Central Limit Theorem and Sample Complexity of Stationary Stochastic Programs

Alexander Shapiro **Yi Cheng**

School of Industrial and Systems Engineering,
Georgia Institute of Technology,
Atlanta, Georgia 30332-0205,
e-mail: ashapiro@isye.gatech.edu, cheng.yi@gatech.edu

Abstract. In this paper we discuss sample complexity of solving stationary stochastic programs by the Sample Average Approximation (SAA) method. We investigate this in the framework of Optimal Control (in discrete time) setting. In particular we derive a Central Limit Theorem type asymptotics for the optimal values of the SAA problems. The main conclusion is that the sample size, required to attain a given relative error of the SAA solution, is not sensitive to the discount factor, even if the discount factor is very close to one. We consider the risk neutral and risk averse settings. The presented numerical experiments confirm the theoretical analysis.

Key Words: multistage programs, optimal control, Central Limit Theorem, dynamic programming, Bellman equation, SDDP algorithm, risk averse approach

1 Introduction

Consider the following optimal control (in discrete time) infinite horizon problem

$$\begin{array}{cl} \min\limits_{u_t \in \mathcal{U}} & \mathbb{E}_P\big[\sum_{t=0}^{\infty} \gamma^t c(x_t, u_t, \xi_t)\big] \\ \text{s.t.} & x_{t+1} = F(x_t, u_t, \xi_t). \end{array} \tag{1.1}$$

Variables $x_t \in \mathbb{R}^n$ represent state of the system, $u_t \in \mathbb{R}^m$ are controls, $\xi_t \in \mathbb{R}^d$, $t = 0, ...,$ is a sequence of independent identically distributed (iid) random vectors (random noise or disturbances) with probability distribution P of ξ_t supported on set $\Xi \subset \mathbb{R}^d$, $c : \mathcal{X} \times \mathbb{R}^m \times \Xi \to \mathbb{R}$ is the cost function, $F : \mathcal{X} \times \mathbb{R}^m \times \Xi \to \mathcal{X}$ is a measurable mapping, $\mathcal{U} \subset \mathbb{R}^m$ and $\mathcal{X} \subset \mathbb{R}^n$ are nonempty closed sets, and $\gamma \in (0, 1)$ is the discount factor. Value x_0 is given (initial conditions). The notation $\mathbb{E}_P$ emphasises that the expectation is taken with respect of the probability distribution P of ξ_t. In such setting, problem (1.1) is the classical formulation of *stationary* optimal control (in discrete time) problem (e.g., [2]).

Problem (1.1) can be also considered in the framework of stochastic programming by viewing $y_t = (x_t, u_t)$ as decision variables (e.g., [6]). In case the problem is convex, it is possible to apply a Stochastic Dual Dynamic Programming (SDDP) cutting plane type algorithm for a numerical solution. For periodical infinite horizon stochastic programming problems such algorithms were recently discussed in [7], problem (1.1) can be viewed as a particular case of the periodical setting with the period of one. In order to solve (1.1) numerically the (generally continuous) distribution of the random process ξ_t should be discretized. The so-called Sample Average Approximation (SAA) method approaches this by generating a random sample of the (marginal) distribution of ξ_t by using Monte Carlo sampling techniques.

This raises the question of the involved sample complexity, i.e., how large should be the sample size N in order for the SAA problem to give an accurate approximation of the original problem. In some applications the discount factor γ is very close to one. It is well known that as the discount factor approaches one, it becomes more difficult to solve problem (1.1). For a given $\gamma \in (0, 1)$, the sample complexity of the discretization is discussed in [7], with the derived upper bound on the sample size N being of order $O((1-\gamma)^{-3}\varepsilon^{-2})$ as a function of the discount factor γ and the error level $\varepsilon > 0$. Since the optimal value of problem (1.1) increases at the rate of $O((1-\gamma)^{-1})$ as γ approaches one, in terms of the relative error $(1-\gamma)^{-1}\varepsilon$, this would imply the required sample size is of order $O((1-\gamma)^{-1})$ as a function of γ. This suggests that increasing γ from 0.99 to 0.999 would require to increase the sample size by the factor of 10 in order to achieve more or less the same relative accuracy of the SAA method. However, the above is just an *upper* bound and some numerical experiments indicate that the relative error of the SAA approach is not much sensitive to increase of the discount factor even when it is very close to one.

In this paper we will investigate the question of sample complexity from a different point of view. We are going to derive a Central Limit Theorem (CLT) type result for the optimal value of the Sample Average Approximation of problem (1.1). We demonstrate that the standard error (standard deviation) of the distribution of the optimal value of the SAA grows more or less at the same rate $O((1-\gamma)^{-1})$ as the respective optimal value. This supports the evidence of numerical experiments that variability of the sample error of the optimal values, measured in terms of the relative error, is not sensitive to increase of the discount factor, even when the discount factor is very close to one. We investigate both the risk neutral and risk averse settings.

The paper is organized as follows. In the next section we present the basic theoretical analysis for risk neutral and risk averse cases. In particular, we show how the statistical upper bound of the SDDP algorithm can be constructed in the risk averse case. In section 3 we discuss in detail the classical inventory model. Finally in section 4 we present results of numerical experiments.

2 General analysis

The (classical) Bellman equation for the value function, associated with problem (1.1), can be written as

$$V(x) = \inf_{u \in \mathcal{U}} \mathbb{E}_P\big[c(x,u,\xi) + \gamma V(F(x,u,\xi))\big], \ x \in \mathcal{X}. \tag{2.1}$$

Consider the following assumptions.

(A1) The cost function is *bounded*, i.e., there is a constant $\kappa > 0$ such that $|c(x,u,\xi)| \le \kappa$ for all $(x,u,\xi) \in \mathcal{X} \times \mathcal{U} \times \Xi$.

(A2) The function $c(\cdot,\cdot,\cdot)$ and the mapping $F(\cdot,\cdot,\cdot)$ are continuous on the set $\mathcal{X} \times \mathcal{U} \times \Xi$.

Let $\mathbb{B}(\mathcal{X})$ be the space of bounded functions $g : \mathcal{X} \to \mathbb{R}$ equipped with the sup-norm $\|g\|_\infty = \sup_{x\in\mathcal{X}} |g(x)|$. Then, under the assumption (A1), $V(\cdot)$ is the fixed point of mapping $\mathcal{T} : \mathbb{B}(\mathcal{X}) \to \mathbb{B}(\mathcal{X})$ defined as

$$\mathcal{T}(g)(x) := \inf_{u \in \mathcal{U}} \mathbb{E}_P\big[c(x,u,\xi) + \gamma g(F(x,u,\xi))\big], \ g \in \mathbb{B}(\mathcal{X}). \tag{2.2}$$

As it is well known, the mapping $\mathcal{T}$ is a contraction mapping for $\gamma < 1$. Thus equations (2.1) have unique solution $\bar V$ (e.g., [2]). The corresponding optimal policy is given by $\bar u_t = \pi(x_t)$, $t = 0, ...,$ with

$$\pi(x) \in \arg\min_{u \in \mathcal{U}} \mathbb{E}_P\big[c(x,u,\xi) + \gamma \bar V(F(x,u,\xi))\big]. \tag{2.3}$$

For a given $x = x_0$ consider $\vartheta(P) := \bar V(x)$ viewed as a function of the probability measure P. Given a sample ξ^j, $j = 1, ..., N$, of the random vector ξ, consider the corresponding empirical measure[1] $\hat P_N = N^{-1}\sum_{j=1}^N \delta_{\xi^j}$. We are interested in the asymptotics of the value function $\hat V_N(x) = \vartheta(\hat P_N)$ of the corresponding SAA problem. That is, we would like to derive a Central Limit Theorem for $N^{1/2}(\hat V_N(x) - \bar V(x))$ for a fixed point $x \in \mathcal{X}$.

We can approach this problem in the following way. For a probability measure Q and $\tau \in [0,1]$, consider probability measure $(1-\tau)P + \tau Q = P + \tau(Q-P)$, and the directional derivative (if it exists)

$$\vartheta'(P, Q-P) := \lim_{\tau \downarrow 0} \frac{\vartheta(P + \tau(Q-P)) - \vartheta(P)}{\tau}. \tag{2.4}$$

[1] By δ_ξ we denote the measure of mass one at ξ.

Then we can use the approximation

$$\vartheta(\hat{P}_N) - \vartheta(P) \approx \vartheta'(P, \hat{P}_N - P). \tag{2.5}$$

This is the approach of Von Mises statistical functionals. It requires to compute the directional derivative (2.4), and consequently uses approximation (2.5) to derive the asymptotics. Even if this directional derivative does exist, the approximation (2.5) is a heuristic (this approach is routinely used in Statistics). In order to justify the obtained asymptotics in a rigorous way often the functional Delta Theorem is employed, we will discuss this later.

To compute the directional derivative (2.4) we proceed as follows. Consider the set of optimal policies

$$\mathfrak{S}(x) := \arg\min_{u \in \mathcal{U}} \mathbb{E}_P\big[c(x, u, \xi) + \gamma \bar{V}(F(x, u, \xi))\big], \tag{2.6}$$

where $\bar{V}(\cdot)$ is the solution of Bellman equation (2.1). Under the assumptions (A1) and (A2) the value function $\bar{V}(\cdot)$ is continuous, and the set $\mathfrak{S}(x)$ is nonempty, provided the set $\mathcal{U}$ is nonempty and compact.

Note that for any $\pi(x) \in \mathfrak{S}(x)$, the value function of the true problem can be written as

$$\bar{V}(x) = \mathbb{E}_P\left[\sum_{t=0}^{\infty} \gamma^t c(x_t, \pi(x_t), \xi_t)\right], \tag{2.7}$$

with

$$x_{t+1} = F_t(x_t, \pi(x_t), \xi_t),\ x_0 = x,\ t \geq 0. \tag{2.8}$$

Consider the following formula for the directional derivative (2.4),

$$\vartheta'(P, Q - P) = \inf_{\pi(x) \in \mathfrak{S}(x)} \mathbb{E}_{Q-P}\left[\sum_{t=0}^{\infty} \gamma^t c(x_t, \pi(x_t), \xi_t)\right], \tag{2.9}$$

with initial value x_0. We will give a proof of formula (2.9) in some cases and discuss difficulties associated with a rigorous derivation of (2.9) for a general setting.

Since

$$\vartheta(P) = \mathbb{E}_P\left[\sum_{t=0}^{\infty} \gamma^t c(x_t, \pi(x_t), \xi_t)\right],\ \text{for } \pi(x) \in \mathfrak{S}(x), \tag{2.10}$$

by (2.5) this leads to the approximation

$$\begin{aligned} \vartheta(\hat{P}_N) - \vartheta(P) &\approx \inf_{\pi(x) \in \mathfrak{S}(x)} \mathbb{E}_{\hat{P}_N - P}\left[\sum_{t=0}^{\infty} \gamma^t c(x_t, \pi(x_t), \xi_t)\right] && (2.11) \\ &= \inf_{\pi(x) \in \mathfrak{S}(x)} \mathbb{E}_{\hat{P}_N}\left[\sum_{t=0}^{\infty} \gamma^t c(x_t, \pi(x_t), \xi_t)\right] - \vartheta(P). && (2.12) \end{aligned}$$

- In particular if $\mathfrak{S}(x) = \{\bar{\pi}(x)\}$ is a singleton for every $x \in \mathcal{X}$, then by the CLT the approximation (2.12) suggests that $N^{1/2}(\vartheta(\hat{P}_N) - \vartheta(P))$ converges in distribution to normal $\mathcal{N}(0, \sigma^2(x_0))$ with

$$\sigma^2(x_0) = \mathrm{Var}\left(\sum_{t=0}^{\infty} \gamma^t c(x_t, \bar{\pi}(x_t), \xi_t)\right). \tag{2.13}$$

Note that in the approximation (2.12) the set of optimal policies $\pi(x)$ is computed with respect to the distribution P of ξ_t, and that the variance in (2.13) is taken with respect to the distribution P and initial value x_0.

Consider an optimal policy $\pi(x) \in \mathfrak{S}(x)$ (for the true problem). Since this policy is feasible we have that

$$\vartheta(P+\tau(Q-P)) \le \mathbb{E}_{P+\tau(Q-P)}\left[\sum_{t=0}^{\infty} \gamma^t c(x_t, \pi(x_t), \xi_t)\right],\ \tau \in [0,1].$$

Together with (2.10) this implies

$$\limsup_{\tau \downarrow 0} \frac{\vartheta(P+\tau(Q-P)) - \vartheta(P)}{\tau} \le \inf_{\pi(x)\in\mathfrak{S}(x)} \mathbb{E}_{Q-P}\left[\sum_{t=0}^{\infty} \gamma^t c(x_t, \pi(x_t), \xi_t)\right]. \tag{2.14}$$

This gives the upper bound for the directional derivative. In order to derive the respective lower bound there is a need for some type of compactness condition.

Consider the set $\mathfrak{P}$ of measurable mappings $\pi : \mathcal{X} \to \mathcal{U}$. Equipped with the distance

$$d(\pi_1, \pi_2) := \sup_{x\in\mathcal{X}} \|\pi_1(x) - \pi_2(x)\|,$$

the set $\mathfrak{P}$ becomes a metric space. We can view any $\pi \in \mathfrak{P}$ as a policy for the considered infinite horizon problem. For a given distribution P, the optimal policy is obtained by choosing $\pi \in \mathfrak{P}$ which minimizes the right hand side of (2.3). Suppose that we can choose a subset $\mathfrak{P}^* \subset \mathfrak{P}$ such that by restricting the optimization to $\pi \in \mathfrak{P}^*$ the corresponding optimal value does not change for all probability measures of the form $P + \tau(Q-P)$, $\tau \in [0,1]$. We refer to such set $\mathfrak{P}^*$ as the *restricted* set, and to the corresponding metric space $(\mathfrak{P}^*, d)$ as the restricted metric space. Of course choice of the restricted set $\mathfrak{P}^*$ is associated with the probability measures P and Q. If we can choose the restricted the metric space $(\mathfrak{P}^*, d)$ to be compact, then we can proceed to the following proof.

Proposition 2.1 *Suppose that the assumptions* (A1) *and* (A2) *are satisfied and there exists the restricted* compact *metric space* $(\mathfrak{P}^*, d)$. *Then formula* (2.9) *holds.*

Proof. For $\tau \in [0,1]$ and policy $\pi \in \mathfrak{P}^*$ consider function

$$h(\tau, \pi) := \mathbb{E}_{P+\tau(Q-P)}\left[\sum_{t=0}^{\infty} \gamma^t c(x_t, \pi(x_t), \xi_t)\right],$$

with $x = x_0$ and $x_{t+1} = F(x_t, \pi(x_t), \xi_t)$ for $t \ge 0$. We have that

$$\frac{\partial h(\tau,\pi)}{\partial \tau} = \mathbb{E}_{Q-P}\left[\sum_{t=0}^{\infty} \gamma^t c(x_t, \pi(x_t), \xi_t)\right]. \tag{2.15}$$

By the Lebesgue dominated convergence theorem, the right hand side of (2.15) is continuous with respect to $\pi \in \mathfrak{P}^*$. Formula (2.9) now follows by Danskin's theorem (e.g., [3, Theorem

4.13]) applied to the function $h(\tau,\pi)$. ∎

The main technical difficulty in applying the above proposition is verification of existence of the restricted *compact* metric space $(\mathfrak{P}^*, d)$. Note that the metric space $(\mathfrak{P}, d)$ is compact if either the set $\mathcal{X}$ is finite and the set $\mathcal{U}$ is compact, or the set $\mathcal{U}$ is finite. In such cases we can take $\mathfrak{P}^* = \mathfrak{P}$.

2.1 Risk averse case

Let $\mathcal{R}$ be a law invariant coherent risk measure (cf., [1]), and consider the corresponding nested formulation of stationary inventory model. In that case Bellman equation can be written, similar to (3.3), as (e.g., [7])

$$V(x) = \inf_{u\in\mathcal{U}} \mathcal{R}\big[c(x,u,\xi) + \gamma V(F(x,u,\xi))\big],\ x \in \mathcal{X}. \tag{2.16}$$

For example we can consider the Average Value-at-Risk measure[2] (also called Conditional Value-at-Risk, Expected Shortfall, Expected Tail Loss)

$$\mathsf{AV@R}_\alpha(Z) = \inf_{\eta\in\mathbb{R}} \mathbb{E}_P\left\{\eta + \alpha^{-1}[Z-\eta]_+\right\},\ \alpha \in (0,1).$$

Then equation (2.16) takes the form

$$V(x) = \inf_{u\in\mathcal{U},\,\eta\in\mathbb{R}} \mathbb{E}_P\left\{\eta + \alpha^{-1}[c(x,u,\xi) + \gamma V(F(x,u,\xi)) - \eta]_+\right\}. \tag{2.17}$$

Let $(\bar\pi(x), \bar\eta(x))$ be an optimal solution of (2.17). Then the optimal value of the corresponding nested infinite horizon problem is given by

$$\mathbb{E}_P\left[\sum_{t=0}^{\infty} \gamma^t\Big(\bar\eta(x_t) + \alpha^{-1}[c(x_t,\bar\pi(x_t),\xi_t) - \bar\eta(x_t)]_+\Big)\right]. \tag{2.18}$$

Suppose that the optimal solution $(\bar\pi(x), \bar\eta(x))$ is unique for all $x \in \mathcal{X}$. By derivations similar to the risk neutral (expected value) case, this suggests that $N^{1/2}(\vartheta(\hat P_N) - \vartheta(P))$ converges in distribution to normal $\mathcal{N}(0, \sigma^2(x_0))$ with

$$\sigma^2(x_0) = \mathrm{Var}\left(\sum_{t=0}^{\infty} \gamma^t\Big(\bar\eta(x_t) + \alpha^{-1}[c(x_t,\bar\pi(x_t),\xi_t) - \bar\eta(x_t)]_+\Big)\right). \tag{2.19}$$

3 Inventory model

Consider the stationary inventory model (cf., [10])

$$\begin{array}{ll} \min\limits_{u_t\ge 0} & \mathbb{E}\left[\sum\limits_{t=0}^{\infty} \gamma^t\big(cu_t + b[D_t - (x_t+u_t)]_+ + h[x_t+u_t-D_t]_+\big)\right] \\ \text{s.t.} & x_{t+1} = x_t + u_t - D_t, \end{array} \tag{3.1}$$

[2] Recall that $[a]_+ = \max\{0, a\}$, for $a \in \mathbb{R}$.

where $c, b, h \in \mathbb{R}_+$ are the ordering cost, backorder penalty cost and holding cost per unit, respectively (with $b > c \geq 0$), x_t is the current inventory level, u_t is the order quantity, and $D_t \in \mathbb{R}_+$ is the demand at time t which is a random iid process. Then the optimal policy is myopic basestock policy $\bar{\pi}(x) = [x^* - x]_+$, where

$$x^* = F^{-1}\left(\frac{b - (1-\gamma)c}{b+h}\right) \tag{3.2}$$

with $F(x) = P(D \leq x)$ being the cdf of the demand (e.g., [10]). The optimal (basestock) policy is $\bar{u}_t = [x^* - x_t]_+$, and $x_{t+1} = x_t + \bar{u}_t - D_t$. That is $\bar{u}_t = x^* - x_t$ if $x_t \leq x^*$, and $\bar{u}_t = 0$ if $x_t \geq x^*$. Consequently $x_{t+1} = x^* - D_t$ if $x_t \leq x^*$, and $x_{t+1} = x_t - D_t$ if $x_t \geq x^*$.

The corresponding Bellman equation can be written as

$$V(x) = \inf_{u \geq 0} \mathbb{E}_P\big[cu + \psi(x+u, D) + \gamma V(x+u-D)\big], \ x \in \mathbb{R}, \tag{3.3}$$

with $D \sim P$ and

$$\psi(x, D) := b[D - x]_+ + h[x - D]_+.$$

Substituting $\bar{u}(x) = [x^* - x]_+$ into the right hand side of (3.3) we obtain,

$$V(x) = -cx + cx^* + \mathbb{E}_P\big[\psi(x^*, D) + \gamma V(x^* - D)\big], \ for \ x \leq x^*, \tag{3.4}$$

$$V(x) = \mathbb{E}_P\big[\psi(x, D) + \gamma V(x - D)\big], \ for \ x \geq x^*. \tag{3.5}$$

Since D is nonnegative we have that $x^* - D \leq x^*$, and hence by (3.4) that

$$V(x^* - D) = cD + \mathbb{E}_P\big[\psi(x^*, D) + \gamma V(x^* - D)\big].$$

It follows that for $x \leq x^*$,

$$V(x) = -cx + cx^* + \mathbb{E}_P\big[\gamma cD + \psi(x^*, D) + \gamma\psi(x^*, D) + \gamma^2 V(x^* - D))\big].$$

By continuing this process we obtain for $x \leq x^*$,

$$V(x) = -cx + (1-\gamma)^{-1}\mathbb{E}_P\big[\gamma cD + (1-\gamma)cx^* + \psi(x^*, D)\big]. \tag{3.6}$$

Note that $x^* \in \mathfrak{V}$, where

$$\mathfrak{V} := \arg\min_{x \in \mathbb{R}} \mathbb{E}_P\big[(1-\gamma)cx + \psi(x, D)\big].$$

Then by [6, Theorem 5.7] we have the following result.

Theorem 3.1 *For $x \leq x^*$ it holds that*

$$\hat{V}_N(x) = -cx + (1-\gamma)^{-1} \inf_{x \in \mathfrak{V}} \mathbb{E}_{\hat{P}_N}\big[\gamma cD + (1-\gamma)cx + \psi(x, D)\big] + o_p(N^{-1/2}). \tag{3.7}$$

In particular if the set $\mathfrak{V}$ is the singleton, i.e. the quantile x^* in (3.2) is unique, then $N^{1/2}(\hat{V}_N(x) - V(x))$ converges in distribution to normal $\mathcal{N}(0, \sigma^2)$ with

$$\sigma^2 = (1-\gamma)^{-2}\mathrm{Var}\big(\gamma cD + \psi(x^*, D)\big). \tag{3.8}$$

The variance in (3.8) is taken with respect to the distribution P of the demand. In the present case it was possible to derive the corresponding asymptotics of the form (2.12) in the rigorous way.

3.0.1 Risk averse case

Let $\mathcal{R}$ be a law invariant coherent risk measure and consider the corresponding nested formulation of stationary inventory model. In that case Bellman equation can be written, similar to (3.3), as

$$V(x) = \inf_{u \geq 0} \mathcal{R}\big[\psi(x, u, D) + \gamma V(x + u - D)\big], \ x \in \mathbb{R}. \tag{3.9}$$

For example we can consider the Average Value-at-Risk measure $\mathcal{R}(\cdot) := \mathsf{AV@R}_\alpha(\cdot)$. The base stock policy is optimal here as well with

$$x^* \in \arg\min_{x \in \mathbb{R}} \mathcal{R}\big(cx + \psi(x, D) + \gamma V(x - D)\big). \tag{3.10}$$

Counterparts of equations (3.4) and (3.5) follow here with the expectation $\mathbb{E}_P$ replaced by the risk measure $\mathcal{R}$.

4 Numerical illustration

In this section, we present numerical illustrations of the sample complexity and CLT for the stationary control problems for different values of the discount factor γ. Numerical experiments are performed on the stationary inventory problem and the Brazilian Interconnected Power System problem with the risk neutral and risk averse formulations (we refer to [8] for the description of the setting of the Brazilian problem).

4.1 Test cases and experimental settings

Inventory problem. The stationary inventory problem has deterministic ordering cost c, holding cost h and backlogging cost b, following the description in section 3, and iid demand process D_t. For the numerical test, the first stage is set to be deterministic with $D_1 = 5.5$ and initial state $x_0 = 10.0$. For the second stage and onwards, the model is stationary with $h = 0.2, b = 2.8, c = \cos(\frac{\pi}{3}) + 1.5$ and the demand is generated by

$$D_t = d + \phi \cdot \xi_t, \tag{4.1}$$

where $d = 9.0$, $\phi = 0.6$ and ξ_t is an iid sequence of random variables, each uniformly distributed on the interval $[0, 1]$.

Hydro-thermal planning problem. The hydro-thermal planning problem has larger scale than the inventory problem. The original problem has total number $T = 120$ of stages, and 4 state variables corresponding to the energy equivalent reservoirs of 4 interconnected regions. The random data process is characterized by the underlying stochastic monthly energy inflows. Specifically, the monthly inflows are sampled from a log-normal distributions estimated from the historical data and are assumed to be stagewise independent. We refer to [8] for more details of the problem. For illustration purpose, we assume that the energy inflows have period one. That is, the distribution of the inflows from the second stage and onwards (first stage is deterministic) are the same. In this way, the considered problem becomes stationary.

For the considered problems we discretize, at each stage, the continuous random variables using Monte Carlo sampling with N realizations per stage. This produces the approximation of the 'true' problem by its SAA counterpart. To illustrate the sample complexity for the stationary programs, we consider different sample sizes: $N = 10, 50, 100$. Besides, we perform numerical tests with different discount factors $\gamma = 0.8, 0.9, 0.9906$ and 0.999.

Note that each SAA problem is a function of the sample (of size N). By randomizing SAA problems M times (i.e., by generating M instances of the SAA problems, each with independently generated sample of size N), we obtain M optimal values of the SAA problems corresponding to different samples. By our analysis we expect that for N large enough, the optimal values of the SAA problems have approximately normal distribution. Variability of the SAA optimal values can be measured by the their standard deviation, which in turn can be estimated from the constructed sample of M repetitions.

That is, let $\hat{V}^{(r)}$ denote the optimal value (computed up to some precision) of the SAA problem related to the r-th sample, for $r = 1, ..., M$. The respective sample standard deviation is computed as

$$\hat{\sigma} := \sqrt{\tfrac{1}{M-1} \sum_{r=1}^{M} \left[\bar{V}_M - \hat{V}^{(r)}\right]^2}, \tag{4.2}$$

where $\bar{V}_M := \frac{1}{M} \sum_{r=1}^{M} \hat{V}^{(r)}$. In the numerical tests, we choose $M = 100$. We will present more details of how to compute $\hat{V}^{(r)}$ later in this section.

For each risk measure, a test instance is determined by selections of N and γ. We conduct the numerical experiments for each test instance in the following three steps. For numerical calculations we use the periodical SDDP type algorithm with period of one (cf., [7]).

1. Run the Periodical SDDP type algorithm to solve M SAA problems and obtain lower bounds for the optimal values $\hat{V}^{(r)}$, $r = 1, \ldots, M$.

2. Construct upper bounds for the SAA problems and compare with the respective lower bounds to check convergence. For risk neutral formulations, dual bounds are accessible for all sample sizes N and discount factors γ (cf., [9]). For risk averse formulations, only statistical upper bounds are available for discount factors $\gamma = 0.8, 0.9$ and all sample sizes N.

3. Compute sample standard deviation of the optimal values of the SAA problems according to (4.2).

4. For inventory problem, compute theoretical standard deviation for risk neutral case by (3.8) and risk averse case by using (3.10). Compare the results with those from step 3.

All implementations were written in `Python 3` using the `MSPPy` solver described in [4] and the `dualsddp` described in [9].

4.2 Risk neutral case

In this section, we report numerical results for the risk neutral formulation of the stationary inventory problem and the hydro-thermal planning problem.

In Table 1, we provide a summary of solving the SAA problem of the stationary inventory problem and the hydro-thermal planning problem for different test instances. The first two columns represent the parameters (sample size and discount factor) of the test case. Column 3 and 4 give the (deterministic) lower bounds (primal bounds) and upper bounds (dual bounds) of the problems. The last column reports the relative gap calculated by $\frac{\text{UB}-\text{LB}}{\text{LB}} \times 100\%$. Observe that for each sample size N, the gaps for different discount factors remain in low level. This shows that increasing the discount factor does not require to increase the sample size, even when the discount factor is very close to one, in order to achieve similar convergence in solving the SAA counterparts of the true problem.

Inventory problem				
N	γ	LB	UB	Gap(%)
10	0.8	67.210	67.238	4.17×10^{-2}
	0.9	158.196	158.283	5.5×10^{-2}
	0.9906	1928.66	1933.93	0.27
	0.999	18227.69	18408.19	0.993
50	0.8	67.941	67.98	5.74×10^{-2}
	0.9	159.84	159.93	5.63×10^{-2}
	0.9906	1947.93	1953.19	0.27
	0.999	18409.09	18629.29	1.19
100	0.8	68.032	68.06	4.12×10^{-2}
	0.9	160.05	160.13	4.998×10^{-2}
	0.9906	1950.32	1956.33	0.31
	0.999	18431.61	18675.72	1.32

Hydro-thermal planning problem				
N	γ	LB ($\cdot10^6$)	UB ($\cdot10^6$)	Gap(%)
50	0.8	1.2259	1.2287	0.23
	0.9	2.4518	2.4961	1.77
	0.9906	26.0858	26.2726	1.03
	0.999	243.5701	257.7629	5.5
100	0.8	1.2259	1.2276	0.14
	0.9	2.4519	2.5591	4.19
	0.9906	26.0863	26.3817	1.12
	0.999	243.5745	257.3031	5.33

Table 1: Risk neutral case: convergence of solving SAA problems.

In Table 2, we present the sample standard deviations computed by using $M = 100$ optimal values of the SAA problems. The first two columns of the table account for the parameters of the test instances. The third column displays the sample standard deviation of 100 optimal values of the SAA problems for each test instance according to (4.2). The last column is the result of multiplying the sample standard deviation by $(1-\gamma)$. Additionally in Table 3, we report for the inventory problem the theoretical standard deviation of the optimal value functions for each discount factor, which is computed according to (3.8). We make the following observations. First, the sample standard deviations of the optimal values of the SAA problems almost proportional to the factor $(1-\gamma)^{-1}$. Evidence can be found in the last column of Table 2, which demonstrates that for each N, the values of $\hat{\sigma}_N \cdot (1-\gamma)$ resemble each other for different discount factors. This is also the case for the theoretical standard deviations derived from the inventory model (see third column of Table 3). Second, the sample standard deviations are close to the theoretical ones. For the inventory problem, comparisons between $\hat{\sigma}_N \cdot (1-\gamma)$ in Table 2 and $\sigma \cdot (1-\gamma)/\sqrt{N}$ in Table 3 for each N and γ support such claim. For the hydro-thermal planning problem, the closed form of standard deviation of the optimal value function is not known, thus we only report the sample standard deviation.

As the empirical results suggest, the standard deviations for different discount factors, are proportional to $(1-\gamma)^{-1}$. It could be seen that convergence of the risk neutral SAA problems with different discount factors do not vary much.

4.3 Risk averse case

Numerical experiments for the risk averse case adopt the risk measure of weighted sum of the expectation and Average Value-at-Risk, with parameters λ (the weight parameter of $\mathsf{AV@R}_\alpha$) and α (the confidence level). For the inventory problem, we choose $\lambda = 0.2$, $\alpha = 0.05$; for the hydro-thermal planning problem, $\lambda = 0.5$, $\alpha = 0.05$. For the selected risk measure, the analogue of formula (2.18) is given by

$$\mathbb{E}_P\left[\sum_{t=0}^{\infty}\gamma^t\Big((1-\lambda)c(x_t,\bar{\pi}(x_t),\xi_t)+\lambda\left(\bar{\eta}(x_t)+\alpha^{-1}[c(x_t,\bar{\pi}(x_t),\xi_t)-\bar{\eta}(x_t)]_+\right)\Big)\right]. \tag{4.3}$$

We apply the risk averse SDDP algorithm with the biased sampling techniques (see [5]) to solve the SAA problems (these biased sampling techniques significantly enhanced rates of convergence of the numerical procedure). To construct the upper bounds, we compute the statistical upper bounds for the expected policy value in (4.3). Specifically, we replace $T = \infty$ with a large value of T in (4.3) to approximate the true policy value. Here, we choose $T = 120$ for the numerical experiments. For discount factors very close to one (e.g. $\gamma = 0.9906, 0.999$), it is very challenging to compute a valid statistical upper bound (see [9]). For this reason, we only provide statistical upper bounds for SAA problems with discount factors $\gamma = 0.8$ and $\gamma = 0.9$ in the risk averse case. When solving SAA problems with larger discounts ($\gamma = 0.9906, 0.999$), we adopt the stopping criteria as when the deterministic bounds (primal lower bounds) become stabilized.

Table 4 presents the lower bounds and 95% confidence intervals for the SAA problems (if applicable). The confidence intervals are computed based on the policy values evaluated on

Inventory problem			
N	γ	$\hat{\sigma}_N$	$\hat{\sigma}_N \cdot (1-\gamma)$
10	0.8	0.52779	0.10556
	0.9	1.05557	0.10556
	0.9906	11.2296	0.10556
	0.999	104.849	0.104849
50	0.8	0.25295	0.05059
	0.9	0.50590	0.05059
	0.9906	5.3819	0.05059
	0.999	50.2507	0.05025
100	0.8	0.16361	0.03272
	0.9	0.32722	0.03272
	0.9906	3.48112	0.03272
	11 0.999	32.5026	0.032503

Hydro-thermal planning problem			
N	γ	$\hat{\sigma}_N$	$\hat{\sigma}_N \cdot (1-\gamma)$
50	0.8	30.9749	6.195
	0.9	63.0804	6.3089
	0.9906	701.624	6.5959
	0.999	6552.6559	6.553
100	0.8	22.9013	4.5803
	0.9	45.3019	4.5302
	0.9906	516.5621	4.8557
	0.999	4852.4276	4.8524

Table 2: Risk neutral case: sample standard deviations of optimal values of $M = 100$ SAA problems.

the policy by generating 1000 sample paths. Gaps are computed via $\frac{\text{UB}-\text{LB}}{\text{LB}} \times 100\%$ where UB denotes the upper end of the confidence interval. In particular, Figure 1 compares the evolution of the lower and upper bounds for the SAA problems of the inventory problem with $\gamma = 0.8$ and different sample sizes N. The confidence intervals displayed in the figure are computed based on the policy values obtained from 6 forward passes per iteration. For illustrating purpose, we truncate the plots to some iteration larger than 0 to avoid displaying super large number of the upper bounds. We can see from the table and the figure that the constructed statistical bounds are indeed valid upper bounds for the risk averse formulations. Besides, for both problems with relatively small discount factors, the gaps are evident to show convergence.

Similar to table 2, table 5 reports the sample standard deviations of the optimal values collected from solving $M = 100$ risk-averse SAA problems for each test instance. Likewise,

γ	σ	$\sigma \cdot (1-\gamma)$	$\sigma \cdot (1-\gamma)/\sqrt{10}$	$\sigma \cdot (1-\gamma)/\sqrt{50}$	$\sigma \cdot (1-\gamma)/\sqrt{100}$
0.8	1.4596	0.2919	0.0923	0.0412	0.02919
0.9	3.1812	0.3181	0.1006	0.04499	0.03497
0.9906	36.9038	0.3469	0.1097	0.04906	0.03469
0.999	349.7119	0.3497	0.1106	0.04946	0.03497

Table 3: Risk neutral case: theoretical standard deviation of the optimal value function for the inventory problem.

for each N, the sample standard deviation almost proportional to $(1-\gamma)^{-1}$, which can be observed from the similar values in the column "$\hat{\sigma}_N \cdot (1-\gamma)$". For the inventory problem, by following the formula in (3.10), we can also compute the closed form of the standard deviation of the optimal value function under the risk measure mentioned above, where the base stock policy is optimal. Table 6 shows such theoretical standard deviations. By comparing values of $\hat{\sigma}_N \cdot (1-\gamma)$ in Table 5 and $\sigma \cdot (1-\gamma)/\sqrt{N}$ for each (γ, N) for the inventory problem, we conclude that the theoretical standard deviations and the sample ones are close to each other. Therefore, we come to the same conclusion as in section 4.2 that the standard deviations of the optimal value function with discount factor γ are almost proportional to the factor $(1-\gamma)^{-1}$.

References

[1] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9:203–228, 1999.

[2] D.P. Bertsekas and S.E. Shreve. *Stochastic Optimal Control, The Discrete Time Case*. Academic Press, New York, 1978.

[3] J. Frédéric Bonnans and Alexander Shapiro. *Perturbation Analysis of Optimization Problems*. Springer Series in Operations Research. Springer, 2000.

[4] L. Ding, S. Ahmed, and A. Shapiro. A python package for multi-stage stochastic programming. *Optimization online*, 2019.

[5] Rui Peng Liu and Alexander Shapiro. Risk neutral reformulation approach to risk averse stochastic programming. *European Journal of Operational Research*, 286(1):21–31, 2020.

[6] A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on Stochastic Programming: Modeling and Theory*. SIAM, Philadelphia, 2nd edition, 2014.

[7] A. Shapiro and L. Ding. Periodical multistage stochastic programs. *SIAM Journal on Optimization*, 2020.

[8] A. Shapiro, W. Tekaya, J.P. da Costa, and M. Pereira Soares. Risk neutral and risk averse stochastic dual dynamic programming method. *European Journal of Operational Research*, 224:375–391, 2013.

Inventory problem				
N	γ	LB	CI	Gap(%)
10	0.8	67.599	[67.58,67.62]	0.023
	0.9	158.995	[158.96,159.02]	0.014
	0.9906	1938.91	-	-
	0.999	18324.308	-	-
50	0.8	68.429	[68.41, 68.46]	0.045
	0.9	160.843	[160.81,160.89]	0.029
	0.9906	1962.44	-	-
	0.999	18529.947	-	-
100	0.8	68.493	[68.47,68.51]	0.029
	0.9	160.993	[160.95, 161.03]	0.023
	0.9906	1950.32	-	-
	0.999	18545.915	-	-
Hydro-thermal planning problem				
N	γ	LB ($\cdot 10^6$)	CI ($\cdot 10^6$)	Gap(%)
50	0.8	1.2259	[1.2268,1.2269]	0.08
	0.9	2.452	[2.453,2.454]	0.1
	0.9906	26.0902	-	-
	0.999	243.6119	-	-
100	0.8	1.226	[1.2272,1.2273]	0.11
	0.9	2.452	[2.454,2.455]	0.1
	0.9906	26.0926	-	-
	0.999	243.6346	-	-

Table 4: Risk averse case: convergence of solving SAA problems.

[9] Alexander Shapiro and Yi Cheng. Dual bounds for periodical stochastic programs. *Optimization online*, 2020.

[10] P.H. Zipkin. *Foundation of inventory management*. McGraw-Hill, 2000.

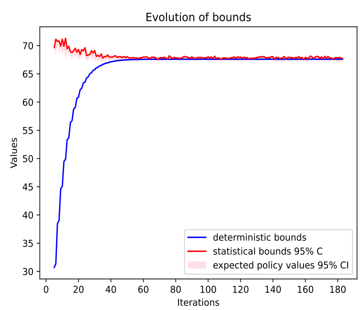
5 Appendix

Inventory problem			
N	γ	$\hat{\sigma}_N$	$\hat{\sigma}_N \cdot (1-\gamma)$
10	0.8	0.512	0.1024
	0.9	1.024	0.1024
	0.9906	10.894	0.1024
	0.999	101.721	0.1017
50	0.8	0.212	0.0424
	0.9	0.424	0.0424
	0.9906	4.508	0.0424
	0.999	42.089	0.0421
100	0.8	0.1354	0.02707
	0.9	0.2707	0.02707
	0.9906	2.8802	0.02707
	0.999	26.8921	0.0269
Hydro-thermal planning problem			
N	γ	$\hat{\sigma}_N$	$\hat{\sigma}_N \cdot (1-\gamma)$
50	0.8	74.856	14.971
	0.9	130.065	13.007
	0.9906	1664.465	15.646
	0.999	15542.136	15.542
100	0.8	50.097	10.019
	0.9	106.75	10.675
	0.9906	1145.777	10.77
	0.999	10698.883	10.7

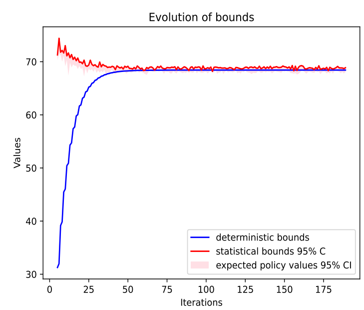
Table 5: Risk averse case: sample standard deviations of optimal values of $M = 100$ SAA problems.

γ	σ	$\sigma \cdot (1-\gamma)$	$\sigma \cdot (1-\gamma)/\sqrt{10}$	$\sigma \cdot (1-\gamma)/\sqrt{50}$	$\sigma \cdot (1-\gamma)/\sqrt{100}$
0.8	1.399	0.279	0.09	0.04	0.028
0.9	2.811	0.281	0.09	0.04	0.028
0.9906	36.122	0.339	0.107	0.048	0.034
0.999	340.041	0.34	0.107	0.048	0.034

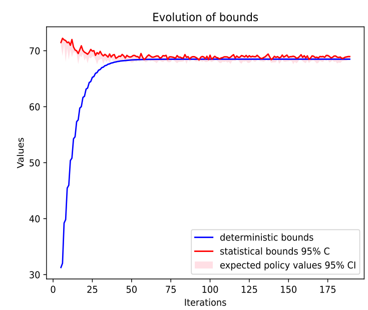
Table 6: Risk averse case: theoretical standard deviation of the optimal value function for the inventory problem.



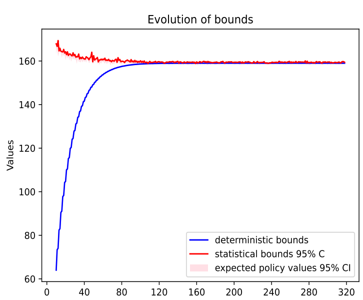
(a) $\gamma = 0.8, N = 10$



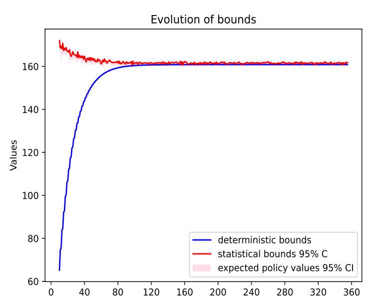
(b) $\gamma = 0.8, N = 50$



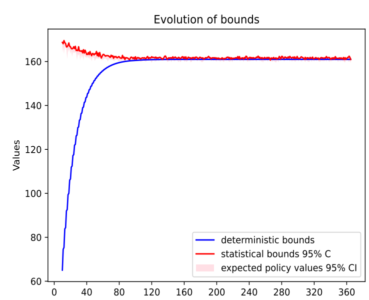
(c) $\gamma = 0.8, N = 100$



(d) $\gamma = 0.9, N = 10$



(e) $\gamma = 0.9, N = 50$



(f) $\gamma = 0.9, N = 100$

Figure 1: Risk averse case: bounds evolution of the SAA problems of the inventory problem.

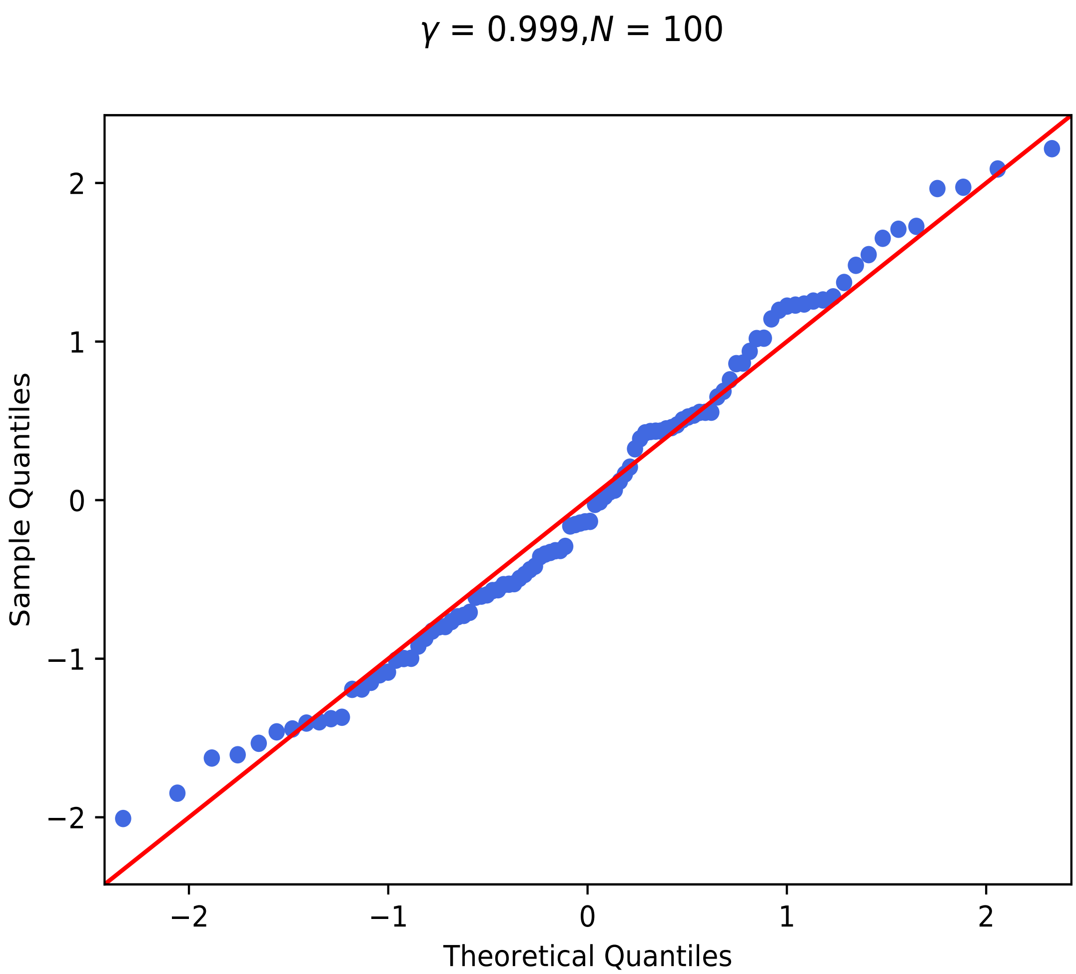


Figure 2: $\gamma = 0.999, N = 100$

Figure 3: Normal probability plot (Q-Q plot) for the optimal values of risk neutral hydro example

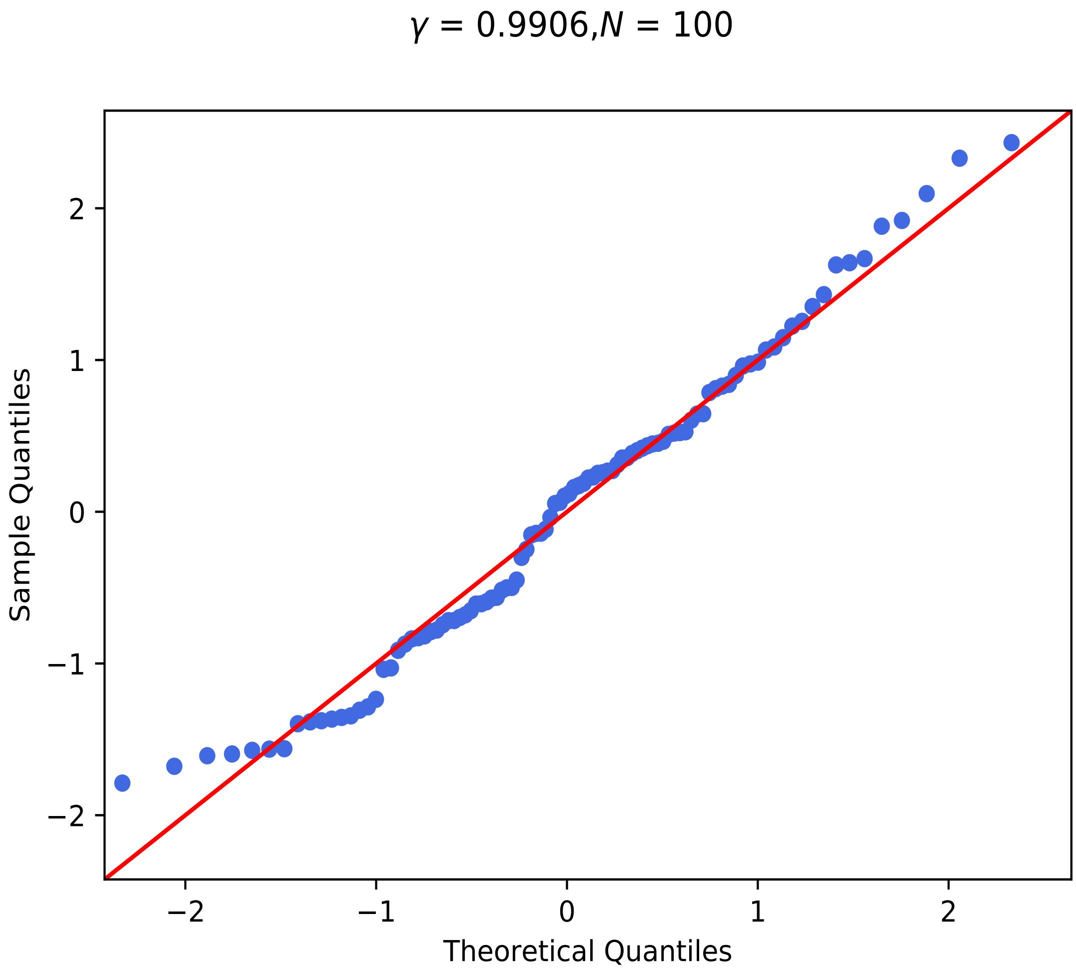


Figure 4: $\gamma = 0.9906, N = 100$

Figure 5: Normal probability plot (Q-Q plot) for the optimal values of risk averse hydro example