j+1}) - \psi(u_j) \right| \le V(\psi)$. This completes the proof. □

This can be extended to a multidimensional setting as follows. Consider a function $\psi : I^d \to \mathbb{R}$. The variation of ψ, in the sense of Vitali, is defined as

$$V^{(d)}(\psi) := \sup_{\mathcal{P} \in \mathcal{J}} \sum_{A \in \mathcal{P}} |\Delta_\psi(A)|, \tag{5.147}$$

where $\mathcal{J}$ denotes the family of all partitions $\mathcal{P}$ of I^d into subintervals, and for $A \in \mathcal{P}$ the notation $\Delta_\psi(A)$ stands for an alternating sum of the values of ψ at the vertices of A (i.e., function values at adjacent vertices have opposite signs). The variation of ψ, in the sense of Hardy and Krause, is defined as

$$V(\psi) := \sum_{k=1}^d \sum_{1 \le i_1 < i_2 < \cdots i_k \le d} V^{(k)}(\psi; i_1, \ldots, i_k), \tag{5.148}$$

where $V^{(k)}(\psi; i_1, \ldots, i_k)$ is the variation in the sense of Vitali of restriction of ψ to the k-dimensional face of I^d defined by $u_j = 1$ for $j \notin \{i_1, \ldots, i_k\}$.

Theorem 5.27 (Hlawka). *If $\psi : I^d \to \mathbb{R}$ has bounded variation $V(\psi)$ on I^d in the sense of Hardy and Krause, then for any $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N \in I^d$ we have*

$$\left| \frac{1}{N} \sum_{j=1}^N \psi(\boldsymbol{u}_j) - \int_{I^d} \psi(\boldsymbol{u}) d\boldsymbol{u} \right| \le V(\psi) \mathcal{D}^*(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N). \tag{5.149}$$

In order to see how good the above error estimates could be, let us consider the one-dimensional case with $u_j := (2j-1)/(2N)$, $j = 1, \ldots, N$. Then $\mathcal{D}^*(u_1, \ldots, u_N) = 1/(2N)$, and hence the estimate (5.146) leads to the error bound $V(\psi)/(2N)$. This error bound gives the correct order $O(N^{-1})$ for the error estimates (provided that ψ has bounded variation), but the involved constant $V(\psi)/2$ typically is far too large for practical calculations. Even worse, the inverse function $H^{-1}(u)$ is monotonically nondecreasing, and hence its variation is given by the difference of the limits $\lim_{u \to +\infty} H^{-1}(u)$ and $\lim_{u \to -\infty} H^{-1}(u)$. Therefore, if one of these limits is infinite, i.e., the support of the corresponding random variable is unbounded, then the associated variation is infinite. Typically, this variation unboundedness will carry over to the function $\psi(u) = F(H^{-1}(u))$. For example, if the function $F(\cdot)$ is monotonically nondecreasing, then

$$V(\psi) = F\left(\lim_{u \to +\infty} H^{-1}(u) \right) - F\left(\lim_{u \to -\infty} H^{-1}(u) \right).$$

This overestimation of the corresponding constant becomes even worse with an increase in the dimension d.

A sequence $\{\boldsymbol{u}_j\}_{j\in\mathbb{N}} \subset I^d$ is called a *low-discrepancy sequence* if $\mathcal{D}^*(u_1, \ldots, u_N)$ is "small" for all $N \geq 1$. We proceed now to a description of classical constructions of low-discrepancy sequences. Let us start with the one-dimensional case. It is not difficult to show that $\mathcal{D}^*(u_1, \ldots, u_N)$ always greater than or equal to $1/(2N)$ and this lower bound is attained for $u_j := (2j-1)/2N$, $j = 1, \ldots, N$. While the lower bound of order $O(N^{-1})$ is attained for some N-element point sets from [0,1], there does not exist a sequence $u_1, \ldots,$ in [0,1] such that $\mathcal{D}^*(u_1, \ldots, u_N) \leq c/N$ for some $c > 0$ and all $N \in \mathbb{N}$. It is possible to show that a best possible for $\mathcal{D}^*(u_1, \ldots, u_N)$, for a sequence of points $u_j \in [0, 1]$, $j = 1, \ldots,$ is of order $O\left(N^{-1} \ln N\right)$. We are now going to construct a sequence for which this rate is attained.

For any integer $n \geq 0$ there is a unique digit expansion

$$n = \sum_{i\geq 0} a_i(n) b^i \tag{5.150}$$

in integer base $b \geq 2$, where $a_i(n) \in \{0, 1, \ldots, b-1\}$, $i = 0, 1, \ldots,$ and $a_i(n) = 0$ for all i large enough, i.e., the sum (5.150) is finite. The associated *radical-inverse function* $\phi_b(n)$, in base b, is defined by

$$\phi_b(n) := \sum_{i\geq 0} a_i(n) b^{-i-1}. \tag{5.151}$$

Note that

$$\phi_b(n) \leq (b-1) \sum_{i=0}^{\infty} b^{-i-1} = 1,$$

and hence $\phi_b(n) \in [0, 1]$ for any integer $n \geq 0$.

Definition 5.28. *For an integer $b \geq 2$, the* van der Corput *sequence in base b is the sequence $u_j := \phi_b(j)$, $j = 0, 1, \ldots$.*

It is possible to show that to every van der Corput sequence $u_1, \ldots,$ in base b, corresponds constant C_b such that

$$\mathcal{D}^*(u_1, \ldots, u_n) \leq C_b\, N^{-1} \ln N \quad \forall N \in \mathbb{N}.$$

A classical extension of van der Corput sequences to multidimensional settings is the following. Let $p_1 = 2, p_2 = 3, \ldots, p_d$ be the first d prime numbers. Then the *Halton sequence*, in the bases $p_1, \ldots, p_d$, is defined as

$$\boldsymbol{u}_j := (\phi_{p_1}(j), \ldots, \phi_{p_d}(j)) \in I^d, \quad j = 0, 1, \ldots\ . \tag{5.152}$$

It is possible to show that for that sequence,

$$\mathcal{D}^*(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N) \leq A_d N^{-1} (\ln N)^d + O(N^{-1} (\ln N)^{d-1}) \quad \forall N \geq 2, \tag{5.153}$$

where $A_d = \prod_{i=1}^{d} \frac{p_i - 1}{2 \ln p_i}$. By bound (5.149) of Theorem 5.27, this implies that the error of the corresponding quasi–Monte Carlo approximation is of order $O\left(N^{-1} (\ln N)^d\right)$, provided that variation $V(\psi)$ is finite. This compares favorably with the bound $O_p(N^{-1/2})$ of the

Monte Carlo sampling. Note, however, that by the prime number theorem we have that $\frac{\ln A_d}{d \ln d}$ tends to 1 as $d \to \infty$. That is, the coefficient A_d, of the leading term in the right-hand side of (5.153), grows superexponentially with increase of the dimension d. This makes the corresponding error bounds useless for larger values of d. It should be noticed that the above are *upper* bounds for the rates of convergence and in practice convergence rates could be much better. It seems that for low dimensional problems, say, $d \leq 20$, quasi–Monte Carlo methods are advantageous over Monte Carlo methods. With increase of the dimension d this advantage becomes less apparent. Of course, all this depends on a particular class of problems and applied quasi–Monte Carlo method. This issue requires a further investigation.

A drawback of (deterministic) quasi–Monte Carlo sequences $\{\boldsymbol{u}_j\}_{j\in\mathbb{N}}$ is that there is no easy way to estimate the error of the corresponding approximations $N^{-1}\sum_{j=1}^{N}\psi(\boldsymbol{u}_j)$. In that respect, bounds like (5.149) typically are too loose and impossible to calculate anyway. A way of dealing with this problem is to use a randomization of the set $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N\}$, of generating points in I^d without destroying its regular structure. Such a simple randomization procedure was suggested by Cranley and Patterson [39]. That is, generate a random point $\boldsymbol{u}$ uniformly distributed over I^d, and use the randomization[25] $\tilde{\boldsymbol{u}}_j := (\boldsymbol{u}_j + \boldsymbol{u}) \mod 1$, $j = 1, \ldots, N$. It is not difficult to show that (marginal) distribution of each random vector $\tilde{\boldsymbol{u}}_j$ is uniform on I^d. Therefore, each $\psi(\tilde{\boldsymbol{u}}_j)$, and hence $N^{-1}\sum_{j=1}^{N}\psi(\tilde{\boldsymbol{u}}_j)$, is an unbiased estimator of the corresponding expectation $\mathbb{E}[\psi(\boldsymbol{U})]$. Variance of the estimator $N^{-1}\sum_{j=1}^{N}\psi(\tilde{\boldsymbol{u}}_j)$ can be significantly smaller than variance of the corresponding Monte Carlo estimator based on samples of the same size. This randomization procedure can be applied in batches. That is, it can be repeated M times for independently generated uniformly distributed vectors $\boldsymbol{u} = \boldsymbol{u}^i$, $i = 1, \ldots, M$, and consequently averaging the obtained replications of $N^{-1}\sum_{j=1}^{N}\psi(\tilde{\boldsymbol{u}}_j)$. Simultaneously, variance of this estimator can be evaluated by calculating the sample variance of the obtained M independent replications of $N^{-1}\sum_{j=1}^{N}\psi(\tilde{\boldsymbol{u}}_j)$.

5.5 Variance-Reduction Techniques

Consider the sample average estimators $\hat{f}_N(x)$. We have that if the sample is iid, then the variance of $\hat{f}_N(x)$ is equal to $\sigma^2(x)/N$, where $\sigma^2(x) := \mathbb{V}\text{ar}[F(x,\xi)]$. In some cases it is possible to reduce the variance of generated sample averages, which in turn enhances convergence of the corresponding SAA estimators. In section 5.4 we discussed quasi–Monte Carlo techniques for enhancing rates of convergence of sample average approximations. In this section we briefly discuss some other variance-reduction techniques which seem to be useful in the SAA method.

5.5.1 Latin Hypercube Sampling

Suppose that the random data vector $\xi = \xi(\omega)$ is one-dimensional with the corresponding cumulative distribution function (cdf) $H(\cdot)$. We can then write

$$\mathbb{E}[F(x,\xi)] = \int_{-\infty}^{+\infty} F(x,\xi)\,dH(\xi). \tag{5.154}$$

[25] For a number $a \in \mathbb{R}$ the notation "$a \mod 1$" denotes the fractional part of a, i.e., $a \mod 1 = a - \lfloor a \rfloor$, where $\lfloor a \rfloor$ denotes the largest integer less than or equal to a. In the vector case, the "modulo 1" reduction is understood coordinatewise.

In order to evaluate the above integral numerically, it will be much better to generate sample points evenly distributed than to use an iid sample. (This was already discussed in section 5.4.) That is, we can generate independent random points[26]

$$U^j \sim U\left[(j-1)/N,\, j/N\right], \quad j = 1, \ldots, N, \tag{5.155}$$

and then construct the random sample of ξ by the inverse transformation $\xi^j := H^{-1}(U^j)$, $j = 1, \ldots, N$ (compare with (5.141)).

Now suppose that j is chosen at random from the set $\{1, \ldots, N\}$ (with equal probability for each element of that set). Then conditional on j, the corresponding random variable U^j is uniformly distributed on the interval $[(j-1)/N,\, j/N]$, and the unconditional distribution of U^j is uniform on the interval $[0, 1]$. Consequently, let $\{j_1, \ldots, j_N\}$ be a random permutation of the set $\{1, \ldots, N\}$. Then the random variables $\xi^{j_1}, \ldots, \xi^{j_N}$ have the same marginal distribution, with the same cdf $H(\cdot)$, and are negatively correlated with each other. Therefore, the expected value of

$$\hat{f}_N(x) = \frac{1}{N}\sum_{i=1}^{N} F(x, \xi^j) = \frac{1}{N}\sum_{s=1}^{N} F(x, \xi^{j_s}) \tag{5.156}$$

is $f(x)$, while

$$\mathbb{V}\mathrm{ar}\left[\hat{f}_N(x)\right] = N^{-1}\sigma^2(x) + 2N^{-2}\sum_{s<t} \mathbb{C}\mathrm{ov}\left(F(x, \xi^{j_s}),\, F(x, \xi^{j_t})\right). \tag{5.157}$$

If the function $F(x, \cdot)$ is monotonically increasing or decreasing, than the random variables $F(x, \xi^{j_s})$ and $F(x, \xi^{j_t})$, $s \neq t$, are also negatively correlated. Therefore, the variance of $\hat{f}_N(x)$ tends to be smaller, and in some cases much smaller, than $\sigma^2(x)/N$.

Suppose now that the random vector $\xi = (\xi_1, \ldots, \xi_d)$ is d-dimensional and that its components ξ_i, $i = 1, \ldots, d$, are distributed independently of each other. Then we can use the above procedure for each component ξ_i. That is, a random sample U^j of the form (5.155) is generated, and consequently N replications of the first component of ξ are computed by the corresponding inverse transformation applied to randomly permuted U^{j_s}. The same procedure is applied to every component of ξ with the corresponding random samples of the form (5.155) and random permutations generated independently of each other. This sampling scheme is called the *Latin hypercube* (LH) sampling.

If the function $F(x, \cdot)$ is decomposable, i.e., $F(x, \xi) := F_1(x, \xi_1) + \cdots + F_d(x, \xi_d)$, then $\mathbb{E}[F(x, \xi)] = \mathbb{E}[F_1(x, \xi_1)] + \cdots + \mathbb{E}[F_d(x, \xi_d)]$, where each expectation is calculated with respect to a one-dimensional distribution. In that case, the LH sampling ensures that each expectation $\mathbb{E}[F_i(x, \xi_i)]$ is estimated in a nearly optimal way. Therefore, the LH sampling works especially well in cases where the function $F(x, \cdot)$ tends to have a somewhat decomposable structure. In any case, the LH sampling procedure is easy to implement and can be applied to SAA optimization procedures in a straightforward way. Since in LH sampling the random replications of $F(x, \xi)$ are correlated with each other, one cannot use variance estimates like (5.21). Therefore, the LH method usually is applied in several independent batches in order to estimate variance of the corresponding estimators.

[26] For an interval $[a, b] \subset \mathbb{R}$, we denote by $U[a, b]$ the uniform probability distribution on that interval.

5.5.2 Linear Control Random Variables Method

Suppose that we have a measurable function $A(x, \xi)$ such that $\mathbb{E}[A(x, \xi)] = 0$ for all $x \in X$. Then, for any $t \in \mathbb{R}$, the expected value of $F(x, \xi) + tA(x, \xi)$ is $f(x)$, while

$$\mathbb{V}\mathrm{ar}\big[F(x, \xi) + tA(x, \xi)\big] = \mathbb{V}\mathrm{ar}\,[F(x, \xi)] + t^2\mathbb{V}\mathrm{ar}\,[A(x, \xi)] + 2t\,\mathbb{C}\mathrm{ov}\big(F(x, \xi), A(x, \xi)\big).$$

It follows that the above variance attains its minimum, with respect to t, for

$$t^* := -\rho_{F,A}(x)\left[\frac{\mathbb{V}\mathrm{ar}(F(x, \xi))}{\mathbb{V}\mathrm{ar}(A(x, \xi))}\right]^{1/2}, \tag{5.158}$$

where $\rho_{F,A}(x) := \mathbb{C}\mathrm{orr}\big(F(x, \xi), A(x, \xi)\big)$, and with

$$\mathbb{V}\mathrm{ar}\big[F(x, \xi) + t^*A(x, \xi)\big] = \mathbb{V}\mathrm{ar}\,[F(x, \xi)]\big[1 - \rho_{F,A}(x)^2\big]. \tag{5.159}$$

For a given $x \in X$ and generated sample $\xi^1, \ldots, \xi^N$, one can estimate, in the standard way, the covariance and variances appearing in the right-hand side of (5.158), and hence construct an estimate $\hat{t}$ of t^*. Then $f(x)$ can be estimated by

$$\hat{f}_N^A(x) := \frac{1}{N}\sum_{j=1}^{N}\big[F(x, \xi^j) + \hat{t}A(x, \xi^j)\big]. \tag{5.160}$$

By (5.159), the *linear control* estimator $\hat{f}_N^A(x)$ has a smaller variance than $\hat{f}_N(x)$ if $F(x, \xi)$ and $A(x, \xi)$ are highly correlated with each other.

Let us make the following observations. The estimator $\hat{t}$, of the optimal value t^*, depends on x and the generated sample. Therefore, it is difficult to apply linear control estimators in an SAA optimization procedure. That is, linear control estimators are mainly suitable for estimating expectations at a fixed point. Also, if the same sample is used in estimating $\hat{t}$ and $\hat{f}_N^A(x)$, then $\hat{f}_N^A(x)$ can be a slightly biased estimator of $f(x)$.

Of course, the above linear control procedure can be successful only if a function $A(x, \xi)$, with mean zero and highly correlated with $F(x, \xi)$, is available. Choice of such a function is problem dependent. For instance, one can use a linear function $A(x, \xi) := \lambda(\xi)^{\mathrm{T}}x$. Consider, for example, two-stage stochastic programming problems with recourse of the form (2.1)–(2.2). Suppose that the random vector $h = h(\omega)$ and matrix $T = T(\omega)$, in the second-stage problem (2.2), are independently distributed, and let $\mu := \mathbb{E}[h]$. Then

$$\mathbb{E}\big[(h - \mu)^{\mathrm{T}}T\big] = \mathbb{E}\,[(h - \mu)]^{\mathrm{T}}\,\mathbb{E}\,[T] = 0,$$

and hence one can use $A(x, \xi) := (h - \mu)^{\mathrm{T}}Tx$ as the control variable.

Let us finally remark that the above procedure can be extended in a straightforward way to a case where several functions $A_1(x, \xi), \ldots, A_m(x, \xi)$, each with zero mean and highly correlated with $F(x, \xi)$, are available.

5.5.3 Importance Sampling and Likelihood Ratio Methods

Suppose that ξ has a continuous distribution with probability density function (pdf) $h(\cdot)$. Let $\psi(\cdot)$ be another pdf such that the so-called *likelihood ratio* function $L(\cdot) := \frac{h(\cdot)}{\psi(\cdot)}$ is

well defined. That is, if $\psi(z) = 0$ for some $z \in \mathbb{R}^d$, then $h(z) = 0$, and by the definition, $0/0 = 0$, i.e., we do not divide a positive number by zero. Then we can write

$$f(x) = \int F(x,\xi)h(\xi)d\xi = \int F(x,\zeta)L(\zeta)\psi(\zeta)d\zeta = \mathbb{E}_\psi[F(x,Z)L(Z)], \quad (5.161)$$

where the integration is performed over the space $\mathbb{R}^d$ and the notation $\mathbb{E}_\psi$ emphasizes that the expectation is taken with respect to the random vector Z having pdf $\psi(\cdot)$.

Let us show that for a fixed x, the variance of $F(x, Z)L(Z)$ attains its minimal value for $\psi(\cdot)$ proportional to $|F(x,\cdot)h(\cdot)|$, i.e., for

$$\psi^*(\cdot) := \frac{|F(x,\cdot)h(\cdot)|}{\int |F(x,\zeta)h(\zeta)|d\zeta}. \quad (5.162)$$

Since $\mathbb{E}_\psi[F(x, Z)L(Z)] = f(x)$ and does not depend on $\psi(\cdot)$, we have that the variance of $F(x, Z)L(Z)$ is minimized if

$$\mathbb{E}_\psi[F(x,Z)^2L(Z)^2] = \int \frac{F(x,\zeta)^2h(\zeta)^2}{\psi(\zeta)}d\zeta \quad (5.163)$$

is minimized. Furthermore, by the Cauchy inequality we have

$$\left(\int |F(x,\zeta)h(\zeta)|d\zeta\right)^2 \le \left(\int \frac{F(x,\zeta)^2h(\zeta)^2}{\psi(\zeta)}d\zeta\right)\left(\int \psi(\zeta)d\zeta\right). \quad (5.164)$$

It remains to note that $\int \psi(\zeta)d\zeta = 1$ and the left-hand side of (5.164) is equal to the expected value of squared $F(x, Z)L(Z)$ for $\psi(\cdot) = \psi^*(\cdot)$.

Note that if $F(x,\cdot)$ is nonnegative valued, then $\psi^*(\cdot) = F(x,\cdot)h(\cdot)/f(x)$ and for that choice of the pdf $\psi(\cdot)$, the function $F(x,\cdot)L(\cdot)$ is identically equal to $f(x)$. Of course, in order to achieve such absolute variance reduction to zero, we need to know the expectation $f(x)$, which was our goal in the first place. Nevertheless, it gives the idea that if we can construct a pdf $\psi(\cdot)$ roughly proportional to $|F(x,\cdot)h(\cdot)|$, then we may achieve a considerable variance reduction by generating a random sample $\zeta^1, \ldots, \zeta^N$ from the pdf $\psi(\cdot)$, and then estimating $f(x)$ by

$$\tilde{f}_N^{\psi}(x) := \frac{1}{N}\sum_{j=1}^{N} F(x,\zeta^j)L(\zeta^j). \quad (5.165)$$

The estimator $\tilde{f}_N^{\psi}(x)$ is an unbiased estimator of $f(x)$ and may have significantly smaller variance than $\hat{f}_N(x)$, depending on a successful choice of the pdf $\psi(\cdot)$.

Similar analysis can be performed in cases where ξ has a discrete distribution by replacing the integrals with the corresponding summations.

Let us remark that the above approach, called *importance sampling*, is extremely sensitive to a choice of the pdf $\psi(\cdot)$ and is notorious for its instability. This is understandable since the likelihood ratio function in the tail is the ratio of two very small numbers. For a successful choice of $\psi(\cdot)$, the method may work very well while even a small perturbation of $\psi(\cdot)$ may be disastrous. This is why a single choice of $\psi(\cdot)$ usually does not work for different points x and consequently cannot be used for a whole optimization procedure.

Note also that $\mathbb{E}_\psi[L(Z)] = 1$. Therefore, $L(\zeta) - 1$ can be used as a linear control variable for the likelihood ratio estimator $\tilde{f}_N^{\psi}(x)$.

In some cases it is also possible to use the likelihood ratio method for estimating first and higher order derivatives of $f(x)$. Consider, for example, the optimal value $Q(x, \xi)$ of the second-stage linear program (2.2). Suppose that the vector q and matrix W are fixed, i.e., not stochastic, and for the sake of simplicity that $h = h(\omega)$ and $T = T(\omega)$ are distributed independently of each other. We have then that $Q(x, \xi) = \mathcal{Q}(h - Tx)$, where

$$\mathcal{Q}(z) := \inf\left\{q^{\mathsf{T}} y : Wy = z,\ y \geq 0\right\}.$$

Suppose, further, that h has a continuous distribution with pdf $\eta(\cdot)$. We have that

$$\mathbb{E}[Q(x, \xi)] = \mathbb{E}_T\left\{\mathbb{E}_{h|T}[Q(x, \xi)]\right\},$$

and by using the transformation $z = h - Tx$, since h and T are independent we obtain

$$\begin{array}{rcl} \mathbb{E}_{h|T}[Q(x, \xi)] & = & \mathbb{E}_h[Q(x, \xi)] \\ & = & \int \mathcal{Q}(h - Tx)\eta(h)dh = \int \mathcal{Q}(z)\eta(z + Tx)dz \\ & = & \int \mathcal{Q}(\zeta)L(x, \zeta)\psi(\zeta)d\zeta = \mathbb{E}_\psi[L(x, Z)\mathcal{Q}(Z)], \end{array} \tag{5.166}$$

where $\psi(\cdot)$ is a chosen pdf and $L(x, \zeta) := \eta(\zeta + Tx)/\psi(\zeta)$. If the function $\eta(\cdot)$ is smooth, then the likelihood ratio function $L(\cdot, \zeta)$ is also smooth. In that case, under mild additional conditions, first and higher order derivatives can be taken inside the expected value in the right-hand side of (5.166) and consequently can be estimated by sampling. Note that the first order derivatives of $Q(\cdot, \xi)$ are piecewise constant, and hence its second order derivatives are zeros whenever defined. Therefore, second order derivatives cannot be taken inside the expectation $\mathbb{E}[Q(x, \xi)]$ even if ξ has a continuous distribution.

5.6 Validation Analysis

Suppose that we are given a feasible point $\bar{x} \in X$ as a candidate for an optimal solution of the true problem. For example, $\bar{x}$ can be an output of a run of the corresponding SAA problem. In this section we discuss ways to evaluate quality of this candidate solution. This is important, in particular, for a choice of the sample size and stopping criteria in simulation based optimization. There are basically two approaches to such validation analysis. We can either try to estimate the optimality gap $f(\bar{x}) - \vartheta^*$ between the objective value at the considered point $\bar{x}$ and the optimal value of the true problem, or to evaluate first order (KKT) optimality conditions at $\bar{x}$.

Let us emphasize that the following analysis is designed for the situations where the value $f(\bar{x})$, of the true objective function at the considered point, is *finite*. In the case of two stage programming this requires, in particular, that the second-stage problem, associated with first-stage decision vector $\bar{x}$, is feasible for almost every realization of the random data.

5.6.1 Estimation of the Optimality Gap

In this section we consider the problem of estimating the optimality gap

$$\mathrm{gap}(\bar{x}) := f(\bar{x}) - \vartheta^* \tag{5.167}$$

associated with the candidate solution $\bar{x}$. Clearly, for any feasible $\bar{x} \in X$, $\text{gap}(\bar{x})$ is non-negative and $\text{gap}(\bar{x}) = 0$ iff $\bar{x}$ is an optimal solution of the true problem.

Consider the optimal value $\hat{\vartheta}_N$ of the SAA problem (5.2). We have that $\vartheta^* \geq \mathbb{E}[\hat{\vartheta}_N]$. (See the discussion following (5.22).) This means that $\hat{\vartheta}_N$ provides a valid *statistical lower bound* for the optimal value ϑ^* of the true problem. The expectation $\mathbb{E}[\hat{\vartheta}_N]$ can be estimated by averaging. That is, one can solve M times sample average approximation problems based on independently generated samples each of size N. Let $\hat{\vartheta}_N^1, \ldots, \hat{\vartheta}_N^M$ be the computed optimal values of these SAA problems. Then

$$\bar{v}_{N,M} := \frac{1}{M} \sum_{m=1}^{M} \hat{\vartheta}_N^m \tag{5.168}$$

is an unbiased estimator of $\mathbb{E}[\hat{\vartheta}_N]$. Since the samples, and hence $\hat{\vartheta}_N^1, \ldots, \hat{\vartheta}_N^M$, are independent and have the same distribution, we have that $\mathbb{V}\text{ar}\left[\bar{v}_{N,M}\right] = M^{-1}\mathbb{V}\text{ar}\left[\hat{\vartheta}_N\right]$, and hence we can estimate variance of $\bar{v}_{N,M}$ by

$$\hat{\sigma}^2_{N,M} := \frac{1}{M}\Big[\underbrace{\frac{1}{M-1} \sum_{m=1}^{M} \left(\hat{\vartheta}_N^m - \bar{v}_{N,M}\right)^2}_{\text{estimate of } \mathbb{V}\text{ar}[\hat{\vartheta}_N]} \Big]. \tag{5.169}$$

Note that the above make sense only if the optimal value ϑ^* of the true problem is finite. Note also that the inequality $\vartheta^* \geq \mathbb{E}[\hat{\vartheta}_N]$ holds and $\hat{\vartheta}_N$ gives a valid statistical lower bound even if $f(x) = +\infty$ for some $x \in X$. Note finally that the samples do not need to be iid (for example, one can use LH sampling); they only should be independent of each other in order to use estimate (5.169) of the corresponding variance.

In general, the random variable $\hat{\vartheta}_N$, and hence its replications $\hat{\vartheta}_N^j$, does not have a normal distribution, even approximately. (See Theorem 5.7 and the discussion that follows.) However, by the CLT, the probability distribution of the average $\bar{v}_{N,M}$ becomes approximately normal as M increases. Therefore, we can use

$$L_{N,M} := \bar{v}_{N,M} - t_{\alpha,M-1}\hat{\sigma}_{N,M} \tag{5.170}$$

as an approximate $100(1-\alpha)\%$ confidence[27] lower bound for the expectation $\mathbb{E}[\hat{\vartheta}_N]$.

We can also estimate $f(\bar{x})$ by sampling. That is, let $\hat{f}_{N'}(\bar{x})$ be the sample average estimate of $f(\bar{x})$, based on a sample of size N' generated independently of samples involved in computing $\bar{x}$. Let $\hat{\sigma}^2_{N'}(\bar{x})$ be an estimate of the variance of $\hat{f}_{N'}(\bar{x})$. In the case of the iid sample, one can use the sample variance estimate

$$\hat{\sigma}^2_{N'}(\bar{x}) := \frac{1}{N'(N'-1)} \sum_{j=1}^{N'} \left[F(\bar{x}, \xi^j) - \hat{f}_{N'}(\bar{x}) \right]^2. \tag{5.171}$$

Then

$$U_{N'}(\bar{x}) := \hat{f}_{N'}(\bar{x}) + z_\alpha \hat{\sigma}_{N'}(\bar{x}) \tag{5.172}$$

[27] Here $t_{\alpha,\nu}$ is the α-critical value of t-distribution with ν degrees of freedom. This critical value is slightly bigger than the corresponding standard normal critical value z_α, and $t_{\alpha,\nu}$ quickly approaches z_α as ν increases.

gives an approximate $100(1-\alpha)\%$ confidence upper bound for $f(\bar{x})$. Note that since N' typically is large, we use here critical value z_α from the standard normal distribution rather than a t-distribution.

We have that

$$\mathbb{E}\left[\hat{f}_{N'}(\bar{x}) - \bar{v}_{N,M}\right] = f(\bar{x}) - \mathbb{E}[\hat{\vartheta}_N] = \mathrm{gap}(\bar{x}) + \vartheta^* - \mathbb{E}[\hat{\vartheta}_N] \geq \mathrm{gap}(\bar{x}),$$

i.e., $\hat{f}_{N'}(\bar{x}) - \bar{v}_{N,M}$ is a biased estimator of the $\mathrm{gap}(\bar{x})$. Also the variance of this estimator is equal to the sum of the variances of $\hat{f}_{N'}(\bar{x})$ and $\bar{v}_{N,M}$, and hence

$$\hat{f}_{N'}(\bar{x}) - \bar{v}_{N,M} + z_\alpha \sqrt{\hat{\sigma}^2_{N'}(\bar{x}) + \hat{\sigma}^2_{N,M}} \tag{5.173}$$

provides a conservative $100(1-\alpha)\%$ confidence upper bound for the $\mathrm{gap}(\bar{x})$. We say that this upper bound is "conservative" since in fact it gives a $100(1-\alpha)\%$ confidence upper bound for the $\mathrm{gap}(\bar{x}) + \vartheta^* - \mathbb{E}[\hat{\vartheta}_N]$, and we have that $\vartheta^* - \mathbb{E}[\hat{\vartheta}_N] \geq 0$.

In order to calculate the estimate $\hat{f}_{N'}(\bar{x})$, one needs to compute the value $F(\bar{x}, \xi^j)$ of the objective function for every generated sample realization ξ^j, $j = 1, \ldots, N'$. Typically it is much easier to compute $F(\bar{x}, \xi)$ for a given $\xi \in \Xi$ than to solve the corresponding SAA problem. Therefore, often one can use a relatively large sample size N' and hence estimate $f(\bar{x})$ quite accurately. Evaluation of the optimal value ϑ^* by employing the estimator $\bar{v}_{N,M}$ is a more delicate problem.

There are two types of error in using $\bar{v}_{N,M}$ as an estimator of ϑ^*, namely, the bias $\vartheta^* - \mathbb{E}[\hat{\vartheta}_N]$ and variability of $\bar{v}_{N,M}$ measured by its variance. Both errors can be reduced by increasing N, and the variance can be reduced by increasing N and M. Note, however, that the computational effort in computing $\bar{v}_{N,M}$ is proportional to M, since the corresponding SAA problems should be solved M times, and to the computational time for solving a single SAA problem based on a sample of size N. Naturally one may ask what is the best way of distributing computational resources between increasing the sample size N and the number of repetitions M. This question is, of course, problem dependent. In cases where computational complexity of SAA problems grows fast with increase of the sample size N, it may be more advantageous to use a larger number of repetitions M. On the other hand, it was observed empirically that the computational effort in solving SAA problems by "good" subgradient algorithms grows only *linearly* with the sample size N. In such cases, one can use a larger N and make only a few repetitions M in order to estimate the variance of $\bar{v}_{N,M}$.

The bias $\vartheta^* - \mathbb{E}[\hat{\vartheta}_N]$ does not depend on M, of course. It was shown in Proposition 5.6 that if the sample is iid, then $\mathbb{E}[\hat{\vartheta}_N] \leq \mathbb{E}[\hat{\vartheta}_{N+1}]$ for any $N \in \mathbb{N}$. It follows that the bias $\vartheta^* - \mathbb{E}[\hat{\vartheta}_N]$ decreases monotonically with an increase of the sample size N. By Theorem 5.7 we have that, under mild regularity conditions,

$$\hat{\vartheta}_N = \inf_{x \in S} \hat{f}_N(x) + o_p(N^{-1/2}). \tag{5.174}$$

Consequently, if the set S of optimal solutions of the true problem is not a singleton, then the bias $\vartheta^* - \mathbb{E}[\hat{\vartheta}_N]$ typically converges to zero, as N increases, at a rate of $O(N^{-1/2})$, and tends to be bigger for a larger set S. (See (5.29) and the following discussion.) On the other hand, in well conditioned problems, where the optimal set S is a singleton, the bias typically is of order $O(N^{-1})$ (see Theorem 5.8), and the bias tends to be of a lesser

problem. Moreover, if the true problem has a sharp optimal solution x^*, then the event $\hat{x}_N = x^*$, and hence the event $\hat{\vartheta}_N = \hat{f}_N(x^*)$, happens with probability approaching one exponentially fast (see Theorem 5.23). Since $\mathbb{E}\big[\hat{f}_N(x^*)\big] = f(x^*)$, in such cases the bias $\vartheta^* - \mathbb{E}[\hat{\vartheta}_N] = f(x^*) - \mathbb{E}[\hat{\vartheta}_N]$ tends to be much smaller.

In the above approach, the upper and lower statistical bounds were computed independently of each other. Alternatively, it is possible to use the same sample for estimating $f(\bar{x})$ and $\mathbb{E}[\hat{\vartheta}_N]$. That is, for M generated samples each of size N, the gap is estimated by

$$\widehat{\mathrm{gap}}_{N,M}(\bar{x}) := \frac{1}{M}\sum_{m=1}^{M}\Big[\hat{f}_N^m(\bar{x}) - \hat{\vartheta}_N^m\Big], \tag{5.175}$$

where $\hat{f}_N^m(\bar{x})$ and $\hat{\vartheta}_N^m$ are computed from the *same* sample $m = 1, \ldots, M$. We have that the expected value of $\widehat{\mathrm{gap}}_{N,M}(\bar{x})$ is $f(\bar{x}) - \mathbb{E}[\hat{\vartheta}_N]$, i.e., the estimator $\widehat{\mathrm{gap}}_{N,M}(\bar{x})$ has the same bias as $\hat{f}_N(\bar{x}) - \bar{v}_{N,M}$. On the other hand, for a problem with sharp optimal solution x^* it happens with high probability that $\hat{\vartheta}_N^m = \hat{f}_N^m(x^*)$ and as a consequence $\hat{f}_N^m(\bar{x})$ tends to be highly positively correlated with $\hat{\vartheta}_N^m$, provided that $\bar{x}$ is close to x^*. In such cases variability of $\widehat{\mathrm{gap}}_{N,M}(\bar{x})$ can be considerably smaller than variability of $\hat{f}_{N'}(\bar{x}) - \bar{v}_{N,M}$. This is the idea of common random number generated estimators.

Remark 16. Of course, in order to obtain a valid statistical lower bound for the optimal value ϑ^* we can use any (deterministic) lower bound for the optimal value $\hat{\vartheta}_N$ of the corresponding SAA problem instead of $\hat{\vartheta}_N$ itself. For example, suppose that the problem is *convex*. By convexity of $\hat{f}_N(\cdot)$ we have that for any $x' \in X$ and $\gamma \in \partial \hat{f}_N(x')$ it holds that

$$\hat{f}_N(x) \geq \hat{f}_N(x') + \gamma^{\mathrm{T}}(x - x'), \quad \forall x \in \mathbb{R}^n. \tag{5.176}$$

Therefore, we can proceed as follows. Choose points $x_1, \ldots, x_r \in X$, calculate subgradients $\hat{\gamma}_{iN} \in \partial \hat{f}_N(x_i)$, $i = 1, \ldots, r$, and solve the problem

$$\operatorname*{Min}_{x \in X} \max_{1 \leq i \leq r} \left\{\hat{f}_N(x_i) + \hat{\gamma}_{iN}^{\mathrm{T}}(x - x_i)\right\}. \tag{5.177}$$

Denote by $\hat{\lambda}_N$ the optimal value of (5.177). By (5.176) we have that $\hat{\lambda}_N$ is less than or equal to the optimal value $\hat{\vartheta}_N$ of the corresponding SAA problem and hence gives a valid statistical lower bound for ϑ^*. A possible advantage of $\hat{\lambda}_N$ over $\hat{\vartheta}_N$ is that it could be easier to solve (5.177) than the corresponding SAA problem. For instance, if the set X is polyhedral, then (5.177) can be formulated as a linear programming problem.

Of course, this approach raises the question of how to choose the points $x_1, \ldots, x_r \in X$. Suppose that the expectation function $f(x)$ is differentiable at the points $x_1, \ldots, x_r$. Then for any choice of $\hat{\gamma}_{iN} \in \partial \hat{f}_N(x_i)$ we have that subgradients $\hat{\gamma}_{iN}$ converge to $\nabla f(x_i)$ w.p. 1. Therefore $\hat{\lambda}_N$ converges w.p. 1 to the optimal value of the problem

$$\operatorname*{Min}_{x \in X} \max_{1 \leq i \leq r} \left\{f(x_i) + \nabla f(x_i)^{\mathrm{T}}(x - x_i)\right\}, \tag{5.178}$$

provided that the set X is bounded. Again by convexity arguments, the optimal value of (5.178) is less than or equal to the optimal value ϑ^* of the true problem. If we can find such

points $x_1, \ldots, x_r \in X$ that the optimal value of (5.178) is less than ϑ^* by a small amount, then it could be advantageous to use $\hat{\lambda}_N$ instead of $\hat{\vartheta}_N$. We also should keep in mind that the number r should be relatively small; otherwise we may loose the advantage of solving the easier problem (5.177).

A natural approach to choosing the required points and hence to applying the above procedure is the following. By solving (once) an SAA problem, find points $x_1, \ldots, x_r \in X$ such that the optimal value of the corresponding problem (5.177) provides us with high accuracy an estimate of the optimal value of this SAA problem. Use some (all) of these points to calculate lower bound estimates $\hat{\lambda}_N^m$, $m = 1, \ldots, M$, probably with a larger sample size N. Calculate the average $\bar{\lambda}_{N,M}$ together with the corresponding sample variance and construct the associated $100(1-\alpha)\%$ confidence lower bound similar to (5.170).

Estimation of Optimality Gap of Minimax and Expectation-Constrained Problems

Consider a minimax problem of the form (5.46). Let ϑ^* be the optimal value of this (true) minimax problem. Clearly for any $\bar{y} \in Y$ we have that

$$\vartheta^* \geq \inf_{x \in X} f(x, \bar{y}). \tag{5.179}$$

Now for the optimal value of the right-hand side of (5.179) we can construct a valid statistical lower bound, and hence a valid statistical lower bound for ϑ^*, as before by solving the corresponding SAA problems several times and averaging calculated optimal values. Suppose, further, that the minimax problem (5.46) has a nonempty set $S_x \times S_y \subset X \times Y$ of saddle points, and hence its optimal value is equal to the optimal value of its dual problem (5.48). Then for any $\bar{x} \in X$ we have that

$$\vartheta^* \leq \sup_{y \in Y} f(\bar{x}, y), \tag{5.180}$$

and the equalities in (5.179) and/or (5.180) hold iff $\bar{y} \in S_y$ and/or $\bar{x} \in S_x$. By (5.180) we can construct a valid statistical upper bound for ϑ^* by averaging optimal values of sample average approximations of the right-hand side of (5.180). Of course, the quality of these bounds will depend on a good choice of the points $\bar{y}$ and $\bar{x}$. A natural construction for the candidate solutions $\bar{y}$ and $\bar{x}$ will be to use optimal solutions of a run of the corresponding minimax SAA problem (5.47).

Similar ideas can be applied to validation of stochastic problems involving constraints given as expected value functions (See (5.11)–(5.13)). That is, consider the problem

$$\operatorname*{Min}_{x \in X_0} f(x) \ \text{ s.t. } g_i(x) \leq 0, \ i = 1, \ldots, p, \tag{5.181}$$

where X_0 is a nonempty subset of $\mathbb{R}^n$, $f(x) := \mathbb{E}[F(x, \xi)]$, and $g_i(x) := \mathbb{E}[G_i(x, \xi)]$, $i = 1, \ldots, p$. We have that

$$\vartheta^* = \inf_{x \in X_0} \sup_{\lambda \geq 0} L(x, \lambda), \tag{5.182}$$

where ϑ^* is the optimal value and $L(x, \lambda) := f(x) + \sum_{i=1}^p \lambda_i g_i(x)$ is the Lagrangian of problem (5.181). Therefore, for any $\bar{\lambda} \geq 0$, we have that

$$\vartheta^* \geq \inf_{x \in X_0} L(x, \bar{\lambda}), \tag{5.183}$$

and the equality in (5.183) is attained if the problem (5.179) is convex and $\bar{\lambda}$ is a Lagrange multipliers vector satisfying the corresponding first order optimality conditions. Of course, a statistical lower bound for the optimal value of the problem in the right-hand side of (5.183) is also a statistical lower bound for ϑ^*.

Unfortunately, an upper bound which can be obtained by interchanging the "inf" and "sup" operators in (5.182) cannot be used in a straightforward way . This is because if, for a chosen $\bar{x} \in X_0$, it happens that $\hat{g}_{iN}(\bar{x}) > 0$ for some $i \in \{1, \ldots, p\}$, then

$$\sup_{\lambda \geq 0} \left\{ \hat{f}_N(\bar{x}) + \sum_{i=1}^{p} \lambda_i \hat{g}_{iN}(\bar{x}) \right\} = +\infty. \tag{5.184}$$

Of course, in such a case the obtained upper bound $+\infty$ is useless. This typically will be the case if $\bar{x}$ is constructed as a solution of an SAA problem and some of the SAA constraints are active at $\bar{x}$. Note also that if $\hat{g}_{iN}(\bar{x}) \leq 0$ for all $i \in \{1, \ldots, p\}$, then the supremum in the left-hand side of (5.184) is equal to $\hat{f}_N(\bar{x})$.

If we can ensure, with a high probability $1-\alpha$, that a chosen point $\bar{x}$ is a feasible point of the true problem 5.179), then we can construct an upper bound by estimating $f(\bar{x})$ using a relatively large sample. This, in turn, can be approached by verifying, for an independent sample of size N', that $\hat{g}_{iN'}(\bar{x}) + \kappa \hat{\sigma}_{iN'}(\bar{x}) \leq 0$, $i = 1, \ldots, p$, where $\hat{\sigma}^2_{iN'}(\bar{x})$ is a sample variance of $\hat{g}_{iN'}(\bar{x})$ and κ is a positive constant chosen in such a way that the probability of $g_i(\bar{x})$ being bigger that $\hat{g}_{iN'}(\bar{x}) + \kappa \hat{\sigma}_{iN'}(\bar{x})$ is less than α/p for all $i \in \{1, \ldots, p\}$.

5.6.2 Statistical Testing of Optimality Conditions

Suppose that the feasible set X is defined by (equality and inequality) constraints in the form

$$X := \big\{ x \in \mathbb{R}^n : g_i(x) = 0,\ i = 1, \ldots, q;\ g_i(x) \leq 0,\ \ i = q+1, \ldots, p \big\}, \tag{5.185}$$

where $g_i(x)$ are smooth (at least continuously differentiable) *deterministic* functions. Let $x^* \in X$ be an optimal solution of the true problem and suppose that the expected value function $f(\cdot)$ is differentiable at x^*. Then, under a constraint qualification, first order (KKT) optimality conditions hold at x^*. That is, there exist Lagrange multipliers λ_i such that $\lambda_i \geq 0$, $i \in \mathcal{I}(x^*)$ and

$$\nabla f(x^*) + \sum_{i \in \mathcal{J}(x^*)} \lambda_i \nabla g_i(x^*) = 0, \tag{5.186}$$

where $\mathcal{I}(x) := \{i : g_i(x) = 0,\ i = q+1, \ldots, p\}$ denotes the index set of inequality constraints active at a point $x \in \mathbb{R}^n$, and $\mathcal{J}(x) := \{1, \ldots, q\} \cup \mathcal{I}(x)$. Note that if the constraint functions are linear, say, $g_i(x) := a_i^{\mathrm{T}} x + b_i$, then $\nabla g_i(x) = a_i$ and the above KKT conditions hold without a constraint qualification. Consider the (polyhedral) cone

$$K(x) := \left\{ z \in \mathbb{R}^n : z = \sum_{i \in \mathcal{J}(x)} \alpha_i \nabla g_i(x),\ \alpha_i \leq 0,\ i \in \mathcal{I}(x) \right\}. \tag{5.187}$$

Then the KKT optimality conditions can be written in the form $\nabla f(x^*) \in K(x^*)$.

Suppose now that $f(\cdot)$ is differentiable at the candidate point $\bar{x} \in X$ and that the gradient $\nabla f(\bar{x})$ can be estimated by a (random) vector $\gamma_N(\bar{x})$. In particular, if $F(\cdot, \xi)$ is differentiable at $\bar{x}$ w.p. 1, then we can use the estimator

$$\gamma_N(\bar{x}) := \frac{1}{N} \sum_{j=1}^{N} \nabla_x F(\bar{x}, \xi^j) = \nabla \hat{f}_N(\bar{x}) \tag{5.188}$$

associated with the generated[28] random sample. Note that if, moreover, the derivatives can be taken inside the expectation, that is,

$$\nabla f(\bar{x}) = \mathbb{E}[\nabla_x F(\bar{x}, \xi)], \tag{5.189}$$

then the above estimator is unbiased, i.e., $\mathbb{E}[\gamma_N(\bar{x})] = \nabla f(\bar{x})$. In the case of two-stage linear stochastic programming with recourse, formula (5.189) typically holds if the corresponding random data have a continuous distribution. On the other hand, if the random data have a discrete distribution with a finite support, then the expected value function $f(x)$ is piecewise linear and typically is nondifferentiable at an optimal solution.

Suppose, further, that $V_N := N^{1/2}[\gamma_N(\bar{x}) - \nabla f(\bar{x})]$ converges in distribution, as N tends to infinity, to multivariate normal with zero mean vector and covariance matrix Σ, written $V_N \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma)$. For the estimator $\gamma_N(\bar{x})$ defined in (5.188), this holds by the CLT if the interchangeability formula (5.189) holds, the sample is iid, and $\nabla_x F(\bar{x}, \xi)$ has finite second order moments. Moreover, in that case the covariance matrix Σ can be estimated by the corresponding sample covariance matrix

$$\hat{\Sigma}_N := \frac{1}{N-1} \sum_{j=1}^{N} \Big[\nabla_x F(\bar{x}, \xi^j) - \nabla \hat{f}_N(\bar{x})\Big]\Big[\nabla_x F(\bar{x}, \xi^j) - \nabla \hat{f}_N(\bar{x})\Big]^{\mathrm{T}}. \tag{5.190}$$

Under the above assumptions, the sample covariance matrix $\hat{\Sigma}_N$ is an unbiased and consistent estimator of Σ.

We have that if $V_N \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma)$ and the covariance matrix Σ is nonsingular, then (given a consistent estimator $\hat{\Sigma}_N$ of Σ) the following holds:

$$N\big(\gamma_N(\bar{x}) - \nabla f(\bar{x})\big)^{\mathrm{T}} \hat{\Sigma}_N^{-1} \big(\gamma_N(\bar{x}) - \nabla f(\bar{x})\big) \xrightarrow{\mathcal{D}} \chi^2_n, \tag{5.191}$$

where χ^2_n denotes chi-square distribution with n degrees of freedom. This allows us to construct the following (approximate) $100(1-\alpha)\%$ confidence region[29] for $\nabla f(\bar{x})$:

$$\left\{ z \in \mathbb{R}^n : \big(\gamma_N(\bar{x}) - z)\big)^{\mathrm{T}} \hat{\Sigma}_N^{-1} \big(\gamma_N(\bar{x}) - z)\big) \le \frac{\chi^2_{\alpha,n}}{N} \right\}. \tag{5.192}$$

Consider the statistic

$$T_N := N \inf_{z \in K(\bar{x})} \big(\gamma_N(\bar{x}) - z\big)^{\mathrm{T}} \hat{\Sigma}_N^{-1} \big(\gamma_N(\bar{x}) - z\big). \tag{5.193}$$

[28] We emphasize that the random sample in (5.188) is generated independently of the sample used to compute the candidate point $\bar{x}$.

[29] Here $\chi^2_{\alpha,n}$ denotes the α-critical value of chi-square distribution with n degrees of freedom. That is, if $Y \sim \chi^2_n$, then $\Pr\{Y \ge \chi^2_{\alpha,n}\} = \alpha$.

Note that since the cone $K(\bar{x})$ is polyhedral and $\hat{\Sigma}_N^{-1}$ is positive definite, the minimization in the right-hand side of (5.193) can be formulated as a quadratic programming problem, and hence can be solved by standard quadratic programming algorithms. We have that the confidence region, defined in (5.192), does not have common points with the cone $K(\bar{x})$ iff $T_N > \chi^2_{\alpha,n}$. We can also use the statistic T_N for testing the hypothesis:

$$H_0 : \ \nabla f(\bar{x}) \in K(\bar{x}) \ \text{ against the alternative } \ H_1 : \ \nabla f(\bar{x}) \notin K(\bar{x}). \tag{5.194}$$

The T_N statistic represents the squared distance, with respect to the norm[30] $\|\cdot\|_{\hat{\Sigma}_N^{-1}}$, from $N^{1/2}\gamma_N(\bar{x})$ to the cone $K(\bar{x})$. Suppose for the moment that only equality constraints are present in the definition (5.185) of the feasible set, and that the gradient vectors $\nabla g_i(\bar{x})$, $i = 1, \ldots, q$, are linearly independent. Then the set $K(\bar{x})$ forms a linear subspace of $\mathbb{R}^n$ of dimension q, and the optimal value of the right-hand side of (5.193) can be written in a closed form. Consequently, it is possible to show that T_N has asymptotically noncentral chi-square distribution with $n - q$ degrees of freedom and the noncentrality parameter[31]

$$\delta := N \inf_{z \in K(\bar{x})} \big(\nabla f(\bar{x}) - z\big)^{\mathrm{T}} \Sigma^{-1} \big(\nabla f(\bar{x}) - z\big). \tag{5.195}$$

In particular, under H_0 we have that $\delta = 0$, and hence the null distribution of T_N is asymptotically central chi-square with $n - q$ degrees of freedom.

Consider now the general case where the feasible set is defined by equality and inequality constraints as in (5.185). Suppose that the gradient vectors $\nabla g_i(\bar{x})$, $i \in \mathcal{J}(\bar{x})$, are linearly independent and that the *strict complementarity* condition holds at $\bar{x}$, that is, the Lagrange multipliers λ_i, $i \in \mathcal{I}(\bar{x})$, corresponding to the active at $\bar{x}$ inequality constraints, are positive. Then for $\gamma_N(\bar{x})$ sufficiently close to $\nabla f(\bar{x})$ the minimizer in the right-hand side of (5.193) will be lying in the linear space generated by vectors $\nabla g_i(\bar{x})$, $i \in \mathcal{J}(\bar{x})$. Therefore, in such case the null distribution of T_N is asymptotically central chi-square with $\nu := n - |\mathcal{J}(\bar{x})|$ degrees of freedom. Consequently, for a computed value T_N^* of the statistic T_N we can calculate (approximately) the corresponding p-value, which is equal to $\Pr\{Y \geq T_N^*\}$, where $Y \sim \chi^2_\nu$. This p-value gives an indication of the quality of the candidate solution $\bar{x}$ with respect to the stochastic precision.

It should be understood that by accepting (i.e., failing to reject) H_0, we do not claim that the KKT conditions hold exactly at $\bar{x}$. By accepting H_0 we rather assert that we cannot separate $\nabla f(\bar{x})$ from $K(\bar{x})$, given precision of the generated sample. That is, statistical error of the estimator $\gamma_N(\bar{x})$ is bigger than the squared $\|\cdot\|_{\Sigma^{-1}}$-norm distance between $\nabla f(\bar{x})$ and $K(\bar{x})$. Also, rejecting H_0 does not necessarily mean that $\bar{x}$ is a poor candidate for an optimal solution of the true problem. The calculated value of the T_N statistic can be large, i.e., the p-value can be small, simply because the estimated covariance matrix $N^{-1}\hat{\Sigma}_N$ of $\gamma_N(\bar{x})$ is "small." In such cases, $\gamma_N(\bar{x})$ provides an accurate estimator of $\nabla f(\bar{x})$ with the corresponding confidence region (5.192) being small. Therefore, the above p-value should be compared with the size of the confidence region (5.192), which in turn is defined by the size of the matrix $N^{-1}\hat{\Sigma}_N$ measured, for example, by its eigenvalues. Note also that it may happen that $|\mathcal{J}(\bar{x})| = n$, and hence $\nu = 0$. Under the strict complementarity condition, this

[30] For a positive definite matrix A, the norm $\|\cdot\|_A$ is defined as $\|z\|_A := (z^{\mathrm{T}}Az)^{1/2}$.

[31] Note that under the alternative (i.e., if $\nabla f(\bar{x}) \notin K(\bar{x})$), the noncentrality parameter δ tends to infinity as $N \to \infty$. Therefore, in order to justify the above asymptotics, one needs a technical assumption known as Pitman's parameter drift.

means that $\nabla f(\bar{x})$ lies in the interior of the cone $K(\bar{x})$, which in turn is equivalent to the condition that $\tilde{f}'(\bar{x}, d) \geq c\|d\|$ for some $c > 0$ and all $d \in \mathbb{R}^n$. Then, by the LD principle (see (7.192) in particular), the event $\gamma_N(\bar{x}) \in K(\bar{x})$ happens with probability approaching one exponentially fast.

Let us remark again that the above testing procedure is applicable if $F(\cdot, \xi)$ is differentiable at $\bar{x}$ w.p. 1 and the interchangeability formula (5.189) holds. This typically happens in cases where the corresponding random data have a continuous distribution.

5.7 Chance Constrained Problems

Consider a chance constrained problem of the form

$$\operatorname*{Min}_{x \in X} f(x) \ \text{ s.t. } p(x) \leq \alpha, \tag{5.196}$$

where $X \subset \mathbb{R}^n$ is a closed set, $f : \mathbb{R}^n \to \mathbb{R}$ is a continuous function, $\alpha \in (0, 1)$ is a given significance level, and

$$p(x) := \Pr\big\{C(x, \xi) > 0\big\} \tag{5.197}$$

is the probability that constraint is violated at point $x \in X$. We assume that ξ is a random vector, whose probability distribution P is supported on set $\Xi \subset \mathbb{R}^d$, and the function $C : \mathbb{R}^n \times \Xi \to \mathbb{R}$ is a Carathéodory function. The chance constraint $p(x) \leq \alpha$ can be written equivalently in the form

$$\Pr\big\{C(x, \xi) \leq 0\big\} \geq 1 - \alpha. \tag{5.198}$$

Let us also remark that several chance constraints

$$\Pr\big\{C_i(x, \xi) \leq 0, \ i = 1, \ldots, q\big\} \geq 1 - \alpha \tag{5.199}$$

can be reduced to one chance constraint (5.198) by employing the max-function $C(x, \xi) := \max_{1 \leq i \leq q} C_i(x, \xi)$. Of course, in some cases this may destroy a nice structure of considered functions. At this point, however, this is not important.

5.7.1 Monte Carlo Sampling Approach

We discuss now a way of solving problem (5.196) by Monte Carlo sampling. For the sake of simplicity we assume that the objective function $f(x)$ is given explicitly and only the chance constraints should be approximated.

We can write the probability $p(x)$ in the form of the expectation,

$$p(x) = \mathbb{E}\left[\mathbf{1}_{(0,\infty)}(C(x, \xi))\right],$$

and estimate this probability by the corresponding SAA function (compare with (5.14)–(5.16))

$$\hat{p}_N(x) := N^{-1} \sum_{j=1}^{N} \mathbf{1}_{(0,\infty)}\left(C(x, \xi^j)\right). \tag{5.200}$$

Recall that $\mathbf{1}_{(0,\infty)}(C(x,\xi))$ is equal to 1 if $C(x,\xi) > 0$, and it is equal 0 otherwise. Therefore, $\hat{p}_N(x)$ is equal to the proportion of times that $C(x,\xi^j) > 0$, $j = 1,\ldots,N$. Consequently we can write the corresponding SAA problem as

$$\operatorname*{Min}_{x\in X} f(x) \ \text{ s.t. } \hat{p}_N(x) \leq \alpha. \tag{5.201}$$

Proposition 5.29. *Let $C(x,\xi)$ be a Carathéodory function. Then the functions $p(x)$ and $\hat{p}_N(x)$ are lower semicontinuous. Suppose, further, that the sample is* iid. *Then $\hat{p}_N \overset{\text{e}}{\to} p$ w.p.* 1. *Moreover, suppose that for every $x \in X$ it holds that*

$$\Pr\{\xi \in \Xi : C(x,\xi) = 0\} = 0, \tag{5.202}$$

i.e., $C(x,\xi) \neq 0$ w.p. 1. *Then the function $p(x)$ is continuous on X and $\hat{p}_N(x)$ converges to $p(x)$ w.p.* 1 *uniformly on any compact subset of X.*

Proof. Consider function $\psi(x,\xi) := \mathbf{1}_{(0,\infty)}\big(C(x,\xi)\big)$. Recall that $p(x) = \mathbb{E}_P[\psi(x,\xi)]$ and $\hat{p}_N(x) = \mathbb{E}_{P_N}[\psi(x,\xi)]$, where $P_N := N^{-1}\sum_{j=1}^N \Delta(\xi^j)$ is the respective empirical measure. Since the function $\mathbf{1}_{(0,\infty)}(\cdot)$ is lower semicontinuous and $C(x,\xi)$ is a Carathéodory function, it follows that the function $\psi(x,\xi)$ is random lower semicontinuous. Lower semicontinuity of $p(x)$ and $\hat{p}_N(x)$ follows by Fatou's lemma (see Theorem 7.42). If the sample is iid, the epiconvergence $\hat{p}_N \overset{\text{e}}{\to} p$ w.p. 1 follows by Theorem 7.51. Note that the dominance condition, from below and from above, holds here automatically since $|\psi(x,\xi)| \leq 1$.

Suppose, further, that condition (5.202) holds. Then for every $x \in X$, $\psi(\cdot,\xi)$ is continuous at x w.p. 1. It follows by the Lebesgue dominated convergence theorem that $p(\cdot)$ is continuous at x (see Theorem 7.43). Finally, the uniform convergence w.p. 1 follows by Theorem 7.48. □

Since the function $p(x)$ is lower semicontinuous and the set X is closed, it follows that the feasible set of problem (5.196) is closed. If, moreover, it is nonempty and bounded, then problem (5.196) has a nonempty set S of optimal solutions. (Recall that the objective function $f(x)$ is assumed to be continuous here.) The same applies to the corresponding SAA problem (5.201). We have here the following consistency properties of the optimal value $\hat{\vartheta}_N$ and the set $\hat{S}_N$ of optimal solutions of the SAA problem (5.201) (compare with Theorem 5.5).

Proposition 5.30. *Suppose that the set X is compact, the function $f(x)$ is continuous, $C(x,\xi)$ is a Carathéodory function, the sample is iid, and the following condition holds:* (a) *there is an optimal solution $\bar{x}$ of the true problem such that for any $\epsilon > 0$ there is $x \in X$ with $\|x - \bar{x}\| \leq \epsilon$ and $p(x) < \alpha$. Then $\hat{\vartheta}_N \to \vartheta^*$ and $\mathbb{D}(\hat{S}_N, S) \to 0$ w.p.* 1 *as $N \to \infty$.*

Proof. By condition (a), the set S is nonempty and there is $x' \in X$ such that $p(x') < \alpha$. By the LLN we have that $\hat{p}_N(x')$ converges to $p(x')$ w.p. 1. Consequently $\hat{p}_N(x') < \alpha$, and hence the SAA problem has a feasible solution, w.p. 1 for N large enough. Since $\hat{p}_N(\cdot)$ is lower semicontinuous, the feasible set of SAA problem is closed and hence compact and thus $\hat{S}_N$ is nonempty w.p. 1 for N large enough. Of course, if x' is a feasible solution of an SAA problem, then $f(x') \geq \hat{\vartheta}_N$, where $\hat{\vartheta}_N$ is the optimal value of that SAA problem.

For a given $\varepsilon > 0$ let $x' \in X$ be a point sufficiently close to $\bar{x} \in S$ such that $\hat{p}_N(x') < \alpha$ and $f(x') \leq f(\bar{x}) + \varepsilon$. Since $f(\cdot)$ is continuous, existence of such point is ensured by condition (a). Consequently,

$$\limsup_{N\to\infty} \hat{\vartheta}_N \leq f(x') \leq f(\bar{x}) + \varepsilon = \vartheta^* + \varepsilon \quad \text{w.p. 1.} \tag{5.203}$$

Since $\varepsilon > 0$ is arbitrary, it follows that

$$\limsup_{N\to\infty} \hat{\vartheta}_N \leq \vartheta^* \quad \text{w.p. 1.} \tag{5.204}$$

Now let $\hat{x}_N \in \hat{S}_N$, i.e., $\hat{x}_N \in X$, $\hat{p}_N(\hat{x}_N) \leq \alpha$ and $\hat{\vartheta}_N = f(\hat{x}_N)$. Since the set X is compact, we can assume by passing to a subsequence if necessary that $\hat{x}_N$ converges to a point $\bar{x} \in X$ w.p. 1. Also by Proposition 5.29 we have that $\hat{p}_N \overset{\text{e}}{\to} p$ w.p. 1, and hence

$$\liminf_{N\to\infty} \hat{p}_N(\hat{x}_N) \geq p(\bar{x}) \quad \text{w.p. 1.}$$

It follows that $p(\bar{x}) \leq \alpha$ and hence $\bar{x}$ is a feasible point of the true problem, and thus $f(\bar{x}) \geq \vartheta^*$. Also $f(\hat{x}_N) \to f(\bar{x})$ w.p. 1, and hence

$$\liminf_{N\to\infty} \hat{\vartheta}_N \geq \vartheta^* \quad \text{w.p. 1.} \tag{5.205}$$

It follows from (5.204) and (5.205) that $\hat{\vartheta}_N \to \vartheta^*$ w.p. 1. It also follows that the point $\bar{x}$ is an optimal solution of the true problem and consequently we obtain that $\mathbb{D}(\hat{S}_N, S) \to 0$ w.p. 1. □

The above condition (a) is essential for the consistency of $\hat{\vartheta}_N$ and $\hat{S}_N$. Think, for example, about a situation where the constraint $p(x) \leq \alpha$ defines just one feasible point $\bar{x}$ such that $p(\bar{x}) = \alpha$. Then arbitrary small changes in the constraint $\hat{p}_N(x) \leq \alpha$ may result in that the feasible set of the corresponding SAA problem becomes empty. Note also that condition (a) was not used in the proof of inequality (5.205). Verification of this condition (a) can be done by ad hoc methods.

We have that under mild regularity conditions, optimal value and optimal solutions of the SAA problem (5.201) converge w.p. 1, as $N \to \infty$, to their counterparts of the true problem (5.196). There are, however, several potential problems with the SAA approach here. In order for $\hat{p}_N(x)$ to be a reasonably accurate estimate of $p(x)$, the sample size N should be significantly bigger than α^{-1}. For small α this may result in a large sample size. Another problem is that typically the function $\hat{p}_N(x)$ is discontinuous and the SAA problem (5.201) is a combinatorial problem which could be difficult to solve. Therefore we consider the following approach.

Convex Approximation Approach

For a generated sample $\xi^1, \ldots, \xi^N$ consider the problem

$$\operatorname*{Min}_{x\in X} f(x) \quad s.t.\ C(x, \xi^j) \leq 0,\ j = 1, \ldots, N. \tag{5.206}$$

Note that for $\alpha = 0$ the SAA problem (5.201) coincides with problem (5.206). If the set X and functions $f(\cdot)$ and $C(\cdot, \xi)$, $\xi \in \Xi$, are convex, then (5.206) is a convex problem and could be efficiently solved provided that the involved functions are given in a closed form and the sample size N is not too large. Clearly, as $N \to \infty$ the feasible set of problem (5.206) will shrink to the set of $x \in X$ determined by the constraints $C(x, \xi) \leq 0$ for a.e. $\xi \in \Xi$, and hence for large N will be overly conservative for the true chance constrained problem (5.196). Nevertheless, it makes sense to ask the question for what sample size N an optimal solution of problem (5.206) is guaranteed to be a feasible point of problem (5.196).

We need the following auxiliary result.

Lemma 5.31. *Suppose that the set X and functions $f(\cdot)$ and $C(\cdot, \xi)$, $\xi \in \Xi$, are convex and let $\bar{x}_N$ be an optimal solution of problem* (5.206)*. Then there exists an index set $J \subset \{1, \ldots, N\}$ such that $|J| \leq n$ and $\bar{x}_N$ is an optimal solution of the problem*

$$\operatorname*{Min}_{x \in X} f(x) \ \ s.t. \ C(x, \xi^j) \leq 0, \ \ j \in J. \tag{5.207}$$

Proof. Consider sets $A_0 := \{x \in X : f(x) < f(\bar{x}_N)\}$ and $A_j := \{x \in X : C(x, \xi^j) \leq 0\}$, $j = 1, \ldots, N$. Since X, $f(\cdot)$ and $C(\cdot, \xi^j)$ are convex, these sets are convex. Now we argue by a contradiction. Suppose that the assertion of this lemma is not correct. Then the intersection of A_0 and any n sets A_j is nonempty. Since the intersection of all sets A_j, $j \in \{1, \ldots, N\}$, is nonempty (these sets have at least one common element $\bar{x}_N$), it follows that the intersection of any $n + 1$ sets of the family A_j, $j \in \{0, 1, \ldots, N\}$, is nonempty. By Helly's theorem (Theorem 7.3) this implies that the intersection of all sets A_j, $j \in \{0, 1, \ldots, N\}$, is nonempty. This, in turn, implies existence of a feasible point $\tilde{x}$ of problem (5.206) such that $f(\tilde{x}) < f(\bar{x}_N)$, which contradicts optimality of $\bar{x}_N$. □

We will use the following assumptions.

(F1) For any $N \in \mathbb{N}$ and any $(\xi_1, \ldots, \xi_N) \in \Xi^N$, problem (5.206) attains the unique optimal solution $\bar{x}_N = \bar{x}(\xi_1, \ldots, \xi_N)$.

Recall that sometimes we use the same notation for a random vector and its particular value (realization). In the above assumption we view $\xi_1, \ldots, \xi_N$ as an element of the set Ξ^N and $\bar{x}_N$ as a function of $\xi_1, \ldots, \xi_N$. Of course, if $\xi_1, \ldots, \xi_N$ is a random sample, then $\bar{x}_N$ becomes a random vector.

Let $\mathcal{J} = \mathcal{J}(\xi^1, \ldots, \xi^N) \subset \{1, \ldots, N\}$ be an index set such that $\bar{x}_N$ is an optimal solution of the problem (5.207) for $J = \mathcal{J}$. Moreover, let the index set $\mathcal{J}$ be minimal in the sense that if any of the constraints $C(x, \xi^j) \leq 0$, $j \in \mathcal{J}$, is removed, then $\bar{x}_N$ is not an optimal solution of the obtained problem. We assume that w.p. 1 such minimal index set is unique. By Lemma 5.31, we have that $|\mathcal{J}| \leq n$. By P^N we denote here the product probability measure on the set Ξ^N, i.e., P^N is the probability distribution of the iid sample $\xi^1, \ldots, \xi^N$.

(F2) There is an integer $\mathfrak{n} \in \mathbb{N}$ such that, for any $N \geq \mathfrak{n}$, w.p. 1 the minimal set $\mathcal{J} = \mathcal{J}(\xi^1, \ldots, \xi^N)$ is uniquely defined and has constant cardinality $\mathfrak{n}$, i.e., $P^N\{|\mathcal{J}| = \mathfrak{n}\} = 1$.

By Lemma 5.31 we have that $\mathfrak{n} \leq n$.

Assumption (F1) holds, for example, if the set X is compact and convex, functions $f(\cdot)$ and $C(\cdot, \xi)$, $\xi \in \Xi$, are convex, and either $f(\cdot)$ or the feasible set of problem (5.206) is strictly convex. Assumption (F2) is more involved; it is needed to show an equality in the estimate (5.209) of the following theorem.

The following result is due to Campi and Garatti [30], building on work of Calafiore and Campi [29]. Denote

$$\mathfrak{b}(k; \alpha, N) := \sum_{i=0}^{k} \binom{N}{i} \alpha^i (1-\alpha)^{N-i}, \quad k = 0, \ldots, N. \tag{5.208}$$

That is, $\mathfrak{b}(k; \alpha, N) = \mathrm{Pr}(W \leq k)$, where $W \sim B(\alpha, N)$ is a random variable having binomial distribution.

Theorem 5.32. *Suppose that the set X and functions $f(\cdot)$ and $C(\cdot, \xi)$, $\xi \in \Xi$, are convex and conditions* (F1) *and* (F2) *hold. Then for $\alpha \in (0, 1)$ and for an* iid *sample $\xi^1, \ldots, \xi^N$ of size $N \geq \mathfrak{n}$ we have that*

$$\mathrm{Pr}\big\{p(\bar{x}_N) > \alpha\big\} = \mathfrak{b}(\mathfrak{n} - 1; \alpha, N). \tag{5.209}$$

Proof. Let $\mathfrak{J}_\mathfrak{n}$ be the family of all sets $J \subset \{1, \ldots, N\}$ of cardinality $\mathfrak{n}$. We have that $|\mathfrak{J}_\mathfrak{n}| = \binom{N}{\mathfrak{n}}$. For $J \in \mathfrak{J}_\mathfrak{n}$ define the set

$$\Sigma_J := \big\{(\xi^1, \ldots, \xi^N) \in \Xi^N : \mathcal{J}(\xi^1, \ldots, \xi^N) = J\big\} \tag{5.210}$$

and denote by $\hat{x}_J = \hat{x}_J(\xi^1, \ldots, \xi^N)$ an optimal solution of problem (5.207) for $\mathcal{J} = J$. By condition (F1), such optimal solution $\hat{x}_J$ exists and is unique, and hence

$$\Sigma_J = \big\{(\xi^1, \ldots, \xi^N) \in \Xi^N : \hat{x}_J = \bar{x}_N\big\}. \tag{5.211}$$

Note that for any permutation of vectors $\xi^1, \ldots, \xi^N$, problem (5.206) remains the same. Therefore, any set from the family $\{\Sigma_J\}_{J \in \mathfrak{J}_\mathfrak{n}}$ can be obtained from another set of that family by an appropriate permutation of its components. Since P^N is the direct product probability measure, it follows that the probability measure of each set Σ_J, $J \in \mathfrak{J}_\mathfrak{n}$, is the same. The sets Σ_J are disjoint and, because of condition (F2), union of all these sets is equal to Ξ^N up to a set of P^N-measure zero. Since there are $\binom{N}{\mathfrak{n}}$ such sets, we obtain that

$$P^N(\Sigma_J) = \frac{1}{\binom{N}{\mathfrak{n}}}. \tag{5.212}$$

Consider the optimal solution $\bar{x}_\mathfrak{n} = \bar{x}(\xi^1, \ldots, \xi^\mathfrak{n})$ for $N = \mathfrak{n}$, and let $H(z)$ be the cdf of the random variable $p(\bar{x}_\mathfrak{n})$, i.e.,

$$H(z) := P^\mathfrak{n}\{p(\bar{x}_\mathfrak{n}) \leq z\}. \tag{5.213}$$

Let us show that for $N \geq \mathfrak{n}$,

$$P^N(\Sigma_J) = \int_0^1 (1-z)^{N-\mathfrak{n}} dH(z). \tag{5.214}$$

Indeed, for $z \in [0, 1]$ and $J \in \mathfrak{J}_{\mathfrak{n}}$ consider the sets

$$\begin{array}{l} \Delta_z := \{(\xi_1, \ldots, \xi_N) : p(\bar{x}_N) \in [z, z + dz]\}, \\ \Delta_{J,z} := \left\{(\xi_1, \ldots, \xi_N) : p(\hat{x}_J) \in [z, z + dz]\right\}. \end{array} \tag{5.215}$$

By (5.211) we have that $\Delta_z \cap \Sigma_J = \Delta_{J,z} \cap \Sigma_J$. For $J \in \mathfrak{J}_{\mathfrak{n}}$ let us evaluate probability of the event$\Delta_{J,z} \cap \Sigma_J$. For the sake of notational simplicity let us take $J = \{1, \ldots, \mathfrak{n}\}$. Note that $\hat{x}_J$ depends on $(\xi^1, \ldots, \xi^{\mathfrak{n}})$ only. Therefore $\Delta_{J,z} = \Delta_z^* \times \Xi^{N-\mathfrak{n}}$, where Δ_z^* is a subset of $\Xi^{\mathfrak{n}}$ corresponding to the event $p(\hat{x}_J) \in [z, z + dz]$. Conditional on $(\xi^1, \ldots, \xi^{\mathfrak{n}})$, the event $\Delta_{J,z} \cap \Sigma_J$ happens iff the point $\hat{x}_J = \hat{x}_J(\xi^1, \ldots, \xi^{\mathfrak{n}})$ remains feasible for the remaining $N - \mathfrak{n}$ constraints, i.e., iff $C(\hat{x}_J, \xi^j) \leq 0$ for all $j = \mathfrak{n} + 1, \ldots, N$. If $p(\hat{x}_J) = z$, then probability of each event "$C(\hat{x}_J, \xi^j) \leq 0$" is equal to $1 - z$. Since the sample is iid, we obtain that conditional on $(\xi^1, \ldots, \xi^{\mathfrak{n}}) \in \Delta_z^*$, probability of the event $\Delta_{J,z} \cap \Sigma_J$ is equal to $(1 - z)^{N-\mathfrak{n}}$. Consequently, the unconditional probability

$$P^N(\Delta_z \cap \Sigma_J) = P^N(\Delta_{J,z} \cap \Sigma_J) = (1 - z)^{N-\mathfrak{n}} dH(z), \tag{5.216}$$

and hence (5.214) follows.

It follows from (5.212) and (5.214) that

$$\binom{N}{\mathfrak{n}} \int_0^1 (1 - z)^{N-\mathfrak{n}} dH(z) = 1, \quad N \geq \mathfrak{n}. \tag{5.217}$$

Let us observe that $H(z) := z^{\mathfrak{n}}$ satisfies (5.217) for all $N \geq \mathfrak{n}$. Indeed, using integration by parts, we have

$$\begin{array}{rl} \binom{N}{\mathfrak{n}} \int_0^1 (1 - z)^{N-\mathfrak{n}} dz^{\mathfrak{n}} = & -\binom{N}{\mathfrak{n}} \frac{\mathfrak{n}}{N-\mathfrak{n}+1} \int_0^1 z^{\mathfrak{n}-1} d(1 - z)^{N-\mathfrak{n}+1} \\ = & \binom{N}{\mathfrak{n}-1} \int_0^1 (1 - z)^{N-\mathfrak{n}+1} dz^{\mathfrak{n}-1} = \cdots = 1. \end{array} \tag{5.218}$$

We also have that (5.217) determine respective moments of random variable $1 - Z$, where $Z \sim H(z)$, and hence (since random variable $p(\bar{x}_{\mathfrak{n}})$ has a bounded support) by the general theory of moments these equations have unique solution. Therefore we conclude that $H(z) = z^{\mathfrak{n}}$, $0 \leq z \leq 1$, is the cdf of $p(\bar{x}_{\mathfrak{n}})$.

We also have by (5.216) that

$$P^N\left\{p(\bar{x}_N) \in [z, z + dz]\right\} = \sum_{J \in \mathfrak{J}_{\mathfrak{n}}} P^N(\Delta_z \cap \Sigma_J) = \binom{N}{\mathfrak{n}} (1 - z)^{N-\mathfrak{n}} dH(z). \tag{5.219}$$

Therefore, since $H(z) = z^{\mathfrak{n}}$ and using integration by parts similar to (5.218), we can write

$$\begin{array}{l} P^N\left\{p(\bar{x}_N) > \alpha\right\} \\ \quad = \binom{N}{\mathfrak{n}} \int_\alpha^1 (1 - z)^{N-\mathfrak{n}} dH(z) = \binom{N}{\mathfrak{n}} \mathfrak{n} \int_\alpha^1 (1 - z)^{N-\mathfrak{n}} z^{\mathfrak{n}-1} dz \\ \quad = \binom{N}{\mathfrak{n}} \frac{\mathfrak{n}}{N-\mathfrak{n}+1} \left[-(1 - z)^{N-\mathfrak{n}+1} z^{\mathfrak{n}-1}\Big|_\alpha^1 + \int_\alpha^1 (1 - z)^{N-\mathfrak{n}+1} dz^{\mathfrak{n}-1}\right] \\ \quad = \binom{N}{\mathfrak{n}-1} (1 - \alpha)^{N-\mathfrak{n}+1} \alpha^{\mathfrak{n}-1} + \binom{N}{\mathfrak{n}-1} \int_\alpha^1 (1 - z)^{N-\mathfrak{n}+1} dz^{\mathfrak{n}-1} \\ \quad = \cdots = \sum_{i=0}^{\mathfrak{n}-1} \binom{N}{i} \alpha^i (1 - \alpha)^{N-i}. \end{array} \tag{5.220}$$

Since $\Pr\left\{p(\bar{x}_N) > \alpha\right\} = P^N\left\{p(\bar{x}_N) > \alpha\right\}$, this completes the proof. □

Of course, the event "$p(\bar{x}_N) > \alpha$" means that $\bar{x}_N$ is not a feasible point of the true problem (5.196). Recall that $\mathfrak{n} \leq n$. Therefore, given $\beta \in (0, 1)$, the inequality (5.209) implies that for sample size $N \geq n$ such that

$$\mathfrak{b}(n-1; \alpha, N) \leq \beta, \tag{5.221}$$

we have with probability at least $1 - \beta$ that $\bar{x}_N$ is a feasible solution of the true problem (5.196).

Recall that

$$\mathfrak{b}(n-1; \alpha, N) = \mathrm{Pr}(W \leq n-1),$$

where $W \sim B(\alpha, N)$ is a random variable having binomial distribution. For "not too small" α and large N, good approximation of that probability is suggested by the CLT. That is, W has approximately normal distribution with mean $N\alpha$ and variance $N\alpha(1-\alpha)$, and hence[32]

$$\mathfrak{b}(n-1; \alpha, N) \approx \Phi\left(\frac{n-1-N\alpha}{\sqrt{N\alpha(1-\alpha)}}\right). \tag{5.222}$$

For $N\alpha \geq n-1$, the Hoeffding inequality (7.188) gives the estimate

$$\mathfrak{b}(n-1; \alpha, N) \leq \exp\left\{-\frac{2(N\alpha - n + 1)^2}{N}\right\}, \tag{5.223}$$

and the Chernoff inequality (7.190) gives

$$\mathfrak{b}(n-1; \alpha, N) \leq \exp\left\{-\frac{(N\alpha - n + 1)^2}{2\alpha N}\right\}. \tag{5.224}$$

The estimates (5.221) and (5.224) show that the required sample size N should be of order $O(\alpha^{-1})$. This, of course, is not surprising since just to estimate the probability $p(x)$, for a given x, by Monte Carlo sampling we will need a sample size of order $O(1/p(x))$. For example, for $n = 100$ and $\alpha = \beta = 0.01$, bound (5.221) suggests estimate $N = 12460$ for the required sample size. Normal approximation (5.222) gives practically the same estimate of N. The estimate derived from the bound (5.223) gives a significantly bigger estimate of $N = 40372$. The estimate derived from the Chernoff inequality (5.224) gives a much better estimate of $N = 13410$.

This indicates that the guaranteed estimates like (5.221) could be too conservative for practical calculations. Note also that Theorem 5.32 does not make any claims about quality of $\bar{x}_N$ as a candidate for an optimal solution of the true problem (5.196); it guarantees only its feasibility.

5.7.2 Validation of an Optimal Solution

We discuss now an approach to a practical validation of a candidate point $\bar{x} \in X$ for an optimal solution of the true problem (5.196). This task is twofold, namely, we need to verify feasibility and optimality of $\bar{x}$. Of course, if a point $\bar{x}$ is feasible for the true problem, then $\vartheta^* \leq f(\bar{x})$, i.e., $f(\bar{x})$ gives an upper bound for the true optimal value.

[32] Recall that $\Phi(\cdot)$ is the cdf of standard normal distribution.

Upper Bounds

Let us start with verification of the feasibility of the point $\bar{x}$. For that we need to estimate the probability $p(\bar{x}) = \Pr\{C(\bar{x}, \xi) > 0\}$. We proceed by employing Monte Carlo sampling techniques. For a generated iid random sample $\xi^1, \ldots, \xi^N$, let $\mathfrak{m}$ be the number of times that the constraints $C(\bar{x}, \xi^j) \leq 0$, $j = 1, \ldots, N$, are violated, i.e.,

$$\mathfrak{m} := \sum_{j=1}^{N} \mathbf{1}_{(0,\infty)}\big(C(\bar{x}, \xi^j)\big).$$

Then $\hat{p}_N(\bar{x}) = \mathfrak{m}/N$ is an unbiased estimator of $p(\bar{x})$, and $\mathfrak{m}$ has Binomial distribution $B\,(p(\bar{x}), N)$.

If the sample size N is significantly bigger than $1/p(\bar{x})$, then the distribution of $\hat{p}_N(\bar{x})$ can be reasonably approximated by a normal distribution with mean $p(\bar{x})$ and variance $p(\bar{x})(1-p(\bar{x}))/N$. In that case, one can consider, for a given confidence level $\beta \in (0, 1/2)$, the following approximate upper bound for the probability[33] $p(\bar{x})$:

$$\hat{p}_N(\bar{x}) + z_\beta \sqrt{\frac{\hat{p}_N(\bar{x})(1 - \hat{p}_N(\bar{x}))}{N}}. \tag{5.225}$$

Let us discuss the following, more accurate, approach for constructing an upper confidence bound for the probability $p(\bar{x})$. For a given $\beta \in (0, 1)$ consider

$$\mathfrak{U}_{\beta,N}(\bar{x}) := \sup_{\rho \in [0,1]} \big\{\rho : \mathfrak{b}(\mathfrak{m}; \rho, N) \geq \beta\big\}. \tag{5.226}$$

We have that $\mathfrak{U}_{\beta,N}(\bar{x})$ is a function of $\mathfrak{m}$ and hence is a random variable. Note that $\mathfrak{b}(\mathfrak{m}; \rho, N)$ is continuous and monotonically decreasing in $\rho \in (0, 1)$. Therefore, in fact, the supremum in the right-hand side of (5.226) is attained, and $\mathfrak{U}_{\beta,N}(\bar{x})$ is equal to such $\bar{\rho}$ that $\mathfrak{b}(\mathfrak{m}; \bar{\rho}, N) = \beta$. Denoting $V := \mathfrak{b}(\mathfrak{m}; p(\bar{x}), N)$, we have that

$$\begin{aligned}\Pr\big\{p(\bar{x}) < \mathfrak{U}_{\beta,N}(\bar{x})\big\} &= \Pr\Big\{V > \overbrace{\mathfrak{b}(\mathfrak{m}; \bar{\rho}, N)}^{\beta}\Big\} \\ &= 1 - \Pr\{V \leq \beta\} = 1 - \sum_{k=0}^{N} \Pr\big\{V \leq \beta \big| \mathfrak{m} = k\big\} \Pr(\mathfrak{m} = k).\end{aligned}$$

Since

$$\Pr\big\{V \leq \beta \big| \mathfrak{m} = k\big\} = \begin{cases} 1 & \text{if } \mathfrak{b}(k; p(\bar{x}), N) \leq \beta, \\ 0 & \text{otherwise,} \end{cases}$$

and $\Pr(\mathfrak{m} = k) = \binom{N}{k} p(\bar{x})^k (1 - p(\bar{x}))^{N-k}$, it follows that

$$\sum_{k=0}^{N} \Pr\big\{V \leq \beta \big| \mathfrak{m} = k\big\} \Pr(\mathfrak{m} = k) \leq \beta,$$

and hence

$$\Pr\big\{p(\bar{x}) < \mathfrak{U}_{\beta,N}(\bar{x})\big\} \geq 1 - \beta. \tag{5.227}$$

[33] Recall that $z_\beta := \Phi^{-1}(1-\beta) = -\Phi^{-1}(\beta)$, where $\Phi(\cdot)$ is the cdf of the standard normal distribution.

That is, $p(\bar{x}) < \mathfrak{U}_{\beta,N}(\bar{x})$ with probability at least $1-\beta$. Therefore we can take $\mathfrak{U}_{\beta,N}(\bar{x})$ as an upper $(1-\beta)$-confidence bound for $p(\bar{x})$. In particular, if $\mathfrak{m} = 0$, then

$$\mathfrak{U}_{\beta,N}(\bar{x}) = 1 - \beta^{1/N} < N^{-1}\ln(\beta^{-1}).$$

We obtain that if $\mathfrak{U}_{\beta,N}(\bar{x}) \leq \alpha$, then $\bar{x}$ is a feasible solution of the true problem with probability at least $1-\beta$. In that case, we can use $f(\bar{x})$ as an upper bound, with confidence $1-\beta$, for the optimal value ϑ^* of the true problem (5.196). Since this procedure involves only calculations of $C(\bar{x}, \xi^j)$, it can be performed with a large sample size N, and hence feasibility of $\bar{x}$ can be verified with a high accuracy provided that α is not too small.

It also could be noted that the bound given in (5.225), in a sense, is an approximation of the upper bound $\bar{\rho} = \mathfrak{U}_{\beta,N}(\bar{x})$. Indeed, by the CLT the cumulative distribution $\mathfrak{b}(k; \bar{\rho}, N)$ can be approximated by $\Phi(\frac{k-\bar{\rho}N}{\sqrt{N\bar{\rho}(1-\bar{\rho})}})$. Therefore, approximately $\bar{\rho}$ is the solution of the equation $\Phi(\frac{\mathfrak{m}-\rho N}{\sqrt{N\rho(1-\rho)}}) = \beta$, which can be written as

$$\rho = \frac{\mathfrak{m}}{N} + z_\beta\sqrt{\frac{\rho(1-\rho)}{N}}.$$

By approximating ρ in the right-hand side of the above equation by $\mathfrak{m}/N$ we obtain the bound (5.225).

Lower Bounds

It is more tricky to construct a valid lower statistical bound for ϑ^*. One possible approach is to apply a general methodology of the SAA method. (See the discussion at the end of section 5.6.1.) We have that for any $\lambda \geq 0$ the following inequality holds (compare with (5.183)):

$$\vartheta^* \geq \inf_{x\in X}\big[f(x) + \lambda(p(x) - \alpha)\big]. \tag{5.228}$$

We also have that expectation of

$$\hat{v}_N(\lambda) := \inf_{x\in X}\big[f(x) + \lambda(\hat{p}_N(x) - \alpha)\big] \tag{5.229}$$

gives a valid lower bound for the right-hand side of (5.228), and hence for ϑ^*. An unbiased estimate of $\mathbb{E}[\hat{v}_N(\lambda)]$ can be obtained by solving the right-hand-side problem of (5.229) several times and averaging calculated optimal values. Note, however, that there are two difficulties with applying this approach. First, recall that typically the function $\hat{p}_N(x)$ is discontinuous and hence it could be difficult to solve these optimization problems. Second, it may happen that for any choice of $\lambda \geq 0$ the optimal value of the right-hand side of (5.228) is smaller than ϑ^*, i.e., there is a gap between problem (5.196) and its (Lagrangian) dual.

We discuss now an alternative approach to construction statistical lower bounds. For chosen positive integers N and M, and constant $\gamma \in [0, 1)$, let us generate M independent samples $\xi^{1,m}, \ldots, \xi^{N,m}$, $m = 1, \ldots, M$, each of size N, of random vector ξ. For each sample, solve the associated optimization problem

$$\operatorname*{Min}_{x\in X} f(x) \ \text{ s.t. } \ \sum_{j=1}^{N} \mathbf{1}_{(0,\infty)}\big(C(x, \xi^{j,m})\big) \leq \gamma N \tag{5.230}$$

and hence calculate its optimal value $\hat{\vartheta}^m_{\gamma,N}$, $m = 1, \ldots, M$. That is, we solve M times the corresponding SAA problem at the significance level γ. In particular, for $\gamma = 0$, problem (5.230) takes the form

$$\operatorname*{Min}_{x \in X} f(x) \ \text{ s.t. } \ C(x, \xi^{j,m}) \leq 0, \ j = 1, \ldots, N. \tag{5.231}$$

It may happen that problem (5.230) is either infeasible or unbounded from below, in which case we assign its optimal value as $+\infty$ or $-\infty$, respectively. We can view $\hat{\vartheta}^m_{\gamma,N}$, $m = 1, \ldots, M$, as an iid sample of the random variable $\hat{\vartheta}_{\gamma,N}$, where $\hat{\vartheta}_{\gamma,N}$ is the optimal value of the respective SAA problem at significance level γ. Next we rearrange the calculated optimal values in the nondecreasing order, $\hat{\vartheta}^{(1)}_{\gamma,N} \leq \cdots \leq \hat{\vartheta}^{(M)}_{\gamma,N}$; i.e., $\hat{\vartheta}^{(1)}_{\gamma,N}$ is the smallest, $\hat{\vartheta}^{(2)}_{\gamma,N}$ is the second smallest, etc., among the values $\hat{\vartheta}^m_{\gamma,N}$, $m = 1, \ldots, M$. By definition, we choose an integer $L \in \{1, \ldots, M\}$ and use the random quantity $\hat{\vartheta}^{(L)}_{\gamma,N}$ as a lower bound of the true optimal value ϑ^*.

Let us analyze the resulting bounding procedure. Let $\tilde{x} \in X$ be a feasible point of the true problem, i.e.,

$$\Pr\{C(\tilde{x}, \xi) > 0\} \leq \alpha.$$

Since $\sum_{j=1}^N \mathbf{1}_{(0,\infty)}\left(C(\tilde{x}, \xi^{j,m})\right)$ has binomial distribution with probability of success equal to the probability of the event $\{C(\tilde{x}, \xi) > 0\}$, it follows that $\tilde{x}$ is feasible for problem (5.230) with probability at least[34]

$$\sum_{i=0}^{\lfloor \gamma N \rfloor} \binom{N}{i} \alpha^i (1-\alpha)^{N-i} = \mathfrak{b}\left(\lfloor \gamma N \rfloor; \alpha, N\right) =: \theta_N.$$

When $\tilde{x}$ is feasible for (5.230), we of course have that $\hat{\vartheta}^m_{\gamma,N} \leq f(\tilde{x})$. Let $\varepsilon > 0$ be an arbitrary constant and $\tilde{x}$ be a feasible point of the true problem such that $f(\tilde{x}) \leq \vartheta^* + \varepsilon$. Then for every $m \in \{1, \ldots, M\}$ we have

$$\theta := \Pr\left\{\hat{\vartheta}^m_{\gamma,N} \leq \vartheta^* + \varepsilon\right\} \geq \Pr\left\{\hat{\vartheta}^m_{\gamma,N} \leq f(\tilde{x})\right\} \geq \theta_N.$$

Now, in the case of $\hat{\vartheta}^{(L)}_{\gamma,N} > \vartheta^* + \varepsilon$, the corresponding realization of the random sequence $\hat{\vartheta}^1_{\gamma,N}, \ldots, \hat{\vartheta}^M_{\gamma,N}$ contains less than L elements which are less than or equal to $\vartheta^* + \varepsilon$. Since the elements of the sequence are independent, the probability of the latter event is $\mathfrak{b}(L-1; \theta, M)$. Since $\theta \geq \theta_N$, we have that $\mathfrak{b}(L-1; \theta, M) \leq \mathfrak{b}(L-1; \theta_N, M)$. Thus, $\Pr\{\hat{\vartheta}^{(L)}_{\gamma,N} > \vartheta^* + \varepsilon\} \leq \mathfrak{b}(L-1; \theta_N, M)$. Since the resulting inequality is valid for any $\varepsilon > 0$, we arrive at the bound

$$\Pr\left\{\hat{\vartheta}^{(L)}_{\gamma,N} > \vartheta^*\right\} \leq \mathfrak{b}(L-1; \theta_N, M). \tag{5.232}$$

We obtain the following result.

Proposition 5.33. *Given $\beta \in (0, 1)$ and $\gamma \in [0, 1)$, let us choose positive integers M, N, and L in such a way that*

$$\mathfrak{b}(L-1; \theta_N, M) \leq \beta, \tag{5.233}$$

[34] Recall that the notation $\lfloor a \rfloor$ stands for the largest integer less than or equal to $a \in \mathbb{R}$.

where $\theta_N := \mathfrak{b}(\lfloor \gamma N \rfloor; \alpha, N)$. *Then*

$$\Pr\left\{\hat{\vartheta}_{\gamma,N}^{(L)} > \vartheta^*\right\} \leq \beta. \tag{5.234}$$

For given sample sizes N and M, it is better to take the largest integer $L \in \{1, \ldots, M\}$ satisfying condition (5.233). That is, for

$$L^* := \max_{1 \leq L \leq M} \left\{L : \mathfrak{b}(L-1; \theta_N, M) \leq \beta\right\},$$

we have that the random quantity $\hat{\vartheta}_{\gamma,N}^{(L^*)}$ gives a lower bound for the true optimal value ϑ^* with probability at least $1-\beta$. If no $L \in \{1, \ldots, M\}$ satisfying (5.233) exists, the lower bound, by definition, is $-\infty$.

The question arising in connection with the outlined bounding scheme is how to choose M, N, and γ. In the convex case it is advantageous to take $\gamma = 0$, since then we need to solve convex problems (5.231), rather than combinatorial problems (5.230). Note that for $\gamma = 0$, we have that $\theta_N = (1-\alpha)^N$ and the bound (5.233) takes the form

$$\sum_{k=0}^{L-1} \binom{M}{k} (1-\alpha)^{Nk} [1-(1-\alpha)^N]^{M-k} \leq \beta. \tag{5.235}$$

Suppose that N and $\gamma \geq 0$ are given (fixed). Then the larger M is, the better. We can view $\hat{\vartheta}_{\gamma,N}^m$, $m = 1, \ldots, M$, as a random sample from the distribution of the random variable $\hat{\vartheta}_N$ with $\hat{\vartheta}_N$ being the optimal value of the corresponding SAA problem of the form (5.230). It follows from the definition that L^* is equal to the (lower) β-quantile of the binomial distribution $B(\theta_N, M)$. By the CLT we have that

$$\lim_{M \to \infty} \frac{L^* - \theta_N M}{\sqrt{M\theta_N(1-\theta_N)}} = \Phi^{-1}(\beta),$$

and L^*/M tends to θ_N as $M \to \infty$. It follows that the lower bound $\hat{\vartheta}_{\gamma,N}^{(L^*)}$ converges to the θ_N-quantile of the distribution of $\hat{\vartheta}_N$ as $M \to \infty$.

In reality, however, M is bounded by the computational effort required to solve M problems of the form (5.230). Note that the effort per problem is larger the larger the sample size N. For $L = 1$ (which is the smallest value of L) and $\gamma = 0$, the left-hand side of (5.235) is equal to $[1-(1-\alpha)^N]^M$. Note that $(1-\alpha)^N \approx e^{-\alpha N}$ for small $\alpha > 0$. Therefore, if αN is large, one will need a very large M to make $[1-(1-\alpha)^N]^M$ smaller than, say, $\beta = 0.01$, and hence to get a meaningful lower bound. For example, for $\alpha N = 7$ we have that $e^{-\alpha N} = 0.0009$, and we will need $M > 5000$ to make $[1-(1-\alpha)^N]^M$ smaller than 0.01. Therefore, for $\gamma = 0$ it is recommended to take N not larger than, say, $2/\alpha$.

5.8 SAA Method Applied to Multistage Stochastic Programming

Consider a multistage stochastic programming problem, in the general form (3.1), driven by the random data process $\xi_1, \xi_2, \ldots, \xi_T$. The exact meaning of this formulation was

discussed in section 3.1.1. In this section we discuss application of the SAA method to such multistage problems.

Consider the following sampling scheme. Generate a sample $\xi_2^1, \ldots, \xi_2^{N_1}$ of N_1 realizations of random vector ξ_2. Conditional on each ξ_2^i, $i = 1, \ldots, N_1$, generate a random sample ξ_3^{ij}, $j = 1, \ldots, N_2$, of N_2 realizations of ξ_3 according to conditional distribution of ξ_3 given $\xi_2 = \xi_2^i$. Conditional on each ξ_3^{ij}, generate a random sample of size N_3 of ξ_4 conditional on $\xi_3 = \xi_3^{ij}$, and so on for later stages. (Although we do not consider such possibility here, it is also possible to generate at each stage conditional samples of different sizes.) In that way we generate a scenario tree with $N = \prod_{t=1}^{T-1} N_t$ number of scenarios each taken with equal probability $1/N$. We refer to this scheme as *conditional sampling*. Unless stated otherwise,[35] we assume that, at the first stage, the sample $\xi_2^1, \ldots, \xi_2^{N_1}$ is iid and the following samples, at each stage $t = 2, \ldots, T-1$, are conditionally iid. If, moreover, all conditional samples at each stage are independent of each other, we refer to such conditional sampling as the *independent conditional sampling*. The multistage stochastic programming problem induced by the original problem (3.1) on the scenario tree generated by conditional sampling is viewed as the sample average approximation (SAA) of the "true" problem (3.1).

It could be noted that in case of stagewise independent process $\xi_1, \ldots, \xi_T$, the independent conditional sampling destroys the stagewise independence structure of the original process. This is because at each stage conditional samples are independent of each other and hence are different. In the stagewise independence case, an alternative approach is to use the same sample at each stage. That is, independent of each other, random samples $\xi_t^1, \ldots, \xi_t^{N_{t-1}}$ of respective ξ_t, $t = 2, \ldots, T$, are generated and the corresponding scenario tree is constructed by connecting every ancestor node at stage $t-1$ with the same set of children nodes $\xi_t^1, \ldots, \xi_t^{N_{t-1}}$. In that way stagewise independence is preserved in the scenario tree generated by conditional sampling. We refer to this sampling scheme as the *identical conditional sampling*.

5.8.1 Statistical Properties of Multistage SAA Estimators

Similar to two-stage programming, it makes sense to discuss convergence of the optimal value and first-stage solutions of multistage SAA problems to their true counterparts as sample sizes $N_1, \ldots, N_{T-1}$ tend to infinity. We denote $\mathcal{N} := \{N_1, \ldots, N_{T-1}\}$ and by ϑ^* and $\hat{\vartheta}_{\mathcal{N}}$ the optimal values of the true and the corresponding SAA multistage programs, respectively.

In order to simplify the presentation let us consider now three-stage stochastic programs, i.e., $T = 3$. In that case, conditional sampling consists of sample ξ_2^i, $i = 1, \ldots, N_1$, of ξ_2 and for each $i = 1, \ldots, N_1$ of conditional samples ξ_3^{ij}, $j = 1, \ldots, N_2$, of ξ_3 given $\xi_2 = \xi_2^i$. Let us write dynamic programming equations for the true problem. We have

$$Q_3(x_2, \xi_3) = \inf_{x_3 \in \mathcal{X}_3(x_2, \xi_3)} f_3(x_3, \xi_3), \tag{5.236}$$

$$Q_2(x_1, \xi_2) = \inf_{x_2 \in \mathcal{X}_2(x_1, \xi_2)} \left\{ f_2(x_2, \xi_2) + \mathbb{E}\left[Q_3(x_2, \xi_3) \middle| \xi_2 \right] \right\}, \tag{5.237}$$

[35] It is also possible to employ quasi–Monte Carlo sampling in constructing conditional sampling. In some situations this may reduce variability of the corresponding SAA estimators. In the following analysis we assume independence in order to simplify statistical analysis.

and at the first stage we solve the problem

$$\operatorname*{Min}_{x_1 \in \mathcal{X}_1} \left\{ f_1(x_1) + \mathbb{E}\left[Q_2(x_1, \xi_2) \right] \right\}. \tag{5.238}$$

If we could calculate values $Q_2(x_1, \xi_2)$, we could approximate problem (5.238) by the sample average problem

$$\operatorname*{Min}_{x_1 \in \mathcal{X}_1} \left\{ \hat{f}_{N_1}(x_1) := f_1(x_1) + \tfrac{1}{N_1} \textstyle\sum_{i=1}^{N_1} Q_2(x_1, \xi_2^i) \right\}. \tag{5.239}$$

However, values $Q_2(x_1, \xi_2^i)$ are not given explicitly and are approximated by

$$\hat{Q}_{2,N_2}(x_1, \xi_2^i) := \inf_{x_2 \in \mathcal{X}_2(x_1, \xi_2^i)} \left\{ f_2(x_2, \xi_2^i) + \tfrac{1}{N_2} \textstyle\sum_{j=1}^{N_2} Q_3(x_2, \xi_3^{ij}) \right\}, \tag{5.240}$$

$i = 1, \ldots, N_1$. That is, the SAA method approximates the first stage problem (5.238) by the problem

$$\operatorname*{Min}_{x_1 \in \mathcal{X}_1} \left\{ \tilde{f}_{N_1,N_2}(x_1) := f_1(x_1) + \tfrac{1}{N_1} \textstyle\sum_{i=1}^{N_1} \hat{Q}_{2,N_2}(x_1, \xi_2^i) \right\}. \tag{5.241}$$

In order to verify consistency of the SAA estimators, obtained by solving problem (5.241), we need to show that $\tilde{f}_{N_1,N_2}(x_1)$ converges to $f_1(x_1) + \mathbb{E}\left[Q_2(x_1, \xi_2) \right]$ w.p. 1 uniformly on any compact subset X of $\mathcal{X}_1$. (Compare with the analysis of section 5.1.1.) That is, we need to show that

$$\lim_{N_1, N_2 \to \infty} \sup_{x_1 \in X} \left| \tfrac{1}{N_1} \textstyle\sum_{i=1}^{N_1} \hat{Q}_{2,N_2}(x_1, \xi_2^i) - \mathbb{E}\left[Q_2(x_1, \xi_2) \right] \right| = 0 \text{ w.p. 1.} \tag{5.242}$$

For that it suffices to show that

$$\lim_{N_1 \to \infty} \sup_{x_1 \in X} \left| \tfrac{1}{N_1} \textstyle\sum_{i=1}^{N_1} Q_2(x_1, \xi_2^i) - \mathbb{E}\left[Q_2(x_1, \xi_2) \right] \right| = 0 \text{ w.p. 1} \tag{5.243}$$

and

$$\lim_{N_1, N_2 \to \infty} \sup_{x_1 \in X} \left| \tfrac{1}{N_1} \textstyle\sum_{i=1}^{N_1} \hat{Q}_{2,N_2}(x_1, \xi_2^i) - \tfrac{1}{N_1} \textstyle\sum_{i=1}^{N_1} Q_2(x_1, \xi_2^i) \right| = 0 \text{ w.p. 1.} \tag{5.244}$$

Condition (5.243) can be verified by applying a version of the uniform Law of Large Numbers (see section 7.2.5). Condition (5.244) is more involved. Of course, we have that

$$\begin{aligned} &\sup_{x_1 \in X} \left| \tfrac{1}{N_1} \textstyle\sum_{i=1}^{N_1} \hat{Q}_{2,N_2}(x_1, \xi_2^i) - \tfrac{1}{N_1} \textstyle\sum_{i=1}^{N_1} Q_2(x_1, \xi_2^i) \right| \\ &\leq \tfrac{1}{N_1} \textstyle\sum_{i=1}^{N_1} \sup_{x_1 \in X} \left| \hat{Q}_{2,N_2}(x_1, \xi_2^i) - Q_2(x_1, \xi_2^i) \right|, \end{aligned}$$

and hence condition (5.244) holds if $\hat{Q}_{2,N_2}(x_1, \xi_2^i)$ converges to $Q_2(x_1, \xi_2^i)$ w.p. 1 as $N_2 \to \infty$ in a certain uniform way. Unfortunately an exact mathematical analysis of such condition could be quite involved. The analysis simplifies considerably if the underline random process is stagewise independent. In the present case this means that random vectors ξ_2 and ξ_3 are independent. In that case distribution of random sample ξ_3^{ij}, $j = 1, \ldots, N_2$, does not depend on i (in both sampling schemes whether samples ξ_3^{ij} are the same for all $i = 1, \ldots, N_1$, or independent of each other), and we can apply Theorem 7.48 to establish that,

under mild regularity conditions, $\frac{1}{N_2}\sum_{j=1}^{N_2} Q_3(x_2, \xi_3^{ij})$ converges to $\mathbb{E}[Q_3(x_2, \xi_3)]$ w.p. 1 as $N_2 \to \infty$ uniformly in x_2 on any compact subset of $\mathbb{R}^{n_2}$. With an additional assumptions about mapping $\mathcal{X}_2(x_1, \xi_2)$, it is possible to verify the required uniform type convergence of $\hat{Q}_{2,N_2}(x_1, \xi_2^i)$ to $Q_2(x_1, \xi_2^i)$. Again a precise mathematical analysis is quite technical and will be left out. Instead, in section 5.8.2 we discuss a uniform exponential convergence of the sample average function $\tilde{f}_{N_1,N_2}(x_1)$ to the objective function $f_1(x_1) + \mathbb{E}[Q_2(x_1, \xi_2)]$ of the true problem.

Let us make the following observations. By increasing sample sizes $N_1, \ldots, N_{T-1}$ of conditional sampling, we eventually reconstruct the scenario tree structure of the original multistage problem. Therefore it should be expected that in the limit, as these sample sizes tend (simultaneously) to infinity, the corresponding SAA estimators of the optimal value and first-stage solutions are consistent, i.e., converge w.p. 1 to their true counterparts. And, indeed, this can be shown under certain regularity conditions. However, consistency alone does not justify the SAA method since in reality sample sizes are always finite and are constrained by available computational resources. Similar to the two-stage case we have here that (for minimization problems)

$$\vartheta^* \geq \mathbb{E}[\hat{\vartheta}_{\mathcal{N}}]. \tag{5.245}$$

That is, the SAA optimal value $\hat{\vartheta}_{\mathcal{N}}$ is a downward biased estimator of the true optimal value ϑ^*.

Suppose now that the data process $\xi_1, \ldots, \xi_T$ is stagewise independent. As discussed above, in that case it is possible to use two different approaches to conditional sampling, namely, to use at every stage independent or the same samples for every ancestor node at the previous stage. These approaches were referred to as the independent and identical conditional samplings, respectively. Consider, for instance, the three-stage stochastic programming problem (5.236)–(5.238). In the second approach of identical conditional sampling we have sample ξ_2^i, $i = 1, \ldots, N_1$, of ξ_2 and sample ξ_3^j, $j = 1, \ldots, N_2$, of ξ_3 independent of ξ_2^i. In that case formula (5.240) takes the form

$$\hat{Q}_{2,N_2}(x_1, \xi_2^i) = \inf_{x_2 \in \mathcal{X}_2(x_1, \xi_2^i)} \left\{ f_2(x_2, \xi_2^i) + \tfrac{1}{N_2}\textstyle\sum_{j=1}^{N_2} Q_3(x_2, \xi_3^j) \right\}. \tag{5.246}$$

Because of independence of ξ_2 and ξ_3 we have that conditional distribution of ξ_3 given ξ_2 is the same as its unconditional distribution, and hence in both sampling approaches $\hat{Q}_{2,N_2}(x_1, \xi_2^i)$ has the same distribution independent of i. Therefore in both sampling schemes $\frac{1}{N_1}\sum_{i=1}^{N_1} \hat{Q}_{2,N_2}(x_1, \xi_2^i)$ has the same expectation, and hence we may expect that in both cases the estimator $\hat{\vartheta}_{\mathcal{N}}$ has a similar bias. Variance of $\hat{\vartheta}_{\mathcal{N}}$, however, could be quite different. In the case of independent conditional sampling we have that $\hat{Q}_{2,N_2}(x_1, \xi_2^i)$, $i = 1, \ldots, N_1$, are independent of each other, and hence

$$\mathbb{V}\mathrm{ar}\left[\tfrac{1}{N_1}\textstyle\sum_{i=1}^{N_1} \hat{Q}_{2,N_2}(x_1, \xi_2^i)\right] = \tfrac{1}{N_1}\mathbb{V}\mathrm{ar}\left[\hat{Q}_{2,N_2}(x_1, \xi_2^i)\right]. \tag{5.247}$$

On the other hand, in the case of identical conditional sampling the right-hand side of (5.246) has the same component $\frac{1}{N_2}\sum_{j=1}^{N_2} Q_3(x_2, \xi_3^j)$ for all $i = 1, \ldots, N_1$. Consequently, $\hat{Q}_{2,N_2}(x_1, \xi_2^i)$ would tend to be positively correlated for different values of i, and as a result

$\hat{\vartheta}_{\mathcal{N}}$ will have a higher variance than in the case of independent conditional sampling. Therefore, from a statistical point of view it is advantageous to use the independent conditional sampling.

Example 5.34 (Portfolio Selection). Consider the example of multistage portfolio selection discussed in section 1.4.2. Suppose for the moment that the problem has three stages, $t = 0, 1, 2$. In the SAA approach we generate sample ξ_1^i, $i = 1, \ldots, N_0$, of returns at stage $t = 1$, and conditional samples ξ_2^{ij}, $j = 1, \ldots, N_1$, of returns at stage $t = 2$. The dynamic programming equations for the SAA problem can be written as follows (see (1.50)–(1.52)). At stage $t = 1$ for $i = 1, \ldots, N_0$, we have

$$\hat{Q}_{1,N_1}(W_1, \xi_1^i) = \sup_{x_1 \geq 0} \left\{ \tfrac{1}{N_1} \textstyle\sum_{j=1}^{N_1} U\big((\xi_2^{ij})^{\mathsf{T}} x_1\big) : e^{\mathsf{T}} x_1 = W_1 \right\}, \tag{5.248}$$

where $e \in \mathbb{R}^n$ is vector of ones, and at stage $t = 0$ we solve the problem

$$\operatorname*{Max}_{x_0 \geq 0} \frac{1}{N_0} \sum_{i=1}^{N_0} \hat{Q}_{1,N_1}\big((\xi_1^i)^{\mathsf{T}} x_0, \xi_1^i\big) \ \text{ s.t. } \ e^{\mathsf{T}} x_0 = W_0. \tag{5.249}$$

Now let $U(W) := \ln W$ be the logarithmic utility function. Suppose that the data process is stagewise independent. Then the optimal value ϑ^* of the true problem is (see (1.58))

$$\vartheta^* = \ln W_0 + \sum_{t=0}^{T-1} \nu_t, \tag{5.250}$$

where ν_t is the optimal value of the problem

$$\operatorname*{Max}_{x_t \geq 0} \mathbb{E}\left[\ln\left(\xi_{t+1}^{\mathsf{T}} x_t\right)\right] \ \text{ s.t. } \ e^{\mathsf{T}} x_t = 1. \tag{5.251}$$

Let the SAA method be applied with the identical conditional sampling, with respective sample ξ_t^j, $j = 1, \ldots, N_{t-1}$, of ξ_t, $t = 1, \ldots, T$. In that case, the corresponding SAA problem is also stagewise independent and the optimal value of the SAA problem

$$\hat{\vartheta}_{\mathcal{N}} = \ln W_0 + \sum_{t=0}^{T-1} \hat{\nu}_{t,N_t}, \tag{5.252}$$

where $\hat{\nu}_{t,N_t}$ is the optimal value of the problem

$$\operatorname*{Max}_{x_t \geq 0} \frac{1}{N_t} \sum_{j=1}^{N_t} \ln\left((\xi_{t+1}^j)^{\mathsf{T}} x_t\right) \ \text{ s.t. } \ e^{\mathsf{T}} x_t = 1. \tag{5.253}$$

We can view $\hat{\nu}_{t,N_t}$ as an SAA estimator of ν_t. Since here we solve a maximization rather than a minimization problem, $\hat{\nu}_{t,N_t}$ is an upward biased estimator of ν_t, i.e., $\mathbb{E}[\hat{\nu}_{t,N_t}] \geq \nu_t$. We also have that $\mathbb{E}[\hat{\vartheta}_{\mathcal{N}}] = \ln W_0 + \sum_{t=0}^{T-1} \mathbb{E}[\hat{\nu}_{t,N_t}]$, and hence

$$\mathbb{E}[\hat{\vartheta}_{\mathcal{N}}] - \vartheta^* = \sum_{t=0}^{T-1} \big(\mathbb{E}[\hat{\nu}_{t,N_t}] - \nu_t\big). \tag{5.254}$$

That is, for the logarithmic utility function and identical conditional sampling, bias of the SAA estimator of the optimal value grows additively with increase of the number of stages. Also because the samples at different stages are independent of each other, we have that

$$\mathbb{V}\mathrm{ar}[\hat{\vartheta}_{\mathcal{N}}] = \sum_{t=0}^{T-1} \mathbb{V}\mathrm{ar}[\hat{v}_{t,N_t}]. \tag{5.255}$$

Let now $U(W) := W^\gamma$, with $\gamma \in (0, 1]$, be the power utility function and suppose that the data process is stagewise independent. Then (see (1.61))

$$\vartheta^* = W_0^\gamma \prod_{t=0}^{T-1} \eta_t, \tag{5.256}$$

where η_t is the optimal value of problem

$$\operatorname*{Max}_{x_t \geq 0} \mathbb{E}\left[\left(\xi_{t+1}^{\mathsf{T}} x_t\right)^\gamma\right] \quad \text{s.t. } e^{\mathsf{T}} x_t = 1. \tag{5.257}$$

For the corresponding SAA method with the identical conditional sampling, we have that

$$\hat{\vartheta}_{\mathcal{N}} = W_0^\gamma \prod_{t=0}^{T-1} \hat{\eta}_{t,N_t}, \tag{5.258}$$

where $\hat{\eta}_{t,N_t}$ is the optimal value of problem

$$\operatorname*{Max}_{x_t \geq 0} \frac{1}{N_t} \sum_{j=1}^{N_t} \left((\xi_{t+1}^j)^{\mathsf{T}} x_t\right)^\gamma \quad \text{s.t. } e^{\mathsf{T}} x_t = 1. \tag{5.259}$$

Because of the independence of the samples, and hence independence of $\hat{\eta}_{t,N_t}$, we can write $\mathbb{E}[\hat{\vartheta}_{\mathcal{N}}] = W_0^\gamma \prod_{t=0}^{T-1} \mathbb{E}[\hat{\eta}_{t,N_t}]$, and hence

$$\mathbb{E}[\hat{\vartheta}_{\mathcal{N}}] = \vartheta^* \prod_{t=0}^{T-1} (1 + \beta_{t,N_t}), \tag{5.260}$$

where $\beta_{t,N_t} := \frac{\mathbb{E}[\hat{\eta}_{t,N_t}] - \eta_t}{\eta_t}$ is the relative bias of $\hat{\eta}_{t,N_t}$. That is, bias of $\hat{\vartheta}_{\mathcal{N}}$ grows with increase of the number of stages in a *multiplicative* way. In particular, if the relative biases β_{t,N_t} are constant, then bias of $\hat{\vartheta}_{\mathcal{N}}$ grows *exponentially* fast with increase of the number of stages. ∎

Statistical Validation Analysis

By (5.245) we have that the optimal value $\hat{\vartheta}_{\mathcal{N}}$ of SAA problem gives a valid statistical lower bound for the optimal value ϑ^*. Therefore, in order to construct a lower bound for ϑ^* one can proceed exactly in the same way as it was discussed in section 5.6.1. Unfortunately, typically the bias and variance of $\hat{\vartheta}_{\mathcal{N}}$ grow fast with increase of the number of stages, which

makes the corresponding statistical lower bounds quite inaccurate already for a mild number of stages.

In order to construct an upper bound we proceed as follows. Let $\boldsymbol{x}_t(\xi_{[t]})$ be a feasible policy. Recall that a policy is feasible if it satisfies the feasibility constraints (3.2). Since the multistage problem can be formulated as the minimization problem (3.3) we have that

$$\mathbb{E}\big[f_1(x_1) + f_2(\boldsymbol{x}_2(\xi_{[2]}), \xi_2) + \cdots + f_T\left(\boldsymbol{x}_T(\xi_{[T]}), \xi_T\right)\big] \geq \vartheta^*, \tag{5.261}$$

and equality in (5.261) holds iff the policy $\boldsymbol{x}_t(\xi_{[t]})$ is optimal. The expectation in the left-hand side of (5.261) can be estimated in a straightforward way. That is, generate random sample $\xi_1^j, \ldots, \xi_T^j$, $j = 1, \ldots, N$, of N realizations (scenarios) of the random data process $\xi_1, \ldots, \xi_T$ and estimate this expectation by the average

$$\frac{1}{N}\sum_{j=1}^{N}\Big[f_1(x_1) + f_2\big(\boldsymbol{x}_2(\xi_{[2]}^j), \xi_2^j\big) + \cdots + f_T\big(\boldsymbol{x}_T(\xi_{[T]}^j), \xi_T^j\big)\Big]. \tag{5.262}$$

Note that in order to construct the above estimator we do not need to generate a scenario tree, say, by conditional sampling; we only need to generate a sample of single scenarios of the data process. The above estimator (5.262) is an unbiased estimator of the expectation in the left-hand side of (5.261) and hence is a valid statistical upper bound for ϑ^*. Of course, the quality of this upper bound depends on a successful choice of the feasible policy, i.e., on how small the optimality gap is between the left- and right-hand sides of (5.261). It also depends on variability of the estimator (5.262), which unfortunately often grows fast with increase of the number of stages.

We also may address the problem of validating a given feasible first-stage solution $\bar{x}_1 \in \mathcal{X}_1$. The value of the multistage problem at $\bar{x}_1$ is given by the optimal value of the problem

$$\begin{array}{ll} \underset{\boldsymbol{x}_2,\ldots,\boldsymbol{x}_T}{\operatorname{Min}} & f_1(\bar{x}_1) + \mathbb{E}\big[f_2(\boldsymbol{x}_2(\xi_{[2]}), \xi_2) + \cdots + f_T\left(\boldsymbol{x}_T(\xi_{[T]}), \xi_T\right)\big] \\ \text{s.t.} & \boldsymbol{x}_t(\xi_{[t]}) \in \mathcal{X}_t(\boldsymbol{x}_{t-1}(\xi_{[t-1]}), \xi_t), \ t = 2, \ldots, T. \end{array} \tag{5.263}$$

Recall that the optimization in (5.263) is performed over feasible policies. That is, in order to validate $\bar{x}_1$ we basically need to solve the corresponding $T-1$ stage problems. Therefore, for $T > 2$, validation of $\bar{x}_1$ can be almost as difficult as solving the original problem.

5.8.2 Complexity Estimates of Multistage Programs

In order to compute value of two-stage stochastic program $\min_{x\in X}\mathbb{E}[F(x,\xi)]$, where $F(x,\xi)$ is the optimal value of the corresponding second-stage problem, at a feasible point $\bar{x} \in X$ we need to calculate the expectation $\mathbb{E}[F(\bar{x},\xi)]$. This, in turn, involves two difficulties. First, the objective value $F(\bar{x},\xi)$ is not given explicitly; its calculation requires solution of the associated second-stage optimization problem. Second, the multivariate integral $\mathbb{E}[F(\bar{x},\xi)]$ cannot be evaluated with a high accuracy even for moderate values of dimension d of the random data vector ξ. Monte Carlo techniques allow us to evaluate $\mathbb{E}[F(\bar{x},\xi)]$ with accuracy $\varepsilon > 0$ by employing samples of size $N = O(\varepsilon^{-2})$. The required sample size N gives, in a sense, an estimate of complexity of evaluation of $\mathbb{E}[F(\bar{x},\xi)]$ since this is how many times we will need to solve the corresponding second-stage problem. It is

remarkable that in order to solve the two-stage stochastic program with accuracy $\varepsilon > 0$, say, by the SAA method, we need a sample size basically of the same order $N = O(\varepsilon^{-2})$. These complexity estimates were analyzed in detail in section 5.3. Two basic conditions required for such analysis are that the problem has relatively complete recourse and that for given x and ξ the optimal value $F(x, \xi)$ of the second-stage problem can be calculated with a high accuracy.

In this section we discuss analogous estimates of complexity of the SAA method applied to multistage stochastic programming problems. From the point of view of the SAA method it is natural to evaluate complexity of a multistage stochastic program in terms of the total number of scenarios required to find a first-stage solution with a given accuracy $\varepsilon > 0$.

In order to simplify the presentation we consider three-stage stochastic programs, say, of the form (5.236)–(5.238). Assume that for every $x_1 \in \mathcal{X}_1$ the expectation $\mathbb{E}[Q_2(x_1, \xi_2)]$ is well defined and finite valued. In particular, this assumption implies that the problem has relatively complete recourse. Let us look at the problem of computing value of the first-stage problem (5.238) at a feasible point $\bar{x}_1 \in \mathcal{X}_1$. Apart from the problem of evaluating the expectation $\mathbb{E}[Q_2(\bar{x}_1, \xi_2)]$, we also face here the problem of computing $Q_2(\bar{x}_1, \xi_2)$ for different realizations of random vector ξ_2. For that we need to solve the two-stage stochastic programming problem given in the right-hand side of (5.237). As discussed, in order to evaluate $Q_2(\bar{x}_1, \xi_2)$ with accuracy $\varepsilon > 0$ by solving the corresponding SAA problem, given in the right-hand side of (5.240), we also need a sample of size $N_2 = O(\varepsilon^{-2})$. Recall that the total number of scenarios involved in evaluation of the sample average $\hat{f}_{N_1,N_2}(\bar{x}_1)$, defined in (5.241), is $N = N_1 N_2$. Therefore we will need $N = O(\varepsilon^{-4})$ scenarios just to compute value of the first-stage problem at a given feasible point with accuracy ε by the SAA method. This indicates that complexity of the SAA method, applied to multistage stochastic programs, grows exponentially with increase of the number of stages.

We now discuss in detail the sample size estimates of the three-stage SAA program (5.239)–(5.241). For the sake of simplicity we assume that the data process is stagewise independent, i.e., random vectors ξ_2 and ξ_3 are independent. Also, similar to assumptions (M1)–(M5) of section 5.3, let us make the following assumptions:

(M′1) For every $x_1 \in \mathcal{X}_1$ the expectation $\mathbb{E}[Q_2(x_1, \xi_2)]$ is well defined and finite valued.

(M′2) The random vectors ξ_2 and ξ_3 are independent.

(M′3) The set $\mathcal{X}_1$ has finite diameter D_1.

(M′4) There is a constant $L_1 > 0$ such that

$$\left|Q_2(x_1', \xi_2) - Q_2(x_1, \xi_2)\right| \leq L_1 \|x_1' - x_1\| \tag{5.264}$$

for all $x_1', x_1 \in \mathcal{X}_1$ and a.e. ξ_2.

(M′5) There exists a constant $\sigma_1 > 0$ such that for any $x_1 \in \mathcal{X}_1$ it holds that

$$M_{1,x_1}(t) \leq \exp\left\{\sigma_1^2 t^2/2\right\}, \quad \forall t \in \mathbb{R}, \tag{5.265}$$

where $M_{1,x_1}(t)$ is the moment-generating function of $Q_2(x_1, \xi_2) - \mathbb{E}[Q_2(x_1, \xi_2)]$.

(M′6) There is a set $\mathcal{C}$ of finite diameter D_2 such that for every $x_1 \in \mathcal{X}_1$ and a.e. ξ_2, the set $\mathcal{X}_2(x_1, \xi_2)$ is contained in $\mathcal{C}$.

(M′7) There is a constant $L_2 > 0$ such that

$$\left|Q_3(x_2', \xi_3) - Q_3(x_2, \xi_3)\right| \leq L_2 \|x_2' - x_2\| \tag{5.266}$$

for all $x_2', x_2 \in \mathcal{C}$ and a.e. ξ_3.

(M′8) There exists a constant $\sigma_2 > 0$ such that for any $x_2 \in \mathcal{X}_2(x_1, \xi_2)$ and all $x_1 \in \mathcal{X}_1$ and a.e. ξ_2 it holds that

$$M_{2,x_2}(t) \leq \exp\left\{\sigma_2^2 t^2/2\right\}, \quad \forall t \in \mathbb{R}, \tag{5.267}$$

where $M_{2,x_2}(t)$ is the moment-generating function of $Q_3(x_2, \xi_3) - \mathbb{E}[Q_3(x_2, \xi_3)]$.

Theorem 5.35. *Under assumptions* (M′1)–(M′8) *and for* $\varepsilon > 0$ *and* $\alpha \in (0, 1)$*, and the sample sizes* N_1 *and* N_2 *(using either independent or identical conditional sampling schemes) satisfying*

$$\left[\tfrac{O(1)D_1L_1}{\varepsilon}\right]^{n_1} \exp\left\{-\tfrac{O(1)N_1\varepsilon^2}{\sigma_1^2}\right\} + \left[\tfrac{O(1)D_2L_2}{\varepsilon}\right]^{n_2} \exp\left\{-\tfrac{O(1)N_2\varepsilon^2}{\sigma_2^2}\right\} \leq \alpha, \tag{5.268}$$

we have that any $\varepsilon/2$*-optimal solution of the SAA problem* (5.241) *is an* ε*-optimal solution of the first stage* (5.238) *of the true problem with probability at least* $1 - \alpha$*.*

Proof. The proof of this theorem is based on the uniform exponential bound of Theorem 7.67. Let us sketch the arguments. Assume that the conditional sampling is identical. We have that for every $x_1 \in \mathcal{X}_1$ and $i = 1, \ldots, N_1$,

$$\left|\hat{Q}_{2,N_2}(x_1, \xi_2^i) - Q_2(x_1, \xi_2^i)\right| \leq \sup_{x_2 \in \mathcal{C}} \left|\tfrac{1}{N_2}\textstyle\sum_{j=1}^{N_2} Q_3(x_2, \xi_3^j) - \mathbb{E}[Q_3(x_2, \xi_3)]\right|,$$

where $\mathcal{C}$ is the set postulated in assumption (M′6). Consequently,

$$\begin{aligned} &\sup_{x_1 \in \mathcal{X}_1} \left|\tfrac{1}{N_1}\textstyle\sum_{i=1}^{N_1} \hat{Q}_{2,N_2}(x_1, \xi_2^i) - \tfrac{1}{N_1}\sum_{i=1}^{N_1} Q_2(x_1, \xi_2^i)\right| \\ &\quad \leq \tfrac{1}{N_1}\textstyle\sum_{i=1}^{N_1} \sup_{x_1 \in \mathcal{X}_1} \left|\hat{Q}_{2,N_2}(x_1, \xi_2^i) - Q_2(x_1, \xi_2^i)\right| \\ &\quad \leq \sup_{x_2 \in \mathcal{C}} \left|\tfrac{1}{N_2}\textstyle\sum_{j=1}^{N_2} Q_3(x_2, \xi_3^j) - \mathbb{E}[Q_3(x_2, \xi_3)]\right|. \end{aligned} \tag{5.269}$$

By the uniform exponential bound (7.217) we have that

$$\begin{aligned} &\Pr\left\{\sup_{x_2 \in \mathcal{C}} \left|\tfrac{1}{N_2}\textstyle\sum_{j=1}^{N_2} Q_3(x_2, \xi_3^j) - \mathbb{E}[Q_3(x_2, \xi_3)]\right| > \varepsilon/2\right\} \\ &\qquad \leq \left[\tfrac{O(1)D_2L_2}{\varepsilon}\right]^{n_2} \exp\left\{-\tfrac{O(1)N_2\varepsilon^2}{\sigma_2^2}\right\}, \end{aligned} \tag{5.270}$$

and hence

$$\begin{aligned} &\Pr\left\{\sup_{x_1 \in \mathcal{X}_1} \left|\tfrac{1}{N_1}\textstyle\sum_{i=1}^{N_1} \hat{Q}_{2,N_2}(x_1, \xi_2^i) - \tfrac{1}{N_1}\sum_{i=1}^{N_1} Q_2(x_1, \xi_2^i)\right| > \varepsilon/2\right\} \\ &\qquad \leq \left[\tfrac{O(1)D_2L_2}{\varepsilon}\right]^{n_2} \exp\left\{-\tfrac{O(1)N_2\varepsilon^2}{\sigma_2^2}\right\}. \end{aligned} \tag{5.271}$$

By the uniform exponential bound (7.217) we also have that

$$\begin{array}{c}\Pr\left\{\sup_{x_1\in\mathcal{X}_1}\left|\frac{1}{N_1}\sum_{i=1}^{N_1}Q_2(x_1,\xi_2^i)-\mathbb{E}[Q_2(x_1,\xi_2)]\right|>\varepsilon/2\right\}\\ \leq\left[\frac{O(1)D_1L_1}{\varepsilon}\right]^{n_1}\exp\left\{-\frac{O(1)N_1\varepsilon^2}{\sigma_1^2}\right\}.\end{array}\tag{5.272}$$

Let us observe that if Z_1, Z_2 are random variables, then

$$\Pr(Z_1+Z_2>\varepsilon)\leq\Pr(Z_1>\varepsilon/2)+\Pr(Z_2>\varepsilon/2).$$

Therefore it follows from (5.271) and (5.271) that

$$\begin{array}{c}\Pr\left\{\sup_{x_1\in\mathcal{X}_1}\left|\tilde{f}_{N_1,N_2}(x_1)-f_1(x_1)-\mathbb{E}[Q_2(x_1,\xi_2)]\right|>\varepsilon\right\}\\ \leq\left[\frac{O(1)D_1L_1}{\varepsilon}\right]^{n_1}\exp\left\{-\frac{O(1)N_1\varepsilon^2}{\sigma_1^2}\right\}+\left[\frac{O(1)D_2L_2}{\varepsilon}\right]^{n_2}\exp\left\{-\frac{O(1)N_2\varepsilon^2}{\sigma_2^2}\right\},\end{array}\tag{5.273}$$

which implies the assertion of the theorem.

In the case of the independent conditional sampling the proof can be completed in a similar way. □

Remark 17. We have, of course, that

$$\left|\hat{\vartheta}_{\mathcal{N}}-\vartheta^*\right|\leq\sup_{x_1\in\mathcal{X}_1}\left|\tilde{f}_{N_1,N_2}(x_1)-f_1(x_1)-\mathbb{E}[Q_2(x_1,\xi_2)]\right|.\tag{5.274}$$

Therefore bound (5.273) also implies that

$$\begin{array}{rcl}\Pr\left\{\left|\hat{\vartheta}_{\mathcal{N}}-\vartheta^*\right|>\varepsilon\right\} & \leq & \left[\frac{O(1)D_1L_1}{\varepsilon}\right]^{n_1}\exp\left\{-\frac{O(1)N_1\varepsilon^2}{\sigma_1^2}\right\}\\ & & +\left[\frac{O(1)D_2L_2}{\varepsilon}\right]^{n_2}\exp\left\{-\frac{O(1)N_2\varepsilon^2}{\sigma_2^2}\right\}.\end{array}\tag{5.275}$$

In particular, suppose that $N_1=N_2$. Then for

$$n:=\max\{n_1,n_2\},\ L:=\max\{L_1,L_2\},\ D:=\max\{D_1,D_2\},\ \sigma:=\max\{\sigma_1,\sigma_2\},$$

the estimate (5.268) implies the following estimate of the required sample size $N_1=N_2$:

$$\left(\frac{O(1)DL}{\varepsilon}\right)^n\exp\left\{-\frac{O(1)N_1\varepsilon^2}{\sigma^2}\right\}\leq\alpha,\tag{5.276}$$

which is equivalent to

$$N_1\geq\frac{O(1)\sigma^2}{\varepsilon^2}\left[n\ln\left(\frac{O(1)DL}{\varepsilon}\right)+\ln\left(\frac{1}{\alpha}\right)\right].\tag{5.277}$$

The estimate (5.277), for three-stage programs, looks similar to the estimate (5.116), of Theorem 5.18, for two-stage programs. Recall, however, that if we use the SAA method with conditional sampling and respective sample sizes N_1 and N_2, then the total number of scenarios is $N=N_1N_2$. Therefore, our analysis indicates that for three-stage problems we need random samples with the total number of scenarios of order of the *square* of the

SPbook
2009/8/20
page 230

corresponding sample size for two-stage problems. This analysis can be extended to T-stage problems with the conclusion that the total number of scenarios needed to solve the true problem with a reasonable accuracy grows *exponentially* with increase of the number of stages T. Some numerical experiments seem to confirm this conclusion. Of course, it should be mentioned that the above analysis does *not* prove in a *rigorous* mathematical sense that complexity of multistage programming grows exponentially with increase of the number of stages. It indicates only that the SAA method, which showed a considerable promise for solving two-stage problems, could be practically inapplicable for solving multistage problems with a large (say, greater than four) number of stages.

5.9 Stochastic Approximation Method

To an extent, this section is based on Nemirovski et al. [133]. Consider the stochastic optimization problem (5.1). We assume that the expected value function $f(x) = \mathbb{E}[F(x, \xi)]$ is well defined, finite valued, and continuous at every $x \in X$ and that the set $X \subset \mathbb{R}^n$ is nonempty, closed, and bounded. We denote by $\bar{x}$ an optimal solution of problem (5.1). Such an optimal solution does exist since the set X is compact and $f(x)$ is continuous. Clearly, $\vartheta^* = f(\bar{x})$. (Recall that ϑ^* denotes the optimal value of problem (5.1).) We also assume throughout this section that the set X is *convex* and the function $f(\cdot)$ is *convex*. Of course, if $F(\cdot, \xi)$ is convex for every $\xi \in \Xi$, then convexity of $f(\cdot)$ follows. We assume availability of the following *stochastic oracle*:

- There is a mechanism which for every given $x \in X$ and $\xi \in \Xi$ returns value $F(x, \xi)$ and a stochastic subgradient, a vector $G(x, \xi)$ such that $g(x) := \mathbb{E}[G(x, \xi)]$ is well defined and is a subgradient of $f(\cdot)$ at x, i.e., $g(x) \in \partial f(x)$.

Remark 18. Recall that if $F(\cdot, \xi)$ is convex for every $\xi \in \Xi$, and x is an interior point of X, i.e., $f(\cdot)$ is finite valued in a neighborhood of x, then

$$\partial f(x) = \mathbb{E}\left[\partial_x F(x, \xi)\right] \tag{5.278}$$

(see Theorem 7.47). Therefore, in that case we can employ a measurable selection $G(x, \xi) \in \partial_x F(x, \xi)$ as a stochastic subgradient. Note also that for an implementation of a stochastic approximation algorithm we only need to employ stochastic subgradients, while objective values $F(x, \xi)$ are used for accuracy estimates in section 5.9.4.

We also assume that we can generate, say, by Monte Carlo sampling techniques, an iid sequence ξ^j, $j = 1, \ldots$, of realizations of the random vector ξ, and hence to compute a stochastic subgradient $G(x_j, \xi^j)$ at an iterate point $x_j \in X$.

5.9.1 Classical Approach

We denote by $\|x\|_2 = (x^{\mathrm{T}}x)^{1/2}$ the Euclidean norm of vector $x \in \mathbb{R}^n$ and by

$$\Pi_X(x) := \arg\min_{z \in X} \|x - z\|_2 \tag{5.279}$$

the metric projection of x onto the set X. Since X is convex and closed, the minimizer in the right-hand side of (5.279) exists and is unique. Note that Π_X is a nonexpanding operator, i.e.,

$$\|\Pi_X(x') - \Pi_X(x)\|_2 \leq \|x' - x\|_2, \quad \forall x', x \in \mathbb{R}^n. \tag{5.280}$$

The classical stochastic approximation (SA) algorithm solves problem (5.1) by mimicking a simple subgradient descent method. That is, for chosen initial point $x_1 \in X$ and a sequence $\gamma_j > 0$, $j = 1, \ldots$, of stepsizes, it generates the iterates by the formula

$$x_{j+1} = \Pi_X(x_j - \gamma_j G(x_j, \xi^j)). \tag{5.281}$$

The crucial question of that approach is how to choose the stepsizes γ_j. Also, the set X should be simple enough so that the corresponding projection can be easily calculated. We now analyze convergence of the iterates, generated by this procedure, to an optimal solution $\bar{x}$ of problem (5.1). Note that the iterate $x_{j+1} = x_{j+1}(\xi_{[j]})$, $j = 1, \ldots$, is a function of the history $\xi_{[j]} = (\xi^1, \ldots, \xi^j)$ of the generated random process and hence is random, while the initial point x_1 is given (deterministic). We assume that there is number $M > 0$ such that

$$\mathbb{E}\left[\|G(x, \xi)\|_2^2\right] \leq M^2, \quad \forall x \in X. \tag{5.282}$$

Note that since for a random variable Z it holds that $\mathbb{E}[Z^2] \geq (\mathbb{E}|Z|)^2$, it follows from (5.282) that $\mathbb{E}\|G(x, \xi)\| \leq M$.

Denote

$$A_j := \tfrac{1}{2}\|x_j - \bar{x}\|_2^2 \quad \text{and} \quad a_j := \mathbb{E}[A_j] = \tfrac{1}{2}\mathbb{E}\big[\|x_j - \bar{x}\|_2^2\big]. \tag{5.283}$$

By (5.280) and since $\bar{x} \in X$ and hence $\Pi_X(\bar{x}) = \bar{x}$, we have

$$\begin{array}{rl} A_{j+1} & = \tfrac{1}{2}\big\|\Pi_X\big(x_j - \gamma_j G(x_j, \xi^j)\big) - \bar{x}\big\|_2^2 \\ & = \tfrac{1}{2}\big\|\Pi_X\big(x_j - \gamma_j G(x_j, \xi^j)\big) - \Pi_X(\bar{x})\big\|_2^2 \\ & \leq \tfrac{1}{2}\big\|x_j - \gamma_j G(x_j, \xi^j) - \bar{x}\big\|_2^2 \\ & = A_j + \tfrac{1}{2}\gamma_j^2\|G(x_j, \xi^j)\|_2^2 - \gamma_j(x_j - \bar{x})^{\mathrm{T}} G(x_j, \xi^j). \end{array} \tag{5.284}$$

Since $x_j = x_j(\xi_{[j-1]})$ is independent of ξ_j, we have

$$\begin{array}{rcl} \mathbb{E}\big[(x_j - \bar{x})^{\mathrm{T}} G(x_j, \xi^j)\big] & = & \mathbb{E}\left\{\mathbb{E}\big[(x_j - \bar{x})^{\mathrm{T}} G(x_j, \xi^j)\,|\xi_{[j-1]}\big]\right\} \\ & = & \mathbb{E}\left\{(x_j - \bar{x})^{\mathrm{T}}\mathbb{E}[G(x_j, \xi^j)\,|\xi_{[j-1]}]\right\} \\ & = & \mathbb{E}\big[(x_j - \bar{x})^{\mathrm{T}} g(x_j)\big]. \end{array}$$

Therefore, by taking expectation of both sides of (5.284) and since (5.282) we obtain

$$a_{j+1} \leq a_j - \gamma_j\mathbb{E}\left[(x_j - \bar{x})^{\mathrm{T}} g(x_j)\right] + \tfrac{1}{2}\gamma_j^2 M^2. \tag{5.285}$$

Suppose, further, that the expectation function $f(x)$ is differentiable and strongly convex on X with parameter $c > 0$, i.e.,

$$(x' - x)^{\mathrm{T}}(\nabla f(x') - \nabla f(x)) \geq c\|x' - x\|_2^2, \quad \forall x, x' \in X. \tag{5.286}$$

SPbook
2009/8/20
page 232

Note that strong convexity of $f(x)$ implies that the minimizer $\bar{x}$ is unique and that because of differentiability of $f(x)$ it follows that $\partial f(x) = \{\nabla f(x)\}$ and hence $g(x) = \nabla f(x)$. By optimality of $\bar{x}$ we have that

$$(x - \bar{x})^{\mathrm{T}} \nabla f(\bar{x}) \geq 0, \quad \forall x \in X, \tag{5.287}$$

which together with (5.286) implies that

$$\begin{array}{rcl} \mathbb{E}\left[(x_j - \bar{x})^T \nabla f(x_j)\right] & \geq & \mathbb{E}\left[(x_j - \bar{x})^T (\nabla f(x_j) - \nabla f(\bar{x}))\right] \\ & \geq & c\mathbb{E}\left[\|x_j - \bar{x}\|_2^2\right] = 2ca_j. \end{array} \tag{5.288}$$

Therefore it follows from (5.285) that

$$a_{j+1} \leq (1 - 2c\gamma_j)a_j + \tfrac{1}{2}\gamma_j^2 M^2. \tag{5.289}$$

In the classical approach to stochastic approximation the employed stepsizes are $\gamma_j := \theta/j$ for some constant $\theta > 0$. Then by (5.289) we have

$$a_{j+1} \leq (1 - 2c\theta/j)a_j + \tfrac{1}{2}\theta^2 M^2/j^2. \tag{5.290}$$

Suppose now that $\theta > 1/(2c)$. Then it follows from (5.290) by induction that for $j = 1, \ldots,$

$$2a_j \leq \frac{\max\left\{\theta^2 M^2 (2c\theta - 1)^{-1}, 2a_1\right\}}{j}. \tag{5.291}$$

Recall that $2a_j = \mathbb{E}\left[\|x_j - \bar{x}\|^2\right]$ and, since x_1 is deterministic, $2a_1 = \|x_1 - \bar{x}\|_2^2$. Therefore, by (5.291) we have that

$$\mathbb{E}\left[\|x_j - \bar{x}\|_2^2\right] \leq \frac{Q(\theta)}{j}, \tag{5.292}$$

where

$$Q(\theta) := \max\left\{\theta^2 M^2 (2c\theta - 1)^{-1}, \|x_1 - \bar{x}\|_2^2\right\}. \tag{5.293}$$

The constant $Q(\theta)$ attains its optimal (minimal) value at $\theta = 1/c$.

Suppose, further, that $\bar{x}$ is an *interior* point of X and $\nabla f(x)$ is Lipschitz continuous, i.e., there is constant $L > 0$ such that

$$\|\nabla f(x') - \nabla f(x)\|_2 \leq L\|x' - x\|_2, \quad \forall x', x \in X. \tag{5.294}$$

Then

$$f(x) \leq f(\bar{x}) + \tfrac{1}{2}L\|x - \bar{x}\|_2^2, \quad \forall x \in X, \tag{5.295}$$

and hence by (5.292)

$$\mathbb{E}\big[f(x_j) - f(\bar{x})\big] \leq \tfrac{1}{2}L\,\mathbb{E}\left[\|x_j - \bar{x}\|_2^2\right] \leq \frac{Q(\theta)L}{2j}. \tag{5.296}$$

We obtain that under the specified assumptions, after j iterations the expected error of the current solution in terms of the distance to the true optimal solution $\bar{x}$ is of order $O(j^{-1/2})$, and the expected error in terms of the objective value is of order $O(j^{-1})$, provided that $\theta > 1/(2c)$. Note, however, that the classical stepsize rule $\gamma_j = \theta/j$ could be very dangerous if the parameter c of strong convexity is overestimated, i.e., if $\theta < 1/(2c)$.

Example 5.36. As a simple example, consider $f(x) := \frac{1}{2}\kappa x^2$ with $\kappa > 0$ and $X := [-1, 1] \subset \mathbb{R}$ and assume that there is no noise, i.e., $G(x, \xi) \equiv \nabla f(x)$. Clearly $\bar{x} = 0$ is the optimal solution and zero is the optimal value of the corresponding optimization (minimization) problem. Let us take $\theta = 1$, i.e., use stepsizes $\gamma_j = 1/j$, in which case the iteration process becomes

$$x_{j+1} = x_j - f'(x_j)/j = \left(1 - \frac{\kappa}{j}\right) x_j. \tag{5.297}$$

For $\kappa = 1$, the above choice of the stepsizes is optimal and the optimal solution is obtained in one iteration.

Suppose now that $\kappa < 1$. Then starting with $x_1 > 0$, we have

$$x_{j+1} = x_1 \prod_{s=1}^{j} \left(1 - \frac{\kappa}{s}\right) = x_1 \exp\left\{-\sum_{s=1}^{j} \ln\left(1 + \frac{\kappa}{s-\kappa}\right)\right\} > x_1 \exp\left\{-\sum_{s=1}^{j} \frac{\kappa}{s-\kappa}\right\}.$$

Moreover,

$$\sum_{s=1}^{j} \frac{\kappa}{s-\kappa} \leq \frac{\kappa}{1-\kappa} + \int_1^j \frac{\kappa}{t-\kappa} dt < \frac{\kappa}{1-\kappa} + \kappa \ln j - \kappa \ln(1-\kappa).$$

It follows that

$$x_{j+1} > O(1)\, j^{-\kappa} \quad \text{and} \quad f(x_{j+1}) > O(1) j^{-2\kappa}, \quad j = 1, \ldots. \tag{5.298}$$

(In the first of the above inequalities the constant $O(1) = x_1 \exp\{-\kappa/(1-\kappa) + \kappa \ln(1-\kappa)\}$, and in the second inequality the generic constant $O(1)$ is obtained from the first one by taking square and multiplying it by $\kappa/2$.) That is, the convergence becomes extremely slow for small κ close to zero. In order to reduce the value x_j (the objective value $f(x_j)$) by factor 10, i.e., to improve the error of current solution by one digit, we will need to increase the number of iterations j by factor $10^{1/\kappa}$ (by factor $10^{1/(2\kappa)}$). For example, for $\kappa = 0.1$, $x_1 = 1$ and $j = 10^5$ we have that $x_j > 0.28$. In order to reduce the error of the iterate to 0.028 we will need to increase the number of iterations by factor 10^{10}, i.e., to $j = 10^{15}$. ∎

It could be added that if $f(x)$ loses strong convexity, i.e., the parameter c degenerates to zero, and hence no choice of $\theta > 1/(2c)$ is possible, then the stepsizes $\gamma_j = \theta/j$ may become completely unacceptable for any choice of θ.

5.9.2 Robust SA Approach

It was argued in section 5.9.1 that the classical stepsizes $\gamma_j = O(j^{-1})$ can be too small to ensure a reasonable rate of convergence even in the no-noise case. An important improvement to the SA method was developed by Polyak [152] and Polyak and Juditsky [153], where longer stepsizes were suggested with consequent averaging of the obtained iterates. Under the outlined classical assumptions, the resulting algorithm exhibits the same optimal $O(j^{-1})$ asymptotical convergence rate while using an easy to implement and "robust" stepsize policy. The main ingredients of Polyak's scheme (long steps and averaging) were, in

a different form, proposed in Nemirovski and Yudin [135] for problems with general-type Lipschitz continuous convex objectives and for convex–concave saddle point problems. Results of this section go back to Nemirovski and Yudin [135], [136].

Recall that $g(x) \in \partial f(x)$ and $a_j = \frac{1}{2}\mathbb{E}\left[\|x_j - \bar{x}\|_2^2\right]$, and we assume the boundedness condition (5.282). By convexity of $f(x)$ we have that $f(x) \geq f(x_j) + (x - x_j)^{\mathrm{T}} g(x_j)$ for any $x \in X$, and hence

$$\mathbb{E}\big[(x_j - \bar{x})^{\mathrm{T}} g(x_j)\big] \geq \mathbb{E}\big[f(x_j) - f(\bar{x})\big]. \tag{5.299}$$

Together with (5.285) this implies

$$\gamma_j \mathbb{E}\big[f(x_j) - f(\bar{x})\big] \leq a_j - a_{j+1} + \tfrac{1}{2}\gamma_j^2 M^2.$$

It follows that whenever $1 \leq i \leq j$, we have

$$\sum_{t=i}^{j} \gamma_t \mathbb{E}\big[f(x_t) - f(\bar{x})\big] \leq \sum_{t=i}^{j} [a_t - a_{t+1}] + \tfrac{1}{2}M^2 \sum_{t=i}^{j} \gamma_t^2 \leq a_i + \tfrac{1}{2}M^2 \sum_{t=i}^{j} \gamma_t^2. \tag{5.300}$$

Denote

$$\nu_t := \frac{\gamma_t}{\sum_{\tau=i}^{j} \gamma_\tau} \quad \text{and} \quad D_X := \max_{x \in X} \|x - x_1\|_2. \tag{5.301}$$

Clearly $\nu_t \geq 0$ and $\sum_{t=i}^{j} \nu_t = 1$. By (5.300) we have

$$\mathbb{E}\left[\sum_{t=i}^{j} \nu_t f(x_t) - f(\bar{x})\right] \leq \frac{a_i + \frac{1}{2}M^2 \sum_{t=i}^{j} \gamma_t^2}{\sum_{t=i}^{j} \gamma_t}. \tag{5.302}$$

Consider points

$$\tilde{x}_{i,j} := \sum_{t=i}^{j} \nu_t x_t. \tag{5.303}$$

Since X is convex, it follows that $\tilde{x}_{i,j} \in X$ and by convexity of $f(\cdot)$ we have

$$f(\tilde{x}_{i,j}) \leq \sum_{t=i}^{j} \nu_t f(x_t).$$

Thus, by (5.302) and in view of $a_1 \leq D_X^2$ and $a_i \leq 4D_X^2$, $i > 1$, we get

$$\mathbb{E}\left[f(\tilde{x}_{1,j}) - f(\bar{x})\right] \leq \frac{D_X^2 + M^2 \sum_{t=1}^{j} \gamma_t^2}{2\sum_{t=1}^{j} \gamma_t} \quad \text{for } 1 \leq j, \tag{5.304}$$

$$\mathbb{E}\left[f(\tilde{x}_{i,j}) - f(\bar{x})\right] \leq \frac{4D_X^2 + M^2 \sum_{t=i}^{j} \gamma_t^2}{2\sum_{t=i}^{j} \gamma_t} \quad \text{for } 1 < i \leq j. \tag{5.305}$$

Based of the above bounds on the expected accuracy of approximate solutions $\tilde{x}_{i,j}$, we can now develop "reasonable" stepsize policies along with the associated efficiency estimates.

Constant Stepsizes and Error Estimates

Assume now that the number of iterations of the method is fixed in advance, say, equal to N, and that we use the *constant* stepsize policy, i.e., $\gamma_t = \gamma$, $t = 1, \ldots, N$. It follows then from (5.304) that

$$\mathbb{E}\left[f(\tilde{x}_{1,N}) - f(\bar{x})\right] \leq \frac{D_X^2 + M^2 N \gamma^2}{2N\gamma}. \tag{5.306}$$

Minimizing the right-hand side of (5.306) over $\gamma > 0$, we arrive at the *constant* stepsize policy

$$\gamma_t = \frac{D_X}{M\sqrt{N}}, \quad t = 1, \ldots, N, \tag{5.307}$$

along with the associated efficiency estimate

$$\mathbb{E}\left[f(\tilde{x}_{1,N}) - f(\bar{x})\right] \leq \frac{D_X M}{\sqrt{N}}. \tag{5.308}$$

By (5.305), with the constant stepsize policy (5.307), we also have for $1 \leq K \leq N$

$$\mathbb{E}\left[f(\tilde{x}_{K,N}) - f(\bar{x})\right] \leq \frac{C_{N,K} D_X M}{\sqrt{N}}, \tag{5.309}$$

where

$$C_{N,K} := \frac{2N}{N - K + 1} + \frac{1}{2}.$$

When $K/N \leq 1/2$, the right-hand side of (5.309) coincides, within an absolute constant factor, with the right-hand side of (5.308). If we change the stepsizes (5.307) by a factor of $\theta > 0$, i.e., use the stepsizes

$$\gamma_t = \frac{\theta D_X}{M\sqrt{N}}, \quad t = 1, \ldots, N, \tag{5.310}$$

then the efficiency estimate (5.309) becomes

$$\mathbb{E}\left[f(\tilde{x}_{K,N}) - f(\bar{x})\right] \leq \max\left\{\theta, \theta^{-1}\right\} \frac{C_{N,K} D_X M}{\sqrt{N}}. \tag{5.311}$$

The expected error of the iterates (5.303), with constant stepsize policy (5.310), after N iterations is $O(N^{-1/2})$. Of course, this is worse than the rate $O(N^{-1})$ for the classical SA algorithm as applied to a smooth strongly convex function attaining minimum at an interior point of the set X. However, the error bound (5.311) is guaranteed independently of any smoothness and/or strong convexity assumptions on $f(\cdot)$. Moreover, changing the stepsizes by factor θ results just in rescaling of the corresponding error estimate (5.311). This is in a sharp contrast to the classical approach discussed in the previous section, when such change of stepsizes can be a disaster. These observations, in particular the fact that there is no necessity in fine tuning the stepsizes to the objective function $f(\cdot)$, explains the adjective "robust" in the name of the method.

It can be interesting to compare sample size estimates derived from the error bounds of the (robust) SA approach with respective sample size estimates of the SAA method discussed in section 5.3.2. By Chebyshev (Markov) inequality we have that for $\varepsilon > 0$,

$$\Pr\left\{f(\tilde{x}_{1,N}) - f(\bar{x}) \geq \varepsilon\right\} \leq \varepsilon^{-1}\mathbb{E}\left[f(\tilde{x}_{1,N}) - f(\bar{x})\right]. \tag{5.312}$$

Together with (5.308) this implies that, for the constant stepsize policy (5.307),

$$\Pr\left\{f(\tilde{x}_{1,N}) - f(\bar{x}) \geq \varepsilon\right\} \leq \frac{D_X M}{\varepsilon\sqrt{N}}. \tag{5.313}$$

It follows that for $\alpha \in (0, 1)$ and sample size

$$N \geq \frac{D_X^2 M^2}{\varepsilon^2 \alpha^2} \tag{5.314}$$

we are guaranteed that $\tilde{x}_{1,N}$ is an ε-optimal solution of the "true" problem (5.1) with probability at least $1 - \alpha$.

Compared with the corresponding estimate (5.126) for the sample size by the SAA method, the estimate (5.314) is of the same order with respect to parameters D_X,M, and ε. On the other hand, the dependence on the significance level α is different: in (5.126) it is of order $O\left(\ln(\alpha^{-1})\right)$, while in (5.314) it is of order $O(\alpha^{-2})$. It is possible to derive better estimates, similar to the respective estimates of the SAA method, of the required sample size by using the large deviations theory; we discuss this further in the next section (see Theorem 5.41 in particular).

5.9.3 Mirror Descent SA Method

The robust SA approach discussed in the previous section is tailored to Euclidean structure of the space $\mathbb{R}^n$. In this section, we discuss a generalization of the Euclidean SA approach allowing to adjust, to some extent, the method to the geometry, not necessary Euclidean, of the problem in question. A rudimentary form of the following generalization can be found in Nemirovski and Yudin [136], from where the name "mirror descent" originates.

In this section we denote by $\|\cdot\|$ a *general* norm on $\mathbb{R}^n$. Its dual norm is defined as

$$\|x\|_* := \sup_{\|y\|\leq 1} y^{\mathrm{T}} x.$$

By $\|x\|_p := \left(|x_1|^p + \cdots + |x_n|^p\right)^{1/p}$ we denote the ℓ_p, $p \in [1, \infty)$, norm on $\mathbb{R}^n$. In particular, $\|\cdot\|_2$ is the Euclidean norm. Recall that the dual of $\|\cdot\|_p$ is the norm $\|\cdot\|_q$, where $q > 1$ is such that $1/p + 1/q = 1$. The dual norm of ℓ_1 norm $\|x\|_1 = |x_1| + \cdots + |x_n|$ is the ℓ_∞ norm $\|x\|_\infty = \max\left\{|x_1|, \cdots, |x_n|\right\}$.

Definition 5.37. *We say that a function $\mathfrak{d} : X \to \mathbb{R}$ is a* distance-generating function *with modulus $\kappa > 0$ with respect to norm $\|\cdot\|$ if the following holds: $\mathfrak{d}(\cdot)$ is convex continuous on X, the set*

$$X^\star := \{x \in X : \partial\, \mathfrak{d}(x) \neq \emptyset\} \tag{5.315}$$

is convex, $\mathfrak{d}(\cdot)$ is continuously differentiable on $X^\star$, and

$$(x' - x)^{\mathrm{T}}(\nabla\mathfrak{d}(x') - \nabla\mathfrak{d}(x)) \geq \kappa\|x' - x\|^2, \quad \forall x, x' \in X^\star. \tag{5.316}$$

Note that the set $X^\star$ includes the relative interior of the set X, and hence condition (5.316) implies that $\mathfrak{d}(\cdot)$ is strongly convex on X with the parameter κ taken with respect to the considered norm $\|\cdot\|$.

A simple example of a distance generating function (with modulus 1 with respect to the Euclidean norm) is $\mathfrak{d}(x) := \frac{1}{2}x^{\mathrm{T}}x$. Of course, this function is continuously differentiable at every $x \in \mathbb{R}^n$. Another interesting example is the *entropy function*

$$\mathfrak{d}(x) := \sum_{i=1}^{n} x_i \ln x_i, \tag{5.317}$$

defined on the standard simplex $X := \{x \in \mathbb{R}^n : \sum_{i=1}^n x_i = 1,\ x \geq 0\}$. (Note that by continuity, $x \ln x = 0$ for $x = 0$.) Here the set $X^\star$ is formed by points $x \in X$ having all coordinates different from zero. The set $X^\star$ is the subset of X of those points at which the entropy function is differentiable with $\nabla\mathfrak{d}(x) = (1 + \ln x_1, \ldots, 1 + \ln x_n)$. The entropy function is strongly convex with modulus 1 on standard simplex with respect to $\|\cdot\|_1$ norm.

Indeed, it suffices to verify that $h^{\mathrm{T}}\nabla^2\mathfrak{d}(x)h \geq \|h\|_1^2$ for every $h \in \mathbb{R}^n$ and $x \in X^\star$. This, in turn, is verified by

$$\begin{array}{ll} \left[\sum_i |h_i|\right]^2 & = \left[\sum_i (x_i^{-1/2}|h_i|)x_i^{1/2}\right]^2 \leq \left[\sum_i h_i^2 x_i^{-1}\right]\left[\sum_i x_i\right] \\ & = \sum_i h_i^2 x_i^{-1} = h^T\nabla^2\mathfrak{d}(x)h, \end{array} \tag{5.318}$$

where the inequality follows by Cauchy inequality.

Let us define function $V : X^\star \times X \to \mathbb{R}_+$ as follows:

$$V(x, z) := \mathfrak{d}(z) - [\mathfrak{d}(x) + \nabla\mathfrak{d}(x)^{\mathrm{T}}(z - x)]. \tag{5.319}$$

In what follows we refer to $V(\cdot,\cdot)$ as the *prox-function*[36] associated with the distance-generating function $\mathfrak{d}(x)$. Note that $V(x,\cdot)$ is nonnegative and is strongly convex with modulus κ with respect to the norm $\|\cdot\|$. Let us define *prox-mapping* $P_x : \mathbb{R}^n \to X^\star$, associated with the distance-generating function and a point $x \in X^\star$, viewed as a parameter, as follows:

$$P_x(y) := \arg\min_{z\in X}\left\{y^{\mathrm{T}}(z - x) + V(x, z)\right\}. \tag{5.320}$$

Observe that the minimum in the right-hand side of (5.320) is attained since $\mathfrak{d}(\cdot)$ is continuous on X and X is compact, and a corresponding minimizer is unique since $V(x,\cdot)$ is strongly convex on X. Moreover, by the definition of the set $X^\star$, all these minimizers belong to $X^\star$. Thus, the prox-mapping is well defined.

For the (Euclidean) distance-generating function $\mathfrak{d}(x) := \frac{1}{2}x^{\mathrm{T}}x$, we have that $P_x(y) = \Pi_X(x - y)$. In that case the iteration formula (5.281) of the SA algorithm can be written as

$$x_{j+1} = P_{x_j}(\gamma_j G(x_j, \xi^j)), \quad x_1 \in X^\star. \tag{5.321}$$

Our goal is to demonstrate that the main properties of the recurrence (5.281) are inherited by (5.321) for any distance-generating function $\mathfrak{d}(x)$.

Lemma 5.38. *For every $u \in X$, $x \in X^\star$ and $y \in \mathbb{R}^n$ one has*

$$V(P_x(y), u) \leq V(x, u) + y^{\mathrm{T}}(u - x) + (2\kappa)^{-1}\|y\|_*^2. \tag{5.322}$$

[36] The function $V(\cdot,\cdot)$ is also called Bregman divergence.

Proof. Let $x \in X^\star$ and $v := P_x(y)$. Note that v is of the form $\operatorname{argmin}_{z\in X}\left[h^\mathsf{T} z + \mathfrak{d}(z)\right]$ and thus $v \in X^\star$, so that $\mathfrak{d}(\cdot)$ is differentiable at v. Since $\nabla_v V(x,v) = \nabla\mathfrak{d}(v) - \nabla\mathfrak{d}(x)$, the optimality conditions for (5.320) imply that

$$(\nabla\mathfrak{d}(v) - \nabla\mathfrak{d}(x) + y)^\mathsf{T}(v-u) \le 0, \quad \forall u \in X. \tag{5.323}$$

Therefore, for $u \in X$ we have

$$\begin{aligned} V(v,u) - V(x,u) &= [\mathfrak{d}(u) - \nabla\mathfrak{d}(v)^\mathsf{T}(u-v) - \mathfrak{d}(v)] - [\mathfrak{d}(u) - \nabla\mathfrak{d}(x)^\mathsf{T}(u-x) - \mathfrak{d}(x)] \\ &= (\nabla\mathfrak{d}(v) - \nabla\mathfrak{d}(x) + y)^\mathsf{T}(v-u) + y^\mathsf{T}(u-v) - [\mathfrak{d}(v) - \nabla\mathfrak{d}(x)^\mathsf{T}(v-x) - \mathfrak{d}(x)] \\ &\le y^\mathsf{T}(u-v) - V(x,v), \end{aligned}$$

where the last inequality follows by (5.323).

For any $a, b \in \mathbb{R}^n$ we have by the definition of the dual norm that $\|a\|_*\|b\| \ge a^\mathsf{T} b$ and hence

$$(\|a\|_*^2/\kappa + \kappa\|b\|^2)/2 \ge \|a\|_*\|b\| \ge a^\mathsf{T} b. \tag{5.324}$$

Applying this inequality with $a = y$ and $b = x - v$ we obtain

$$y^\mathsf{T}(x-v) \le \frac{\|y\|_*^2}{2\kappa} + \frac{\kappa}{2}\|x-v\|^2.$$

Also due to the strong convexity of $V(x,\cdot)$ and since $V(x,x) = 0$ we have

$$\begin{aligned} V(x,v) &\ge V(x,x) + (x-v)^\mathsf{T}\nabla_v V(x,v) + \tfrac{1}{2}\kappa\|x-v\|^2 \\ &= (x-v)^\mathsf{T}(\nabla\mathfrak{d}(v) - \nabla\mathfrak{d}(x)) + \tfrac{1}{2}\kappa\|x-v\|^2 \\ &\ge \tfrac{1}{2}\kappa\|x-v\|^2, \end{aligned} \tag{5.325}$$

where the last inequality holds by convexity of $\mathfrak{d}(\cdot)$. We get

$$\begin{aligned} V(v,u) - V(x,u) &\le y^\mathsf{T}(u-v) - V(x,v) = y^\mathsf{T}(u-x) + y^\mathsf{T}(x-v) - V(x,v) \\ &\le y^\mathsf{T}(u-x) + (2\kappa)^{-1}\|y\|_*^2, \end{aligned}$$

as required in (5.322). □

Using (5.322) with $x = x_j$, $y = \gamma_j G(x_j, \xi^j)$, and $u = \bar{x}$, and noting that by (5.321) $x_{j+1} = P_x(y)$ here, we get

$$\gamma_j(x_j - \bar{x})^\mathsf{T} G(x_j,\xi^j) \le V(x_j,\bar{x}) - V(x_{j+1},\bar{x}) + \frac{\gamma_j^2}{2\kappa}\|G(x_j,\xi^j)\|_*^2. \tag{5.326}$$

Let us observe that for the Euclidean distance-generating function $\mathfrak{d}(x) = \frac{1}{2}x^\mathsf{T}x$, one has $V(x,z) = \frac{1}{2}\|x-z\|_2^2$ and $\kappa = 1$. That is, in the Euclidean case (5.326) becomes

$$\tfrac{1}{2}\|x_{j+1} - \bar{x}\|_2^2 \le \tfrac{1}{2}\|x_j - \bar{x}\|_2^2 + \tfrac{1}{2}\gamma_j^2\|G(x_j,\xi^j)\|_2^2 - \gamma_j(x_j - \bar{x})^\mathsf{T} G(x_j,\xi^j). \tag{5.327}$$

The above inequality is exactly the relation (5.284), which played a crucial role in the developments related to the Euclidean SA. We are about to process, in a similar way, the relation (5.326) in the case of a general distance-generating function, thus arriving at the mirror descent SA.

Specifically, setting

$$\Delta_j := G(x_j, \xi^j) - g(x_j), \tag{5.328}$$

we can rewrite (5.326), with j replaced by t, as

$$\gamma_t (x_t - \bar{x})^{\mathrm{T}} g(x_t) \le V(x_t, \bar{x}) - V(x_{t+1}, \bar{x}) - \gamma_t \Delta_t^{\mathrm{T}} (x_t - \bar{x}) + \frac{\gamma_t^2}{2\kappa} \|G(x_t, \xi^t)\|_*^2. \tag{5.329}$$

Summing up over $t = 1, \ldots, j$, and taking into account that $V(x_{j+1}, u) \ge 0$, $u \in X$, we get

$$\sum_{t=1}^{j} \gamma_t (x_t - \bar{x})^{\mathrm{T}} g(x_t) \le V(x_1, \bar{x}) + \sum_{t=1}^{j} \frac{\gamma_t^2}{2\kappa} \|G(x_t, \xi^t)\|_*^2 - \sum_{t=1}^{j} \gamma_t \Delta_t^{\mathrm{T}} (x_t - \bar{x}). \tag{5.330}$$

Set $\nu_t := \frac{\gamma_t}{\sum_{\tau=1}^{j} \gamma_\tau}$, $t = 1, \ldots, j$, and

$$\tilde{x}_{1,j} := \sum_{t=1}^{j} \nu_t x_t. \tag{5.331}$$

By convexity of $f(\cdot)$ we have $f(x_t) - f(\bar{x}) \le (x_t - \bar{x})^{\mathrm{T}} g(x_t)$, and hence

$$\begin{array}{rcl} \sum_{t=1}^{j} \gamma_t (x_t - \bar{x})^{\mathrm{T}} g(x_t) & \ge & \sum_{t=1}^{j} \gamma_t [f(x_t) - f(\bar{x})] \\ & = & \left(\sum_{t=1}^{j} \gamma_t\right) \left[\sum_{t=1}^{j} \nu_t f(x_t) - f(\bar{x})\right] \\ & \ge & \left(\sum_{t=1}^{j} \gamma_t\right) \left[f(\tilde{x}_{1,j}) - f(\bar{x})\right]. \end{array}$$

Combining this with (5.330) we obtain

$$f(\tilde{x}_{1,j}) - f(\bar{x}) \le \frac{V(x_1, \bar{x}) + \sum_{t=1}^{j} (2\kappa)^{-1} \gamma_t^2 \|G(x_t, \xi^t)\|_*^2 - \sum_{t=1}^{j} \gamma_t \Delta_t^{\mathrm{T}} (x_t - \bar{x})}{\sum_{t=1}^{j} \gamma_t}. \tag{5.332}$$

- Assume from now on that the procedure starts with the minimizer of $\mathfrak{d}(\cdot)$, that is,

$$x_1 := \operatorname{argmin}_{x \in X} \mathfrak{d}(x). \tag{5.333}$$

Since by the optimality of x_1 we have that $(u - x_1)^{\mathrm{T}} \nabla \mathfrak{d}(x_1) \ge 0$ for any $u \in X$, it follows from the definition (5.319) of the function $V(\cdot, \cdot)$ that

$$\max_{u \in X} V(x_1, u) \le D_{\mathfrak{d},X}^2, \tag{5.334}$$

where

$$D_{\mathfrak{d},X} := \left[\max_{u \in X} \mathfrak{d}(u) - \min_{x \in X} \mathfrak{d}(x)\right]^{1/2}. \tag{5.335}$$

Together with (5.332) this implies

$$f(\tilde{x}_{1,j}) - f(\bar{x}) \le \frac{D_{\mathfrak{d},X}^2 + \sum_{t=1}^{j} (2\kappa)^{-1} \gamma_t^2 \|G(x_t, \xi^t)\|_*^2 - \sum_{t=1}^{j} \gamma_t \Delta_t^{\mathrm{T}} (x_t - \bar{x})}{\sum_{t=1}^{j} \gamma_t}. \tag{5.336}$$

We also have (see (5.325)) that $V(x_1, u) \geq \frac{1}{2}\kappa\|x_1 - u\|^2$, and hence it follows from (5.334) that for all $u \in X$,

$$\|x_1 - u\| \leq \sqrt{\frac{2}{\kappa}} D_{\mathfrak{d},X}. \tag{5.337}$$

Let us assume, as in the previous section (see (5.282)), that there is a positive number M_* such that

$$\mathbb{E}\left[\|G(x,\xi)\|_*^2\right] \leq M_*^2, \quad \forall x \in X. \tag{5.338}$$

Proposition 5.39. *Let* $x_1 := \operatorname{argmin}_{x\in X} \mathfrak{d}(x)$ *and suppose that condition* (5.338) *holds. Then*

$$\mathbb{E}\left[f(\tilde{x}_{1,j}) - f(\bar{x})\right] \leq \frac{D_{\mathfrak{d},X}^2 + (2\kappa)^{-1} M_*^2 \sum_{t=1}^j \gamma_t^2}{\sum_{t=1}^j \gamma_t}. \tag{5.339}$$

Proof. Taking expectations of both sides of (5.336) and noting that (i) x_t is a deterministic function of $\xi_{[t-1]} = (\xi^1, \ldots, \xi^{t-1})$, (ii) conditional on $\xi_{[t-1]}$, the expectation of Δ_t is 0, and (iii) the expectation of $\|G(x_t, \xi^t)\|_*^2$ does not exceed M_*^2, we obtain (5.339). □

Constant Stepsize Policy

Assume that the total number of steps N is given in advance and the constant stepsize policy $\gamma_t = \gamma$, $t = 1, \ldots, N$, is employed. Then (5.339) becomes

$$\mathbb{E}\left[f(\tilde{x}_{1,j}) - f(\bar{x})\right] \leq \frac{D_{\mathfrak{d},X}^2 + (2\kappa)^{-1} M_*^2 N \gamma^2}{N\gamma}. \tag{5.340}$$

Minimizing the right-hand side of (5.340) over $\gamma > 0$ we arrive at the constant stepsize policy

$$\gamma_t = \frac{\sqrt{2\kappa} D_{\mathfrak{d},X}}{M_*\sqrt{N}}, \quad t = 1, \ldots, N, \tag{5.341}$$

and the associated efficiency estimate

$$\mathbb{E}\left[f(\tilde{x}_{1,N}) - f(\bar{x})\right] \leq D_{\mathfrak{d},X} M_* \sqrt{\frac{2}{\kappa N}}. \tag{5.342}$$

This can be compared with the respective stepsize (5.307) and efficiency estimate (5.308) for the robust Euclidean SA method. Passing from the stepsizes (5.341) to the stepsizes

$$\gamma_t = \frac{\theta\sqrt{2\kappa} D_{\mathfrak{d},X}}{M_*\sqrt{N}}, \quad t = 1, \ldots, N, \tag{5.343}$$

with rescaling parameter $\theta > 0$, the efficiency estimate becomes

$$\mathbb{E}\left[f(\tilde{x}_{1,N}) - f(\bar{x})\right] \leq \max\left\{\theta, \theta^{-1}\right\} D_{\mathfrak{d},X} M_* \sqrt{\frac{2}{\kappa N}}, \tag{5.344}$$

similar to the Euclidean case. We refer to the SA method based on (5.321), (5.331), and (5.343) as the mirror descent SA algorithm with constant stepsize policy.

Comparing (5.308) to (5.342), we see that for both the Euclidean and the mirror descent SA algorithms, the expected inaccuracy, in terms of the objective values of the approximate solutions, is $O(N^{-1/2})$. A benefit of the mirror descent over the Euclidean algorithm is in potential possibility to reduce the constant factor hidden in $O(\cdot)$ by adjusting the norm $\|\cdot\|$ and the distance generating function $\mathfrak{d}(\cdot)$ to the geometry of the problem.

Example 5.40. Let $X := \{x \in \mathbb{R}^n : \sum_{i=1}^n x_i = 1,\ x \geq 0\}$ be the standard simplex. Consider two setups for the mirror descent SA, namely, the *Euclidean setup*, where the considered norm $\|\cdot\| := \|\cdot\|_2$ and $\mathfrak{d}(x) := \frac{1}{2}x^{\mathsf{T}}x$, and ℓ_1*-setup*, where $\|\cdot\| := \|\cdot\|_1$ and $\mathfrak{d}(\cdot)$ is the *entropy function* (5.317). The Euclidean setup, leads to the Euclidean robust SA, which is easily implementable. Note that the Euclidean diameter of X is $\sqrt{2}$ and hence is independent of n. The corresponding efficiency estimate is

$$\mathbb{E}\left[f(\tilde{x}_{1,N}) - f(\bar{x})\right] \leq O(1)\max\left\{\theta, \theta^{-1}\right\} MN^{-1/2} \tag{5.345}$$

with $M^2 = \sup_{x\in X}\mathbb{E}\left[\|G(x,\xi)\|_2^2\right]$.

The ℓ_1-setup corresponds to $X^\star = \{x \in X : x > 0\}$, $D_{\mathfrak{d},X} = \sqrt{\ln n}$,

$$x_1 := \operatorname*{argmin}_{x\in X} \mathfrak{d}(x) = n^{-1}(1,\ldots,1)^{\mathsf{T}},$$

$\|x\|_* = \|x\|_\infty$, and $\kappa = 1$ (see (5.318) for verification that $\kappa = 1$). The associated mirror descent SA is easily implementable. The prox-function here is

$$V(x,z) = \sum_{i=1}^n z_i \ln\frac{z_i}{x_i},$$

and the prox-mapping $P_x(y)$ is given by the explicit formula

$$[P_x(y)]_i = \frac{x_i e^{-y_i}}{\sum_{k=1}^n x_k e^{-y_k}},\quad i = 1,\ldots,n.$$

The respective efficiency estimate of the ℓ_1-setup is

$$\mathbb{E}\left[f(\tilde{x}_{1,N}) - f(\bar{x})\right] \leq O(1)\max\left\{\theta, \theta^{-1}\right\} (\ln n)^{1/2} M_* N^{-1/2} \tag{5.346}$$

with $M_*^2 = \sup_{x\in X}\mathbb{E}\left[\|G(x,\xi)\|_\infty^2\right]$, provided that the constant stepsizes (5.343) are used.

To compare (5.346) and (5.345), observe that $M_* \leq M$, and the ratio M_*/M can be as small as $n^{-1/2}$. Thus, the efficiency estimate for the ℓ_1-setup is never much worse than the estimate for the Euclidean setup, and for large n can be *far better* than the latter estimate. That is,

$$\sqrt{\frac{1}{\ln n}} \leq \frac{M}{\sqrt{\ln n}M_*} \leq \sqrt{\frac{n}{\ln n}},$$

with both the upper and lower bounds being achievable. Thus, when X is a standard simplex of large dimension, we have strong reasons to prefer the ℓ_1-setup to the usual Euclidean one. ∎

Comparison with the SAA Approach

Similar to (5.312)–(5.314), by using Chebyshev (Markov) inequality, it is possible to derive from (5.344) an estimate of the sample size N which guarantees that $\tilde{x}_{1,N}$ is an ε-optimal solution of the true problem with probability at least $1-\alpha$. It is possible, however, to obtain much finer bounds on deviation probabilities when imposing more restrictive assumptions on the distribution of $G(x,\xi)$. Specifically, assume that there is constant $M_* > 0$ such that

$$\mathbb{E}\left[\exp\left\{\|G(x,\xi)\|_*^2/M_*^2\right\}\right] \le \exp\{1\}, \quad \forall x \in X. \tag{5.347}$$

Note that condition (5.347) is stronger than (5.338). Indeed, if a random variable Y satisfies $\mathbb{E}[\exp\{Y/a\}] \le \exp\{1\}$ for some $a > 0$, then by Jensen inequality

$$\exp\{\mathbb{E}[Y/a]\} \le \mathbb{E}[\exp\{Y/a\}] \le \exp\{1\},$$

and therefore $\mathbb{E}[Y] \le a$. By taking $Y := \|G(x,\xi)\|_*^2$ and $a := M^2$, we obtain that (5.347) implies (5.338). Of course, condition (5.347) holds if $\|G(x,\xi)\|_* \le M_*$ for all $(x,\xi) \in X \times \Xi$.

Theorem 5.41. *Suppose that condition* (5.347) *is fulfilled. Then for the constant stepsizes* (5.343)*, the following holds for any* $\Theta \ge 0$:

$$\Pr\left\{ f(\tilde{x}_{1,N}) - f(\bar{x}) \ge \frac{C(1+\Theta)}{\sqrt{\kappa N}} \right\} \le 4\exp\{-\Theta\}, \tag{5.348}$$

where $C := (\max\left\{\theta, \theta^{-1}\right\} + 8\sqrt{3})M_* D_{\mathfrak{d},X}/\sqrt{2}$.

Proof. By (5.336) we have

$$f(\tilde{x}_{1,N}) - f(\bar{x}) \le A_1 + A_2, \tag{5.349}$$

where

$$A_1 := \frac{D_{\mathfrak{d},X}^2 + (2\kappa)^{-1}\sum_{t=1}^N \gamma_t^2 \|G(x_t,\xi^t)\|_*^2}{\sum_{t=1}^N \gamma_t} \quad \text{and} \quad A_2 := \sum_{t=1}^N \nu_t \Delta_t^{\mathsf{T}}(\bar{x} - x_t).$$

Consider $Y_t := \gamma_t^2\|G(x_t,\xi^t)\|_*^2$ and $c_t := M_*^2\gamma_t^2$. Note that by (5.347),

$$\mathbb{E}\left[\exp\{Y_i/c_i\}\right] \le \exp\{1\}, \quad i = 1,\ldots,N. \tag{5.350}$$

Since $\exp\{\cdot\}$ is a convex function we have

$$\exp\left\{\frac{\sum_{i=1}^N Y_i}{\sum_{i=1}^N c_i}\right\} = \exp\left\{\sum_{i=1}^N \frac{c_i(Y_i/c_i)}{\sum_{i=1}^N c_i}\right\} \le \sum_{i=1}^N \frac{c_i}{\sum_{i=1}^N c_i}\exp\{Y_i/c_i\}.$$

By taking expectation of both sides of the above inequality and using (5.350) we obtain

$$\mathbb{E}\left[\exp\left\{\frac{\sum_{i=1}^N Y_i}{\sum_{i=1}^N c_i}\right\}\right] \le \exp\{1\}.$$

Consequently by Chebyshev's inequality we have for any number Θ

$$\Pr\left[\exp\left\{\frac{\sum_{i=1}^N Y_i}{\sum_{i=1}^N c_i}\right\} \geq \exp\{\Theta\}\right] \leq \frac{\exp\{1\}}{\exp\{\Theta\}} = \exp\{1-\Theta\},$$

and hence

$$\Pr\left\{\sum_{i=1}^N Y_i \geq \Theta \sum_{i=1}^N c_i\right\} \leq \exp\{1-\Theta\} \leq 3\exp\{-\Theta\}. \tag{5.351}$$

That is, for any Θ,

$$\Pr\left\{\sum_{t=1}^N \gamma_t^2 \|G(x_t,\xi^t)\|_*^2 \geq \Theta M_*^2 \sum_{t=1}^N \gamma_t^2\right\} \leq 3\exp\{-\Theta\}. \tag{5.352}$$

For the constant stepsize policy (5.343), we obtain by (5.352) that

$$\Pr\left\{A_1 \geq \max\{\theta,\theta^{-1}\}\frac{M_* D_{\mathfrak{d},X}(1+\Theta)}{\sqrt{2\kappa N}}\right\} \leq 3\exp\{-\Theta\}. \tag{5.353}$$

Consider now the random variable A_2. By (5.337) we have that

$$\|\bar{x}-x_t\| \leq \|x_1-\bar{x}\| + \|x_1-x_t\| \leq 2\sqrt{2}\kappa^{-1/2} D_{\mathfrak{d},X},$$

and hence

$$\left|\Delta_t^{\mathrm{T}}(\bar{x}-x_t)\right|^2 \leq \|\Delta_t\|_*^2 \|\bar{x}-x_t\|^2 \leq 8\kappa^{-1} D_{\mathfrak{d},X}^2 \|\Delta_t\|_*^2.$$

We also have that

$$\mathbb{E}\left[(\bar{x}-x_t)^{\mathrm{T}}\Delta_t \middle| \xi_{[t-1]}\right] = (\bar{x}-x_t)^{\mathrm{T}}\mathbb{E}\left[\Delta_t \middle| \xi_{[t-1]}\right] = 0 \ \text{ w.p. 1},$$

and by condition (5.347) that

$$\mathbb{E}\left[\exp\left\{\|\Delta_t\|_*^2/(4M_*^2)\right\} \middle| \xi_{[t-1]}\right] \leq \exp\{1\} \ \text{ w.p. 1}.$$

Consequently, by applying inequality (7.194) of Proposition 7.64 with $\phi_t := \nu_t \Delta_t^{\mathrm{T}}(\bar{x}-x_t)$ and $\sigma_t^2 := 32\kappa^{-1} M_*^2 D_{\mathfrak{d},X}^2 \nu_t^2$, we obtain for any $\Theta \geq 0$

$$\Pr\left\{A_2 \geq 4\sqrt{2}\kappa^{-1/2} M_* D_{\mathfrak{d},X}\Theta\sqrt{\sum_{t=1}^N \nu_t^2}\right\} \leq \exp\left\{-\Theta^2/3\right\}. \tag{5.354}$$

Since for the constant stepsize policy we have that $\nu_t = 1/N$, $t = 1,\ldots,N$, by changing variables $\Theta^2/3$ to Θ and noting that $\Theta^{1/2} \leq 1+\Theta$ for any $\Theta \geq 0$, we obtain from (5.354) that for any $\Theta \geq 0$

$$\Pr\left\{A_2 \geq \frac{8\sqrt{3}M_* D_{\mathfrak{d},X}(1+\Theta)}{\sqrt{2\kappa N}}\right\} \leq \exp\{-\Theta\}. \tag{5.355}$$

Finally, (5.348) follows from (5.349), (5.353), and (5.355). □

By setting $\varepsilon = \frac{C(1+\Theta)}{\sqrt{\kappa N}}$, we can rewrite the estimate (5.348) in the form[37]

$$\Pr\left\{f(\tilde{x}_{1,N}) - f(\bar{x}) > \varepsilon\right\} \leq 12\exp\left\{-\varepsilon C^{-1}\sqrt{\kappa N}\right\}. \tag{5.356}$$

[37] The constant 12 in the right-hand side of (5.356) comes from the simple estimate $4\exp\{1\} < 12$.

For $\varepsilon > 0$ this gives the following estimate of the sample size N which guarantees that $\tilde{x}_{1,N}$ is an ε-optimal solution of the true problem with probability at least $1 - \alpha$:

$$N \geq O(1)\varepsilon^{-2}\kappa^{-1}M_*^2 D_{\mathfrak{d},X}^2 \ln^2\left(12/\alpha\right). \tag{5.357}$$

This estimate is similar to the respective estimate (5.126) of the sample size for the SAA method. However, as far as complexity of solving the problem numerically is concerned, the SAA method requires a solution of the generated optimization problem, while an SA algorithm is based on computing a single subgradient $G(x_j, \xi^j)$ at each iteration point. As a result, for the same sample size N it typically takes considerably less computation time to run an SA algorithm than to solve the corresponding SAA problem.

5.9.4 Accuracy Certificates for Mirror Descent SA Solutions

We discuss now a way to estimate lower and upper bounds for the optimal value of problem (5.1) by employing SA iterates. This will give us an accuracy certificate for obtained solutions. Assume that we run an SA procedure with respective iterates $x_1, \ldots, x_N$ computed according to formula (5.321). As before, set

$$\nu_t := \frac{\gamma_t}{\sum_{\tau=1}^N \gamma_\tau}, \ t = 1, \ldots, N, \quad \text{and } \tilde{x}_{1,N} := \sum_{t=1}^N \nu_t x_t.$$

We assume now that the stochastic objective value $F(x, \xi)$ as well as the stochastic subgradient $G(x, \xi)$ are computable at a given point $(x, \xi) \in X \times \Xi$.

Consider

$$f_*^N := \min_{x \in X} f^N(x) \ \text{ and } \ f^{*N} := \sum_{t=1}^N \nu_t f(x_t), \tag{5.358}$$

where

$$f^N(x) := \sum_{t=1}^N \nu_t \left[f(x_t) + g(x_t)^\mathsf{T}(x - x_t)\right]. \tag{5.359}$$

Since $\nu_t > 0$ and $\sum_{t=1}^N \nu_t = 1$, by convexity of $f(x)$ we have that the function $f^N(x)$ underestimates $f(x)$ everywhere on X, and hence[38] $f_*^N \leq \vartheta^*$. Since $\tilde{x}_{1,N} \in X$ we also have that $\vartheta^* \leq f(\tilde{x}_{1,N})$ and by convexity of f that $f(\tilde{x}_{1,N}) \leq f^{*N}$. It follows that $\vartheta^* \leq f^{*N}$. That is, for any realization of the random process $\xi^1, \ldots$, we have that

$$f_*^N \leq \vartheta^* \leq f^{*N}. \tag{5.360}$$

It follows, of course, that $\mathbb{E}[f_*^N] \leq \vartheta^* \leq \mathbb{E}[f^{*N}]$ as well.

Along with the "unobservable" bounds f_*^N, f^{*N}, consider their observable (computable) counterparts

$$\begin{array}{l} \underline{f}^N := \min_{x \in X} \left\{\sum_{t=1}^N \nu_t [F(x_t, \xi^t) + G(x_t, \xi^t)^\mathsf{T}(x - x_t)]\right\}, \\ \overline{f}^N := \sum_{t=1}^N \nu_t F(x_t, \xi^t), \end{array} \tag{5.361}$$

[38] Recall that ϑ^* denotes the optimal value of the true problem (5.1).

which will be referred to as *online* bounds. The bound $\overline{f}^N$ can be easily calculated while running the SA procedure. The bound $\underline{f}^N$ involves solving the optimization problem of minimizing a linear in x objective function over set X. If the set X is defined by linear constraints, this is a linear programming problem.

Since x_t is a function of $\xi_{[t-1]}$ and ξ^t is independent of $\xi_{[t-1]}$, we have that

$$\mathbb{E}\big[\overline{f}^N\big] = \sum_{t=1}^N \nu_t \mathbb{E}\big\{\mathbb{E}[F(x_t,\xi^t)|\xi_{[t-1]}]\big\} = \sum_{t=1}^N \nu_t \mathbb{E}\left[f(x_t)\right] = \mathbb{E}[f^{*N}]$$

and

$$\begin{aligned}
\mathbb{E}\big[\underline{f}^N\big] &= \mathbb{E}\Big[\mathbb{E}\big\{\min_{x\in X}\big\{\textstyle\sum_{t=1}^N \nu_t[F(x_t,\xi^t)+G(x_t,\xi^t)^{\mathrm{T}}(x-x_t)]\big\}\big|\xi_{[t-1]}\big\}\Big]\\
&\le \mathbb{E}\Big[\min_{x\in X}\big\{\mathbb{E}\big[\textstyle\sum_{t=1}^N \nu_t[F(x_t,\xi^t)+G(x_t,\xi^t)^{\mathrm{T}}(x-x_t)]\big]\big|\xi_{[t-1]}\big\}\Big]\\
&= \mathbb{E}\big[\min_{x\in X} f^N(x)\big] = \mathbb{E}\big[f_*^N\big].
\end{aligned}$$

It follows that

$$\mathbb{E}\big[\underline{f}^N\big] \le \vartheta^* \le \mathbb{E}\big[\overline{f}^N\big]. \tag{5.362}$$

That is, on average $\underline{f}^N$ and $\overline{f}^N$ give, respectively, a lower and an upper bound for the optimal value ϑ^* of the optimization problem (5.1).

In order to see how good the bounds $\underline{f}^N$ and $\overline{f}^N$ are, let us estimate expectations of the corresponding errors. We will need the following result.

Lemma 5.42. *Let $\zeta_t \in \mathbb{R}^n$, $v_1 \in X^\star$, and $v_{t+1} = P_{v_t}(\zeta_t)$, $t = 1,\ldots,N$. Then*

$$\sum_{t=1}^N \zeta_t^{\mathrm{T}}(v_t - u) \le V(v_1,u) + (2\kappa)^{-1}\sum_{t=1}^N \|\zeta_t\|_*^2, \quad \forall u \in X. \tag{5.363}$$

Proof. By the estimate (5.322) of Lemma 5.38 with $x = v_t$ and $y = \zeta_t$ we have that the following inequality holds for any $u \in X$:

$$V(v_{t+1},u) \le V(v_t,u) + \zeta_t^{\mathrm{T}}(u - v_t) + (2\kappa)^{-1}\|\zeta_t\|_*^2. \tag{5.364}$$

Summing this over $t = 1,\ldots,N$, we obtain

$$V(v_{N+1},u) \le V(v_1,u) + \sum_{t=1}^N \zeta_t^{\mathrm{T}}(u - v_t) + (2\kappa)^{-1}\sum_{t=1}^N \|\zeta_t\|_*^2. \tag{5.365}$$

Since $V(v_{N+1},u) \ge 0$, (5.363) follows. □

Consider again condition (5.338), that is,

$$\mathbb{E}\left[\|G(x,\xi)\|_*^2\right] \le M_*^2, \quad \forall x \in X, \tag{5.366}$$

and the following condition: there is a constant $Q > 0$ such that

$$\mathbb{V}\mathrm{ar}[F(x,\xi)] \le Q^2, \quad \forall x \in X. \tag{5.367}$$

Note that, of course, $\mathbb{V}\mathrm{ar}[F(x,\xi)] = \mathbb{E}\big[(F(x,\xi) - f(x))^2\big]$.

Theorem 5.43. *Suppose that conditions* (5.366) *and* (5.367) *hold. Then*

$$\mathbb{E}\big[f^{*N} - f_*^N\big] \le \frac{2D_{\mathfrak{d},X}^2 + \frac{5}{2}\kappa^{-1}M_*^2\sum_{t=1}^N \gamma_t^2}{\sum_{t=1}^N \gamma_t}, \tag{5.368}$$

$$\mathbb{E}\Big[\big|\overline{f}^N - f^{*N}\big|\Big] \le Q\sqrt{\sum_{t=1}^N \nu_t^2}\,, \tag{5.369}$$

$$\mathbb{E}\Big[\big|\underline{f}^N - f_*^N\big|\Big] \le \Big(Q + 4\sqrt{2}\kappa^{-1/2}M_*D_{\mathfrak{d},X}\Big)\sqrt{\sum_{t=1}^N \nu_t^2} + \frac{D_{\mathfrak{d},X}^2 + 2\kappa^{-1}M_*^2\sum_{t=1}^N \gamma_t^2}{\sum_{t=1}^N \gamma_t}. \tag{5.370}$$

Proof. If in Lemma 5.42 we take $v_1 := x_1$ and $\zeta_t := \gamma_t G(x_t,\xi^t)$, then the corresponding iterates v_t coincide with x_t. Therefore, we have by (5.363) and since $V(x_1,u) \le D_{\mathfrak{d},X}^2$ that

$$\sum_{t=1}^N \gamma_t(x_t - u)^{\mathsf{T}}G(x_t,\xi^t) \le D_{\mathfrak{d},X}^2 + (2\kappa)^{-1}\sum_{t=1}^N \gamma_t^2\|G(x_t,\xi^t)\|_*^2, \quad \forall u \in X. \tag{5.371}$$

It follows that for any $u \in X$ (compare with (5.330)),

$$\begin{aligned}&\sum_{t=1}^N \nu_t\big[-f(x_t) + (x_t - u)^{\mathsf{T}}g(x_t)\big] + \sum_{t=1}^N \nu_t f(x_t)\\ &\le \frac{D_{\mathfrak{d},X}^2 + (2\kappa)^{-1}\sum_{t=1}^N \gamma_t^2\|G(x_t,\xi^t)\|_*^2}{\sum_{t=1}^N \gamma_t} + \sum_{t=1}^N \nu_t\Delta_t^{\mathsf{T}}(x_t - u),\end{aligned}$$

where $\Delta_t := G(x_t,\xi^t) - g(x_t)$. Since

$$f^{*N} - f_*^N = \sum_{t=1}^N \nu_t f(x_t) + \max_{u\in X}\sum_{t=1}^N \nu_t\big[-f(x_t) + (x_t - u)^{\mathsf{T}}g(x_t)\big],$$

it follows that

$$f^{*N} - f_*^N \le \frac{D_{\mathfrak{d},X}^2 + (2\kappa)^{-1}\sum_{t=1}^N \gamma_t^2\|G(x_t,\xi^t)\|_*^2}{\sum_{t=1}^N \gamma_t} + \max_{u\in X}\sum_{t=1}^N \nu_t\Delta_t^{\mathsf{T}}(x_t - u). \tag{5.372}$$

Let us estimate the second term in the right-hand side of (5.372). By using Lemma 5.42 with $v_1 := x_1$ and $\zeta_t := \gamma_t\Delta_t$, and the corresponding iterates $v_{t+1} = P_{v_t}(\zeta_t)$, we obtain

$$\sum_{t=1}^N \gamma_t\Delta_t^{\mathsf{T}}(v_t - u) \le D_{\mathfrak{d},X}^2 + (2\kappa)^{-1}\sum_{t=1}^N \gamma_t^2\|\Delta_t\|_*^2, \quad \forall u \in X. \tag{5.373}$$

Moreover,

$$\Delta_t^{\mathrm{T}}(v_t - u) = \Delta_t^{\mathrm{T}}(x_t - u) + \Delta_t^{\mathrm{T}}(v_t - x_t),$$

and hence it follows by (5.373) that

$$\max_{u\in X} \sum_{t=1}^{N} \nu_t \Delta_t^{\mathrm{T}}(x_t - u) \le \sum_{t=1}^{N} \nu_t \Delta_t^{\mathrm{T}}(x_t - v_t) + \frac{D_{\mathfrak{d},X}^2 + (2\kappa)^{-1} \sum_{t=1}^{N} \gamma_t^2 \|\Delta_t\|_*^2}{\sum_{t=1}^{N} \gamma_t}. \quad (5.374)$$

Moreover, $\mathbb{E}\left[\Delta_t | \xi_{[t-1]}\right] = 0$ and v_t and x_t are functions of $\xi_{[t-1]}$, and hence

$$\mathbb{E}\big[(x_t - v_t)^{\mathrm{T}} \Delta_t\big] = \mathbb{E}\big\{(x_t - v_t)^{\mathrm{T}} \mathbb{E}[\Delta_t | \xi_{[t-1]}]\big\} = 0. \quad (5.375)$$

In view of condition (5.366), we have that $\mathbb{E}\left[\|\Delta_t\|_*^2\right] \le 4M_*^2$, and hence it follows from (5.374) and (5.375) that

$$\mathbb{E}\left[\max_{u\in X} \sum_{t=1}^{N} \nu_t \Delta_t^{\mathrm{T}}(x_t - u)\right] \le \frac{D_{\mathfrak{d},X}^2 + 2\kappa^{-1} M_*^2 \sum_{t=1}^{N} \gamma_t^2}{\sum_{t=1}^{N} \gamma_t}. \quad (5.376)$$

Therefore, by taking expectation of both sides of (5.372) and using (5.366) together with (5.376), we obtain (5.368).

In order to prove (5.369), let us observe that

$$\overline{f}^N - f^{*N} = \sum_{t=1}^{N} \nu_t (F(x_t, \xi^t) - f(x_t)),$$

and that for $1 \le s < t \le N$,

$$\begin{aligned}
&\mathbb{E}\big[(F(x_s, \xi^s) - f(x_s))(F(x_t, \xi^t) - f(x_t))\big] \\
&= \mathbb{E}\big\{\mathbb{E}\big[(F(x_s, \xi^s) - f(x_s))(F(x_t, \xi^t) - f(x_t))|\xi_{[t-1]}\big]\big\} \\
&= \mathbb{E}\big\{(F(x_s, \xi_s) - f(x_s))\mathbb{E}\big[(F(x_t, \xi^t) - f(x_t))|\xi_{[t-1]}\big]\big\} = 0.
\end{aligned}$$

Therefore

$$\begin{aligned}
\mathbb{E}\left[\left(\overline{f}^N - f^{*N}\right)^2\right] &= \sum_{t=1}^{N} \nu_t^2 \mathbb{E}\left[\left(F(x_t, \xi^t) - f(x_t)\right)^2\right] \\
&= \sum_{t=1}^{N} \nu_t^2 \mathbb{E}\left\{\mathbb{E}\left[\left(F(x_t, \xi^t) - f(x_t)\right)^2 | \xi_{[t-1]}\right]\right\} \\
&\le Q^2 \sum_{t=1}^{N} \nu_t^2,
\end{aligned} \quad (5.377)$$

where the last inequality is implied by condition (5.367). Since for any random variable Y we have that $\sqrt{\mathbb{E}[Y^2]} \ge \mathbb{E}[|Y|]$, the inequality (5.369) follows from (5.377).

Let us now look at (5.370). Denote

$$\tilde{f}^N(x) := \sum_{t=1}^{N} \nu_t [F(x_t, \xi^t) + G(x_t, \xi^t)^{\mathrm{T}}(x - x_t)].$$

Then

$$\left|\underline{f}^N - f_*^N\right| = \left|\min_{x\in X} \tilde{f}^N(x) - \min_{x\in X} f^N(x)\right| \le \max_{x\in X} \left|\tilde{f}^N(x) - f^N(x)\right|$$

and

$$\tilde{f}^N(x) - f^N(x) = \overline{f}^N - f^{*N} + \textstyle\sum_{t=1}^N \nu_t \Delta_t^\top (x_t - x),$$

and hence

$$\left|\underline{f}^N - f_*^N\right| \le \left|\overline{f}^N - f^{*N}\right| + \left|\max_{x\in X} \sum_{t=1}^N \nu_t \Delta_t^\top (x_t - x)\right|. \tag{5.378}$$

For $\mathbb{E}[|\overline{f}^N - f^{*N}|]$ we already have the estimate (5.369).

By (5.373) we have

$$\left|\max_{x\in X} \sum_{t=1}^N \nu_t \Delta_t^\top (x_t - x)\right| \le \left|\sum_{t=1}^N \nu_t \Delta_t^\top (x_t - v_t)\right| + \frac{D_{\mathfrak{d},X}^2 + (2\kappa)^{-1} \sum_{t=1}^N \gamma_t^2 \|\Delta_t\|_*^2}{\sum_{t=1}^N \gamma_t}. \tag{5.379}$$

Let us observe that for $1 \le s < t \le N$

$$\begin{aligned}\mathbb{E}\big[(\Delta_s^\top (x_s - v_s))(\Delta_t^\top (x_t - v_t))\big] &= \mathbb{E}\Big\{\mathbb{E}\big[(\Delta_s^\top (x_s - v_s))(\Delta_t^\top (x_t - v_t))|\xi_{[t-1]}\big]\Big\} \\ &= \mathbb{E}\Big\{(\Delta_s^\top (x_s - v_s))\mathbb{E}\big[(\Delta_t^\top (x_t - v_t))|\xi_{[t-1]}\big]\Big\} = 0.\end{aligned}$$

Therefore, by condition (5.366) we have

$$\begin{aligned}&\mathbb{E}\left[\left(\textstyle\sum_{t=1}^N \nu_t \Delta_t^\top (x_t - v_t)\right)^2\right] = \textstyle\sum_{t=1}^N \nu_t^2 \mathbb{E}\left[\left|\Delta_t^\top (x_t - v_t)\right|^2\right] \\ &\le \textstyle\sum_{t=1}^N \nu_t^2 \mathbb{E}\left[\|\Delta_t\|_*^2 \|x_t - v_t\|^2\right] = \textstyle\sum_{t=1}^N \nu_t^2 \mathbb{E}\left[\|x_t - v_t\|^2 \mathbb{E}[\|\Delta_t\|_*^2 | \xi_{[t-1]}\right] \\ &\le 4M_*^2 \textstyle\sum_{t=1}^N \nu_t^2 \mathbb{E}\left[\|x_t - v_t\|^2\right] \le 32\kappa^{-1} M_*^2 D_{\mathfrak{d},X}^2 \textstyle\sum_{t=1}^N \nu_t^2,\end{aligned}$$

where the last inequality follows by (5.337). It follows that

$$\mathbb{E}\left[\left|\textstyle\sum_{t=1}^N \nu_t \Delta_t^\top (x_t - v_t)\right|\right] \le 4\sqrt{2}\kappa^{-1/2} M_* D_{\mathfrak{d},X} \sqrt{\textstyle\sum_{t=1}^N \nu_t^2}. \tag{5.380}$$

Putting together (5.378), (5.379), (5.380), and (5.369), we obtain (5.370). □

For the constant stepsize policy (5.343), all estimates given in the right-hand sides of (5.368), (5.369), and (5.370) are of order $O(N^{-1/2})$. It follows that under the specified conditions, the difference between the upper $\overline{f}^N$ and lower $\underline{f}^N$ bounds converges on average to zero, with increase of the sample size N, at a rate of $O(N^{-1/2})$. It is also possible to derive respective large deviations rates of convergence (Lan, Nemirovski, and Shapiro [114]).

Remark 19. The lower SA bound $\underline{f}^N$ can be compared with the respective SAA bound $\hat{\vartheta}_N$ obtained by solving the corresponding SAA problem (see section 5.6.1). Suppose that the same sample $\xi^1, \ldots, \xi^N$ is employed for both the SA and the SAA method, that $F(\cdot, \xi)$ is convex for all $\xi \in \Xi$, and $G(x, \xi) \in \partial_x F(x, \xi)$ for all $(x, \xi) \in X \times \Xi$. By convexity of $F(\cdot, \xi)$ and definition of $\underline{f}^N$, we have

$$\begin{aligned}\hat{\vartheta}_N &= \min_{x\in X}\left\{N^{-1}\textstyle\sum_{t=1}^N F(x, \xi^t)\right\} \\ &\ge \min_{x\in X}\left\{\textstyle\sum_{t=1}^N \nu_t\big[F(x_t, \xi^t) + G(x_t, \xi^t)^\top (x - x_t)\big]\right\} = \underline{f}^N.\end{aligned} \tag{5.381}$$

Therefore, for the same sample, the SA lower bound $\underline{f}^N$ is weaker than the SAA lower bound $\hat{\vartheta}_N$. However, it should be noted that the SA lower bound can be computed much faster than the respective SAA lower bound.

Exercises

5.1. Suppose that set X is defined by constraints in the form (5.11) with constraint functions given as expectations as in (5.12) and the set X_N defined in (5.13). Show that if sample average functions $\hat{g}_{iN}$ converge uniformly to g_i w.p. 1 on a neighborhood of x and g_i are continuous, $i = 1, \ldots, p$, then condition (a) of Theorem 5.5 holds.

5.2. Specify regularity conditions under which equality (5.29) follows from (5.25).

5.3. Let $X \subset \mathbb{R}^n$ be a closed convex set. Show that the multifunction $x \mapsto \mathcal{N}_X(x)$ is closed.

5.4. Prove the following extension of Theorem 5.7. Let $g : \mathbb{R}^m \to \mathbb{R}$ be a continuously differentiable function, $F_i(x, \xi)$, $i = 1, \ldots, m$, be a random lower semicontinuous functions, $f_i(x) := \mathbb{E}[F_i(x, \xi)]$, $i = 1, \ldots, m$, $f(x) = (f_1(x), \ldots, f_m(x))$, X be a nonempty compact subset of $\mathbb{R}^n$, and consider the optimization problem

$$\operatorname*{Min}_{x \in X} g\left(f(x)\right). \tag{5.382}$$

Moreover, let $\xi^1, \ldots, \xi^N$ be an iid random sample, $\hat{f}_{iN}(x) := N^{-1} \sum_{j=1}^N F_i(x, \xi^j)$, $i = 1, \ldots, m$, $\hat{f}_N(x) = (\hat{f}_{1N}(x), \ldots, \hat{f}_{mN}(x))$ be the corresponding sample average functions, and

$$\operatorname*{Min}_{x \in X} g\left(\hat{f}_N(x)\right) \tag{5.383}$$

be the associated SAA problem. Suppose that conditions (A1) and (A2) (used in Theorem 5.7) hold for every function $F_i(x, \xi)$, $i = 1, \ldots, m$. Let ϑ^* and $\hat{\vartheta}_N$ be the optimal values of problems (5.382) and (5.383), respectively, and S be the set of optimal solutions of problem (5.382). Show that

$$\hat{\vartheta}_N - \vartheta^* = \inf_{x \in S} \left(\sum_{i=1}^m w_i(x)\big[\hat{f}_{iN}(x) - f_i(x)\big] \right) + o_p(N^{-1/2}), \tag{5.384}$$

where

$$w_i(x) := \frac{\partial g(y_1, \ldots, y_m)}{\partial y_i}\Big|_{y = f(x)}, \ i = 1, \ldots, m.$$

Moreover, if $S = \{\bar{x}\}$ is a singleton, then

$$N^{1/2}\left(\hat{\vartheta}_N - \vartheta^*\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2), \tag{5.385}$$

where $\bar{w}_i := w_i(\bar{x})$ and

$$\sigma^2 = \mathbb{V}\mathrm{ar}\big[\textstyle\sum_{i=1}^m \bar{w}_i F_i(\bar{x}, \xi)\big]. \tag{5.386}$$

Hint: Consider function $V : C(X) \times \cdots \times C(X) \to \mathbb{R}$ defined as $V(\psi_1, \ldots, \psi_m) := \inf_{x \in X} g(\psi_1(x), \ldots, \psi_m(x))$, and apply the functional CLT together with the Delta and Danskin theorems.

5.5. Consider matrix $\left[\begin{smallmatrix} H & A \\ A^\top & 0 \end{smallmatrix}\right]$ defined in (5.44). Assuming that matrix H is positive definite and matrix A has full column rank, verify that

$$\begin{bmatrix} H & A \\ A^\top & 0 \end{bmatrix}^{-1} = \begin{bmatrix} H^{-1} - H^{-1}A(A^\top H^{-1}A)^{-1}A^\top H^{-1} & H^{-1}A(A^\top H^{-1}A)^{-1} \\ (A^\top H^{-1}A)^{-1}A^\top H^{-1} & -(A^\top H^{-1}A)^{-1} \end{bmatrix}.$$

Using this identity write the asymptotic covariance matrix of $N^{1/2}\left[\begin{smallmatrix} \hat{x}_N - \bar{x} \\ \hat{\lambda}_N - \bar{\lambda} \end{smallmatrix}\right]$, given in (5.45), explicitly.

5.6. Consider the minimax stochastic problem (5.46), the corresponding SAA problem (5.47), and let

$$\Delta_N := \sup_{x \in X, y \in Y} \left| \hat{f}_N(x, y) - f(x, y) \right|. \tag{5.387}$$

(i) Show that $|\hat{\vartheta}_N - \vartheta^*| \leq \Delta_N$, and that if $\hat{x}_N$ is a δ-optimal solution of the SAA problem (5.47), then $\hat{x}_N$ is a $(\delta + 2\Delta_N)$-optimal solution of the minimax problem (5.46).
(ii) By using Theorem 7.65 conclude that, under appropriate regularity conditions, for any $\varepsilon > 0$ there exist positive constants $C = C(\varepsilon)$ and $\beta = \beta(\varepsilon)$ such that

$$\Pr\left\{ |\hat{\vartheta}_N - \vartheta^*| \geq \varepsilon \right\} \leq Ce^{-N\beta}. \tag{5.388}$$

(iii) By using bounds (7.216) and (7.217) derive an estimate, similar to (5.116), of the sample size N which guarantees with probability at least $1 - \alpha$ that a δ-optimal solution $\hat{x}_N$ of the SAA problem (5.47) is an ε-optimal solution of the minimax problem (5.46). Specify required regularity conditions.

5.7. Consider the multistage SAA method based on iid conditional sampling. For corresponding sample sizes $\mathcal{N} = (N_1, \ldots, N_{T-1})$ and $\mathcal{N}' = (N'_1, \ldots, N'_{T-1})$, we say that $\mathcal{N}' \succeq \mathcal{N}$ if $N'_t \geq N_t$, $t = 1, \ldots, T-1$. Let $\hat{\vartheta}_{\mathcal{N}}$ and $\hat{\vartheta}_{\mathcal{N}'}$ be respective optimal (minimal) values of SAA problems. Show that if $\mathcal{N}' \succeq \mathcal{N}$, then $\mathbb{E}[\hat{\vartheta}_{\mathcal{N}'}] \geq \mathbb{E}[\hat{\vartheta}_{\mathcal{N}}]$.

5.8. Consider the chance constrained problem

$$\operatorname*{Min}_{x \in X} f(x) \ \text{ s.t. } \Pr\big\{T(\xi)x + h(\xi) \in C\big\} \geq 1 - \alpha, \tag{5.389}$$

where $X \subset \mathbb{R}^n$ is a closed convex set, $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function, $C \subset \mathbb{R}^m$ is a convex closed set, $\alpha \in (0, 1)$, and matrix $T(\xi)$ and vector $h(\xi)$ are functions of random vector ξ. For example, if

$$C := \left\{ z : z = -Wy - w, \ y \in \mathbb{R}^\ell, \ w \in \mathbb{R}^m_+ \right\}, \tag{5.390}$$

then, for a given $x \in X$, the constraint $T(\xi)x + h(\xi) \in C$ means that the system $Wy + T(\xi)x + h(\xi) \leq 0$ has a feasible solution. Extend the results of section 5.7 to the setting of problem (5.389).

5.9. Consider the following extension of the chance constrained problem (5.196):

$$\operatorname*{Min}_{x\in X} f(x) \ \text{ s.t. } p_i(x) \le \alpha_i, \ i = 1, \ldots, p, \tag{5.391}$$

with several (individual) chance constraints. Here $X \subset \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$, $\alpha_i \in (0, 1)$, $i = 1, \ldots, p$, are given significance levels, and

$$p_i(x) = \Pr\{C_i(x, \xi) > 0\}, \quad i = 1, \ldots, p,$$

with $C_i(x, \xi)$ being Carathéodory functions.
Extend the methodology of constructing lower and upper bounds, discussed in section 5.7.2, to the above problem (5.391). Use SAA problems based on *independent* samples. (See Remark 6 on page 162 and (5.18) in particular.) That is, estimate $p_i(x)$ by

$$\hat{p}_{iN_i}(x) := \frac{1}{N_i} \sum_{j=1}^{N_i} \mathbf{1}_{(0,\infty)}\big(C_i(x, \xi^{ij})\big), \ i = 1, \ldots, p.$$

In order to verify feasibility of a point $\bar{x} \in X$, show that

$$\Pr\big\{p_i(\bar{x}) < U_i(\bar{x}), \ i = 1, \ldots, p\big\} \ge \prod_{i=1}^{p}(1 - \beta_i),$$

where $\beta_i \in (0, 1)$ are chosen constants and

$$U_i(\bar{x}) := \sup_{\rho\in[0,1]} \{\rho : \mathfrak{b}(\mathfrak{m}_i; \rho, N_i) \ge \beta_i\}, \quad i = 1, \ldots, p,$$

with $\mathfrak{m}_i := \hat{p}_{iN_i}(\bar{x})$.
In order to construct a lower bound, generate M independent realizations of the corresponding SAA problems, each of the same sample size $\mathcal{N} = (N_1, \ldots, N_p)$ and significance levels $\gamma_i \in [0, 1)$, $i = 1, \ldots, p$, and compute their optimal values $\hat{\vartheta}^1_{\gamma,\mathcal{N}}, \ldots, \hat{\vartheta}^M_{\gamma,\mathcal{N}}$. Arrange these values in the increasing order $\hat{\vartheta}^{(1)}_{\gamma,\mathcal{N}} \le \cdots \le \hat{\vartheta}^{(M)}_{\gamma,\mathcal{N}}$. Given significance level $\beta \in (0, 1)$, consider the following rule for choice of the corresponding integer L:

- Choose the largest integer $L \in \{1, \ldots, M\}$ such that

$$\mathfrak{b}(L - 1; \theta_{\mathcal{N}}, M) \le \beta, \tag{5.392}$$

where $\theta_{\mathcal{N}} := \prod_{i=1}^{p} \mathfrak{b}(r_i; \alpha_i, N_i)$ and $r_i := \lfloor \gamma_i N_i \rfloor$.

Show that with probability at least $1 - \beta$, the random quantity $\hat{\vartheta}^{(L)}_{\gamma,\mathcal{N}}$ gives a lower bound for the true optimal value ϑ^*.

5.10. Consider the SAA problem (5.241) giving an approximation of the first stage of the corresponding three stage stochastic program. Let

$$\tilde{\vartheta}_{N_1,N_2} := \inf_{x_1\in\mathcal{X}_1} \tilde{f}_{N_1,N_2}(x_1)$$

be the optimal value and $\tilde{x}_{N_1,N_2}$ be an optimal solution of problem (5.241). Consider asymptotics of $\tilde{\vartheta}_{N_1,N_2}$ and $\tilde{x}_{N_1,N_2}$ as N_1 tends to infinity while N_2 is *fixed*. Let $\vartheta^*_{N_2}$ be the optimal value and $\mathcal{S}_{N_2}$ be the set of optimal solutions of the problem

$$\operatorname*{Min}_{x_1\in\mathcal{X}_1}\left\{f_1(x_1)+\mathbb{E}\big[\hat{Q}_{2,N_2}(x_1,\xi_2^i)\big]\right\}, \tag{5.393}$$

where the expectation is taken with respect to the distribution of the random vector $\big(\xi_2^i,\xi_3^{i1},\ldots,\xi_3^{iN_2}\big)$.

(i) By using results of section 5.1.1 show that $\tilde{\vartheta}_{N_1,N_2}\to\vartheta^*_{N_2}$ w.p. 1 and distance from $\tilde{x}_{N_1,N_2}$ to $\mathcal{S}_{N_2}$ tends to 0 w.p. 1 as $N_1\to\infty$. Specify required regularity conditions.

(ii) Show that, under appropriate regularity conditions,

$$\tilde{\vartheta}_{N_1,N_2}=\inf_{x_1\in\mathcal{S}_{N_2}}\tilde{f}_{N_1,N_2}(x_1)+o_p\big(N_1^{-1/2}\big). \tag{5.394}$$

Conclude that if, moreover, $\mathcal{S}_{N_2}=\{\bar{x}_1\}$ is a singleton, then

$$N_1^{1/2}\big(\tilde{\vartheta}_{N_1,N_2}-\vartheta^*_{N_2}\big)\xrightarrow{\mathcal{D}}\mathcal{N}\big(0,\sigma^2(\bar{x}_1)\big), \tag{5.395}$$

where $\sigma^2(\bar{x}_1):=\mathbb{V}\mathrm{ar}\big[\hat{Q}_{2,N_2}(x_1,\xi_2^i)\big]$. *Hint*: Use Theorem 5.7.