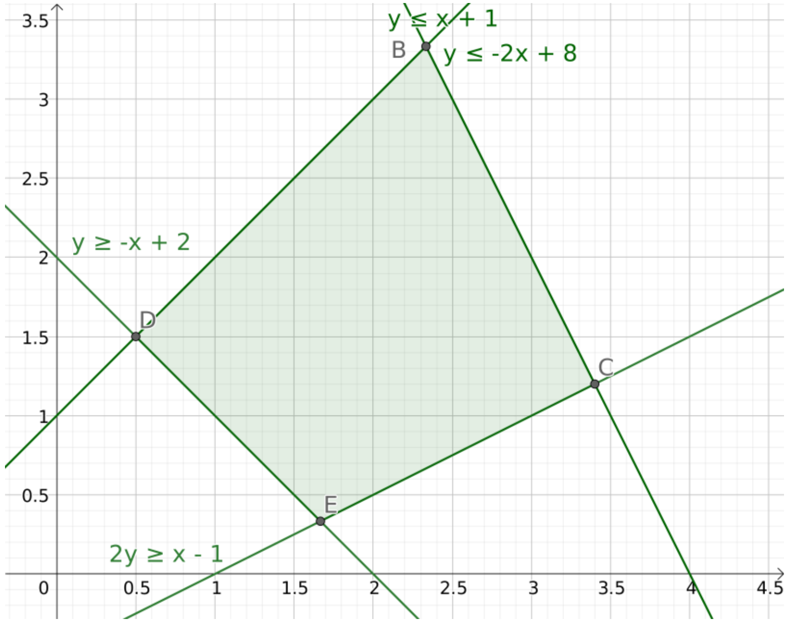
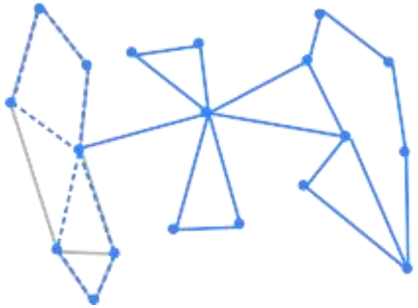
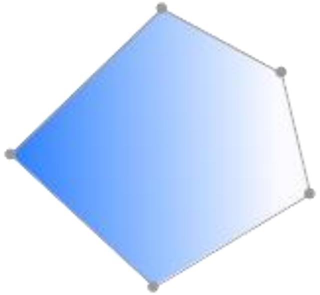
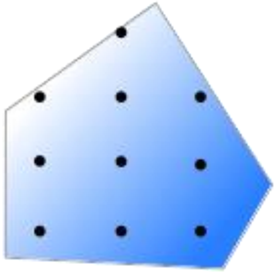
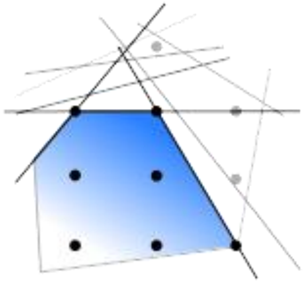
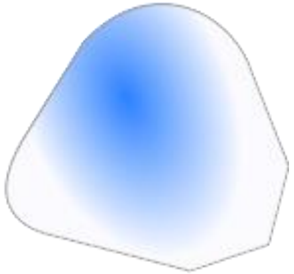
Stochastic Programming

COMP4691 / 8691



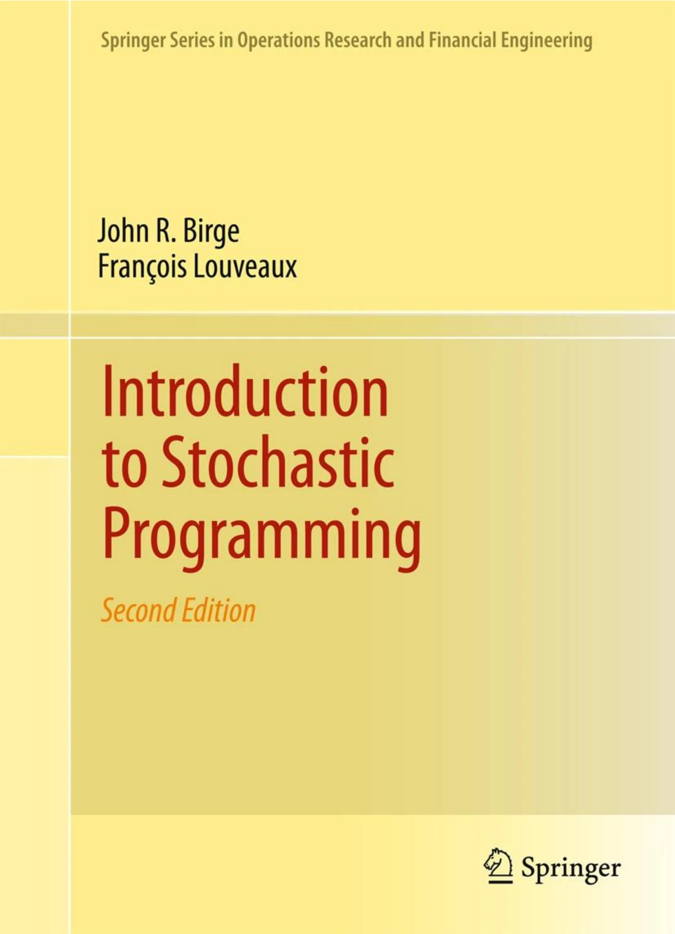


Outline

- Farmer’s example
- Two-Stage Stochastic Programming
- L-Shaped Method
- Chance Constraints
- Multi-Stage Stochastic Programming

Available as e-book and physical copy through the ANU library



Farmer’s Example

- A farmer has 500 acres and need to decide how to split the land for 3 different crops: wheat, corn and sugar beets
- Constraints
  - Land usage (500 acres)
  - At least 200 tons (T) of wheat and 240 T of corn for cattle
- Parameters per crop:
  - yield (T/acre)
  - planting cost ($/acre)
  - selling price ($/T)
  - purchase price ($/T)

  in case it is cheaper to buy wheat/corn instead of producing
- Extra constraint: cap on amount of sugar beets sold

Farmer’s Example – Data

	Wheat	Corn	Beets
Yield (T/acre)	2.5	3	20
Plating cost ($/acre)	150	230	260
Selling price ($/T)	170	150	36 under 6000 T 10 above 6000 T
Purchase price ($/T)	238	210	–
Min. requirement (T)	200	240	–

- All this values comes from historical data the farmer collected over the years

Farmer's Example – Variables

	Wheat	Corn	Beets
Acres allocated	x_1	x_2	x_3
Amount sold	w_1	w_2	w_3 for under 6000T w_4 for over 6000T
Amount purchased	y_1	y_2	–

- Goal: minimize loss (negative loss == profit)

Famer's Problem LP

$$\min 150x_1 + 230x_2 + 238x_3 + 238y_1 + 210y_2 - 170w_1 - 150w_2 - 36w_3 - 10w_4$$

$$\begin{aligned} \text{s.t.} \quad & x_1 + x_2 + x_3 \le 500 \\ & 2.5x_1 + y_1 - w_1 \ge 200 \\ & 3x_2 + y_2 - w_2 \ge 240 \\ & w_3 + w_4 \le 20x_3 \\ & w_3 \le 6000 \end{aligned}$$

x_i: land allocated
w_i: amount sold
y_i: amount purchased
wheat → 1
corn → 2
beets → 3 (up to quota)
4 (above)

- Recall: minimize loss == maximize profit (negative loss)

Solution

	Wheat	Corn	Beets
Acres allocated	120	80	300
Amount sold	300	240	6000 (exactly the quota)
Amount purchased	0	0	–

Total Profit: $118,600

- This is problem as described so far can be solved as a Knapsack problem over profitability
  - 460 $/acre for beets up to quota, 275 $/acre for wheat, 220 $/acre for corn, and -60 $/acre for beets after 6000 T
- What is the underlying issue with this model? Would you use it in your farm?

The Effect of the Weather

- Consider 2 scenarios: -20% and +20% change in yield due weather
- Opt. solution for each one of the cases and the previous average case:

	-20% yield			average yield			+20% yield		
	Wheat	**Corn**	**Beets**	**Wheat**	**Corn**	**Beets**	**Wheat**	**Corn**	**Beets**
Acres allocated	100	25	375	120	80	300	183	66	250
Amount sold	0	0	6000	300	240	6000	350	0	6000
Amount purchased	0	180	–	0	0	–	0	0	–
Total Profit			$59,950			$118,600			$167,667

- In each solution, we allocate enough to sell 6000 T of beets, then handle wheat and buy corn if needed
- Can we use these 3 solutions?

Two-stage Stochastic Program with Recourse

- The farmer's problem has:
  - **two-stages**: decide the land allocation (x) **then** we observe the weather/yield, and
  - **recourse**: we can buy the missing (y) wheat or corn if our production is below the requirements
- This is the class of problems we will focus in this lecture
- Expected profit if scenario has probability 1/3
  - Using the oracle: $115,406
  - Using the average case solution: $107,240
    - Under produces beets in the -20% case: $86,600
    - Over produces beets in the +20% case: $107,683
- Can we do better than using the average case in all weathers?

Stochastic Programming

- Let $g(x,\varepsilon)$ represent the Famer's problem for a yield ε
  - Here **ε is a random variable** representing the uncertainty in the yield
  - $P(\varepsilon = -20\% \text{ yield}) = 1/3$
- So far we have:

Planning with an Oracle		**Planning for the Future**		**Planning for the avg case**
$E_\varepsilon[\min_x g(x,\varepsilon)]$	$\leq$	$\min_x E_\varepsilon[g(x,\varepsilon)]$	$\leq$	$\min_x g(x,E_\varepsilon[\varepsilon])$

Other names:

- Wait-and-See (WS)
- Recourse-Problem (RP)
- Expected-Value Prob. (EV)

Recourse-Problem

- What does it mean $\min_x E_\varepsilon[g(x,\varepsilon)]$?
  - We need to **decide on x and then observe the uncertainty ε**
  - Nothing is preventing us from **planning for contingencies**
    - E.g., If ε is -20% yield, then and only then I will buy corn
- We exploit the recourse actions to find a good decision for x that we can fix it later if needed
- We do so by separating each scenario after we observe ε
  - y_i becomes y_{i1}, y_{i2}, y_{i3} for scenarios 1 (+20%), 2 (avg) and 3 (-20%)
  - same with w_i

Recourse-Problem LP

x_i: land allocated
w_{ik}: amount sold
y_{ik}: amount purchased
wheat → 1
corn → 2
beets → 3 (up to quota)
4 (above)
k → scenario index

$$\begin{aligned}\min\quad & 150x_1 + 230x_2 + 238x_3 \\ & + 1/3 * (238y_{11} + 210y_{21} - 170w_{11} - 150w_{21} - 36w_{31} - 10w_{41}) \\ & + 1/3 * (238y_{12} + 210y_{22} - 170w_{12} - 150w_{22} - 36w_{32} - 10w_{42}) \\ & + 1/3 * (238y_{13} + 210y_{23} - 170w_{13} - 150w_{23} - 36w_{33} - 10w_{43})\end{aligned}$$

s.t. $x_1 + x_2 + x_3 \leq 500$

Scenario 1 **(+20%)**

$3x_1 + y_{11} - w_{11} \geq 200$
$3.6x_2 + y_{21} - w_{21} \geq 240$
$w_{31} + w_{41} \leq 24x_3$
$w_{31} \leq 6000$

Scenario 2 **(avg)**

$2.5x_1 + y_{12} - w_{12} \geq 200$
$3x_2 + y_{22} - w_{22} \geq 240$
$w_{32} + w_{42} \leq 20x_3$
$w_{32} \leq 6000$

Scenario 1 **(-20%)**

$2x_1 + y_{13} - w_{13} \geq 200$
$2.4x_2 + y_{23} - w_{23} \geq 240$
$w_{33} + w_{43} \leq 16x_3$
$w_{33} \leq 6000$

Recourse-Problem Solution

		Wheat	Corn	Beets
First Stage (x)	**Acres allocated**	**170**	**80**	**250**
scenario 1 (+20% yield)	Yield (T)	510	288	6000
	Sold/Purchased (T)	310	48	6000
scenario 2 (avg yield)	Yield (T)	425	240	5000
	Sold/Purchased (T)	225	0	5000
scenario 3 (-20% yield)	Yield (T)	340	192	4000
	Sold/Purchased (T)	140	-48	4000

Total Profit: $108,390

- Key differences:
  - Allocate land for **beets to reach quota at** **best case**
  - Allocate land for corn to meet constraint in the average case
  - Left over land for wheat

Comparing Solutions

Wait-and-See (WS)		Recourse-Problem (RP)		Expected-Value Prob.
$E_\varepsilon[\min_x g(x,\varepsilon)]$	$\le$	$\min_x E_\varepsilon[g(x,\varepsilon)]$	$\le$	$\min_x g(x,E_\varepsilon[\varepsilon])$
-$115,406		-$108,390		-$107,240

- How much should we pay for a perfect prediction of the future?
  - WS - RP = -115,406 -(-108,390) = -$7,016
  - Known as the Expected Value of Perfect Information (EVPI)
- How much is Stochastic Programming helping us?
  - RP - E[EV] = -108,390 - (-107,240) = -$1,150
  - Known as the Value of the Stochastic Solution (VSS)
- Stochastic Programming leverages information about the distribution of the uncertain outcomes to improve the quality of the solution

Sampling

- What if
  1) we have a complicated or black-box model, e.g., weather forecast?
  2) we have a continuous distribution?
- **Sampling can solve both**
  - In some cases, (2) can be solved analytically
- The samples is treated as scenarios of equal probability
  - Referred as the **sample average approximation (SAA)**
- Better results with more samples
- Different sampling methods can also improve the solution:
  - Importance sampling
  - Quasi-Monte Carlo
  - Conditional Sampling

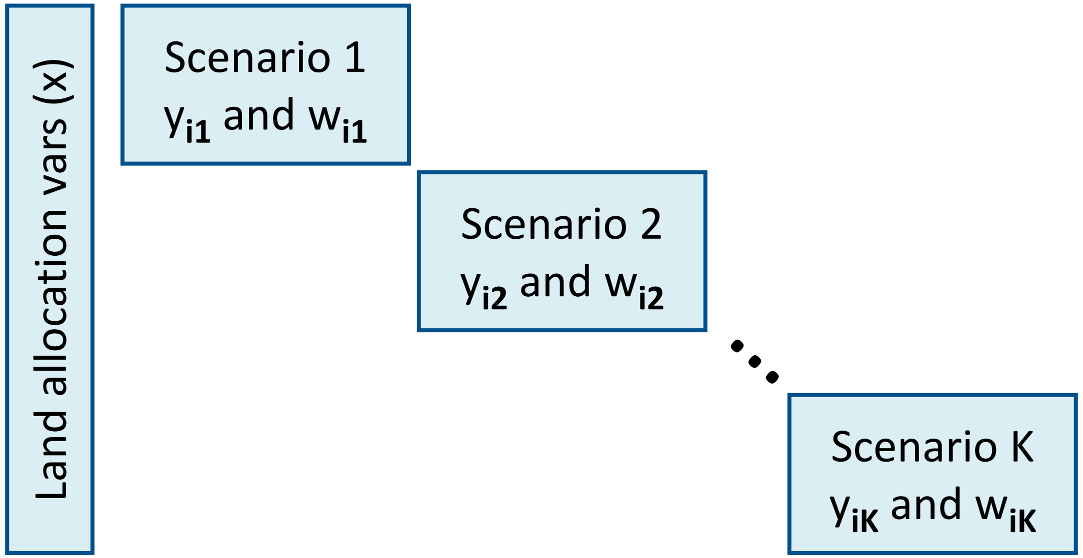
Evaluating Candidate Solutions

- More samples increase the size of the LP
- However, **evaluating a solution is much faster than solving the LP**:
  - Take the solutions of the first-stage x (e.g., land allocation)
  - Sample a scenario (e.g., -50% yield)
  - Compute the recourse-actions y (e.g., how corn and wheat to buy)
- Compute a solution using N samples and evaluate it on M different samples (M >> N)
- Using the Central Limit Theorem, we can get a confidence interval bound on the solution:
  - mean(g(x,ε)) ± z_α sem(g(x,ε))
    - sem is the standard error of the mean → sample standard dev/$\sqrt{n}$
    - 95% confidence interval for z_α = 1.96

Handling Large Problem

- If you have several scenarios or samples in the farmer's example, the LHS of our constraints will look like this:

x_i: land allocated for i
w_{ik}: amount sold
y_{ik}: amount purchased
k → scenario index
i → wheat, corn, beets



- Each scenario is only connected through the first-stage variables (x)
- Where have we seen this before?

L-Shaped Method (aka Benders Decomp)

- Use Benders Decomposition
  - Instead of one big sub-problem, we have K independent small subproblems
- From Benders Lecture adapted notation:

General Recourse Problem

$$\begin{aligned} \min_{x,z} \quad & f^\intercal x + \sum_{k=1}^{K} c_k^\intercal z_k \\ \text{s.t.} \quad & Ax = b \\ & B_k x + D z_k = d_k \quad \forall k \in \{1, \ldots, K\} \\ & x, z \geq 0 \end{aligned}$$

RMP

K sub-problems

Benders Reduced Master Problem

$$\begin{aligned} \min_{x,\eta} \quad & f^\intercal x + \eta \\ \text{s.t.} \quad & Ax = b \\ & \eta \geq \pi_e^\intercal (d_k - B_k x) \quad \forall k, \pi_e \in \bar{E}_k \\ & 0 \geq r_q^\intercal (d_k - B_k x) \quad \forall k, r_q \in \bar{Q}_k \\ & x \geq 0 \end{aligned}$$

(Primal) Sub-Problem k

$$\min_{z_k \geq 0} \{ c_k^\intercal z_k : D z_k = d_k - B_k \hat{x} \}$$

Complete Recourse

- A problem has **complete recourse** when, for all possible observations of the uncertainty ε, there is a recourse action that makes the problem feasible.
- This implies that all sub-problems in the Benders decomposition are feasible regardless of the value of x
- The farmer's example has complete recourse:
  1. Can by as much wheat and corn to satisfy constraint of at least 200 T and 240 T each
  2. **Everything produced can be sold**

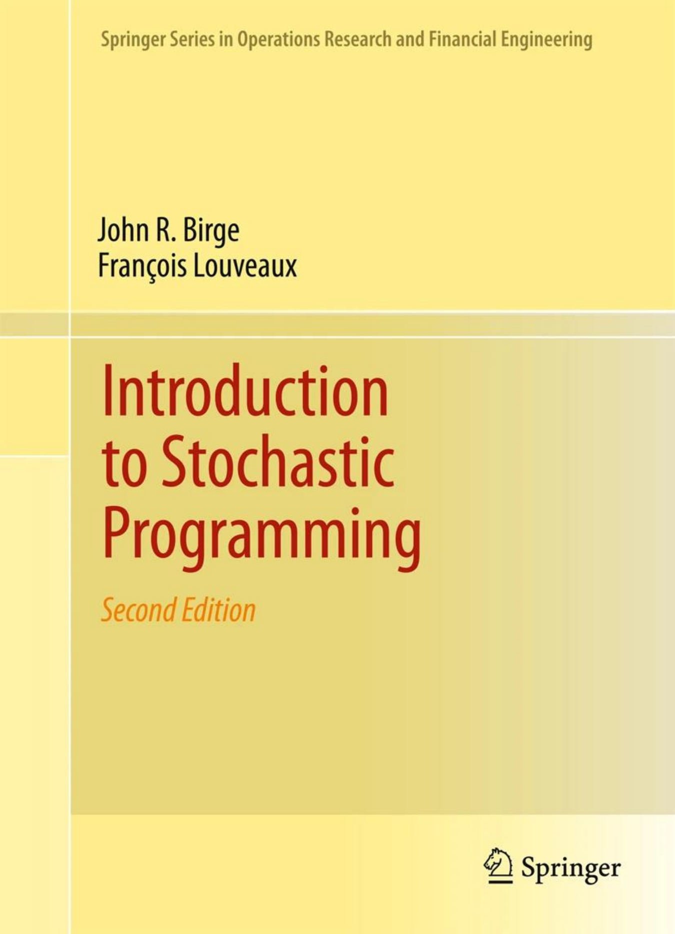
RMP for Complete Recourse Problems

$$
\begin{aligned}
\min_{x,\eta} \quad & f^\intercal x + \eta \\
\text{s.t.} \quad & Ax = b \\
& \eta \geq \pi_e^\intercal (d_k - B_k x) \quad \forall k, \pi_e \in \bar{E}_k \\
& x \geq 0
\end{aligned}
$$

Outline

- Farmer's example
- Two-Stage Stochastic Programming
- L-Shaped Method
- **Chance Constraints**
- Multi-Stage Stochastic Programming

Available as e-book and physical copy through the ANU library



Chance Constraints

- So far in the course, we have seen:
  - **(hard) constraints**: must be satisfied
  - **soft constraints**: penalize if not satisfied
- **Chance constraints**: a probabilistic constraint

$$P(a^T x \leq b) \geq \alpha$$

where either a or b depends on a random variable

- Famer's problem example:
  - P(producing less than 6000 T of beets) ≤ 0.25
  - P(buy 20 T or less of corn and wheat) ≥ 0.8

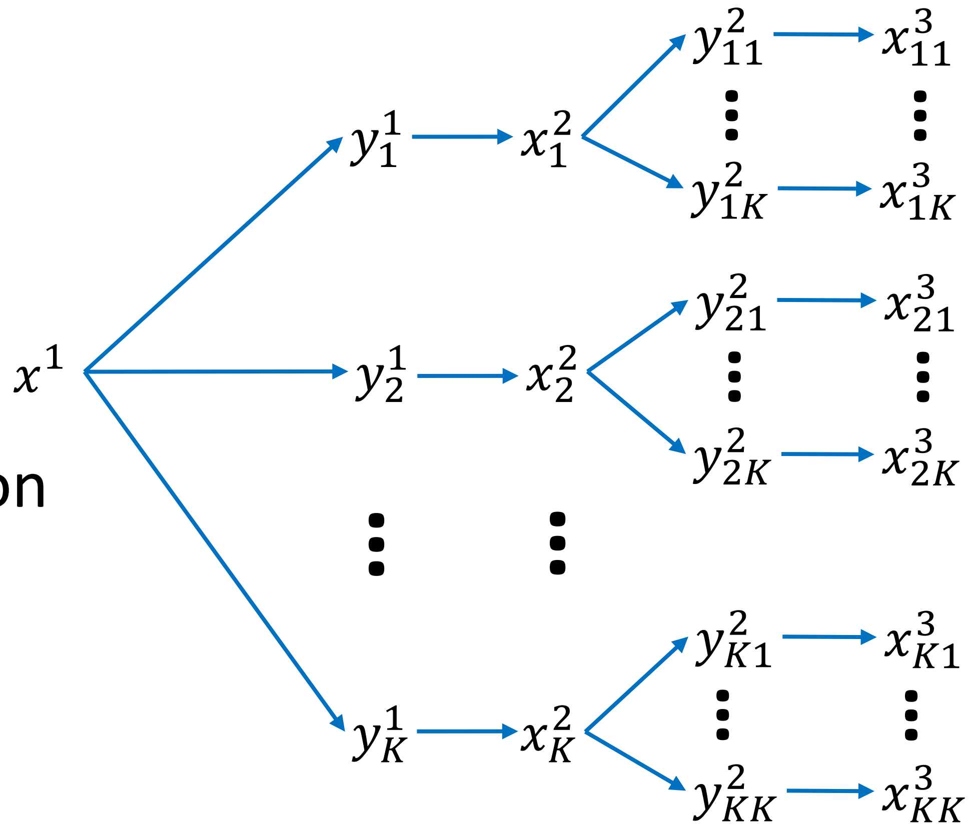
Modeling Chance Constraints

- For discrete distributions and sampling:
  1. use binary variables to count the constraint violations
  2. constraint the sum of scenario probability where violation occurred
- Famer's problem example: P(buy 20 T or less of corn and wheat) ≥ 0.8
  1. for each scenario k:

     $z_k \in \{0,1\}$: **constraint violated implies $z_k = 1$**
     - What is the maximum amount of corn and wheat needed?
       440 from the specification (200 wheat, 240 corn)
     - Use this tight upper bound for a Big-M constraint
       $\mathbf{y_{1k} + y_{2k} \leq 20 + 420z_k}$
  2. in the main problem: $\mathbf{\Sigma_k\, p_k z_k \leq 0.2}$
     - Note that we modeled the complement, i.e., 1 - P(buy 20T or less) ≤ 1 - 0.8

Multi-Stage Stochastic Programming

- Multi-stage is a series of two-stage problems:
  - Superscript denotes discrete time step
- In the farmer's example:
  - crop rotation: rotate field every year t
  - beets production quota over multiple seasons
- Issue: **curse of dimensionality**
  - exponential growth of scenarios wrt horizon
- Key techniques:
  - Nested Benders Decomposition
  - Better Sampling



Stochastic Programming

- More realistic decision making
  - Model uncertainty and the sequential decisions
- Can be used with any model: LP, MIPs, QPs, Convex Programs, etc.
  - There are special branch-and-bound techniques for it
- Successful in multiple industries
  - Used in Tasmania by energy operator
- Key challenges:
  - Curse of dimensionality / good sampling
  - Handling large problems

Available as e-book and physical copy through the ANU library