Stochastic programming

M2 MODO

Exam with solutions - January 15, 2024

Students can write their answers in either French or English.
If students suspect that a question contains an error, they shall indicate it explicitly on their paper, and proceed while taking this error into account.
Duration: 2 hours.
Contents: Three exercises on pages 2-4.
Allowed documents: One A4 sheet of handwritten or typed notes (two-sided).

Foreword

The exercises are meant to be independent, though notations may be carried out from one exercise to another.

- Scalars are denoted by lowercase letters: $a, b, c, \alpha, \beta, \gamma$.
- Vectors are denoted by **bold** lowercase letters: $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$.
- Matrices are denoted by **bold** uppercase letters: $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$.
- Sets are denoted by uppercase cursive letters : $\mathcal{A}, \mathcal{B}, \mathcal{C}$.
- Dimensions of vectors and matrices are always assumed to be greater than or equal to 1.
- Given a vector $\boldsymbol{x} \in \mathbb{R}^n$, the ith coordinate of this vector is denoted by $[\boldsymbol{x}]_i$.
- Given a vector $\boldsymbol{x} \in \mathbb{R}^n$, we use $\boldsymbol{x} \geq \mathbf{0}$ as a shortcut for $x_i \geq 0 \ \forall i = 1, \dots, n$.
- Given a random variable Z and a function φ, we use $\mathbb{E}_Z[\varphi(Z)]$ and $\mathbb{V}_Z[\varphi(Z)]$ to denote the expected value and the variance of $\varphi(Z)$, respectively.

Exercise 1: Two-stage linear stochastic programming

This exercise is concerned with two-stage linear stochastic programming problems of the form

$$\begin{array}{ll} \text{minimize}_{\boldsymbol{x}\in\mathbb{R}^n} & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \mathbb{E}_{\boldsymbol{\xi}}\left[Q(\boldsymbol{x},\boldsymbol{\xi})\right] \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \ge 0, \end{array} \tag{1}$$

where $\boldsymbol{\xi}$ is a random vector with values in Ξ, $\boldsymbol{c} \in \mathbb{R}^n$, $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$, and $Q : \mathbb{R}^n \times \Xi \to \mathbb{R}$ is defined by

$$Q(\boldsymbol{x},\boldsymbol{\xi}) = \min_{\boldsymbol{y}\in\mathbb{R}^{n_1}} \left\{\boldsymbol{q}^{\mathrm{T}}\boldsymbol{y} \text{ subject to } \boldsymbol{T}\boldsymbol{y} + \boldsymbol{W}\boldsymbol{x} = \boldsymbol{h}, \boldsymbol{y} \ge 0\right\},$$

where $\boldsymbol{q} \in \mathbb{R}^{n_1}$, $\boldsymbol{T} \in \mathbb{R}^{m_1\times n_1}$, $\boldsymbol{W} \in \mathbb{R}^{m_1\times n}$ and $\boldsymbol{h} \in \mathbb{R}^{m_1}$ all depend on the random vector $\boldsymbol{\xi}$. For simplicity, we will identify $\boldsymbol{\xi}$ with the t-uple $(\boldsymbol{q},\boldsymbol{T},\boldsymbol{W},\boldsymbol{h})$ in the exercise.

a) Justify that $\boldsymbol{x}$ is called the *here-and-now* decision while $\boldsymbol{y}$ is called the *wait-and-see* decision.

b) Rewrite problem (1) as a single, or one-stage, linear program.

c) Suppose that we use scenarios $\boldsymbol{\xi}^1,\dots,\boldsymbol{\xi}^K$ to represent the uncertainty in the problem, with associated probabilities $p^1,\dots,p^K$ so that $\sum_{k=1}^K p^k = 1$. Write down the resulting reformulation to (1) as a single linear program using one vector of wait-and-see variables per scenario and a single vector $\boldsymbol{x}$ of here-and-now variables.

d) The formulation from question c) is typically solved using the L-shaped method. What is the main idea behing this method?

e) An alternate formulation to that of question c) consists in having one copy of the here-and-now variables per scenario. In that case, what constraints need to be added to the problem, and why?

f) Name an algorithm that is applicable for solving the formulation of question e).

g) We now consider a multistage extension of the above framework with $T+1 > 2$ stages, which we write under the form

$$\begin{array}{ll} \text{minimize}_{\boldsymbol{x}_0\in\mathbb{R}^{n_0}} & \boldsymbol{c}_0^{\mathrm{T}}\boldsymbol{x}_0 + \mathbb{E}_{\boldsymbol{\xi}_1}\left[Q_0(\boldsymbol{x}_0,\boldsymbol{\xi}_1)\right] \\ \text{subject to} & \boldsymbol{A}_0\boldsymbol{x}_0 = \boldsymbol{b}_0 \\ & \boldsymbol{x}_0 \ge 0, \end{array} \tag{2}$$

where

$$\begin{array}{rcl} Q_0(\boldsymbol{x}_0,\boldsymbol{\xi}_1) &=& \displaystyle\min_{\boldsymbol{x}_1\in\mathbb{R}^{n_1}} \left\{\boldsymbol{c}_1^{\mathrm{T}}\boldsymbol{x}_1 + \mathbb{E}_{\boldsymbol{\xi}_2}\left[Q_1(\boldsymbol{x}_1,\boldsymbol{\xi}_2)\right] \text{ s.t } \boldsymbol{B}_1\boldsymbol{x}_0 + \boldsymbol{A}_1\boldsymbol{x}_1 = \boldsymbol{b}_1,\ \boldsymbol{x}_1 \ge 0\right\}, \\ Q_{t-1}(\boldsymbol{x}_{t-1},\boldsymbol{\xi}_t) &=& \displaystyle\min_{\boldsymbol{x}_t\in\mathbb{R}^{n_t}} \left\{\boldsymbol{c}_t^{\mathrm{T}}\boldsymbol{x}_t + \mathbb{E}_{\boldsymbol{\xi}_{t+1}}\left[Q_t(\boldsymbol{x}_t,\boldsymbol{\xi}_{t+1})\right] \text{ s.t. } \boldsymbol{B}_t\boldsymbol{x}_{t-1} + \boldsymbol{A}_t\boldsymbol{x}_t = \boldsymbol{b}_t,\ \boldsymbol{x}_t \ge 0\right\} \\ && \forall t = 2,\dots,T-1, \\ Q_{T-1}(\boldsymbol{x}_{T-1},\boldsymbol{\xi}_T) &=& \displaystyle\min_{\boldsymbol{x}_T\in\mathbb{R}^{n_T}} \left\{\boldsymbol{c}_T^{\mathrm{T}}\boldsymbol{x}_T \text{ s.t. } \boldsymbol{B}_T\boldsymbol{x}_{T-1} + \boldsymbol{A}_T\boldsymbol{x}_T = \boldsymbol{b}_T,\ \boldsymbol{x}_T \ge 0\right\}. \end{array}$$

i) What is the main computational difficulty with a scenario reformulation of problem (2)? How are those scenarios built?

ii) Since $T > 1$, what additional constraints must be added to these reformulations, and why?

Exercise 2: From expected value to risk measures

In this exercise, we consider a discrete random variable ξ with values in $\Xi = \{\xi_1, \xi_2\}$ such that $\mathbb{P}(\xi = \xi_1) = p$ and $\mathbb{P}(\xi = \xi_2) = 1 - p$ with $p \in (0,1)$. We then define the function $F : \mathbb{R}^2 \times \Xi \to \mathbb{R}$ by

$$\forall \boldsymbol{x} \in \mathbb{R}^2, \quad F(\boldsymbol{x}, \xi) = \begin{cases} -x_1 & \text{if } \xi = \xi_1 \\ 0 & \text{otherwise.} \end{cases}$$

a) Show that $\mathbb{E}_\xi[F(\boldsymbol{x},\xi)] = -p\,x_1$ and $\mathbb{V}_\xi[F(\boldsymbol{x},\xi)] = p(1-p)x_1^2$ for any $\boldsymbol{x} \in \mathbb{R}^2$.
Hint: Recall that $\mathbb{V}[Z] = \mathbb{E}[Z^2] - \mathbb{E}[Z]^2$ *for any random variable* Z.

b) In this question, we consider the problem

$$\begin{array}{ll} \text{minimize}_{\boldsymbol{x}\in\mathbb{R}^2} & \mathbb{E}_\xi[F(\boldsymbol{x},\xi)] \\ \text{subject to} & [\boldsymbol{x}]_1 + [\boldsymbol{x}]_2 = 1 \\ & \boldsymbol{x} \geq \mathbf{0}. \end{array} \tag{3}$$

i) Explain why problem (3) is called a *risk neutral* formulation.

ii) Justify that the solution of problem (3) is $\boldsymbol{x}^E = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

iii) Compute the expected wait-and see (EWS) and the expected value of perfect information (EVPI) associated with problem (3).

iv) What conclusion can you draw from the value of the EVPI?

c) In this question, we investigate the problem

$$\begin{array}{ll} \text{minimize}_{\boldsymbol{x}\in\mathbb{R}^2} & \mathbb{V}_\xi[F(\boldsymbol{x},\xi)] \\ \text{subject to} & [\boldsymbol{x}]_1 + [\boldsymbol{x}]_2 = 1 \\ & \boldsymbol{x} \geq \mathbf{0}. \end{array} \tag{4}$$

i) Justify that the solution to problem (4) is $\boldsymbol{x}^V = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

ii) Explain how this illustrates that problem (4) is a *risk-averse* formulation.

d) Finally, we set consider the family of problems parameterized by $\lambda > 0$:

$$\begin{array}{ll} \text{minimize}_{\boldsymbol{x}\in\mathbb{R}^2} & \rho_\lambda(F(\boldsymbol{x},\xi)) = \mathbb{E}_\xi[F(\boldsymbol{x},\xi)] + \lambda\,\mathbb{V}_\xi[F(\boldsymbol{x},\xi)] \\ \text{subject to} & [\boldsymbol{x}]_1 + [\boldsymbol{x}]_2 = 1 \\ & \boldsymbol{x} \geq \mathbf{0}. \end{array} \tag{5}$$

i) Show that $F(\boldsymbol{x}^E,\xi) \leq F(\boldsymbol{x}^V,\xi)$ for any $\xi \in \Xi$. but that $\rho_\lambda(F(\boldsymbol{x}^E,\xi)) > \rho_\lambda(F(\boldsymbol{x}^V,\xi))$ when $\lambda \geq \frac{1}{1-p}$.

ii) Using the expressions from question a), one can show that the solution of problem (5) is given by

$$\boldsymbol{x}^\lambda = \begin{cases} \begin{bmatrix} \frac{1}{2\lambda(1-p)} \\ 1 - \frac{1}{2\lambda(1-p)} \end{bmatrix} & \text{if } \lambda \geq \frac{1}{1-p} \\ \boldsymbol{x}^E & \text{otherwise.} \end{cases} \tag{6}$$

Compute the associated optimal value for problem (5).

iii) Show that this optimal value lies between that of problems (3) and (4). Could we have expected such a result?

Exercise 3: Chance constraints

In this exercise, we consider a chance-constrained problem of the form

$$\begin{array}{ll} \text{minimize}_{\boldsymbol{x}\in\mathbb{R}^n} & \mathbb{E}_\xi\left[F(\boldsymbol{x},\xi)\right] \\ \text{subject to} & \mathbb{P}\left(F(\boldsymbol{x},\xi)\geq\gamma\right)\leq 1-\alpha \\ & \boldsymbol{x}\in\mathcal{X}, \end{array} \tag{7}$$

where ξ is a random variable with values in $\Xi \subset \mathbb{R}$, $F : \mathbb{R}^n \times \Xi \to \mathbb{R}$, $\gamma > 0$, $\alpha \in (0, 1)$ and $\mathcal{X} \subset \mathbb{R}^n$ is a (deterministic) convex set.

a) Is the probabilistic constraint of problem (7) joint or individual?

b) *For this question only*, we suppose that $F(\boldsymbol{x}, \xi) = f(\boldsymbol{x}) - \xi$ where $f : \mathbb{R}^n \to \mathbb{R}$ and ξ is a random Gaussian variable.

   i) The probability density of ξ is log-concave. What does this imply on the feasible set of problem (7)?

   ii) Explain how the chance constraint of problem (7) can be reformulated as a deterministic constraint.

   iii) Can problem (7) be reformulated as a deterministic optimization problem without any random quantity? Justify your answer.

c) We now come back to the general setting, and consider scenarios $\xi^1, \ldots, \xi^K$ for the random variable ξ.

   i) Suppose first that all scenarios are associated with the same probability $\frac{1}{K}$. What quantity can then be used to approximate $\mathbb{P}\left(F(\boldsymbol{x}, \xi) \geq \gamma\right)$?

   ii) Using the quantity from the previous question, what class of (deterministic) optimization problems can (7) be reformulated into?

   iii) Suppose now that the scenarios are drawn at random, and that we replace the probabilistic constraint by the constraints

$$F(\boldsymbol{x}, \xi^k) \geq \gamma \qquad \forall k = 1, \ldots, K.$$

   Let $\boldsymbol{x}(\xi^{1:K})$ be a solution of the resulting optimization problem. For K large enough, what can we say about the feasibility of $\boldsymbol{x}(\xi^{1:K})$ with respect to the probabilistic constraint?

d) The specific form of the probabilistic constraint in problem (7) can be connected to a risk measure, namely the value-at-risk. Indeed, one has

$$\mathbb{P}\left(F(\boldsymbol{x}, \xi) \geq \gamma\right) \leq 1 - \alpha \qquad \Longleftrightarrow \qquad \mathrm{VaR}_\alpha\left[F(\boldsymbol{x}, \xi)\right] \leq \gamma.$$

   i) What is the difference between the value-at-risk and the conditional value-at-risk?

   ii) Recall the way of computing the value-at-risk described in class.

   iii) In the special case of question b) (i.e. $F(\boldsymbol{x}, \xi) = f(\boldsymbol{x}) - \xi$), compare the computing technique of question d)ii) with that described in question b)ii).

Solutions

Solutions for Exercise 1

a) The vector $\boldsymbol{x}$ represents a decision that must be made prior to knowing the value of the uncertainty. By contrast, the vector $\boldsymbol{y}$ is computed once the uncertainty is known.

b) The problem can be rewritten as

$$\begin{array}{ll} \text{minimize}_{\substack{\boldsymbol{x}\in\mathbb{R}^n \\ \boldsymbol{y}\in\mathbb{R}^{n_1}}} & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \mathbb{E}_{\boldsymbol{\xi}}\left[\boldsymbol{q}^{\mathrm{T}}\boldsymbol{y}\right] \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq 0 \\ & \boldsymbol{T}\boldsymbol{y} + \boldsymbol{W}\boldsymbol{x} = \boldsymbol{h} \\ & \boldsymbol{y} \geq \boldsymbol{0}. \end{array}$$

NB: Recall that the constraints involve uncertain coefficients, which is not apparent from this formulation.

c) Letting $\boldsymbol{y}_1, \dots, \boldsymbol{y}_k$ denote the wait-and-see variables for scenarios 1 to K, we obtain the formulation

$$\begin{array}{ll} \text{minimize}_{\substack{\boldsymbol{x}\in\mathbb{R}^n \\ \boldsymbol{y}_1,\dots,\boldsymbol{y}_K\in\mathbb{R}^{n_1}}} & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \sum_{k=1}^{K} p_k \boldsymbol{q}_k^{\mathrm{T}}\boldsymbol{y}_k \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq 0 \\ & \boldsymbol{T}_k\boldsymbol{y}_k + \boldsymbol{W}_k\boldsymbol{x} = \boldsymbol{h}_k \qquad k = 1, \dots, K \\ & \boldsymbol{y}_k \geq \boldsymbol{0} \qquad k = 1, \dots, K. \end{array}$$

d) The main idea behind the L-shaped method is to solve a manager problem over $\boldsymbol{x}$, then K subproblems in $\boldsymbol{y}_k$, that are independent when $\boldsymbol{x}$ is fixed. Depending on the solution of these problems (or lack thereof), constraints are added to the manager problem in $\boldsymbol{x}$, and the process is repeated until convergence.

e) If $\boldsymbol{x}_1, \dots, \boldsymbol{x}_K$ represent the copies of the here-and-now decision variables, one must enforce equality between those copies. This is done by adding nonanticipativity constraints (e.g. $\boldsymbol{x}_1 = \dots = \boldsymbol{x}_K$) to the problem.

f) Progressive hedging is applicable to solving the problem.

g) (Multistage)

i) The number of scenarios grows very rapidly as more stages are added. These scenarios are built through a scenario tree, and the number of scenarios thus tends to grow exponentially fast (a phenomenon sometimes referred to as "combinatorial explosion").

ii) To guarantee consistency between scenarios that are identical in early stages, nonanticipativy constraints must be added to the problem. For example, for any scenarios $\boldsymbol{\xi} = \{\boldsymbol{\xi}_t\}$ and $\hat{\boldsymbol{\xi}} = \{\hat{\boldsymbol{\xi}}_t\}$ such that the uncertainties on the first two stages are identical, we must have $\boldsymbol{x}_1(\boldsymbol{\xi}) = \boldsymbol{x}_1(\hat{\boldsymbol{\xi}})$ and $\boldsymbol{x}_2(\boldsymbol{\xi}) = \boldsymbol{x}_2(\hat{\boldsymbol{\xi}})$.

Solutions for Exercise 2

a) By using the definitions of expected value and variance, we have

$$\mathbb{E}_{\xi}\left[F(\boldsymbol{x},\xi)\right] = \mathbb{P}\left(\xi=\xi_1\right)F(\boldsymbol{x},\xi_1) + \mathbb{P}\left(\xi=\xi_2\right)F(\boldsymbol{x},\xi_2) = p\times(-x_1) + (1-p)\times 0 = -px_1$$

and

$$\mathbb{V}_{\xi}\left[F(\boldsymbol{x},\xi)\right] = \mathbb{E}_{\xi}\left[F(\boldsymbol{x},\xi)^2\right] - \mathbb{E}_{\xi}\left[F(\boldsymbol{x},\xi)\right]^2 = px_1^2 - p^2x_1^2 = p(1-p)x_1^2.$$

b) i) The problem (3) consists in minimizing the expected value, which does not account for extreme values of the quantity $F(\boldsymbol{x},\xi)$, which are viewed as risks (typically in finance). For this reason, it is called a *risk neutral* formulation.

ii) Problem (3) is a linear program. By either looking at the optimality conditions or solving graphically, one obtains that $\boldsymbol{x}^E = \begin{bmatrix}1\\0\end{bmatrix}$.

iii) The value of the recourse problem is the optimal value of problem (3), namely

$$RP = \mathbb{E}_{\xi}\left[F(\boldsymbol{x}^E,\xi)\right] = -p\,x_1^E = -p.$$

To compute the expected wait-and-see cost, we solve the problem once the uncertainty is known:

- If $\xi=\xi_1$, the problem becomes

$$\begin{array}{ll} \text{minimize}_{\boldsymbol{x}\in\mathbb{R}^2} & F(\boldsymbol{x},\xi_1) = -x_1 \\ \text{subject to} & [\boldsymbol{x}]_1 + [\boldsymbol{x}]_2 = 1 \\ & \boldsymbol{x} \geq \mathbf{0}, \end{array}$$

  whose solution (by the same argument than in the previous question) is $\boldsymbol{x}^E$, implying that the optimal value in that case is $WS_1 = -1$.

- If $\xi=\xi_2$, the wait-and-see problem is

$$\begin{array}{ll} \text{minimize}_{\boldsymbol{x}\in\mathbb{R}^2} & F(\boldsymbol{x},\xi_2) = 0 \\ \text{subject to} & [\boldsymbol{x}]_1 + [\boldsymbol{x}]_2 = 1 \\ & \boldsymbol{x} \geq \mathbf{0}, \end{array}$$

  and thus the optimal value of this (feasible!) problem is $WS_2 = 0$.

Overall, the expected wait-and-see cost is

$$WS = \mathbb{P}\left(\xi=\xi_1\right)\times WS_1 + \mathbb{P}\left(\xi=\xi_2\right)\times WS_2 =$$

As a result, the EVPI is $EVPI = WS - RP = 0$.

iv) Since the EVPI is 0, there is no interest in solving the problem prior to knowing the value of the uncertainty.

c) i) Problem (4) is a quadratic program, that can again be solved by inspection. Indeed, for any feasible point $\boldsymbol{x}$, we have that

$$\mathbb{V}_{\xi}\left[F(\boldsymbol{x},\xi)\right] = p(1-p)x_1^2 \geq 0 = \mathbb{V}_{\xi}\left[F(\boldsymbol{x}^V,\xi)\right].$$

Since $\boldsymbol{x}^V$ is feasible, it is a solution of the problem (and it is even unique since $p(1-p)x_1^2 = 0$ if and only if $x_1 = 0$).

ii) By choosing $\boldsymbol{x}^V$, one is guaranteed that $F(\boldsymbol{x}^V, \xi) = 0$ regardless of the uncertainty. This solution is thus of zero variance compared to that of problem (3).

d) i) We have

$$F(\boldsymbol{x}^E, \xi_1) = -p \leq F(\boldsymbol{x}^V, \xi_1) = 0 \quad \text{and} \quad F(\boldsymbol{x}^E, \xi_2) = 0 \leq F(\boldsymbol{x}^V, \xi_2) = 0.$$

On the other hand, we have

$$\rho_\lambda(F(\boldsymbol{x}, \xi)) = -px_1 + \lambda p(1-p)x_1^2,$$

hence

$$\rho_\lambda(F(\boldsymbol{x}^E, \xi)) = -p + \lambda p(1-p) \quad \text{and} \quad \rho_\lambda(F(\boldsymbol{x}^V, \xi)) = 0.$$

Since $\lambda > \frac{1}{1-p}$, we have $\rho_\lambda(F(\boldsymbol{x}^E, \xi)) > \rho_\lambda(F(\boldsymbol{x}^V, \xi))$.

ii) When $\lambda \geq \frac{1}{1-p}$, we have

$$\rho_\lambda(F(\boldsymbol{x}^\lambda, \xi))) = -p\frac{1}{2\lambda(1-p)} + \lambda p(1-p)\frac{1}{4\lambda^2(1-p)^2} = -\frac{p}{4\lambda(1-p)}.$$

Otherwise, we have

$$\rho_\lambda(F(\boldsymbol{x}^\lambda, \xi)) = \rho_\lambda(F(\boldsymbol{x}^E, \xi)) = -p + \lambda p(1-p)$$

iii) When $\lambda \leq \frac{1}{1-p}$, $\rho_\lambda(F(\boldsymbol{x}^\lambda, \xi)) \in (-p, 0]$, hence it lies between the optimal values of problems (3) and (4). When $\lambda > \frac{1}{1-p}$, we have

$$\rho_\lambda(F(\boldsymbol{x}^\lambda, \xi)) = -\frac{p}{4\lambda(1-p)} < 0,$$

and $\lambda > \frac{1}{1-p}$ implies $-\frac{p}{4\lambda(1-p)} > -\frac{p}{4} > -p$, hence the desired result holds.
Since the objective ρ_λ combines that of problems (3) and (4), this could have been expected.

Solutions for Exercise 3: Chance constraints

a) The probabilistic constraint is both joint and individual, since it consists of a single constraint.

b) i) Since the probability density of ξ is log-concave, the feasible set of problem (7) is convex.

ii) By using the cumulative density function, say π, of ξ, we have

$$\mathbb{P}\left(f(\boldsymbol{x}) - \xi \geq \gamma\right) \leq 1 - \alpha \quad \Leftrightarrow \quad \pi(f(\boldsymbol{x}) - \gamma) \leq 1 - \alpha.$$

iii) From the previous question, the feasible set of problem (7) can be reformulated as a deterministic set. However, the objective function involves an expected value, and thus this uncertainty cannot be removed a priori.

c) i) It is possible to approximate $\mathbb{P}\left(F(\boldsymbol{x}, \xi) \geq \gamma\right)$ by

$$\frac{1}{K}\left|\, k \in \{1, \ldots, K\} \,\middle|\, F(\boldsymbol{x}, \xi^k) \geq \gamma \,\right|.$$

ii) Using the quantity from the previous question, problem (7) can be reformulated as a mixed-integer program.

iii) For K large enough, the vector $\boldsymbol{x}(\xi^{1:K})$ is feasible with respect to the probabilistic constraint with high probability.

d) i) The value-at-risk represents a threshold beyond which larger values occur with probability less than $1-\alpha$. The conditional value-at-risk is the average value among those greater than the value-at-risk.

ii) The value-at-risk is given as the solution to an optimization problem, namely

$$\mathrm{VaR}_\alpha\left[F(\boldsymbol{x},\xi)\right] = \operatorname*{argmin}_{\gamma\in\mathbb{R}} \left\{\gamma + \frac{1}{1-\alpha}\,\mathbb{E}_Y\left[\max\{Y-\gamma, 0\}\right]\right\}.$$

iii) In the special case of question b), the technique of question d)ii) involves solving a linear two-stage stochastic program with continuous variables, which is typically less challenging than solving the mixed-integer program of question b)ii).