Optimization Modeling — Part 4/4

Stochastic models

Parin Chaipunya
KMUTT
└ Mathematics @ Faculty of Science
└ The Joint Graduate School of Energy and Environments

Areas of research:

- Multi-agent optimization: Bilevel programs, Game theory
- Optimization modeling: mainly focused on energy and environmental applications

A coordinated course with Metropolitan Electricity Authority (MEA) on STOCHASTIC OPTIMIZATION

by Parin CHAIPUNYA (KMUTT, Thailand) and Michel DE LARA (ENPC, France)

17 Nov – 18 Dec 2025

parin.cha@kmutt.ac.th
parinchaipunya.com
github.com/parinchaipunya

Overview

The main objective of this lecture is prepare the **thought process for stochastic optimization modeling**.

Table of contents

Section 1

Stochastic optimization

Subsection 1

The penetration of uncertainty

The penetration of uncertainty

In addition to the ingredients of an optimization model, now we introduce uncertainty.

Ingredient	Intuition	Expression
Decision variable	What you control.	$u \in \mathcal{U}$
Optimization space	The scope of what you control.	$\mathcal{U}$
Randomness	Uncertain part with known distribution.	$w \in \mathcal{W}$
Randomness space	The scope of randomness.	$(\mathcal{W}, \mathcal{F}, \mathbb{P})$
Objective function	The quantitative evaluation (outcome) of the decision.	$j : \mathcal{U} \times \mathcal{W} \to \mathbb{R}$
Constraints	Limitations, represented as a set of admissible decisions.	$U^{\mathrm{ad}} \subset \mathcal{U}$

Modeling randomness

Fundamentally, the randomness is expressed as a probability space $(\mathcal{W}, \mathcal{F}, \mathbb{P})$.

The randomness is usually interpreted as **nature**. This could be the natural phenomena like temperature, wind, sunlight, rain and resources. But it could as well be the uncertain human behaviors like demand.

Many times, we insist that there is only one true nature $(\mathcal{W}, \mathcal{F}, \mathbb{P})$ which is unknown to us. However, we may observe some portions of nature (temperature, wind, sunlight, etc.) using random variables and estemate their probability laws through data.

Objective function

With randomness taken into account, the objective function is now written as

$$j : \mathcal{U} \times \mathcal{W} \to \mathbb{R}.$$

It explicitly depends on two variables, a decision u and a realization of randomness w.

We intentionally use the lower case symbol j, rather than the earlier J, to highlight that $j(u, w)$ only represents the objective (*e.g.* cost or profit) at a single scenario, not the evaluation over all possible outcomes.

Subsection 2

Scenarios handling approaches

Scenarios handling approaches

Now that we have all the ingredients and aware of several scenarios, what to do with our modeling ?

Let us fix an idea with the problem of minimizing a cost function j where the randomness is a random variable W (and its generic value w) representing uncertain demand.

What problem are you going to solve ?

- Minimize $j(\cdot, w)$ by selecting w with highest probability. (Most-likely scenario approach)
- Minimize $j(\cdot, w)$ with all (or many) choices of w. (Scenario-based optimization)
- Minimize $\max_w j(\cdot, w)$. (Robust / worst-case optimization)
- Minimize $j(\cdot, \mathbb{E}[W])$. (Mean-value deterministic approximation)
- Minimize $\mathbb{E}_w[j(\cdot, w)]$. (Stochastic program with expectation)

Most-likely scenario approach

Minimize $j(\cdot, w)$ by selecting w with highest probability.

Intuition

You behave as if the world will realize its most probable state

Pros

- Simple and computationally cheap.
- Works quite well when the probability distribution has a sharp peak.

Cons

- Completely ignores low-probability but high-impact events.
- Fragile: one unexpected scenario breaks the solution.
- Not appropriate for safety-critical systems.

Where it is used

- Forecast-based operations (best-guess demand, best-guess price).
- When practitioners have a strong belief in the modal forecast.

Scenario-based optimization

Minimize $j(\cdot, w)$ with all (or many) choices of w.

Intuition
You compute solutions for all scenarios separately to compare them, not to choose one robust decision.

Pros
- You gain the understanding and see what happens in each scenarios.

Cons
- It is impossible to choose a decision among the cases.

Where it is used
- It is used to see and compare all scenarios.
- Possibly useful for $N-1$, $N-2$ contingency planning.

Robust / worst-case optimization

$$\text{Minimize } \max_w j(\cdot, w).$$

Intuition

Assume nature is adversarial and realize the worst possible scenario. The optimizer then chooses the action that performs best against this worst case.

Pros

- Strong protection against all risk.
- Guarantees performance in extreme cases.

Cons

- Could be too pessimistic.
- Ignores the likelihood of bad situations.
- Usually expensive in both practical and computational aspects.

Where it is used

- Robust design against extreme conditions.
- When the decision is made once and for all.

Mean-value deterministic approximation

$$\text{Minimize } j(\cdot, \mathbb{E}[W]).$$

Intuition

Replace the random variable by its mean. You optimize as if randomness averages out.

Pros

- Very common and easy to communicate.
- Simple model.
- Works quite will if the model is nearly linear.

Cons

- Could be badly wrong when the model is nonlinear.
- Ignores the distribution of the costs, variability and tail behaviors.
- Many times understood as if the randomness has been correctly adjusted for.

Where it is used

- Practitioners doing expected demand planning.
- When the model is nearly linear.

Stochastic program with expectatation

$$\text{Minimize } \mathbb{E}_w[j(\cdot, w)].$$

Intuition

Choose the decision that minimizes the average cost, weighted by scenario probabilities.

Pros

- Balances all scenarios proportionally to their likelihood.
- Probabilistically sound for risk-neutral decision-makers.
- Supported by the law of large numbers when repeating.

Cons

- May ignore rare but catastrophic events (unless weighted properly).
- Not suitable for risk-averse behavior unless modified (risk measures, chance constraints).

Where it is used

- Energy systems (unit commitment, hydro scheduling).
- Inventory management (like the newsvendor under expectation cost).
- Finance (portfolio optimization in the risk-neutral sense).

A note on stochastic program with expectation

In this approach, we must ensure that for each decision u, the function

$$j(u, \cdot) : \mathcal{W} \to \mathbb{R}$$

is a random variable. Hence, the expectation

$$J(u) = \mathbb{E}_w[j(u, w)]$$

is defined for each u.

It is this $J : \mathcal{U} \to \mathbb{R}$ that is the objective function of our stochastic program.

Property

If $j(\cdot, w)$ is convex for all w, then $J(\cdot) = \mathbb{E}_w[j(\cdot, w)]$ is convex.

Section 2

Modeling examples/exercises

Subsection 1

Blood testing

Blood testing — A natural stochastic problem

Story
A large number N (say, $N = 1000$) of possibly sick individuals are subjected to a blood test. Let the probability that an individual has a disease is $0 < p < 1$ (say, $p = 0.01$) and the sickness of individuals are independent.

Testing method:

- The blood **sample of u individuals are pooled together** and analyzed once.
  - If the pool test is **negative**, then the u individuals in the group are clean. Hence there is no further test in this scenario.
  - If the pool test is **positive**, then the u individuals in the group are tested separately. In this scenario, we require $u + 1$ tests.

What is the **optimal group size** u minimizes the **expected number of tests** ?

Blood testing — A natural stochastic problem

Notation	Description
Decision variables	
$u \in \{1, 2, \ldots, N\}$	The group size.
Randomness	
w	The sickness status of all individuals.
Objective function	
$j(u, w)$	The number of tests, given that each group contains u individuals (total of $\approx N/u$ groups) and given the sickness status w of all individuals.
Data	
N	Number of individuals.
p	Probability of having the disease.

Computing $\mathbb{E}_w[j(u, w)]$

Writing down $j(u, w)$ is a bit complicate, but its expectation is easier:

$$\widetilde{J}(u) = \mathbb{E}_w[j(u, w)] \approx \frac{N}{u} \times \Big(\underbrace{(1-p)^u}_{\text{probability that a pool is tested negative}} \times 1 \quad + \quad \overbrace{(1-(1-p)^u)}^{\text{probability that a pool is tested positive}} \times (u+1) \Big).$$

This $\widetilde{J}$ function is badly nonlinear.

For small p, we may use the binomial approximation

$$(1-p)^u \approx 1 - pu$$

to simplify $\widetilde{J}$ into

$$\widetilde{J}(u) \approx N \times \left(\frac{1}{u} + pu \right) =: J(u).$$

Solving for an optimal solution

If we simplify further by relaxing the integrality of u, simple calculus gives the minimizer $\bar{u}$ of J:

$$\bar{u} = \frac{1}{\sqrt{p}}.$$

This yields the approximated expected number of tests

$$J(\bar{u}) = 2\sqrt{p}N \ll N.$$

More analysis on Session 5.

Subsection 2

News boy problem

News boy problem

Story

A news vendor would purchase the newspapers in the morning at a unitary cost c and then resale it in the afternoon at a higher unitary price $p > c$.

If there are leftover newspapers, they are thrown away.

The demand each day is uncertain, but he made a record of the sales during the past month.

How can he decide optimally how many copies he should buy in the morning ?

News boy problem

Notation	Description
Decision variables	
$u \geq 0$	The copies to purchase.
Randomness	
$w \in \{w_1, \ldots, w_S\}$	Finite demand scenarios.
π_s	The probability of the outcome w_s.
Objective function	
$j(u, w) = cu - p \min\{u, w\}$	The negative revenue: cost $-$ sales.
Data	
c	Purchase cost.
p	Selling price.

Working out the objective function

The original objective function reads

$$j(u, w) = \underbrace{cu}_{\text{purchase cost}} - \underbrace{p \min\{u, w\}}_{\text{actual sales}}.$$

This is meaningful, but it is difficult to continue with this form.

Hence we do some calculation

$$\begin{aligned} j(u, w) &= cu - p \min\{u, w\} \\ &= cu - p \times \big(- \max\{-u, -w\}\big) \\ &= cu + p \max\{-u, -w\} \\ &= \max\Big\{cu + p \times (-u),\ cu + p \times (-w)\Big\} \\ &= \max\Big\{cu - pu,\ cu - pw\Big\}. \end{aligned}$$

This $j(\cdot, w)$ is a convex function, for any fixed w, as it is a maximum between two affine functions (in u).

Working out the objective function (even more)

Now that we have an expression

$$j(u, w) = \max\left\{cu - pu,\ cu - pw\right\}$$

for convexity, we can further reformulate it linearly with the help of two additional linear inequality constraints:

$$\begin{aligned} j(u, w) &= \max\left\{cu - pu,\ cu - pw\right\} \\ &= \min\left\{v \mid v \geq cu - pu,\ v \geq cu - pw\right\}. \end{aligned}$$

Working out the objective function (even more)

The optimization problem becomes

$$
\begin{aligned}
\min_{u \geq 0} \mathbb{E}[j(u, w)] &= \min_{u \geq 0} \mathbb{E}\left[\min\left\{v \mid v \geq cu - pu,\ v \geq cu - pw\right\}\right] \\
&= \min_{u \geq 0} \sum_{s=1}^{S} \pi_s \min\left\{v_s \mid v_s \geq cu - pu,\ v_s \geq cu - pw_s\right\} \\
&= \min_{u \geq 0}\ \min\left\{\sum_{s=1}^{S} \pi_s v_s \;\middle|\; v_s \geq cu - pu,\ v_s \geq cu - pw_s\right\} \\
&= \min\left\{\sum_{s=1}^{S} \pi_s v_s \;\middle|\; \begin{array}{l} v_s \geq cu - pu \\ v_s \geq cu - pw_s \\ u \geq 0 \end{array}\right\}
\end{aligned}
$$

Eventually, we have turned a stochastic program with convex objective into a linear program.

Subsection 3

Stochastic economic dispatch (single time-step)

Economic dispatch revisited.

First, let us look at the economic dispatch problem from Part 2.

Notation	Quantifier	Description
Sets		
I		The set of technologies, indexed with $i \in I$.
Parameters (Data)		
D		The known demand (MW).
C_i	$\forall i \in I$	Cost per MWh of technology i.
$P_i^{\max}$	$\forall i \in I$	Maximum capacity (MW).
Decision variables		
$u_i \geq 0$	$\forall i \in I$	How much the technology i outputs (MW).
Constraints		
$\sum_{i \in I} u_i = D$		The production meets exactly the demand.
$0 \leq u_i \leq P_i^{\max}$	$\forall i \in I$	Capacity limits of each technology.
Objective function		
$\sum_{i \in I} C_i u_i \leftarrow$ Minimize		The total generation cost summing up $C_i u_i = \underbrace{C_i}_{\text{cost per MWh}} \times \underbrace{u_i}_{\text{power output}} \times \underbrace{1}_{\text{one hour}}$

Economic dispatch revisited

From the table, we come up with the following optimization problem.

$$\begin{cases} \min\limits_{u} & \sum\limits_{i \in I} C_i u_i \\ \text{s.t.} & \sum\limits_{i \in I} u_i = D \\ & 0 \leq u_i \leq P_i^{\max} \quad \forall i \in I. \end{cases}$$

Is the stochastic counterpart immediate ? — **No.**

Next, we want to be more realistic by changing the fixed demand D into a random variable W. We assume that W has a finite image $\mathcal{W} = \{w_1, \dots, w_S\}$ with probability π_s assigned to w_s ($s \in \mathcal{S} = \{1, \dots, S\}$).
The **demand constraint** becomes

$$\sum_{i \in I} u_i = W,$$

which means

$$\sum_{i \in I} u_i = w_s, \quad \forall s \in \mathcal{S}.$$

This constraint is **impossible**, and the stochastic problem becomes infeasible before we even write a stochastic objective function.

Inequality relaxation ?

A natural reaction could be to **relax the equality constraints into equalities**:

$$\sum_{i \in I} u_i \geq w_s, \quad \forall s \in \mathcal{S}.$$

This will just implies that

$$\sum_{i \in I} u_i \geq \max_{s \in \mathcal{S}} w_s,$$

bringing us to the **worst-case scenario**.

Relax it with a recourse

A better remedy is to allow a recourse. In this problem, for each $s \in S$, we introduce recourse variables: $y_s^+ \geq 0$ for **upward recourse** (import) and $y_s^- \geq 0$ for **downward recourse** (curtail).

!! Note that the recourse variables are defined with uncertainty index s. This is because the recourse decisions are made **after** a randomness is realized.

The power balance now becomes

$$\sum_{i \in I} u_i + y_s^+ - y_s^- = w_s, \quad \forall s \in \mathcal{S}.$$

Be convinced that the recourses depend on the uncertainty (characterized by $s \in \mathcal{S}$).
Through this lens, they are **unknown random variables** $y^+, y^- : \mathcal{S} \to \mathbb{R}$.
To determine them, it is equivalent to fixing (y_s^+, y_s^-) for every $s \in \mathcal{S}$.

Adjustment of the objective function

The recourses $y = (y^+, y^-)$ are not for free. In fact, they are supposed to be **costly** with the cost vector (C^+, C^-).
Our new scenario-wise objective function is given by

$$j(u, y, w) = \sum_{i \in I} C_i u_i + C^+ y^+(w) + C^- y^-(w),$$

so that we have the stochastic objective

$$\begin{aligned} J(u, (y)_{s \in \mathcal{S}}) &= \mathbb{E}_w \left[j(u, y, w) \right] \\ &= \sum_{i \in I} C_i u_i + \mathbb{E}_w \left[C^+ y^+(w) + C^- y^-(w) \right] \\ &= \sum_{i \in I} C_i u_i + \sum_{s \in \mathcal{S}} \pi_s [C^+ y_s^+ + C^- y_s^-]. \end{aligned}$$

The stochastic economic dispatch model

Finally, we get the complete model for stochastic economic dispatch

$$
\begin{cases}
\min\limits_{(u_i)_{i \in I},\ (y_s)_{s \in \mathcal{S}}} & \sum\limits_{i \in I} C_i u_i + \sum\limits_{s \in \mathcal{S}} \pi_s [C^+ y_s^+ + C^- y_s^-] & \\
\text{s.t.} & \sum\limits_{i \in I} u_i + y_s^+ - y_s^- = w_s & \forall s \in \mathcal{S} \\
 & 0 \leq u_i \leq P_i^{\max} & \forall i \in I \\
 & 0 \leq y_s^+, y_s^- & \forall s \in \mathcal{S}.
\end{cases}
$$

Deterministic (averaged demand input) vs stochastic economic dispatch

The optimal solution from the deterministic model with averaged demand input suggests the following dispatch:

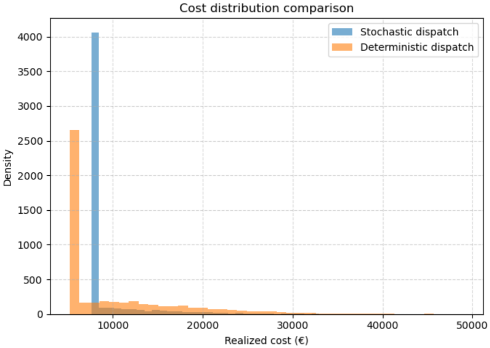
$$\text{Gas} = 60 \text{ MW}, \quad \text{Coal} = 40 \text{ MW}, \quad \text{Peaker} = 20 \text{ MW}.$$

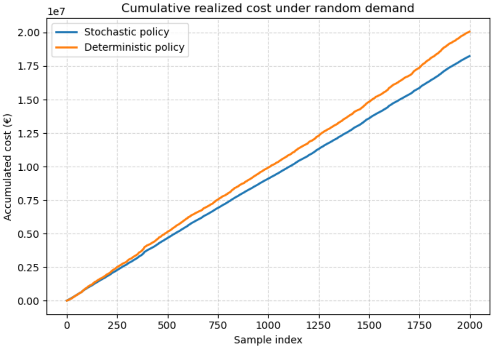
On the other hand, the stochastic model suggests the following higher production:

$$\text{Gas} = 60 \text{ MW}, \quad \text{Coal} = 40 \text{ MW}, \quad \text{Peaker} = 40 \text{ MW}.$$

Deterministic (averaged demand input) vs stochastic economic dispatch

The following Monte Carlo simulations show that the deterministic policy has a wide and heavy-tail cost distribution, while the **stochastic policy has tighter cost distribution**. Moreover, the **stochastic policy produces more savings** in accumulation compared to the deterministic one.





-» That's all of the 4 parts! «-