DM872

Mathematical Optimization at Work

Optimization under Uncertainty

Marco Chiarandini

Department of Mathematics & Computer Science
University of Southern Denmark

Outline

Multiple Plant Models

	Factory A		Factory B	
	Standard	Deluxe	Standard	Deluxe
(Machine 1) Grinding	4	2	5	3
(Machine 2) Polishing	2	5	5	6

$$
\begin{array}{llrcrcl}
\text{Maximize} & \text{Profit} & 10x_1 & + & 15x_2 & & \\
\text{Subject to} & \text{Raw A} & 4x_1 & + & 4x_2 & \leq & 75 \\
 & \text{Grinding A} & 4x_1 & + & 2x_2 & \leq & 80 \\
 & \text{Polishing A} & 2x_1 & + & 5x_2 & \leq & 60 \\
 & & & & x_1, x_2 & \geq & 0
\end{array}
$$

$$
\begin{array}{llrcrcl}
\text{Maximize} & \text{Profit} & 10x_3 & + & 15x_4 & & \\
\text{Subject to} & \text{Raw B} & 4x_3 & + & 4x_4 & \leq & 45 \\
 & \text{Grinding B} & 5x_3 & + & 3x_4 & \leq & 60 \\
 & \text{Polishing B} & 5x_3 & + & 6x_4 & \leq & 75 \\
 & & & & x_3, x_4 & \geq & 0
\end{array}
$$

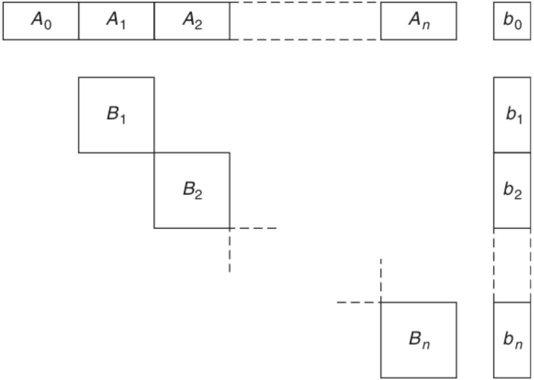
Multiple Plant Models

$$
\begin{array}{llrcrcrcrcr}
\text{Maximize} & \text{Profit} & 10x_1 & + & 15x_2 & + & 10x_3 & + & 15x_4 & & \\
\text{Subject to} & \text{Raw} & 4x_1 & + & 4x_2 & + & 4x_3 & + & 4x_4 & \leq & 120 \\
 & \text{Grinding A} & 4x_1 & + & 2x_2 & + & & & & \leq & 80 \\
 & \text{Polishing A} & 2x_1 & + & 5x_2 & + & & & & \leq & 60 \\
 & \text{Grinding B} & & & & + & 5x_3 & + & 3x_4 & \leq & 60 \\
 & \text{Polishing B} & & & & + & 5x_3 & + & 6x_4 & \leq & 75 \\
 & & & & & & \multicolumn{3}{r}{x_1, x_2, x_3, x_4} & \geq & 0
\end{array}
$$

10	15	10	15		
4	4	4	4	≤	120
4	2			≤	80
2	5			≤	60
		5	3	≤	60
		5	6	≤	75

allocation problems between plants + decision making within plants.

Block Angular Structure

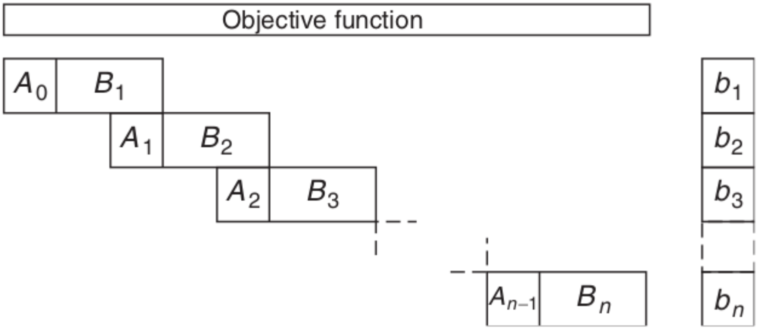


The rows $A_0, \ldots, A_n$ are known as common rows.
The diagonally placed blocks are known as submodels.

Staircase Structure

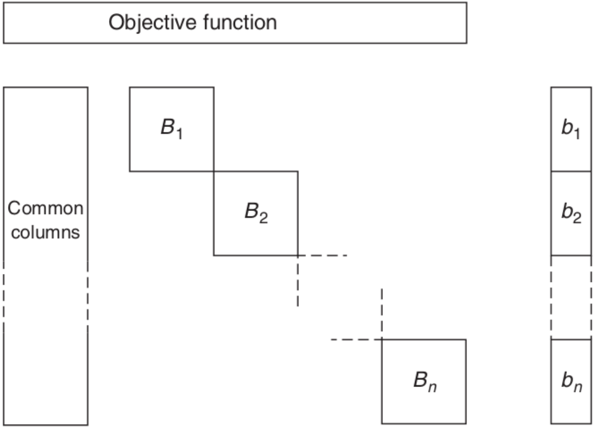
Multi-product and multi-period models lead also to staircase structures:

Amount in store at end of period $(t-1)$ + amount bought in period t =
amount used in period t + amount in store at end of period t



It can be converted into a block angular structure: alternate 'steps' such as $(A_0, B_1), (A_2, B_3)$ can be treated as subproblem constraints and the intermediate 'steps', eg. (A_1, B_1), as common rows.

Block Angular Structure



It can be seen as the dual of the common row structure. However, this structure arises often in stochastic programming cases and it can be treated in its own way.

Outline

Optimization under Uncertainty

Planning under uncertainty when data not known with certainty:

- inaccuracy of data (because of laziness of ignorance)
- multi-stage models where certain events, which need to be modelled, have not yet occurred.

Alternative approaches:

- sensitivity analysis, how solutions change with limited changes to data
- robust optimization, when we cannot quantify the uncertainty and the related risk. Stable solutions
- risk-averse (maximin, conditional value-at-risk): make the worst possible result as little bad as possible
- stochastic optimization, when uncertainty can be quantified.

Examples

Typical examples:

- News vendor problem
- Energy production
- Portfolio optimization
- Multi-period production planning

Stochastic programming

Stochastic programming (SP) is mathematical (i.e. linear, integer, mixed-integer, nonlinear) programming but with a stochastic element present in the data.

- in deterministic mathematical programming the data (coefficients) are known numbers
- in stochastic programming data are unknown, instead we may have a probability distribution present.

We consider two distinct stochastic programming problems:

- probabilistic constraints
- recourse problems.

The following slides are based on John E Beasley's OR-Notes on Stochastic Programming people.brunel.ac.uk/~mastjjb/jeb/or/sp.html)

Learn more about SP at https://www.stoprog.org/.

Outline

Probabilistic constraints

Suppose that we have two six-sided dice.
Die one gives a result a_1 when thrown and die two a result a_2.
Assuming the dice are fair we have discrete probability distributions for a_1 and a_2 as:

$$\begin{array}{lll} a_1 = i & (i = 1, \ldots, 6) & \text{with probability } 1/6 \\ a_2 = j & (j = 1, \ldots, 6) & \text{with probability } 1/6 \end{array}$$

Consider a simple LP with two variables and one constraint:

$$\begin{array}{lcccl} \text{minimise} & 5x & + & 6y & \\ \text{subject to:} & a_1 x & + & a_2 y & \geq 3 \\ & & & x, y & \geq 0 \end{array}$$

What does this LP mean?

Interpretation 1

One interpretation could be that we wish the constraint $a_1 x + a_2 y \geq 3$ to hold for all possible values of a_1 and a_2. Then we simply have a deterministic LP with two variables and 36 constraints:

$$\begin{array}{llcll} \text{minimise} & 5x & + & 6y & \\ \text{subject to:} & ix & + & jy & \geq 3 \qquad i = 1, \ldots, 6 \quad j = 1, \ldots, 6 \\ & & & x, y & \geq 0 \end{array}$$

Interpretation 2

Suppose now that we insist that the constraint $a_1x + a_2y \geq 3$ holds only with a specified probability $1 - \alpha$ (where $0 < \alpha < 1$).

For example $\alpha = 0.05$ would mean that we want the constraint $a_1x + a_2y \geq 3$ to hold with probability 0.95.
Chance constraint: A constraint needs not always be true now, rather it needs only be true, eg, 95% of the time.

$$
\begin{array}{ll}
\text{minimise} & 5x + 6y \\
\text{subject to:} & \text{Prob}(a_1x + a_2y \geq 3) \geq 1 - \alpha \\
 & x, y \geq 0
\end{array}
$$

Here, a_1 and a_2 are unknown, we merely have probability distribution information for them. We are required to choose values for x and y such that the objective function is minimised and the probability that the constraint $a_1x + a_2y \geq 3$ is satisfied is at least $1 - \alpha$.

Is this problem well-defined?

For each pair of values (a_1, a_2) we have an associated joint probability (1/36 in this simple case) then: given values for $x \geq 0$ and $y \geq 0$ we can easily check by enumeration whether the constraint is true with probability $1 - \alpha$.

Eg, for $x = 0, y = 1$ and $\alpha = 0.05$:

a_1	a_2	Is $a_1 0 + a_2 1 \geq 3$?	Probability
1	1	No	1/36
2	1	No	1/36
...			

We already have a probability of $2/36 = 0.0555$ that the constraint is infeasible. Hence, it is impossible for the constraint to be feasible with probability 0.95 (since 1-0.0555 $= 0.9445$). Hence, $x = 0, y = 1$ is not a solution to the problem.

Conceptually, we could simply enumerate all possible values for x and y and choose those values that minimise $5x + 6y$.
Hence, the problem is well defined.

Chance-constrained LP

This problem is an example of a stochastic (linear) program with probabilistic constraints. Such problems are also sometimes called chance-constrained linear programs:

- mix of probabilistic and deterministic coefficients in the same problem
- mix of probabilistic and deterministic constraints in the same problem.

To solve SP's with probabilistic constraints we transform them into an equivalent deterministic program. Note here however that even if the original SP is linear the equivalent deterministic program may not be.

Solving SP's with probabilistic constraints

Define zero-one variables z_{ij} using:

$$\begin{aligned} z_{ij} &= 1 \quad \text{if when } a_1 \text{ takes the value } i \ (i = 1, ..., 6) \text{ and} \\ &\qquad a_2 \text{ takes the value } j \ (j = 1, ..., 6) \ ix + jy \geq 3 \\ &= 0 \quad \text{otherwise} \end{aligned}$$

Let p_{ij} be the probability that a_1 takes the value i $(i = 1, ..., 6)$ and a_2 takes the j $(j = 1, ..., 6)$. That is, $p_{ij} = 1/36$.

The deterministic equivalent is:

$$\text{minimise } Mz_{ij} + (5x + 6y) \tag{1}$$

$$\text{subject to: } z_{ij} \geq [(ix + jy) - 3 + \delta]/M \qquad i = 1, ..., 6 \quad j = 1, ..., 6 \tag{2}$$

$$\sum_{i=1}^{6} \sum_{j=1}^{6} p_{ij} z_{ij} \geq 1 - \alpha \tag{3}$$

$$z_{ij} \in \{0, 1\} \qquad i = 1, ..., 6 \quad j = 1, ..., 6 \tag{4}$$

$$x, y \geq 0 \tag{5}$$

Continuous distributions

Suppose now that:

a_1 has a normal distribution with mean A_1 and standard deviation D_1, i.e. $\mathcal{N}(A_1, (D_1)^2)$;
a_2 has a normal distribution with mean A_2 and standard deviation D_2, i.e. $\mathcal{N}(A_2, (D_2)^2)$

and a_1 and a_2 independent.

$a_1 x + a_2 y \sim \mathcal{N}(A_1 x + A_2 y, [(D_1 x)^2 + (D_2 y)^2]^{1/2})$ because sum of normal distrs.

Hence, $\text{Prob}(a_1 x + a_2 y \geq 3) \geq 1 - \alpha$ can be addressed in the standard way for normal distribution probability calculations.

Let K be the value of the standard normal distribution $\mathcal{N}(0, 1)$ which has a probability of exactly α of being exceeded (e.g. if $\alpha = 0.025$ then $K = 1.96$). Such values are easily obtained from statistical tables.

$$\frac{3 - (A_1 x + A_2 y)}{\sqrt{(D_1 x)^2 + (D_2 y)^2}} \geq K$$

So our SP becomes a non linear program:

$$\begin{aligned} \text{minimise } & 5x + 6y \\ \text{subject to: } & 3 - (A_1 x + A_2 y) \geq K\sqrt{(D_1 x)^2 + (D_2 y)^2} \\ & x, y \geq 0 \end{aligned}$$

Outline

Recourse problems

In the simplest model of this type we have two stages:

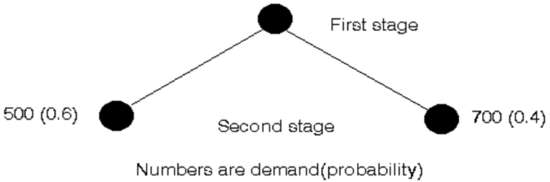
- in the first stage we make a decision
- in the second stage we see a realisation of the stochastic elements of the problem BUT are allowed to make further decisions to avoid the constraints of the problem becoming infeasible.

In the second stage the decisions that we make will be dependent upon the particular realisation of the stochastic elements observed.

Example

- We produce product X
- each unit of X that we make costs us 20 kr.
- X is made to meet demand from customers in the next time period.
- demand is stochastic, with a discrete probability distribution: demand $= D_s$ with probability p_s $(s = 1, ..., S)$. Informally, we can think of having S scenarios for possible future demand.
- customer demand must be met.
- we have the flexibility to buy in the product from an external supplier to meet observed customer demand but this costs us 30 kr per unit (i.e. we have recourse to an additional source of supply if demand exceeds production).
- How much should we choose to make now before we know what the customer demand is?

$S = 2$ and $D_1 = 500$, $p_1 = 0.6$; $D_2 = 700$, $p_2 = 0.4$.



If we were to produce 600 then if demand is 500 we are OK,
if demand is 700 we need recourse to an extra 100 units to meet it.

Two-stage model:

- action, make a decision (amount to produce)
- observation, observe a realisation of the stochastic elements (demand that occurs)
- reaction (recourse), further decisions, depending upon the realisation observed (extra production to meet demand if necessary)

Model

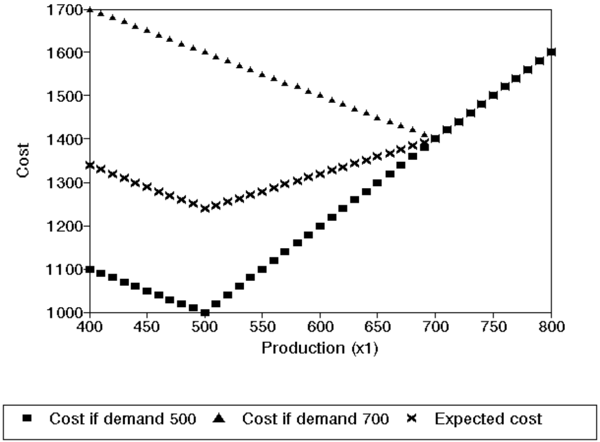
Let $y_{2s} \geq 0$ be the number of units of X to buy from the external supplier at the second stage in scenario s when the stochastic realisation of the demand is D_s $(s = 1, ..., S)$.

Goal: minimise total expected cost

$$\begin{aligned}
\text{minimise } & 2x_1 + \sum_{s=1}^{S} p_s(3y_{2s}) \\
\text{subject to } \quad & x_1 + y_{2s} \geq D_s \qquad s = 1, ..., S \\
& x_1 \geq 0 \\
& y_{2s} \geq 0 \qquad s = 1, ..., S
\end{aligned}$$

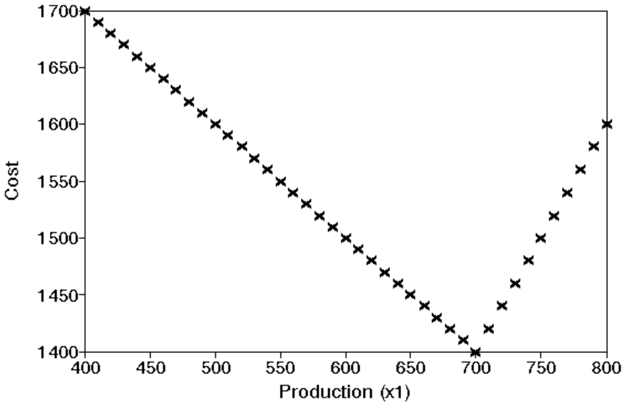
It is a deterministic program. We could require x_1 and y_{2s} to be integer.

Cost incurred



The production quantity that minimises expected cost is $x_1 = 500$.

Optimize the worst case

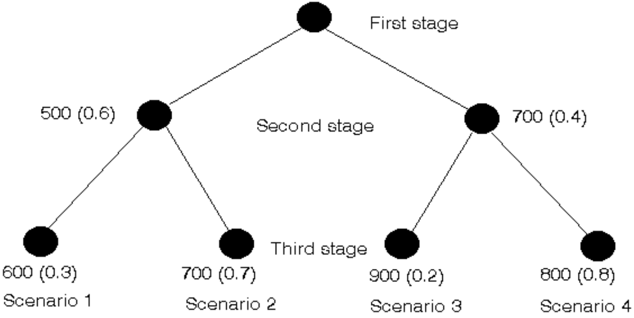


To minimise our worst case cost we should produce 700 now.

Multi-stage Optimization

- the stochastic elements have a discrete distribution
- the realisations of the stochastic elements are represented as a number of future scenarios

We look forward two periods into the future in planning production.



Two level (three stage) binary scenario tree

- We initially make a decision about how much to produce.
- At the second stage have two possible realisations of the stochastic demand:
  - a demand of 500 with probability 0.6
  - a demand of 700 with probability 0.4
- After this realisation we make a decision as to how much to produce to meet demand in the next period (the third-stage).
- At the third stage we again have two possible realisations of the stochastic demand, but these are different depending upon the realisation at the second stage. If second stage was 500 then:
  - a demand of 600 with probability 0.3
  - a demand of 700 with probability 0.7
- (at each level in the scenario tree the appropriate probabilities must sum to one)

This two-level scenario tree actually represents $2^2 = 4$ possible scenarios of the future:

Scenario	Second stage	Third stage	Probability
1	500	600	0.6(0.3) = 0.18
2	500	700	0.6(0.7) = 0.42
3	700	900	0.4(0.2) = 0.08
4	700	800	0.4(0.8) = 0.32

We have the following order of events:

- in the first stage a decision as to how much to produce; then
- in the second stage a realisation of the stochastic element (demand); then
- a decision as to the values of the recourse variables; then
- in the second stage a decision as to how much to produce; then
- in the third stage a realisation of the stochastic element (demand); and finally
- a decision as to the values of the recourse variables.

Model

- $x_1 \geq 0$ be the number of units of X to produce now (at the first stage)
- $y_{2s} \geq 0$ be the number of units of X to buy from the external supplier at the second stage in scenario s ($s = 1, ..., 4$)
- $x_{2s} \geq 0$ be the number of units of X to produce at the second stage in scenario s (s=1,...,4)
- $y_{3s} \geq 0$ be the number of units of X to buy from the external supplier at the third stage in scenario s (s=1,...,4)

At the first stage, the constraints to ensure customer demand is satisfied are:

```
x1 + y2s >= 500    (s=1,2)
x1 + y2s >= 700    (s=3,4)
```

At the second stage we will have units left over (i.e. inventory) to help meet future demand. This inventory level will be:

```
x1 + y2s - 500    (s=1,2)
x1 + y2s - 700    (s=3,4)
```

To ensure that demand is met in the third stage we have:
inventory $+$ amount produced $+$ amount bought externally $\geq$ demand

```
x1 + y2s - 500 + x2s + y3s >= 600   (s=1)
x1 + y2s - 500 + x2s + y3s >= 700   (s=2)
x1 + y2s - 700 + x2s + y3s >= 900   (s=3)
x1 + y2s - 700 + x2s + y3s >= 800   (s=4)
```

non-anticipativity constraints, scenarios with a common history must have the same set of decisions:

```
scenarios 1 and 2, second stage:
     y21=y22
     x21=x22
scenarios 3 and 4, second stage:
     y23=y24
     x23=x24
```

objective function: minimize expected costs

```
Scenario  Probability  Cost
1         0.18         2x21 + 3y21 + 3y31
2         0.42         2x22 + 3y22 + 3y32
3         0.08         2x23 + 3y23 + 3y33
4         0.32         2x24 + 3y24 + 3y34
```

Weighting each scenario cost by the associated scenario probability will give the expected cost.

```
minimise

2x1 + 0.18(2x21 + 3y21 + 3y31) + 0.42(2x22 + 3y22 + 3y32)
    + 0.08(2x23 + 3y23 + 3y33) + 0.32(2x24 + 3y24 + 3y34)
```

The full model


```
minimise

2x1 + 0.18(2x21 + 3y21 + 3y31) + 0.42(2x22 + 3y22 + 3y32)
    + 0.08(2x23 + 3y23 + 3y33) + 0.32(2x24 + 3y24 + 3y34)

subject to

x1 + y2s >= 500    (s=1,2)
x1 + y2s >= 700    (s=3,4)

x1 + y2s - 500 + x2s + y3s >= 600   (s=1)
x1 + y2s - 500 + x2s + y3s >= 700   (s=2)
x1 + y2s - 700 + x2s + y3s >= 900   (s=3)
x1 + y2s - 700 + x2s + y3s >= 800   (s=4)

y21=y22
x21=x22

y23=y24
x23=x24

all variables >=0
```

General Case

After taking a first stage decision, a random outcome (scenario) occurring with probability p_s involving one or more of the future data is observed. Then, an optimal second stage decision (recourse action) depending on the first stage and the scenario s is taken

Example:

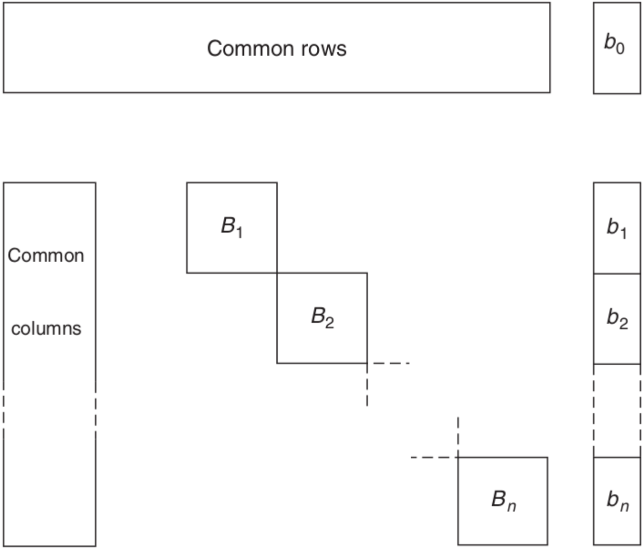
- (Stage 1): decide production before the demand and future prices (uncertain) are known.
- (Stage 2): decide whether to sell any excess production at a lower price or extra produce to make up a shortfall at a higher cost.

(stage 1 variables) Production decisions: $x_1, x_2, \ldots, x_n$.
(stage 2 variables) Excess production or shortfall: $y_1, y_2, \ldots, y_n$, $z_1, z_2, \ldots, z_n$
stage 2 variables will be replicated m times according to each of the possible demand levels $d_j^{(1)}, d_j^{(2)}, \ldots, d_j^{(m)}$ with given probabilities p_s to occurr.
c_j production costs
e_j excess costs (eg, storage)
f_j shortfall costs (missed opportunity)

Two-Stage Stochastic Program with Recursion

$$\begin{aligned}
\text{Minimize } & \sum_j c_j x_j + \sum_s p_s \left(\sum_j e_j y_j^{(s)} + \sum_j f_j z_j^{(s)} \right) & \\
\text{subject to } & \sum_j a_{ij} x_j \leq b_i & \text{for all production constraints } i \\
& x_j - y_j^{(s)} + z_j^{(s)} = d_j^{(s)} & \text{for all } j \text{ and } s \\
& x_j, y_j^{(s)}, z_j^{(s)} \geq 0 & \text{for all } j \text{ and } s
\end{aligned}$$

Structure



Alternative objectives

Minimise the maximum cost we would ever have to pay (minimise the maximum scenario cost).

```
Z >= 2x1 + (2x21 + 3y21 + 3y31)   scenario 1
Z >= 2x1 + (2x22 + 3y22 + 3y32)   scenario 2
Z >= 2x1 + (2x23 + 3y23 + 3y33)   scenario 3
Z >= 2x1 + (2x24 + 3y24 + 3y34)   scenario 4
```

The objective function would then become minimise Z

Alternative objectives

After minimising Z with scenarios that cost less than this maximum cost we may have flexibility about variable values.

Hence if Z* is the minimum value of Z from this formulation it is appropriate to then solve a further program:

```
minimise:    total scenario cost

subject to:  Z <= Z* and the same constraints as above

i.e. minimise

4(2x1) + (2x21 + 3y21 + 3y31) + (2x22 + 3y22 + 3y32)
       + (2x23 + 3y23 + 3y33) + (2x24 + 3y24 + 3y34)

subject to:   Z <= Z* and the same constraints as above
```

Further Material

- Tutorial by Giovanni Pantuso (more theoretical) https://pantuso.sites.ku.dk/talks/
- Example on VRP with stochastic demand
- Talk from gurobi, including robust, value-at-risk and conditional value-at-risk formulations https://www.gurobi.com/events/solving-simple-stochastic-optimization-problems-with-gurobi/
- Stochastic Programming community https://www.stoprog.org/