Introduction to Stochastic Optimization

Michel De Lara
Cermics, École des Ponts ParisTech
Université Paris-Est
France

École des Ponts ParisTech

November 16, 2014

Outline of the presentation

1. Working out classical examples
2. Framing stochastic optimization problems
3. Optimization with finite scenario space
4. Solving stochastic optimization problems by decomposition methods

Outline of the presentation

Working out classical examples

We will work out classical examples in Stochastic Optimization

- the blood-testing problem
  static, only risk
- the newsvendor problem
  static, only risk
- as a startup for stock management problems
  risk and time, with fixed information flow
- the secretary problem
  risk and time, with handleable information flow

Outline of the presentation

The blood-testing problem (R. Dorfman) is a static stochastic optimization problem

- A large number N of individuals are subject to a blood test
- The probability that the test is positive is p, the same for all people
- Individuals are stochastically independent
- The blood samples of k individuals are pooled and analyzed together
  - If the test is negative, this one test suffices for the k people
  - If the test is positive, each of the k persons must be tested separately, and $k + 1$ tests are required, in all
- Find the value of k which minimizes the expected number of tests
- Find the minimal expected number of tests

In army practice, R. Dorfman achieved savings up to 80%

- For the first pool $\{1, \dots, k\}$, the test is
  - negative with probability $(1-p)^k$ (by independence) $\to$ 1 test
  - positive with probability $1-(1-p)^k \to k+1$ tests
- When the pool size k is small compared to the number N of individuals, the blood samples $\{1, \dots, N\}$ are split in approximately N/k groups, so that the expected number of tests is

$$J(k) \approx \frac{N}{k}[(1-p)^k + (k+1)(1-(1-p)^k)]$$

- For small p, the optimal solution is $k^\star \approx 1/\sqrt{p}$
- The minimal expected number of tests is about $J^\star \approx 2N\sqrt{p} < N$
- William Feller reports that, in army practice,
  R. Dorfman achieved savings up to 80%, compared to making N tests
  (take $p = 1/100$, giving $k^\star \approx 10$ and $J^\star \approx N/5$)

Outline of the presentation

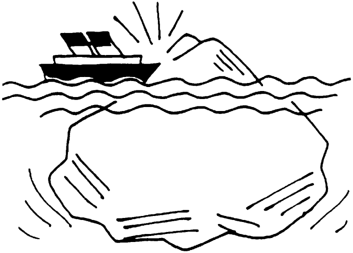
The (single-period) newsvendor problem stands as a classic in stochastic optimization

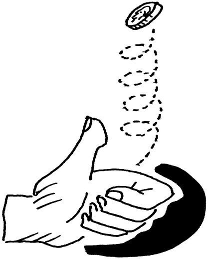
- Traditionally known under the terminology "news*boy* problem", it is now coined the "news*vendor* problem" ;-)
- Each morning, the newsvendor must decide how many copies $u \in \mathbb{U} = \{0, 1, \dots\}$ of the day's paper to order
- The newsvendor will meet an uncertain demand $w \in \mathbb{W} = \{0, 1, \dots\}$
- The newsvendor faces an economic tradeoff
  - she pays the unitary purchasing cost c per copy, when she orders stock
  - she sells a copy at price p
  - if she remains with an unsold copy, it is worthless (perishable good)
- Therefore, the newsvendor's profit is uncertain,

$$\texttt{Payoff}(u, w) = - \underbrace{cu}_{\text{purchasing}} + \underbrace{p \min\{u, w\}}_{\text{selling}}$$

because it depends on the uncertain demand w

For you, Nature is rather random or hostile?





The newsvendor reveals her attitude towards risk in how she aggregates profit with respect to uncertainty

We formulate a problem of profit maximization

- In the robust or pessimistic approach,
  the newsvendor maximizes the worst payoff

$$\max_{u \in \mathbb{U}} \underbrace{\min_{w \in \overline{\mathbb{W}}} \texttt{Payoff}(u, w)}_{\text{worst payoff}}$$

  as if Nature were malevolent

- In the stochastic or expected approach, the newsvendor solves

$$\max_{u \in \mathbb{U}} \underbrace{\mathbb{E}_w \big[\texttt{Payoff}(u, w) \big]}_{\text{expected payoff}}$$

  as if Nature played stochastically

If the newsvendor maximizes the worse profit

- We suppose that
  - the demand w belongs to a set $\overline{\mathbb{W}} = [\![w^\flat, w^\sharp]\!]$
  - the newsvendor knows the set $[\![w^\flat, w^\sharp]\!]$
- The worse profit is

$$J(u) = \min_{w \in [\![w^\flat, w^\sharp]\!]} [-cu + p\min\{u, w\}] = -cu + p\min\{u, w^\flat\}$$

- Show that the order $u^\star = w^\flat$ maximizes the above expression $J(u)$
- Once the newsvendor makes the optimal order $u^\star = w^\flat$, the optimal profit is $w \mapsto (p-c)w^\flat$ which, here, is no longer uncertain

If the newsvendor maximizes the expected profit

- We suppose that
  - the demand w is a random variable
  - the newsvendor knows the probability distribution $\mathbb{P}$ of w

$$\pi_0 = \mathbb{P}(w = 0), \ \pi_1 = \mathbb{P}(w = 1) \ \ldots$$

- The expected profit is

$$J(u) = \mathbb{E}_w[-cu + p \min\{u, w\}] = -cu + p\mathbb{E}[\min\{u, w\}]$$

- Find an order $u^\star$ which maximizes the above expression $J(u)$
  - by calculating $J(u+1) - J(u)$
  - then using the *decumulative distribution function* $d \mapsto \mathbb{P}(w > d)$

Here stand some steps of the computation

$$\begin{aligned} J(u) &= -cu + p\mathbb{E}[\min\{u, w\}] \\ \min\{u, w\} &= u\mathbf{1}_{u<w} + w\mathbf{1}_{u\geq w} \\ \min\{u+1, w\} &= (u+1)\mathbf{1}_{u+1\leq w} + w\mathbf{1}_{u+1>w} \\ &= (u+1)\mathbf{1}_{u<w} + w\mathbf{1}_{u\geq w} \\ \min\{u+1, w\} - \min\{u, w\} &= \mathbf{1}_{u<w} \end{aligned}$$

$$J(u+1) - J(u) = -c + p\mathbb{E}[\mathbf{1}_{u<w}] = -c + p\mathbb{P}(w > u) \downarrow \text{ with } u$$

▷ An optimal decision $u^\star$ satisfies

$$\mathbb{P}(w > u^\star) \approx \frac{c}{p} = \frac{\text{cost}}{\text{price}}$$

▷ Once the newsvendor makes the optimal order $u^\star$, the optimal profit is the random variable $w \mapsto -cu^\star + p\min\{u^\star, w\}$

Where do we stand after having worked out two examples?

- When you move from deterministic optimization to optimization under uncertainty, you come accross the issue of risk attitudes
- Risk attitudes materialize in the a priori knowledge on the uncertainties
  - either probabilistic/stochastic
    - independence and Bernoulli distributions in the blood test example
    - uncertain demand faced by the newsvendor modeled as a random variable
  - or set-membership
    - uncertain demand faced by the newsvendor modeled by a set

Where do we stand after having worked out two examples?

- ▷ When you move from deterministic optimization to optimization under uncertainty, you come accross the issue of risk attitudes
- ▷ Risk attitudes materialize in the a priori knowledge on the uncertainties
  - either probabilistic/stochastic
    - independence and Bernoulli distributions in the blood test example
    - uncertain demand faced by the newsvendor modeled as a random variable
  - or set-membership
    - uncertain demand faced by the newsvendor modeled by a set
- ▷ In addition, when you make a succession of decisions, you need to specify what you know (of the uncertainties) before each decision, and what you know before each decision may depend or not on your previous actions
- ▷ Let us turn to the inventory problem

Outline of the presentation

Inventory control dynamical model

Consider the control dynamical model

$$x(t+1) = x(t) + u(t) - w(t)$$

where

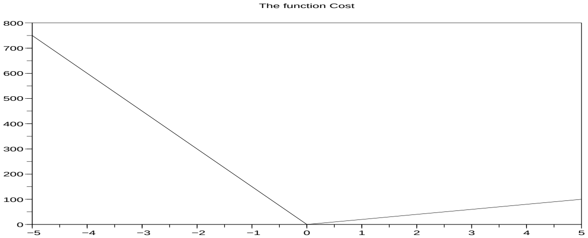
- time $t \in \{t_0, \ldots, T\}$ is discrete (days, weeks or months, etc.)
- $x(t)$ is the stock at the beginning of period t, belonging to $\mathbb{X} = \mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$
- $u(t)$ is the stock ordered at the beginning of period t, belonging to $\mathbb{U} = \mathbb{N} = \{0, 1, 2, \ldots\}$
- $w(t)$ is the uncertain demand during the period t, belonging to $\mathbb{W} = \mathbb{N}$

(When $x(t) < 0$, this corresponds to a *backlogged demand*, supposed to be filled immediately once inventory is again available)

Inventory optimization criterion

- The costs incurred in period t are
  - purchasing costs: $cu(t)$
  - shortage costs: $b \max\{0, -\big(x(t) + u(t) - w(t)\big)\}$
  - holding costs: $h \max\{0, x(t) + u(t) - w(t)\}$
- On the period from t_0 to T, the costs sum up to

$$\sum_{t=t_0}^{T-1} [\underbrace{cu(t)}_{\text{purchasing}} + \overbrace{\underbrace{b \max\{0, -\big(x(t) + u(t) - w(t)\big)\}}_{\text{shortage}} + \underbrace{h \max\{0, x(t) + u(t) - w(t)\}}_{\text{holding}}}^{\texttt{Cost}(x(t)+u(t)-w(t))}]$$



Probabilistic assumptions and risk neutral formulation of the inventory stochastic optimization problem

- We suppose that the sequence of demands $w(t_0), \ldots, w(T-1)$ is a stochastic process with distribution $\mathbb{P}$
- We consider the inventory sochastic optimization problem

$$\min_{u(\cdot)} \mathbb{E} \sum_{t=t_0}^{T-1} \big[c u(t) + \mathtt{Cost}\big(x(t) + u(t) - w(t) \big) \big]$$

Information flow and closed-loop formulation of the inventory stochastic optimization problem

- Let $u(\cdot) = u(t_0), \ldots, u(T-1)$ and consider

$$\underbrace{\min_{u(\cdot)}}_{\text{meaning what?}} \quad \mathbb{E} \sum_{t=t_0}^{T-1} \big[c u(t) + \mathtt{Cost}\big(x(t) + u(t) - w(t) \big) \big]$$

- The decision $u(t)$ at time t belongs to the control set $\mathbb{U}$
- $u(t)$ is a random variable, like are all demands $w(t_0), \ldots, w(T-1)$
- and like are all states $x(t)$ by the dynamics $x(t+1) = x(t) + u(t) - w(t)$

We express that the decision $u(t)$ at time t depends on the past $w(t_0), \ldots, w(t)$

$$u(t) \text{ is measurable w.r.t. } \underbrace{\big(w(t_0), \ldots, w(t)\big)}_{\text{past}}$$

Where do we stand?

- ▷ In addition to risk, we have to pay attention to the information flow
- ▷ When we make a succession of decisions, we need to specify what we know (of the uncertainties) before each decision, and this information may depend or not on our previous actions
- ▷ Let us now turn to the secretary problem

Outline of the presentation

The secretary problem stands as a classic optimal stopping problem

- A firm has opened a single secretarial position to fill (or a princess will only accept one "fiancé")
- Secretary applicants (Alice, Bob, Claire, etc.) can be compared by their absolute rank, corresponding to his/her quality for the position (Alice is 7, Bob is 15, Claire has top rank 1, etc.)
- The interviewer does not know the absolute rank
- The interviewer screens N applicants one-by-one in random order (Bob, then Claire, then Alice, etc.)
- The interviewer is able to rank the applicants interviewed so far (for the job, Claire is better than Alice, who is better than Bob, etc.)
- After each interview, the interviewer decides
  - either to select the applicant (and the process stops)
  - or to reject the applicant (and the process goes on), knowing that, once rejected, an applicant cannot be recalled

Here, a strategy is a stopping rule

- There are N applicants for the position
- The value of N is known
- A strategy provides the number $\nu \in \{1, \ldots, N\}$ of applicants interviewed, as a fonction of the relative ranking of the applicants interviewed so far
- A stopping time is a random variable ν, such that, for any $n = 1, \ldots, N$, the event $\{\nu = n\}$ depends at most upon what happened before interview n
- The interviewer maximizes the probability to select the best applicant, among all strategies

Open-loop strategies yield a probability $1/N$

- An open-loop strategy does not use the information collected up to applicant n, except for the clock n
- Therefore, for any $n = 1, \ldots, N$, the event $\{\nu = n\}$ depends only on n, and not on what happened before interview n
- Thus, an open-loop strategy is a deterministic stopping time ν
- For instance, $\nu = 1$ (constant stopping time) is an open-loop strategy: you select the first applicant
- If you adopt the strategy $\nu = 1$, the probability of selecting the best applicant is $1/N$
- For a fixed $k \in \{1, \ldots, N\}$, the strategy $\nu = k$ also yields probability $1/N$

The best closed loop strategy yields a probability $\approx 1/e$

- A candidate is an applicant who, when interviewed, is better than all the applicants interviewed previously
- For a fixed $k \in \{1, \ldots, N\}$, consider the strategy ν_k:
  - select the first candidate popping up after k applicants have been interviewed
  - or select the last applicant N in case no candidate appears
- We will now show that, when the number N of applicants is large, the best among the strategies ν_k, $k = 1, \ldots, N$, is achieved for

$$k^\star \approx \frac{N}{e} , \quad \text{the so-called 37\% rule}$$

- The probability of selecting the best applicant is $\approx 1/e$

$$\underbrace{\frac{1}{e}}_{\text{closed loop}} > \underbrace{\frac{1}{N}}_{\text{open loop}}$$

Here stand some steps of the computation (1)

We denote $p(k)$ the probability to select the best applicant with strategy ν_k

$$
\begin{aligned}
p(k) &= \sum_{m=k}^{n} \mathbb{P}(\text{ applicant m is selected } \mid \text{ applicant m is the best }) \\
&\qquad \times \mathbb{P}(\text{ applicant m is the best }) \\
&= \sum_{m=k}^{n} \mathbb{P}(\text{ applicant m is selected } \mid \text{ applicant m is the best }) \times \frac{1}{n}
\end{aligned}
$$

- If applicant m is the best applicant, then m is selected if and only if the best applicant among the first $m-1$ applicants is among the first $k-1$ applicants that were rejected
- Deduce that, when $m \geq k$,

$$
\mathbb{P}(\text{ applicant m is selected } \mid \text{ applicant m is the best }) = \frac{k-1}{m-1}
$$

Here stand some steps of the computation (2)

- Sum over $m \geq k$ and obtain

$$p(k) = \sum_{m=k}^{n} \frac{k-1}{m-1} \times \frac{1}{n} = \frac{k-1}{n} \sum_{m=k}^{n} \frac{1}{m-1}$$

- Compute the difference

$$\begin{aligned} n[p(k+1) - p(k)] &= \sum_{m=k+1}^{n} \frac{1}{m-1} - 1 \\ &= \sum_{m=k+1}^{n} \frac{1}{m-1} - 1 \\ &\approx \log n - \log k - 1 \\ &= \log(\frac{n}{ke}) \end{aligned}$$

The optimal strategy is called the 37% rule

- What is the $k^\star$ that maximizes $p(k)$? The 37% rule:

$$k^\star \approx \frac{N}{e} \text{ where } \log e = 1$$

- What is $p(k^\star)$ when N runs to $+\infty$?

$$p(k^\star) \approx \frac{1}{e} \approx 37\%$$

Where do we stand after having worked out the secretary problem?

- In a stopping time problem, as long as you do not stop, you collect information
- This information is valuable for forthcoming decisions
- For Markov decision problems, information is condensed in a state
- Stochastic control problems display trade-off between exploration and exploitation

Many decision problems illustrate the trade-off between exploration and exploitation

- ▷ deciding where to dig
- ▷ animal foraging
- ▷ job search
- ▷ devoting resources to research

The interplay between information and decision makes stochastic control problems especially tricky and difficult

- Decision $\to$ information $\to$ decision $\to$ information $\to \cdots$
- Decisions generally induce a dual effect,
  a terminology which tries to convey the idea that present decisions have two, often conflicting, effects or objectives:
  - directly contributing to optimizing the cost function,
    on the one hand
  - modifying the future information available for forthcoming decisions,
    on the other hand
- Problems with dual effect are among the most difficult decision-making problems

Summary

- Stochastic optimization = risk + information
- Risk is in the eyes of the beholder ;-)
- Information can be either revealed progressively
  - in a fixed way
  - or depending on past decisions
- Now, we turn to the mathematical framing of stochastic optimization problems

Outline of the presentation

Let us work out a toy example of economic dispatch as a cost-minimization problem under supply-demand balance

- Production: consider two energy production units
  - a "cheap" limited one with which we can produce quantity q_0, with $0 \leq q_0 \leq q_0^\sharp$, at cost $c_0 q_0$
  - an "expensive" unlimited one with which we can produce quantity q_1, with $0 \leq q_1$, at cost $c_1 q_1$, with $c_1 > c_0$
- Consumption: the demand is $D \geq 0$
- Balance: ensuring at least the demand

$$D \leq q_0 + q_1$$

- Optimization: total costs minimization

$$\min_{q_0, q_1} \underbrace{c_0 q_0 + c_1 q_1}_{\text{total costs}}$$

When the demand D is deterministic, the optimization problem is well posed

- The deterministic demand D is a single number, and we consider

$$\min_{q_0, q_1} c_0 q_0 + c_1 q_1$$

$$\text{under the constraints} \quad \begin{array}{rl} 0 & \leq q_0 \leq q_0^\sharp \\ 0 & \leq q_1 \\ D & \leq q_0 + q_1 \end{array}$$

- The solution is $q_0^\star = \min\{q_0^\sharp, D\}$, $\quad q_1^\star = [D - q_0^\sharp]_+$, that is,
  - if the demand D is below the capacity $q_0^\sharp$ of the "cheap" energy source

$$D \leq q_0^\sharp \Rightarrow q_0^\star = D \, , \quad q_1^\star = 0$$

  - if the demand D is above the capacity $q_0^\sharp$ of the "cheap" energy source,

$$D > q_0^\sharp \Rightarrow q_0^\star = q_0^\sharp \, , \quad q_1^\star = D - q_0^\sharp$$

- Now, what happens when the demand D is no longer deterministic?

If we know the demand beforehand, the optimization problem is deterministic

- We suppose that the demand is a random variable $D : \Omega \to \mathbb{R}_+$
- If we solve the problem for each possible value $D(\omega)$ of the random variable D, when $\omega \in \Omega$, we obtain

$$q_0(\omega) = \min\{q_0^\sharp, D(\omega)\}, \quad q_1(\omega) = [D(\omega) - q_0^\sharp]_+$$

  and we face an informational issue
- Indeed, we treat the demand D as if it were observed before making the decisions q_0 and q_1
- When the demand D is not observed, how can we do?

What happens if we replace the uncertain value D of the demand by its mean $\overline{D}$ in the deterministic solution?

- If we suppose that the demand D is a random variable $D : \Omega \to \mathbb{R}_+$, with mathematical expectation $\mathbb{E}(D) = \overline{D}$
- and that we propose the "deterministic solution"

$$q_0^{(\overline{D})} = \min\{q_0^\sharp, \overline{D}\} \ , \ \ q_1^{(\overline{D})} = [\overline{D} - q_0^\sharp]_+$$

- we cannot assure the inequality

$$\underbrace{D(\omega)}_{\text{uncertain}} \leq \underbrace{q_0 + q_1}_{\text{deterministic}} \ , \ \forall \omega \in \Omega$$

because $\max_{\omega \in \Omega} D(\omega) > \overline{D} = q_0^{(\overline{D})} + q_1^{(\overline{D})}$

- Are there better solutions among the deterministic ones?

When the demand D is bounded above, the robust optimization problem has a solution

- In the robust optimization problem, we minimize

$$\min_{q_0, q_1} c_0 q_0 + c_1 q_1$$

$$\text{under the constraints} \quad \begin{array}{rl} 0 & \leq q_0 \leq q_0^{\sharp} \\ 0 & \leq q_1 \\ D(\omega) & \leq q_0 + q_1 \ , \quad \forall \omega \in \Omega \end{array}$$

- When $D^{\sharp} = \max_{\omega \in \Omega} D(\omega) < +\infty$, the solution is $q_0^{\star} = \min\{q_0^{\sharp}, D^{\sharp}\} \ , \quad q_1^{\star} = [D^{\sharp} - q_0^{\sharp}]_{+}$
- Now, the total cost $c_0 q_0^{\star} + c_1 q_1^{\star}$ is an increasing function of the upper bound $D^{\sharp}$ of the demand
- Is it not too costly to optimize under the worst-case situation?

Where do we stand?

- When the demand D is deterministic, the optimization problem is well posed
- If we know the demand beforehand, the optimization problem is deterministic
- If we replace the uncertain value D of the demand by its mean $\overline{D}$ in the deterministic solution, we remain with a feasability issue
- When the demand D is bounded above, the robust optimization problem has a solution, but it is costly

Where do we stand?

- When the demand D is deterministic, the optimization problem is well posed
- If we know the demand beforehand, the optimization problem is deterministic
- If we replace the uncertain value D of the demand by its mean $\overline{D}$ in the deterministic solution, we remain with a feasability issue
- When the demand D is bounded above, the robust optimization problem has a solution, but it is costly

To overcome the above difficulties, we propose to introduce stages

$$\underbrace{D(\omega)}_{\text{uncertain}} \leq \underbrace{q_0}_{\text{deterministic}} + \underbrace{q_1(\omega)}_{\text{uncertain}}, \quad \forall \omega \in \Omega$$

- the decision q_0 is made before observing the demand $D(\omega)$
- the decision $q_1(\omega)$ is made after observing the demand $D(\omega)$

To overcome the above difficulties, we turn to stochastic optimization

- We suppose that the demand D is a random variable, and minimize

$$\min_{q_0, q_1} \mathbb{E}[c_0 q_0 + c_1 q_1]$$

$$\text{under the constraints} \quad \begin{array}{ll} 0 & \leq q_0 \leq q_0^{\sharp} \\ 0 & \leq q_1 \\ D & \leq q_0 + q_1 \\ q_1 & \text{depends upon } D \end{array}$$

and we emphasize two issues, new with respect to the deterministic case

- explicititing online information issue:
  the decision q_1 depends upon the random variable D
- expliciting risk attitudes:
  we aggregate the total costs with respect to all possible values by taking the expectation $\mathbb{E}[c_0 q_0 + c_1 q_1]$

Turning to stochastic optimization forces one to specify online information

- We suppose that the demand D is a random variable, and minimize

$$\min_{q_0, q_1} \mathbb{E}[c_0 q_0 + c_1 q_1]$$

$$\text{under the constraints} \quad \begin{array}{ll} 0 & \leq q_0 \leq q_0^\sharp \\ 0 & \leq q_1 \\ D & \leq q_0 + q_1 \\ q_1 & \text{depends upon } D \end{array}$$

- specifying that the decision q_1 depends upon the random variable D, whereas q_0 does not, forces to consider two stages and a so-called non-anticipativity constraint (more on that later)
  - first stage: q_0 does not depend upon the random variable D
  - second stage: q_1 depends upon the random variable D

Turning to stochastic optimization forces one to specify risk attitudes

▷ We suppose that the demand D is a random variable, and minimize

$$\min_{q_0, q_1} \mathbb{E}[c_0 q_0 + c_1 q_1]$$

$$\text{under the constraints} \quad \begin{array}{ll} 0 & \leq q_0 \leq q_0^\sharp \\ 0 & \leq q_1 \\ D & \leq q_0 + q_1 \\ q_1 & \text{depends upon } D \end{array}$$

▷ Now that q_1 depends upon the random variable D,
it is also a random variable, and so is the total cost $c_0 q_0 + c_1 q_1$;
therefore, we have to aggregate the total costs with respect to all possible values, and we chose to do it by taking the expectation $\mathbb{E}[c_0 q_0 + c_1 q_1]$

In the uncertain framework, two additional questions must be answered with respect to the deterministic case

Question (expliciting risk attitudes)

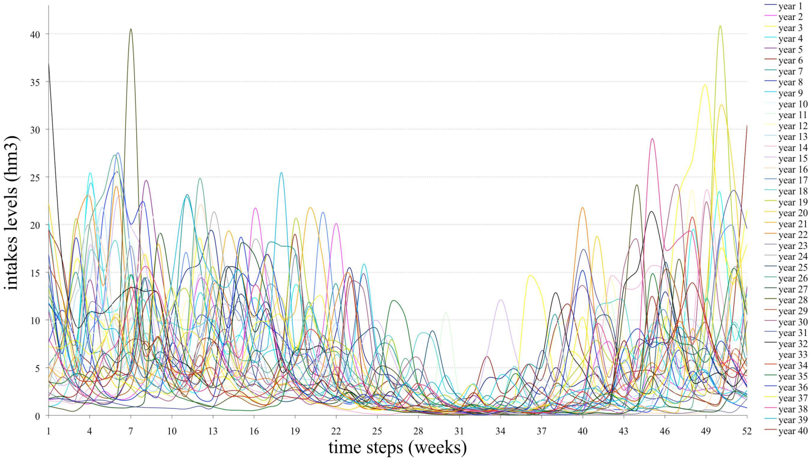
How are the uncertainties taken into account
in the payoff criterion and in the constraints?

Question (expliciting available online information)

Upon which online information are decisions made?

Outline of the presentation

Water inflows historical scenarios

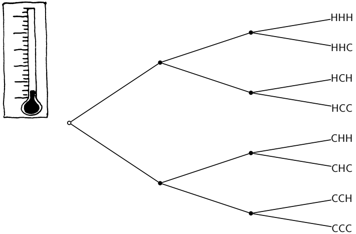


We call scenario a temporal sequence of uncertainties

Scenarios are special cases of "states of Nature"

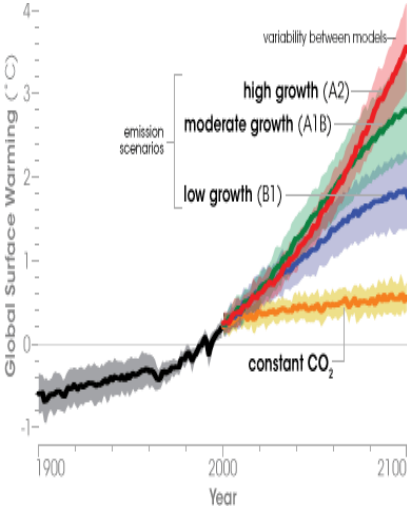
A scenario (pathway, chronicle) is a sequence of uncertainties

$$w(\cdot) := \big(w(t_0), \dots, w(T-1)\big) \in \Omega := \mathbb{W}^{T-t_0}$$



El tiempo se bifurca perpetuamente hacia innumerables futuros
(Jorge Luis Borges, *El jardín de senderos que se bifurcan*)

Beware! Scenario holds a different meaning in other scientific communities

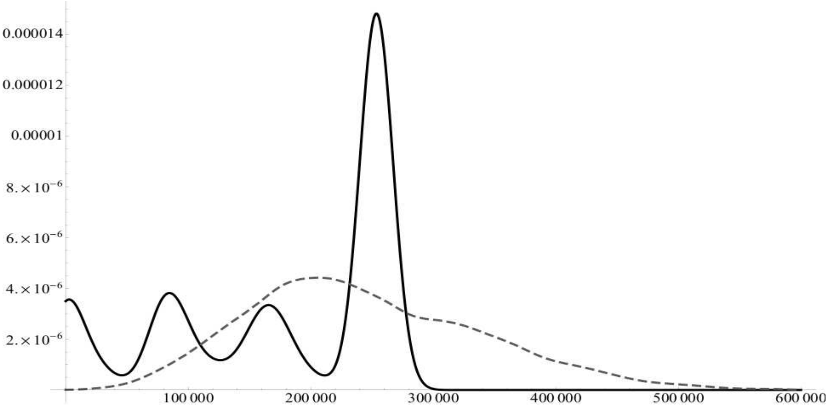


- In practice, what modelers call a "scenario" is a mixture of
  - a sequence of uncertain variables (also called a pathway, a chronicle)
  - a policy `Pol`
  - and even a static or dynamical model
- In what follows

scenario = pathway = chronicle

Outline of the presentation

The output of a stochastic optimization problem is a random variable. How can we rank random variables?



How are the uncertainties taken into account in the payoff criterion and in the constraints?

In a probabilistic setting, where uncertainties are random variables, a classical answer is

- to take the mathematical expectation of the payoff (risk-neutral approach)

$$\mathbb{E}(\text{payoff})$$

- and to satisfy all (physical) constrainsts almost surely that is, practically, for all possible issues of the uncertainties (robust approach)

$$\mathbb{P}(\text{constrainsts}) = 1$$

But there are many other ways to handle risk: robust, worst case, risk measures, in probability, almost surely, by penalization, etc.

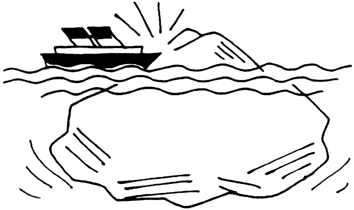
A policy and a criterion yield a real-valued payoff

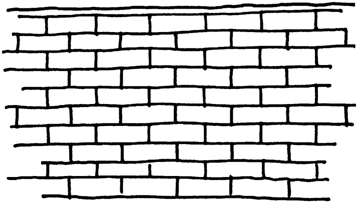
Given an admissible policy $\texttt{Pol} \in \mathcal{U}^{ad}$ and a scenario $w(\cdot) \in \Omega$,
we obtain a payoff

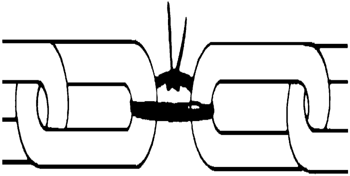
$$\texttt{Payoff}\big(\texttt{Pol}, w(\cdot)\big)$$

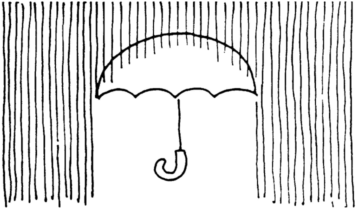
Policies/Scenarios	$w^A(\cdot) \in \Omega$	$w^B(\cdot) \in \Omega$	...
$\texttt{Pol}_1 \in \mathcal{U}^{ad}$	$\texttt{Payoff}\big(\texttt{Pol}_1, w^A(\cdot)\big)$	$\texttt{Payoff}\big(\texttt{Pol}_1, w^B(\cdot)\big)$	...
$\texttt{Pol}_2 \in \mathcal{U}^{ad}$	$\texttt{Payoff}\big(\texttt{Pol}_2, w^A(\cdot)\big)$	$\texttt{Payoff}\big(\texttt{Pol}_2, w^B(\cdot)\big)$	...
...	...	...	...

In the robust or pessimistic approach, Nature is supposed to be malevolent, and the DM aims at protection against all odds









In the robust or pessimistic approach, Nature is supposed to be malevolent

- In the robust approach, the DM considers the worst payoff

$$\underbrace{\min_{w(\cdot)\in\Omega} \mathtt{Payoff}\big(\mathtt{Pol}, w(\cdot)\big)}_{\text{worst payoff}}$$

- Nature is supposed to be malevolent,
  and specifically selects the worst scenario:
  the DM plays after Nature has played, and maximizes the worst payoff

$$\max_{\mathtt{Pol}\in\mathcal{U}^{ad}} \min_{w(\cdot)\in\Omega} \mathtt{Payoff}\big(\mathtt{Pol}, w(\cdot)\big)$$

- Robust, pessimistic, worst-case, maximin, minimax (for costs)

Guaranteed energy production

In a dam, the minimal energy production in a given period, corresponding to the worst water inflow scenario

The robust approach can be softened with plausibility weighting

- Let $\Theta : \Omega \to \mathbb{R} \cup \{-\infty\}$ be a a plausibility function.
- The higher, the more plausible:
  totally implausible scenarios are those for which $\Theta\big(w(\cdot)\big) = -\infty$
- Nature is malevolent, and specifically selects the worst scenario, but weighs it according to the plausibility function Θ
- The DM plays after Nature has played, and solves

$$\max_{\texttt{Pol} \in \mathfrak{U}^{ad}} \left[\min_{w(\cdot) \in \Omega} \left(\texttt{Payoff}\big(\texttt{Pol}, w(\cdot)\big) - \underbrace{\Theta\big(w(\cdot)\big)}_{\text{plausibility}} \right) \right]$$

In the optimistic approach, Nature is supposed to benevolent

Future. That period of time in which our affairs prosper, our friends are true and our happiness is assured.

Ambrose Bierce

- Instead of maximizing the worst payoff as in a robust approach, the optimistic focuses on the most favorable payoff

$$\underbrace{\max_{w(\cdot)\in\Omega} \texttt{Payoff}\big(\texttt{Pol}, w(\cdot)\big)}_{\text{best payoff}}$$

- Nature is supposed to benevolent, and specifically selects the best scenario: the DM plays after Nature has played, and solves

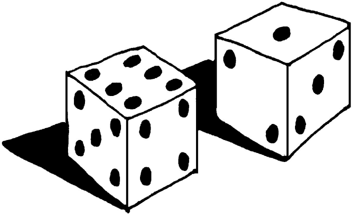
$$\max_{\texttt{Pol}\in\mathcal{U}^{ad}} \max_{w(\cdot)\in\Omega} \texttt{Payoff}\big(\texttt{Pol}, w(\cdot)\big)$$

The Hurwicz criterion reflects an intermediate attitude between optimistic and pessimistic approaches

A proportion $\alpha \in [0, 1]$ graduates the level of prudence

$$\max_{\texttt{Pol} \in \mathcal{U}^{ad}} \Big\{ \alpha \overbrace{\min_{w(\cdot) \in \Omega} \texttt{Payoff}\big(\texttt{Pol}, w(\cdot)\big)}^{\text{pessimistic}} + (1-\alpha) \underbrace{\max_{w(\cdot) \in \Omega} \texttt{Payoff}\big(\texttt{Pol}, w(\cdot)\big)}_{\text{optimistic}} \Big\}$$

In the stochastic or expected approach, Nature is supposed to play stochastically



In the stochastic or expected approach, Nature is supposed to play stochastically

- The expected payoff is

$$\overbrace{\mathbb{E}\Big[\texttt{Payoff}\big(\texttt{Pol}, w(\cdot)\big)\Big]}^{\text{mean payoff}} = \sum_{w(\cdot)\in\Omega} \mathbb{P}\{w(\cdot)\}\texttt{Payoff}\big(\texttt{Pol}, w(\cdot)\big)$$

- Nature is supposed to play stochastically, according to distribution $\mathbb{P}$: the DM plays after Nature has played, and solves

$$\max_{\texttt{Pol}\in\mathfrak{U}^{ad}} \mathbb{E}\Big[\texttt{Payoff}\big(\texttt{Pol}, w(\cdot)\big)\Big]$$

- The discounted expected utility is the special case

$$\mathbb{E}\Big[\sum_{t=t_0}^{+\infty} \delta^{t-t_0} \texttt{L}\big(x(t), u(t), w(t)\big)\Big]$$

The expected utility approach distorts payoffs before taking the expectation

- We consider a utility function $\mathtt{L}$ to assess the utility of the payoffs (for instance a CARA exponential utility function)
- The expected utility is

$$\underbrace{\mathbb{E}\left[\mathtt{L}\Big(\mathtt{Payoff}\big(\mathtt{Pol}, w(\cdot)\big)\Big)\right]}_{\text{expected utility}} = \sum_{w(\cdot)\in\Omega} \mathbb{P}\{w(\cdot)\}\mathtt{L}\Big(\mathtt{Payoff}\big(\mathtt{Pol}, w(\cdot)\big)\Big)$$

- The expected utility maximizer solves

$$\max_{\mathtt{Pol}\in\mathcal{U}^{ad}} \mathbb{E}\left[\mathtt{L}\Big(\mathtt{Payoff}\big(\mathtt{Pol}, w(\cdot)\big)\Big)\right]$$

The ambiguity or multi-prior approach combines robust and expected criterion

- Different probabilities $\mathbb{P}$, termed as beliefs or priors and belonging to a set $\mathcal{P}$ of admissible probabilities on Ω
- The multi-prior approach combines robust and expected criterion by taking the worst beliefs in terms of expected payoff

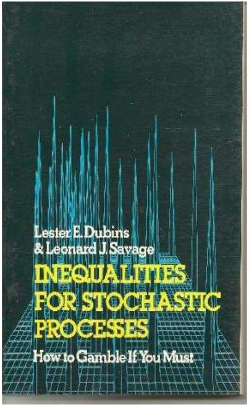
$$\max_{\texttt{Pol} \in \mathfrak{U}^{ad}} \underbrace{\min_{\mathbb{P} \in \mathcal{P}} \overbrace{\mathbb{E}^{\mathbb{P}} \Big[\texttt{Payoff}\big(\texttt{Pol}, w(\cdot)\big) \Big]}^{\text{mean payoff}}}_{\text{pessimistic over probabilities}}$$

Convex risk measures cover a wide range of risk criteria

- Different probabilities $\mathbb{P}$, termed as beliefs or priors and belonging to a set $\mathcal{P}$ of admissible probabilities on Ω
- To each probability $\mathbb{P}$ is attached a plausibility $\Theta(\mathbb{P})$

$$\max_{\texttt{Pol}\in\mathcal{U}^{ad}} \underbrace{\min_{\mathbb{P}\in\mathcal{P}} \overbrace{\mathbb{E}^{\mathbb{P}}\Big[\texttt{Payoff}\big(\texttt{Pol}, w(\cdot)\big)\Big]}^{\text{mean payoff}} - \overbrace{\Theta(\mathbb{P})}^{\text{plausibility}}}_{\text{pessimistic over probabilities}}$$

Non convex risk measures can lead to non diversification



How to gamble if you must,
L.E. Dubbins and L.J. Savage,
1965

Imagine yourself at a casino with \$1,000. For some reason, you desperately need \$10,000 by morning; anything less is worth nothing for your purpose.

The only thing possible is to gamble away your last cent, if need be, in an attempt to reach the target sum of \$10,000.

- ▷ The question is how to play, not whether. What ought you do? How should you play?
  - ▹ Diversify, by playing 1 \$ at a time?
  - ▹ Play boldly and concentrate, by playing 10,000 \$ only one time?
- ▷ What is your decision criterion?

Savage's minimal regret criterion... "Had I known"

$$\min_{\texttt{Pol} \in \mathfrak{U}^{ad}} \Big\{ \overbrace{\max_{w(\cdot)\in\Omega} \underbrace{\Big[\max_{\text{anticipative policies } \overline{\texttt{Pol}}} \texttt{Payoff}\big(\overline{\texttt{Pol}}, w(\cdot)\big) - \texttt{Payoff}\big(\texttt{Pol}, w(\cdot)\big) \Big]}_{\text{regret}}}^{\text{worst regret}} \Big\}$$

- If the DM knows the future in advance, she solves $\max_{\text{anticipative policies } \overline{\texttt{Pol}}} \texttt{Payoff}\big(\overline{\texttt{Pol}}, w(\cdot)\big)$, for each scenario $w(\cdot) \in \Omega$
- The regret attached to a non-anticipative policy $\texttt{Pol} \in \mathfrak{U}^{ad}$ is the loss due to not being visionary
- The best a non-visionary DM can do with respect to regret is minimizing it

Outline of the presentation

Upon which online information are decisions made?

We navigate between two stumbling blocks: rigidity and wizardry

- On the one hand, it is suboptimal to restrict oneself, as in the deterministic case, to open-loop controls depending only upon time, thereby ignoring the available information at the moment of making a decision
- On the other hand, it is impossible to suppose that we know in advance what will happen for all times: clairvoyance is impossible as well as look-ahead solutions

The in-between is **non-anticipativity constraint**

There are two ways to express the non-anticipativity constraint

Denote the uncertainties at time t by $w(t)$, and the control by $u(t)$

- ▷ Functional approach
  The control $u(t)$ may be looked after under the form

$$u(t) = \phi_t\big(\underbrace{w(t_0), \ldots, w(t-1)}_{\text{past}}\big)$$

where ϕ_t is a function, called policy, strategy or decision rule

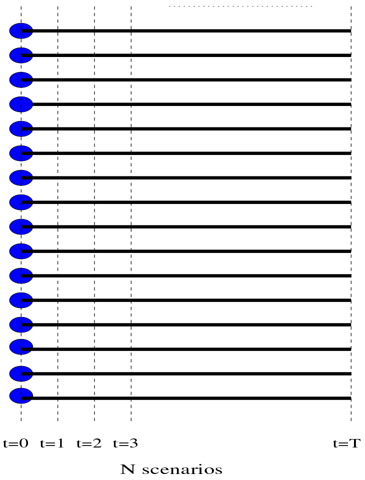
- ▷ Algebraic approach
  When uncertainties are considered as random variables (measurable mappings), the above formula for $u(t)$ expresses the measurability of the control variable $u(t)$ with respect to the past uncertainties, also written as

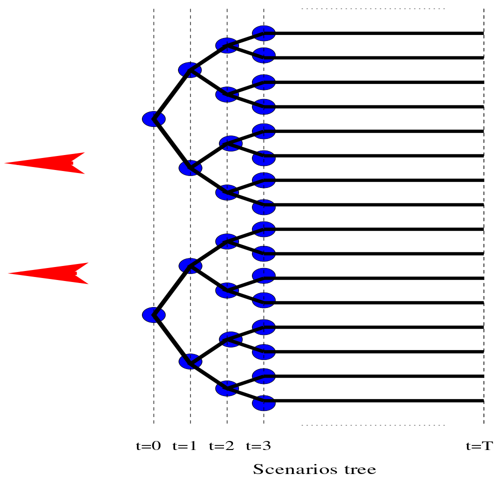
$$\underbrace{\sigma(u(t))}_{\sigma\text{-algebra}} \subset \sigma\big(\underbrace{w(t_0), \ldots, w(t-1)}_{\text{past}}\big)$$

What is a solution at time t?

- In deterministic control, the solution $u(t)$ at time t is a single vector
- In stochastic control, the solution $u(t)$ at time t is a random variable expressed
  - either as $u(t) = \phi_t\big(w(t_0), \ldots, w(t-1)\big)$, where $\phi_t : \mathbb{W}^{t-t_0} \to \mathbb{R}$
  - or as $u(t) : \Omega \to \mathbb{R}$ with measurability constraint $\sigma(u(t)) \subset \sigma\big(w(t_0), \ldots, w(t-1)\big)$ or
    $$u(t) = \mathbb{E}\Big(u(t) \;\Big|\; w(t_0), \ldots, w(t-1)\Big)$$
- Now, as time t goes on, the domain of the function ϕ_t expands, and so do the conditions $\sigma(u(t)) \subset \sigma\big(w(t_0), \ldots, w(t-1)\big)$
- Therefore, for numerical reasons, the information $\big(w(t_0), \ldots, w(t-1)\big)$ has to be compressed or approximated

Scenarios can be organized like a comb or like a tree





There are two classical ways to compress information

- State-based functional approach
  In the special case of the Markovian framework with $\big(w(t_0), \dots, w(T)\big)$ white noise, there is no loss of optimality to look for solutions as

$$u(t) = \psi_t \underbrace{\big(x(t)\big)}_{\text{state}} \quad \text{where} \quad \underbrace{x(t) \in \mathbb{X}}_{\text{fixed space}}, \quad \underbrace{x(t+1) = F_t(x(t), u(t), w(t))}_{\text{dynamical equation}}$$

- Scenario-based measurability approach
  Scenarios are approximated by a finite family $\big(w^s(t_0), \dots, w^s(T)\big)$, $s \in S$
  - Either solutions $u^s(t)$ are indexed by $s \in S$ with the constraint that

$$\big(w^s(t_0), \dots, w^{s'}(t-1)\big) = \big(w^{s'}(t_0), \dots, w^{s'}(t-1)\big) \Rightarrow u^s(t) = u^{s'}(t)$$

  - Or — in the case of the scenario tree approach, where the scenarios $\big(w^s(t_0), \dots, w^s(T)\big)$, $s \in S$, are organized in a tree — solutions $u^n(t)$ are indexed by nodes n on the tree

More on what is a solution at time t State-based approach $u(t) = \psi_t(x(t))$

- ▷ The mapping ψ_t can be computed in advance (that is, at initial time t_0) and evaluated at time t on the available online information at that time t
  - ▷ either exactly (for example, by dynamic programming)
  - ▷ or approximately (for example, among linear decision rules) because the computational burden of finding *any* function is heavy
- ▷ The value $u(t) = \psi_t(x(t))$ can be computed at time t
  - ▷ either exactly by solving a proper optimization problem, which raises issues of dynamic consistency
  - ▷ or approximately (for example, by assuming that controls from time t on are open-loop)

More on what is a solution at time t Scenario-based approach

- An optimal "solution" can be computed scenario by scenario, with the problem that we obtain solutions such that

$$\big(w^s(t_0), \dots, w^s(t-1)\big) = \big(w^{s'}(t_0), \dots, w^{s'}(t-1)\big) \text{ and } u^s(t) \neq u^{s'}(t)$$

- Optimal solutions can be computed scenario by scenario and then merged (for example, by Progressive Hedging) to be forced to satisfy

$$\big(w^s(t_0), \dots, w^s(t-1)\big) = \big(w^{s'}(t_0), \dots, w^{s'}(t-1)\big) \Rightarrow u^s(t) = u^{s'}(t)$$

- The value $u(t)$ can be computed at time t depending on $\big(w^s(t_0), \dots, w^s(t-1)\big)$
  - either exactly by solving a proper optimization problem, which raises issues of dynamic consistency
  - or approximately (for example, by a sequence of two-stages problems)

Outline of the presentation

Where do we stand?

- ▷ How one frames the non-anticipativity constraint impacts numerical resolution methods
- ▷ On a finite scenario space, one obtains large (deterministic) optimization problems either on a tree or on a comb
- ▷ Else, one resorts to state-based formulations, with solutions as policies (dynamic programming)

Optimization approaches to attack complexity

Linear programming

- linear equations and inequalities
- no curse of dimension

Stochastic programming

- no special treatment of time and uncertainties
- no independence assumption
- decisions are indexed by a scenario tree
- what if information is not a node in the tree?

State-based dynamic optimization

- nonlinear equations and inequalities
- curse of dimensionality
- independence assumption on uncertainties
- special treatment of time (dynamic programming equation)
- decisions are indexed by an information state (feedback synthesis)
- an information state summarizes past controls and uncertainties
- decomposition-coordination methods to overcome the curse of dimensionality?

Summary

- *Stochastic* optimization highlights risk attitudes tackling
- Stochastic *dynamic* optimization emphasizes the handling of online information
- Many issues are raised, because
  - many ways to represent risk (criterion, constraints)
  - many information structures
  - tremendous numerical obstacles to overcome
- Each method has its numerical wall
  - in dynamic programming, the bottleneck is the dimension of the state (no more than 3)
  - in stochastic programming, the bottleneck is the number of stages (no more than 2)

Outline of the presentation

From linear to stochastic programming

- ▷ The linear program

$$\begin{aligned} \min \langle c \,, x \rangle & \\ x & \geq 0 \\ Ax + b & \geq 0 \end{aligned}$$

- ▷ becomes a stochastic program

$$\begin{aligned} \min \mathbb{E}(\langle c(\boldsymbol{\xi}) \,, x \rangle) & \\ x & \geq 0 \\ A(\boldsymbol{\xi})x + b(\boldsymbol{\xi}) & \geq 0 \end{aligned}$$

  where $\boldsymbol{\xi} : \Omega \to \Xi$ is a finite random variable

- ▷ so that there are as many inequalities as there are possible values for $\boldsymbol{\xi}$

$$A\big(\boldsymbol{\xi}(\omega)\big)x + b\big(\boldsymbol{\xi}(\omega)\big) \geq 0 \,, \ \ \forall \omega \in \Omega$$

  and these inequality constraints may define an empty domain for optimization

Recourse variables need be introduced for feasability issues

- We denote by $\xi \in \Xi$ any possible value of the random variable $\boldsymbol{\xi}$
- and we introduce a recourse variable $y = (y(\xi), \xi \in \Xi)$ and the program

$$\min \sum_{\xi \in \Xi} \mathbb{P}\{\xi\} \Big(\langle c(\xi) \, , x \rangle + \langle p(\xi) \, , y(\xi) \rangle \Big)$$

$$\begin{aligned} x &\geq 0 \\ y(\xi) &\geq 0 \, , \quad \forall \xi \in \Xi \\ A(\xi)x + b(\xi) - y(\xi) &\geq 0 \, , \quad \forall \xi \in \Xi \end{aligned}$$

- so that the inequality $A(\xi)x + b(\xi) - y(\xi) \geq 0$ is now possible, at (unitary recourse) price vector $p = (p(\xi), \xi \in \Xi)$
- As there are as many inequalities $A(\xi)x + b(\xi) - y(\xi) \geq 0$ as there are possible values for ξ, hence stochastic programs are huge problems, but can remain linear

Two-step stochastic programs with recourse can become deterministic non-smooth convex problems

- Define

$$Q(\xi, x) = \min\{\langle p(\xi) \, , y \rangle \, , A(\xi)x + b(\xi) - y \geq 0\}$$

  which is a convex function of x, non-smooth

- so that the original two-step stochastic program with recourse

$$\min \sum_{\xi \in \Xi} \mathbb{P}\{\xi\} \langle c(\xi) \, , x \rangle + \langle p(\xi) \, , y(\xi) \rangle$$

$$\begin{aligned} x &\geq 0 \\ y(\xi) &\geq 0 \, , \quad \forall \xi \in \Xi \\ A(\xi)x + b(\xi) - y(\xi) &\geq 0 \, , \quad \forall \xi \in \Xi \end{aligned}$$

- now becomes the deterministic non-smooth convex problem

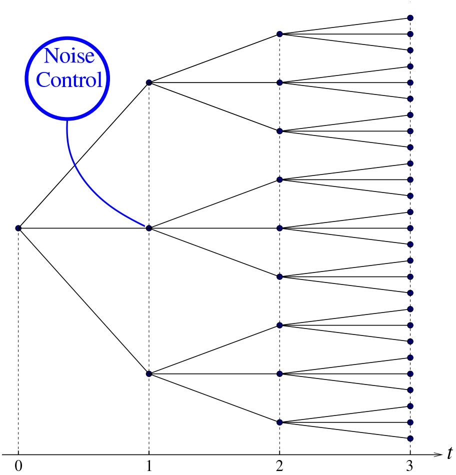
$$\min \langle c \, , x \rangle + \sum_{\xi \in \Xi} \mathbb{P}\{\xi\} Q(\xi, x)$$

$$x \geq 0$$

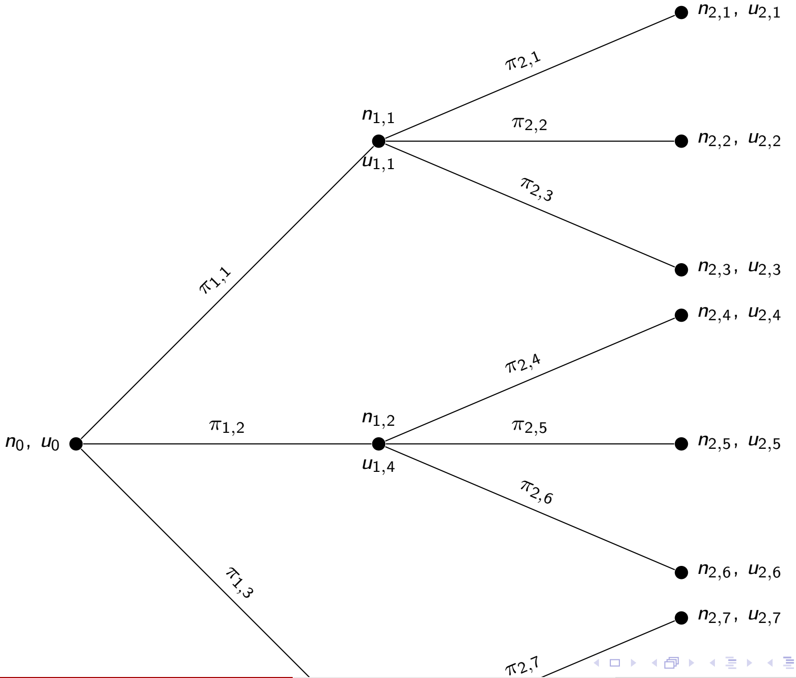
Roger Wets example

http://cermics.enpc.fr/~delara/ENSEIGNEMENT/
CEA-EDF-INRIA_2012/Roger_Wets1.pdf

Solutions of multi-stage stochastic optimization problems, without dual effect, can be indexed by a tree



- Conditional probabilities given on the arcs, probabilities on the leafs
- Solutions indexed by the nodes of the tree



$n_{2,1}$, $u_{2,1}$
$\pi_{2,1}$
$n_{1,1}$
$\pi_{2,2}$
$n_{2,2}$, $u_{2,2}$
$u_{1,1}$
$\pi_{2,3}$
$\pi_{1,1}$
$n_{2,3}$, $u_{2,3}$
$n_{2,4}$, $u_{2,4}$
$\pi_{2,4}$
$n_{1,2}$
$\pi_{1,2}$
$\pi_{2,5}$
n_0, u_0
$n_{2,5}$, $u_{2,5}$
$u_{1,4}$
$\pi_{2,6}$
$\pi_{1,3}$
$n_{2,6}$, $u_{2,6}$
$n_{2,7}$, $u_{2,7}$
$\pi_{2,7}$

Outline of the presentation

1. Working out classical examples
   - The blood-testing problem
   - The newsvendor problem
   - The inventory problem
   - The secretary problem
2. Framing stochastic optimization problems
   - Working out a toy example
   - Scenarios are temporal sequence of uncertainties
   - Expliciting risk attitudes
   - Handling online information
   - Discussing framing and resolution methods
3. Optimization with finite scenario space
4. Solving stochastic optimization problems by decomposition methods
   - A bird's eye view of decomposition methods
   - Progressive Hedging
   - Dynamic Programming

Outline of the presentation

Outline of the presentation

1. Working out classical examples
   - The blood-testing problem
   - The newsvendor problem
   - The inventory problem
   - The secretary problem
2. Framing stochastic optimization problems
   - Working out a toy example
   - Scenarios are temporal sequence of uncertainties
   - Expliciting risk attitudes
   - Handling online information
   - Discussing framing and resolution methods
3. Optimization with finite scenario space
4. Solving stochastic optimization problems by decomposition methods
   - A bird's eye view of decomposition methods
   - Progressive Hedging
   - Dynamic Programming